CSCE 2023 BOOK of ABSTRACTS

The 2023 World Congress in Computer Science, Computer Engineering, and Applied Computing
CSCE 2023

https://www.american-cse.org/csce2023/
July 24-27, 2023
Luxor Hotel (MGM Property), 3900 Las Vegas Blvd. South, Las Vegas, 89109, USA

Table of Contents

Note that the titles of papers, the authors' names and the abstracts that appear in this book ("Book of Abstracts") were extracted from the papers that were submitted to the EVALUATION web portal (i.e., extracted from the first draft submissions). The official published proceedings/book will have any and all changes/revisions that authors may have done.

KEYNOTE ADDRESSES

(The Keynote lectures are open to all participants)

CONGRESS WELCOME REMARKS

Professor Emeritus Hamid R. Arabnia
(Chair, Steering Committee & Coordinator), School of Computing, University of Georgia, USA;
Editor-in-Chief, The Journal of Supercomputing (Springer);
Editor, Transactions of Computational Science & Computational Intelligence (Springer Nature)
Fellow, Center of Excellence in Terrorism, Resilience, Intelligence & Organized Crime Research (CENTRIC)

From Digital Twins to the Metaverse: Exploring the Evolution of Responsive Systems and Social Interaction in Industry

Prof. Diego Galar, PhD
Professor of Condition Monitoring, Division of Operation and Maintenance Engineering,
Lulea University of Technology, Lulea, Sweden

Abstract - A digital twin is not a static model but a responsive system that connects physical and digital systems. It has numerous potential applications in industry, and is also known as a digital shadow, digital mirror, digital model, or digital avatar. However, industries require a more attractive proposition than just a one-to-one digital twinning process. They demand the creation of a virtual scenario where virtual instances gather together with the twins of real ones. We call this scenario the 'metaverse' - a digital replica of our reality where all physical assets are twinned along with entities that only exist in the digital dimension. The metaverse is the digital dimension where the digital entities interact, and is the expansion of digital twins with more content and social meaning. Importantly, the metaverse has a scalable environment that can accommodate many entities, thus reinforcing social meaning among the digital twins coexisting in the space and participating in it. It is crucial for maintainers to keep an eye on the evolution of digital twins and how they transfer those replicas to the metaverse. This is because degradation mechanisms, maintenance plans, and prognostics will seriously affect the digital twin once it starts interacting in the digital arena with its counterparts.

Artificial Intelligence and Chatbots

Dr. James A. Crowder
Systems Fellow, CAES Advanced Program Development, USA

Abstract - There has been much talk over the last few years about the perils of the use of artificial intelligence in virtually everything we touch. From our phones to our cars, and everything in between, artificial intelligence is an integral part of our existence. Many prominent people, like Elon Musk and Stephen Hawking, have warned about the potential for machines to take over and cause havoc in the lives and very existence of humans. Hollywood has made untold billions of dollars painting doom-and-gloom scenarios about artificial intelligence and robots within our society today and in the future. But what is the true reality? We continually push to create increasingly intelligent online entities (Chatbots) that seemingly learn, think, and reason like humans. Most attempts fail miserably. Given the recent "issues" with AI Chatbots like ChatGPT^{Ò} we enter a brave new world of interactions with AI-like entities. These AI Chatbots have little real cognitive functions and do not understand human emotions. Here we will discuss the issues associated with the proliferation of online Chatbots and their effects on society and human relations, both good and bad

Challenges of Integrating ChatGPT with Software Applications and Solving it with Object Messaging Model as a New Design Paradigm

Massoud Alibakhsh
Founder & CEO
Xeba Technologies
Atlanta Georgia, USA
www.xeba.tech

Abstract - The proliferation of AI tools, such as ChatGPT and other Large Language Models (LLMs), has raised expectations regarding their widespread adoption in the workplace. However, this paper delves into the limitations that hinder the deep integration of LLMs with existing software applications used in business environments. While LLMs excel in processing natural language, conventional software applications mostly communicate structured information in a rigid fashion. To bridge the gap between human usage of natural language and the structured data utilized by existing applications, communication tools such as email and Slack emerged. While these tools filled the void by accommodating the informality of chat-like interactions, they struggled to incorporate the applications' structured data. The prevalent form-based models employed by existing software applications present a fundamental obstacle for a meaningful AI integration. This paper examines the challenges of integrating LLMs with form-based software, shedding light on the associated complexities and introduces a new design paradigm, Object Messaging Model (OMM). By leveraging this model, natural language and structured data become logically combined while enabling seamless integration of LLM capabilities within the software application. This paper provides insights into the potential of the OMM and paves the way for unlocking the transformative power of LLMs, resulting in highly enhanced productivity within the contemporary business landscape. The Object Messaging Model offers an elegant path for deep integration of LLMs and software applications with structured data, while optimizing communication among humans in the workplace producing humans and AI collective intelligence.

On the Way to Realize the 5th Industrial Revolution: Achievements, Challenges and Research Areas

Prof. Dr. Eng. George Dimitrakopoulos
Harokopio University of Athens, Greece
Infineon Technologies AG, Munich, Germany

Abstract - The industrial world is rapidly moving towards its 5th industrial revolution (a.k.a. .Industry5.0.). Industry5.0 is reflected on the digital sovereignty in comprehensively sustainable production, through adopting, extending and implementing AI-enabled hardware, as well as AI tools & methods and semiconductors technology across the whole industrial value chain. It is expected that in doing so, manufacturing costs will be decreased, while at the same time, product quality will be increased through AI-enabled innovation, time-to-market will be shortened and user acceptance of versatile technology offerings will be achieved and global supply chains stabilized. The above will, in turn, foster a sustainable development, in an economical, ecological and societal sense. This keynote will discuss on the fundamental research areas relevant to further advancing the digitalized industry, its achievements so far, as well as the challenges confronted, in order to boost industrial competitiveness through interdisciplinary innovations, establishing sustainable value chains and therefore contribute to the Digital Sovereignty.

Quantum Network Basics: How Do Quantum Computers Communicate?

Prof. Michael R. Grimaila
Head, Systems Engineering & Management Department, US Air Force Institute of Technology (AFIT), USA

Abstract: TBA

TUTORIAL: Explaining and Applying Grover's Quantum Algorithm

Prof. Leon Deligiannidis, Ph.D.
School of Computing and Data Science
Wentworth Institute of Technology, Massachusetts, USA

Abstract - Quantum Computing is a groundbreaking technology that exploits the laws and properties of quantum mechanics. It can dramatically speed up computations of certain problems and surpass the computation power of modern supercomputers. One notable quantum algorithm that is used to search an unstructured database is Grover's algorithm. It provides quadratic speedup compared to its classical counterpart. In this tutorial we will explain and illustrate the theoretical and implementational challenges and provide examples and solutions.

Prediction + Explanation: Explainable Machine Learning in Healthcare

Dr. Ahmad P. Tafti
Director, Pitt HexAI Research Laboratory,
Department of Health Information Management,
School of Health and Rehabilitation Sciences,
(Secondary appointment: Intelligent Systems Program (ISP), School of Computing and Information)
University of Pittsburgh, PA, USA

Abstract - While machine learning algorithms have demonstrated very successful applications in a variety of real-world scenarios, the adoption and implementation of machine learning methods in healthcare is a little bit behind from what we are seeing in other high-stake settings, such as airplane navigation and automotive industry. One of the major concerns beside regularity frameworks, patient private data, and healthcare policies, is explainability and interpretability of machine learning algorithms. With the use of machine learning explainability, we are thus opening the black box in a way our clinicians, surgeons, and even our patients can clearly understand how machine learning methods are operating, how they are making a decision, and why! In this talk, we will be exploring explainable machine learning in healthcare settings, particularly within stage-specific cancer survivability prediction.

Cyber Security, Cyberwar, Hacking, Privacy, and Governmental/Personal Data Breaches

Prof. Levent Ertaul
Chair, Department of Computer Science,
California State University, East Bay, California, USA

Abstract - Cyber security, cyberwar, hacking, privacy, and governmental/personal data breaches. We keep hearing these with increasing frequency over and over again. This creates a cyber anxiety everywhere. And on top of that we, as ordinary people started to learn that corporations and governments all around the world keep track of our personal data. For example, mobile phones constantly provide information about our location to service providers. Google knows what we are thinking about from our personal online searches. Facebook can see who our friends are. Yahoo knows the type of news we are interested in. Our online shopping patterns are recorded. Governments are launching surveillance programs to collect our personal data on the cyber space. And the list goes on... It is as if we are all living in glass houses in which we don.t have any privacy or can.t keep any secrets anymore. Cyber security issues affect everyone. Most of all they affect us individuals. That is why ignorance is not bliss in cyber security. Every day we face new questions, new challenges from our rights and responsibilities as citizens of the cyber world to how to protect ourselves, if we can, from new types of security threats. In this talk, I will try to explain vulnerabilities and security issues in the cyber space along with what we can and cannot do to protect ourselves.

What Can we Learn from Student Course-Grade Data?

Prof. Gary Weiss
Computer & Information Science Department, Fordham University, USA

Abstract - Data science has made great inroads into many areas of our daily life, plays a key role in many industries, and is beginning to play an increasing role in education. In this talk I will describe the variety of knowledge that can be gleaned by analyzing undergraduate course-grade data (i.e., data that list all courses a student took and their corresponding grade). I will show how this data can be used to: identify strong and weak instructors and easy and hard graders; characterize grading patterns at the student, course, instructor, and department level; characterize how students sequence their courses and the impact of this sequencing on student learning; identify majors in which a student will perform well; and group courses based on student course co-enrollments or similar patterns in student performance.

The 7th International Conference on Applied Cognitive Computing (ACC'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-ACC
https://www.american-cse.org/csce2023/

Cognitive Cluster Identification

Ainslee Heim, Ken Ferens
Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, MB, Canada

Abstract - This paper's focus is on using cognition in the form of the Variance Fractal Dimension (VFD) to identify malware and benign clusters in an Artificial Neural Network. This paper proposes a novel method of deep learning that extracts higher order features from clusters formed by a Self-Organized Map (SOM) through a VFD calculation to assign a label of malware or benign to each cluster.

Artifcial Emotional Intelligence Testing for AI Avatars

James A. Crowder, John Carbone
CAES Advanced Program Development, Colorado Springs, Colorado, USA;
Global Governments & Critical Infrastructure, USA

Abstract - There has been significant dialogue over the last few years regarding the utilization perils of artificial intelligence in virtually everything we touch. From our phones to our cars, and everything in between, artificial intelligence has become an integral part of our existence. Nikola Tesla described the dangers of unrestricted AI when he expressed that a "borrowed mind" built from human minds may not be able to contain their inherent flaws or biases. Prominent figures, like Elon Musk and Stephen Hawking, have also warned about the potential for machines to take over and wreak havoc in the lives and existence of humans. Hollywood has made untold billions of dollars painting doom-and-gloom scenarios about artificial intelligence and robots within our society today and in the future. But what is the true reality? We continually push to create increasingly intelligent systems/machines that attempt to learn, think, and reason like humans. Most attempts fail miserably. Given the recent "issues" with AI avatar chatbots like ChatGPT, we enter a brave new world of interactions with AI-like entities. These AI avatar entities have no cognitive functions and do not understand human emotion, only human language and primarily only English, currently ignoring the other ~3000 written and ~10000 spoken languages. Here we discuss the need and a framework for understanding human emotions in order to create better human-AI communications, and we present an artificial emotional intelligence test that can be used to determine an AI avatar's ability to, and improve the capacity to, functionally communicate with people and between other AI entities.

Graph-Oriented Modelling of Process Event Activity for the Detection of Malware

Kenneth Brezinski, Ken Ferens
Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada

Abstract - This paper presents an approach to malware detection using Graph Neural Networks (GNN) to capture the complex relationships and dependencies between different components of an operating system (OS). Traditional methods for malware detection rely on known signatures of malware and may fail to detect new or modified malware variants. GNNs offer a promising solution by analyzing graph-structured data and identifying malicious behavior patterns. Specifically, this paper investigates the use of GNNs for malware detection based on the API call sequences of different event types, including File System, Registry, and File and Thread activity. The paper presents a representative dataset of host process activity of malware collected in a custom sandbox environment, comprising over 239 malware executions with randomly executed benignware samples. The paper then

describes the GNN model trained on the dynamic process behavior generated from process execution graphs, with independent models developed based on each category of API events. Finally, the paper presents a trained model that maximizes the generalization performance of the model, demonstrating the applicability of GNNs for malware detection. This paper presents one of the first applications of GNN classification based on process hierarchy during malware execution that includes interaction with benignware as well.

Detection of Twitter Spam with Language Models: a Case Study on how to use BERT to Protect Kids from Spam on Twitter

Bianca Jones, Marwan Omar
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA

Abstract - With the increasing use of social media platforms like Twitter, the problem of spam is becoming more prevalent. This is especially concerning when it comes to children who use Twitter, as they may be exposed to harmful content or scams disguised as legitimate tweets. In this paper, we present a case study on how to use Bidirectional Encoder Representations from Transformers (BERT), a pre-trained language model, to detect and protect children from spam on Twitter. We collected a large data set of tweets that are relevant to children and used BERT to build a spam detection model. Using our approach, we were able to accurately detect spam tweets with a high degree of precision and recall. We also evaluated the effectiveness of our approach using various metrics and found that it outperforms several baseline models. Our study demonstrates the potential of using state-of-the-art language models like BERT to protect children from spam on Twitter, and our findings provide insights on how to develop effective spam detection models for social media platforms.

Modified K-means Clustering Algorithms for Feature Selection

Ayeasha Akhter, Ken Ferens
Department of Electrical and Computer Engineering, University of Manitoba, Winnipeg, Canada

Abstract - Computational effort is difficult when dealing with high dimensional data that has hundreds or thousands of features. Features that don't significantly influence class predictions throughout the classification process increase the computing load. By eliminating unnecessary, redundant, or noisy features from the original features, feature selection, as a dimensionality reduction strategy, tries to pick a small subset of the important features from the original features. Two new feature selection methods are described in this study in relation to the effectiveness of k-means-based clustering methods. This research project aims to reduce the number of different features by clustering the D features into k ($k < D$) clusters, determining the cluster center to represent its members by finding the closest feature to the cluster center or selecting the highest weighted features among the cluster members, and performing feature selection. After removing 41.4% of the features from the VIRUS-MNIST dataset, we are able to deliver accuracy equivalent to the original dataset using both of our suggested methods in a shorter amount of time. Our proposed methods outperform sparse k-means, PCA, LLE, and wk-means-based feature selection method for clustering by ANN following feature reduction in the Wine dataset. With fewer features than the modified k-means feature selection method, our second method performs more accurately on the CNAE dataset.

Exploring the Journey to Drug Overdose: Applying the Journey to Crime Framework to Drug Sales Locations and Overdose Death Locations

Murat Ozer, Ismail Onat, Halil Akbas, Nelly Elsayed, Zag Elsayed, Said Varlioglu
University of Cincinnati, Cincinnati, USA;
University of Scranton, Scranton, Pennsylvania;
Troy University, Troy, Alabama, USA

Abstract - Drug overdose is a significant public health issue in the United States, with tens of thousands of deaths occurring annually. This study uses the Journey to Crime (JTC) framework borrowed from environmental criminology to explore the relationship between drug sales locations and overdose death locations. The JTC framework is based on the distance decay concept,

which suggests that crimes are likely to occur closer to offenders' residential places. The study aims to understand the journey of overdose victims to overdose locations to improve overdose service distributions/interventions. The researchers applied the JTC framework to drug offenses, but to their knowledge, no study has applied it to drug overdose deaths. The study explores the likely relationship between overdose deaths and drug sales locations using data from Hamilton County, OH Coroners, and the Cincinnati Police Department. The researchers argue that a comprehensive strategy is needed to prevent overdose deaths, including targeting and reducing the opportunities to access illicit drugs, improving responses to overdoses through a collaborative multidisciplinary approach, and access to data to inform strategies and evaluate outcomes.

Performance Evaluation of Open-Set Classification in Human Activity Recognition via a Residual Neural Network Architecture

Daniel Rodriguez Gonzalez, Yanzhen Qu
Colorado Technical University, Colorado Springs, Colorado, USA

Abstract - A novel technique using residual neural networks is introduced in this research to address the open-set classification problem in the field of Human Activity Classification (HAR). The open-set issue presents several problems when performing classification tasks due to the removal or lack of labels during the model training phase. Limited amounts of research has been conducted to address this issue directly despite its recurrence in real-world applications. The Time Series Classification Residual Network (TSCResNet) deep learning model introduced in this study is evaluated with the benchmark PAMAP2 Physical Activity Monitoring Dataset. By using the same quantitative methods and data preprocessing protocols as previous research in the field of open-set HAR, results show that the TSCResNet model architecture is able achieve improved classification results.

The 24th International Conference on Bioinformatics & Computational Biology (BIOCOMP'23: July 24-27, 2023; Las Vegas, USA)
https://american-cse.org/csce2023/conferences-BIOCOMP

The 9th International Conference on Biomedical Engineering & Sciences (BIOENG'23: July 24-27, 2023; Las Vegas, USA)
https://american-cse.org/csce2023/conferences-BIOENG

Exploring Deep Brain Stimulation Effects in a Hyperbolic-Based Computational Model of Neuronal Spiking Patterns

Ardavan Vakil, Miad Faezipour
School of Engineering Technology, Electrical and Computer Engineering Technology,
Purdue University, West Lafayette, Indiana, USA

Abstract - Deep brain stimulation (DBS) is a device that relies on methods to desynchronize adverse neurodegenerative activities within the brain. It is a well-known treatment method for people diagnosed with Parkinson.s disease (PD). The adjustment of a precise definition of DBS signal parameters is critical to breaking the synchrony and burst of the increased neuronal activity, while preventing tissue or electrode damage, and also considering reducing the energy consumption for prolonging the battery life of the DBS device. In this research in progress, we study a pulse DBS waveform with delay between its cathodic and anodic parts to disrupt and desynchronize the firing/bursting patterns of Parkinsonian neuronal cell activities. We use a hyperbolic neuronal spiking pattern model to represent the signal activity of the brain. Specifically, we model a neuron infected with bursting behavior similar to that of PD, and apply DBS current to observe the effect, computationally. The expected output of applying the proposed pulse delayed DBS waveform is to break the bursting pattern and convert the bursting spike into a regular tonic spike. In addition, the promising results of pulse delayed waveforms in terms of eliciting action potential, desynchronization of the bursting behavior, and reduced energy consumption, can potentially enhance the performance of the applied DBS.

Predicting Drug Response Using Two Factors from Cell Lines-Drug Sensitivity and Basal Gene Expression

Chuan Wang, Yuan Tian, Wen Shi, Ying Zhou, Yi Zhou
Vanderbilt University Medical Center, Nashville, Tennessee, USA;
Department of Chemical Engineering, Department of Biosystems Engineering, Auburn University, Auburn, Alabama, USA;
Department of Economics, TSYS School of Computer Science, Columbus State University, Columbus, Georgia, USA

Abstract - Patient drug response plays a critical impact in clinical pharmacology. To accurately predict clinical drug response, we propose a prediction method using two-factor cell lines-drug sensitivity and basal gene expression. We input the baseline gene expression data from a large panel of cell lines and use several machine learning models (ridge regression, random forest, SVM) to predict chemotherapeutic response in patients. Specifically, first, we fit 80% of the gene expression and drug sensitivity data into our models for training. Next, we evaluate the accuracy and effectiveness of our prediction method by applying the rest 20% data to our trained models, which yields a good drug sensitivity prediction. Last, we reduce the number of genes with two different types of feature selection, dramatically decreasing the training time without significant changes in prediction accuracy. Our experimental results demonstrate that our models with feature selection are good tools for drug response prediction.

A Research Agenda on the Feasibility of the Proposed 5G Enabled Use Cases on Healthcare

Akshay Monga
Harrisburg University of Science and Technology, Pittsburgh, Pennsylvania, USA

Abstract - This paper is a research agenda on the feasibility of the proposed 5G enabled use cases in healthcare. This paper aims to look at the research that has been published on the use cases involving 5G in healthcare and determine if the use cases are feasible. The paper will then aim to consolidate the disparate research and provide an agenda for future research to help make the proposed use cases into reality. This paper will go on to define what is 5G, define the potential use cases of 5G in healthcare and summarize the future potential direction of research.

Synthetic Image Sequence Generation for Endothelium in situ Simulator

Marcial Sanchez-Tecuatl, Francesco Moccia, Roberto Berra-Romani, Jorge F. Martinez-Carballido
Electronics Department, National Institute of Astrophysics, Optics and Electronics, Puebla, Mexico;
Laboratory of General Physiology, Department of Biology and Biotechnology "Lazzaro Spallanzani", University of Pavia, Pavia, Italy; Biomedicine School, Benemerita Universidad Autonoma de Puebla, Puebla, Mexico

Abstract - Synthetic image generation (SIG) has played a relevant role in the cell physiology field to validate image processing algorithms, and its usage may be extended to simulator development, helping to consolidate knowledge, reduce the usage of experimental animals, and allow research progress without highly specialized and costly instrumentation. Endothelial cells (EC) control critical functions of the entire cardiovascular system through a complex mixture of sensing and communication capabilities, and there is vast evidence that ECs functionality is highly governed by their intracellular calcium concentration ([Ca^2+]i), thus its impairment may lead to life-threatening complications. We present an approach for SIG of ECs towards a simulator that mimics the variability of [Ca^2+]i at in in situ endothelium injury scenarios. Our SIG is composed of three sequential stages: 1) feature extraction measurements from in situ ECs; 2) randomized phantom generation; and 3) behavioral model assignment. Three different instances of behavioral models are shown. The embryonic approach of our SIG was demonstrated to be functional and suitable for any given behavioral model representation.

Vision Transformers and Bi-LSTM for Alzheimer's Disease Diagnosis from 3D MRI

Taymaz Akan, Sait Alp, Mohammad Alfrad Nobel Bhuiyan
Department of Medicine, Louisiana State University Health Sciences Center at Shreveport, Shreveport, Louisiana, USA;
Department of Computer Engineering, Erzurum Technical University, Erzurum, Turkey

Abstract - Alzheimer's is a brain disease that worsens over time and affects memory, thinking, and behavior. Alzheimer's disease (AD) can be treated and managed if diagnosed early, slowing the progression of symptoms and improving quality of life. In this study, we suggested using the Visual Transformer (ViT) and bi-LSTM to process MRI images for diagnosing Alzheimer's disease. We used ViT to extract features from the MRI and then map them to a feature sequence. Then, we used Bi-LSTM sequence modeling to keep the interdependencies between related features. In addition, we evaluated the performance of the proposed model for the binary classification of AD patients using data from the Alzheimer's Disease Neuroimaging Initiative (ADNI). Finally, we evaluated our method against other deep learning models in the literature. The proposed method performs well in terms of accuracy, precision, F-score, and recall for the diagnosis of AD.

Intelligent Upper Limbs Prosthetics with 1D Convolutional Neural Networks and Quick Training

Lucas Givens, Md Julfiker Al Jewel, Md Muminul Hossain, Upali Karunathilake, Fred Wu
West Virginia State University, Institute, West Virginia, USA;
University of Memphis, Tennessee, USA

Abstract - Upper limb amputation can severely restrict the mobility and ability of amputees to perform daily activities. In addressing this issue, deep learning algorithms and electromyography pattern recognition have emerged as promising clinical solutions for functional upper-limb prosthetics. This article presents the use of EMG sensors to capture muscle movement signals and applies the pattern recognition function of a 1D convolutional neural network to identify these signals and control the movement of prosthetics. Experimental results demonstrate that the convolutional neural network exhibits fast training and high-precision recognition capabilities enabling it to accurately identify muscle signals and effectively control prosthetic movements.

Development of Embedded Machine Learning Finger Number Recognition Application Using Edge Impulse Platform

Chun-Ki Kwon
Department of Medical IT Engineering, Soonchunhyang University, Asan, ChungNam, South Korea

Abstract - This work demonstrates the use of a user-friendly AI platform.Edge Impulse.to develop and evaluate Tiny Machine Learning (TinyML) models. Edge Impulse is a platform for building TinyML models that are optimized to run efficiently on any tiny embedded device, and it therefore has tough resource constraints, such as a memory size of a few hundred kilobytes and ultra-low power consumption. To achieve our stated purpose, we focused on five finger number gestures representing the numbers from one to five, and we then used the Edge Impulse platform to build optimized machine learning applications to recognize these five selected finger number gestures. The Arduino Nano 33 BLE Sense is the target device on which a TinyML generated by an Edge Impulse platform is assumed to be running. In terms of performance, the TinyML was able to correctly recognize 92% of the finger number gestures from one to five.

Using Machine Learning Model in Human Body Recognition for Hospital Caregivers Immobile Patients Turning

Chris Cheng Zhang, Kevin Zhang, Yu Shen
Canada Youth Robotics Club, Canada; University of British Columbia, Canada;
JiangSu People's Hospital, Nanjing, Jiangsu, China

Abstract - This report is about the implementation of a device that helps with turning a patient on bed for a caregiver which the objective is to observe and determine whether the caretaker's action fulfills all the required criteria. It will cover the two methods, Human Recognition and Object Detection, introduction of these two methods and explanation on how they would play a role in this project. Moreover, this paper would also introduce what Machine Learning is and how doesthis technology involve in Human Recognition and Object Detection. At the end, this report will show the result of how the program the team designed works, including the steps, UI and the physical look of it.

The 21st International Conference on Scientific Computing (CSC'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-CSC
https://www.american-cse.org/csce2023/

From Floats to Posits: A Conversion Framework

Andrin Vuthaj, Elie Alhajjar
Computer Science, United States Military Academy, West Point, New York, USA

Abstract - Modern Artificial Intelligence (AI) applications for mobile and embedded applications are shifting away from the IEEE floating point system due to its many inefficiencies. Recently, a new data type called posits was introduced as a potential replacement for IEEE-754 Standard floating point numbers (floats). Posits are a core component of the third generation universal numbers (type III unums). They provide compelling advantages over floats, including larger dynamic range, higher accuracy, better closure, bitwise identical results across systems, simpler hardware, and simpler exception handling. In this paper, we aim to provide a framework for an efficient conversion between the two systems. Namely, we devise a novel conversion algorithm that compares accuracy between representations in terms of computational error and highlights the superiority of the new type over the currently used number systems.

Pairwise Adaptive Timesteps for Soft-Sphere Collisions in Ring Simulations

Jonathan Rotter, Mark C. Lewis
Department of Computer Science, Trinity University, San Antonio, Texas, USA

Abstract - This paper looks at integrating soft-sphere collisions using a pair-wise adaptive time step approach. One of the challenges of soft-sphere collisions, compared to hard-sphere collisions, is that a smaller time step must be used to ensure that interactions are properly integrated. As wall-clock time scales as 1/delta (t), this can be prohibitively expensive. In this paper, we examine the feasibility of using a pair-wise adaptive time step for integrating collisions both to improve the accuracy and the speed of these calculations. The implementation accuracy is tested in small runs of two or three particles, while performance tests use large configurations that mimic local-cell planetary ring simulations. We find the approach to be remarkably successful in both areas. Using an adaptive time step allows us to accurately integrate collisions across a broad range of particle sizes and impact velocities. It also gives a significant speed improvement over a fixed time step using a time step that is small enough to be reasonably accurate.

Model Based and Model Free Offline Reinforcement Learning Combat Simulation Scenario

Indu Shukla, Althea C. Henslee, Haley R. Dozier, Brandon E. Hanson, Joseph E. Jabour, Ian D. Dettwiller
Information Technology Laboratory, United States Army Corps of Engineers, Vicksburg, Mississippi, USA

Abstract - Reinforcement learning (RL) is machine learning (ML) paradigm that has produced systems capable of performing at or above a professional-human level. This research explored the ability of RL to train AI agents to achieve best possible offensive behavior in small tactical engagements resembling a simple 1D military simulation. Battlefield environment is complex domain, therefore, planning and building combat simulation is challenging. Therefore, it has been an ongoing interest in offline learning

approaches. Previously, we found that the performance of offline model-based and model-free methods to be profoundly effective in Cartpole environment. In this study, we implemented model-based and model-free offline RL in incremental approach in 1-D, aggregate-level military constructive simulation. We performed extensive experiments across several RL methods to find a good policy from previously collected dataset. The consistent improvements of our approach were measured in terms of both state dynamics prediction and eventual reward. Future work will seek to validate RL performance in larger and more complex combat scenarios. We have discussed the potential of offline RL to enable new approaches for developing complex defense planning.

Numerical Diffusion of Material in Radiation Hydrodynamics Codes

William Dai
Los Alamos National Laboratory, Los Alamos, New Mexico, USA

Abstract - This presentation is to report our investigation for numerical diffusion of material in numerical simulations through the multi-physics codes, xRAGE, in which a few different numerical approaches for hydrodynamics are used, particularly, typical dimensionally split and unsplit algorithms, and techniques of interface steepening for material interface.

Minimizing Turns in Watchman Robot Navigation: Strategies and Solutions

Hamid Hoorfar, Sara Moshtaghi Largani, Reza Rahimi, Alireza Bagheri
Department of Epidemiology and Public Health, University of Maryland, Baltimore, Maryland, USA;
Department of Computer Science, University of Cincinnati, Cincinnati, Ohio, USA;
Department of Mathematics, University of Ottawa, Ottawa, Canada;
Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran

Abstract - The Orthogonal Watchman Route Problem (OWRP) entails the search for the shortest path, known as the watchman route, that a robot must follow within a polygonal environment. The primary objective is to ensure that every point in the environment remains visible from at least one point on the route, allowing the robot to survey the entire area in a single, continuous sweep. This research places particular emphasis on reducing the number of turns in the route, as it is crucial for optimizing navigation in watchman routes within the field of robotics. The cost associated with changing direction is of significant importance, especially for specific types of robots. This paper introduces an efficient linear-time algorithm for solving the OWRP under the assumption that the environment is monotone. The findings of this study contribute to the progress of robotic systems by enabling the design of more streamlined patrol robots. These robots are capable of efficiently navigating complex environments while minimizing the number of turns. This advancement enhances their coverage and surveillance capabilities, making them highly effective in various real-world applications.

Simulation Framework and Big Data for Decision-Making Analyses in Sustainable Development - Higher Education Case Study

Anatoly Kurkovsky
University System of Georgia, Georgia Gwinnett College, Greater Atlanta, Georgia, USA

Abstract - Attempts to include some decision-making functionality in sustainable development analysis of higher education are limited by at least three fundamental problems: a) the complexity of the subject domain, b) the need for specific models to support decision-making, and c) the need for a methodological base to reuse various already created models together with future models along with big data. This paper proposes an approach that can potentially reduce the mentioned problems using a simulation umbrella as a united methodological base of the sustainable development paradigm implementation. Described a simulation case study to demonstrate the proposed approach for the decision-making support for the educational programs. quality as a core component of sustainability associated with one of the US higher education institutions.

Securing Pathways with Orthogonal Robots

Hamid Hoorfar, Faraneh Fathi, Sara Moshtaghi Largani, Alireza Bagheri
Department of Epidemiology and Public Health, University of Maryland, Baltimore, Maryland, USA;
Department of Biomedical Engineering, University of Kentucky, Lexington, Kentucky, USA;
Department of Computer Science, University of Cincinnati, Cincinnati, Ohio, USA;
Department of Computer Engineering, Amirkabir University of Technology, Tehran, Iran

Abstract - The protection of pathways holds immense significance across various domains, including urban planning, transportation, surveillance, and security. This article introduces a groundbreaking approach to safeguarding pathways by employing orthogonal robots. The study specifically addresses the challenge of efficiently guarding orthogonal areas with the minimum number of orthogonal robots. The primary focus is on orthogonal pathways, characterized by a path-like dual graph of vertical decomposition. It is demonstrated that determining the minimum number of orthogonal robots for pathways can be achieved in linear time. However, it is essential to note that the general problem of finding the minimum number of robots for simple polygons with general visibility, even in the orthogonal case, is known to be NP-hard. Emphasis is placed on the flexibility of placing robots anywhere within the polygon, whether on the boundary or in the interior.

Overview of Empirical Distribution Function Computational Framework to Explore Chaotic Dynamic Systems

Alexander Chtcheprov, William P. Rice, Andrei Chtcheprov
Georgia Institute of Technology, Atlanta, Georgia, USA;
Independent Researcher, Amarillo, Texas, USA;
Independent Researcher, Chapel Hill, North Carolina, USA

Abstract - This paper describes a working methodology of implementation of empirical distribution functions used as a powerful tool for a theoretical exploration and a visualization of chaotic dynamic systems. To illustrate the framework, a dynamic process generated by a generalized logistic map called LM(½, ½, 1) is considered. The dynamic sequence is represented as a stochastic process with invariant distributions. Mathematical equations for invariant transformations are derived and numerically solved. It is shown that the algorithm may be very sensitive to the size of the computational grid. The paper also describes a computational procedure for identifying unstable periodic orbits and demonstrates convergence of their distributions to real distribution functions of dynamic regimes. Such behavior is postulated as a property of chaos.

Skin Condition Classification by Automatic Image Machine Learning

Lichuan Liu, Wei Li
Department of Electrical Engineering, CEET, Northern Illinois University, DeKalb, Illinois, USA;
Applied AI R&D Laboratory, Intellikey Innovation LTD, Burnaby, Canada

Abstract—In this paper, we proposed a novel Artificial intelligence (AI) based approach which could classify various images of human skin into appropriate categories based on associated features. The proposed system uses K-nearest neighbor classifiers, a neural network, Naïve Bayes (NB), and Gradient Decent / Ascent Classification to classify the features. In the experiment, the results show an improvement over existing approaches in accuracy. The experimental AI algorithm can achieve 50% accuracy with KNN, 60% with NN, 77% with naïve Bayes and 90% with Gradient Descent method.

Data Products Using Data Mesh Principles

David Keever
Technical Fellow, Leidos, Inc., Reston, Virginia, USA

Abstract--The volume and variety of data in organizations continues undiminished. Current architectures are not necessarily designed to keep pace with the speed, scalability, and security of data and data-centric analytics. This poster presentation describes the technical trends and motivation for data products, the four data mesh principles, and a multi-layered approach to offering data products founded on data mesh principles. The key idea is to shift the responsibility for data from a centralized data team to the individual teams and domains that own and use the data. This break-down of data silos creates a decentralized, self-serve data infrastructure that empowers teams to manage their own data under their assigned authorities. The capabilities of Large Language Models (LLM) to accelerate convergence of distributed data asset definitions and metadata are explored.

A Survey of Applications for Deaf-Mute People Based on IoT and Cloud Computing Services

Hector Caballero-Hernandez, Vianney Munoz-Jimenez, Marco A. Ramos-Corchado
Faculty of Engineering, Autonomous University of State of Mexico, Mexico

Abstract - In the world there are more than 400 million people with disabilities associated with speech and hearing, which represents more than 6% of the population worldwide, these people have various problems to be able to integrate into routine activities, as well as to access work sources. This research presents a compilation on technological developments based on IoT devices, as well as cloud computing services dedicated to helping people deaf and mute in their integration into society and the use of information technologies to closing the gap with people who do not have some type of disability. Among the most important investigations are those that use neural networks applied to video to interpret movements of people who use sign language, and in this way be able to generate the translation of text and speech, on the other hand, the outstanding investigations in IoT can interpret the movements of sign language for translate to text and voice using smartphones and other electronic devices. This paper includes a proposal focused on an evolutionary system to resolving some of the vulnerabilities that sign language translation systems present.

Human and Cognitive Factors Involved in Phishing Detection - a Literature Review

Diana Arevalo, Dario Valarezo, Walter Fuertes, Maria Fernanda Cazares, Roberto O. Andrade, Mayra Macas
Universidad de las Fuerzas Armadas ESPE, Ecuador;
Universidad Politecnica Salesiana, Ecuador;
Escuela Politecnica Nacional, Quito, Ecuador

Abstract - Human and cognitive factors considerably influence social engineering attacks. Cybercriminals take advantage of the innocence, carelessness, stress, lack of knowledge, and other aspects that make human beings vulnerable. Also, there exists difficulty for users to identify an email with phishing. However, the causes and solutions are not only technological; they also depend on human perception. Within this context, in this paper, we perform a systematic literature review using the PRISMA guidelines of the recent studies applying security and cognitive psychology, aiming to identify the human and cognitive factors that are part of a Phishing attack. The main findings of this research are focused on developing future research in cybersecurity, which we believe should go in hand with human cognitive and psychological factors.

A Systematic Review of Approaches for Reviewing Security-Related Aspects in Agile Requirements Specification of Web Applications

Aqsa Munir, Mohammed Assiri, Saba Naveed Alam, Muzna Khan, Wasi Haider Butt, Mamoona Humayun
Department of Computer and Software Engineering, College of Electrical and Mechanical Engineering (EME), National University of Science and Technology (NUST), Islamabad, Pakistan;
Department of CS, College of Sciences and Humanities, Aflaj, Prince Sattam bin Abdulaziz University, Saudi Arabia;
Department of SE, College of Computer Engineering and Sciences, Prince Sattam bin Abdulaziz University, Saudi Arabia;
Department of Information Systems, College of Computer and Information Sciences, Jouf University, Al-Jouf, Saudi Arabia

Abstract - Imperfections in requirement specification can cause serious issues during software development life cycle. It might bring about inferior quality products due to missing attributes, for example, security. Specifically, Web Applications are considered as obvious target for getting significant information. Security aspects have gotten hard to manage in web applications because security requirements are not regularly seen appropriately and often details are missing which lead into ill-characterized security-related aspects. With the help of systemic literature review (SLR), we have identified 32 major research works which were published in period of year 2010 to 2020. We identified 38 security-related aspects of web applications and also 13 techniques and tools for security related aspects in agile requirement specification. We have also analyzed 22 challenges in reviewing security related aspects in agile development life cycle.

Towards Modeling Functional Requirements from Tacit Knowledge

Rameen Jafar, Maram Fahaad Almufareh, Sundas Ashraf, Bushra Khan, Wasi Haider Butt, Mamoona Humayun
Department of Computer and Software Engineering, National University of Sciences and Technology, Pakistan;
Department of Information Systems, College of Computer and Information Sciences, Jouf University, Sakakah, Saudi Arabia

Abstract - Requirements Elicitation is the process of identifying system needs by talking with stakeholders who have a direct or indirect effect on the requirements. Requirements may be derived from several sources and are one of the most important phases of requirement engineering. We need to perform different elicitation techniques to find out the user's needs. Pre-requirements tracing is the process of determining the origins of a specific demand. Typically, certain needs arise that have no apparent source, but stakeholders will testify to their importance. Such criteria, however, are most likely dependent on tacit or tacit-like information entrenched in the issue domain. Tacit knowledge is one of the most common challenges for requirement analysts. This study contributes to constructing a systematic literature review for exploring multiple ways to model functional requirements using tacit knowledge. We have identified and analyzed 10 studies. In these research papers, different authors have discussed different methods to gather functional requirements from tacit knowledge.

Identifying Fake and Real Images by Using Masked Face Periocular Region

Udayasri Nannuri, Kaushik Roy, Jinsheng Xu, Tony Gwyn, Bianca T. Govan
North Carolina Agricultural and Technical state University, Greensboro, North Carolina, USA

Abstract - In this paper, we focus on the face spoofing of masked images to determine whether a masked person is real or fake. We developed a dataset of spoofed masked images generated using the DALL.E 2 tool, performed the ROI extraction using the CNN-DLib detector, and extracted the features using BoVW-SIFT. We applied deep learning and machine learning algorithms. XGBoost and Xception achieved the highest accuracy of 92% and 94% to determine whether the images were real or fake. The approach was tested on the realworld masked face recognition dataset (RMFRD). This shows that periocular information can predict whether the masked image is real or fake.

Mitigating Risk in Machine Learning-Based Portfolio Management: A Deep Reinforcement Learning Approach

Fidel Esteves do Nascimento, Paulo Andre Lima de Castro
Institute of Aeronauthics and Technology, Sao Jose dos Campos, Brazil

Abstract - Many researchers when using machine learning approaches for investing focus on predicting returns, but overlook risk management or portfolio allocation. That may lead to risk management problems or even uncontrolled strategy risk. In order to address such vulnerability, we propose a framework for training an allocation strategy that models the investment process as a Markov decision process. We use a risk decomposition technique and a deep reinforcement learning agent model to compute variances and update portfolio variables efficiently. We tested our framework with a LSTM-based agent for predicting stock movements and found that it was able to offset systematic risk and achieve reasonable returns. The proposed low-dimensional framework may contribute to more effective portfolio management based on reinforcement learning.

The Gap Between the Need and the Realities in Mentoring Computer Science Students by Faculty

Sherri Turner, Anne Hinderliter, Arshia Khan
Department of Educational Psychology, University of Minnesota - Twin Cities Minneapolis, Minnesota, USA;
Department of Chemistry & Biochemistry, University of Minnesota - Duluth, Duluth, Minnesota, USA;
Department of Computer Science, University of Minnesota - Duluth, Duluth, Minnesota, USA

Abstract - This paper describes the gap between the need for mentoring computer science students and available faculty resources. The critical role that mentorship plays in workforce development and in the educational progress and career success of students is highlighted.

Automated Machine Learning Model Selection Analysis

Joseph E. Jabour, Althea C. Henslee, Haley R. Dozier, Indu Shukla, Brandon Hansen, Abderahim Salhi, Ian Dettwiller
United States Army Corps of Engineers, Vicksburg, Mississippi, USA

Abstract - Machine learning (ML) is a branch of artificial intelligence (AI) in which a program uses large sets of data to .learn. methods and patterns. Following the learning process, a model is created based on the learned information, which can then be applied to new datasets to automatically perform tasks. Automated Machine Learning (AutoML) is a new tool which can automate the mundane tasks involved in machine learning, such as data pre-processing, model selection, and hyper parameter tuning. Model selection is one of the most important aspects in creating an effective ML model as it affects the algorithms used to perform ML. AutoML will automate the process of trial and error needed to find the most effective model for a task, however, there are different methods used to select a model. This paper will investigate the various tools used to perform AutoML, taking note of the methods used to for model selection to find the current state of model selection in AutoML.

Predictive Modeling of Diabetes Onset and Survival Analysis Among Diabetic Patients

Mohammad Zarak Shah Ji, Mahdi Ebrahimi
Computer Science Department, California State University Northridge, California, USA

Abstract - This paper aims to develop and compare predictive models to detect diabetes early using machine learning algorithms and presents a survival analysis of patients with diabetes having heart failure using the random survival forest (RSF) algorithm. Gradient Boosting, Logistic Regression, and Random Forest are the three models selected for the diabetes prediction study. Two datasets have been used, one containing demographic, clinical, and laboratory variables of patients diagnosed with diabetes and healthy individuals. The second is a comprehensive dataset, which includes demographic, clinical, and laboratory data of patients diagnosed with diabetes and heart failure. The problem that is being solved is firstly to predict whether a person is suffering from

diabetes or not. After determining the presence of diabetes, the study aims to solve a more significant question of whether a person being diabetic can have heart failure, the reason being diabetes. Experimental results show that all three models used for diabetes prediction have high predictive performance, with Gradient Boosting achieving the highest accuracy and AUC-ROC values. The RSF model is trained and evaluated based on its ability to predict the survival outcomes of patients. The results demonstrate that RSF provides high prediction accuracy and outperforms other survival analysis techniques. The findings of this study have significant implications for improving the early diagnosis and treatment of diabetes, which can eventually improve patient outcomes and reduce healthcare expenses. It can also help healthcare providers develop personalized treatment plans and interventions for improving the survival results of patients with both diabetes and heart failure.

Challenges and Solutions for Integrating Neural Networks in Multiphysics CFD Simulations

Aymeric Millan, David Lugato, Paul Novello
Maison de la Simulation, CEA Saclay, Gif-sur-Yvette, France;
CEA CESTA, Le Barp, France;
DEEL, ANITI, IRT Saint Exupery, Toulouse, France

Abstract—Scientific machine learning (SciML) is a major research area of artificial intelligence (AI) and computational science that involves performing machine learning for scientific computing applications such as physics, biology or engineering. In this work, we focus on its application to multiphysics simulations by studying a hybrid simulation based on the coupling between a neural network approximating chemical reactions and a CFD solver. With this approach, we achieved in previous work a 18.7× speedup compared to a traditional solver simulating the same phenomenon. This paper focuses on the computational challenges for high-performance computing software associated with this approach and proposes three integration methods using two tools: ML frameworks and ML compilers. We also propose a methodology to extend the use of a Neural Network for any input mesh size, irrespective to the training batch size. We demonstrate how the speedup factor can be further improved up to 30× compared to the baseline.

K-Hopped Link Prediction with Graph Embedding

Tonni Das Jui, Erich Baker, Mary Lauren Benton
Computer Science, Baylor University, Waco, Texas, USA

Abstract - Graph embedding approaches aim to provide a low-dimensional latent representation of the graph with minimal reconstruction error. In addition, these approaches attempt to capture local and global topological neighborhood information and data distribution in the latent representation. The primary purpose of the graph's latent representation is simply implementing them into various straightforward machine learning models for graph prediction tasks such as link prediction, clustering, and visualization. Among these tasks, link prediction is a critical task in which researchers mainly analyze the performance of the embeddings on the information of adjacent nodes. Although many embedding techniques claim to capture the hopped neighborhood into the embedding, researchers need to pay more attention to analyzing the performance of the graph embeddings on hopped link prediction. Hopped link prediction demonstrates the performance of capturing a global view of the neighborhood. Our proposed framework develops k-hopped graph topological and feature information to analyze six widely recognized graph embeddings: ARGE, ARVGE, Node2vec, Attri2Vec, GraphSage, and GCN with k-hopped link prediction. We experiment with three graph datasets and show that k-hopped link prediction performance significantly increases for 1-hopped graph information and continuously depletes after 1-hop, demonstrating the importance of embedding performance analysis with k-hopped link prediction.

On the Solutions of a Fifth Order Nonlinear Partial Differential Equation of Fluids

Chaudry Masood Khalique
Material Science, Innovation & Modelling, School of Mathematical and Statistical Sciences,
North-West University, Mafikeng Campus, Mmabatho, Republic of South Africa

Abstract - This paper studies a fifth order nonlinear partial differential equation that describes nonlinear waves in fluids and oceans. Exact solutions are obtained for this equation using the travelling wave variable. This transforms the nonlinear partial differential equation into a nonlinear ordinary differential equation. Thereafter, by invoking the Kudryashov's method and the simplest equation method we seek exact solutions. The solutions gained are the travelling wave solutions. Finally, the dynamics of the obtained solutions are displayed in 3D and 2D graphs.

Segmentation of Skin Cancer and Intensity Classification Using Deep Convolutional Neural Network

Hanan Aljuaid, Rafaf Alnoghaimshi, Arwa Alhammad, Fay Aldhafer, Noura Alsakran
Department of Computer Sciences, College of Computer and Information Sciences, Princess Nourah bint Abdulrahman University (PNU), Riyadh, Saudi Arabia

Abstract - Skin cancer is an excessively common type of cancer. It occurs when mutations appear in the DNA of skin cells. The four main forms of skin cancer are Basal cell carcinoma (BCC), Squamous cell carcinoma (SCC), Merkel cell carcinoma (MCC), and the deadliest form Melanoma. Skin cancer is usually diagnosed late because of the unnoticeable symptoms. Therefore, Reliable automatic detection of skin tumors is needed to help increase the accuracy and efficiency of pathologists. In this paper the DCNN method is used which is designed to perform complex analysis of 2594 images and 2594 of corresponding ground truth (response masks) for training and 1000 images for testing of data using image segmentation and classification for creating model that detect skin cancer in early stages. The models testing produced positive outcome with accuracy 0.95 for classification and 0.895 for segmentation. The results are promising for future enhancement.

Performance Evaluation of Anomaly Detection Algorithms in Machine Learning

James Scheibmeir
NorthWest Arkansas Community College, Bella Vista, Arkansas, USA;
Gartner Inc., USA

Abstract - In this study, we investigate the performance of various anomaly detection algorithms. We compared nine different models based on their speed and robustness on a collection of tweets and their sentiment analysis scores. Our results demonstrate the critical importance of efficient anomaly detection techniques for maintaining high-quality predictions and optimal system performance. Among the nine models we examined, the subspace outlier detection (SOD) model had the best execution time performance against our dataset. While the iforest algorithm found the greatest number of anomalies on average. This suggests that algorithm selection may differ based upon the performance and robustness requirements of the system being built.

Sarcasm Detection on Twitter: a Comparative Survey

Hajar Abdulrahmman Alghofaily, Fethi Fkih
Department of Computer Science, College of Computer, Qassim University, Buraydah, Saudi Arabia

Abstract - Sarcasm is an expression or style of utterance that has the opposite meaning from the written words. With this ambiguity, sarcasm detection is a difficult task even for humans. Twitter is one of the most popular microblog platforms on which users can publish their thoughts and opinions and Sarcasm detection on Twitter is a field that has recently attracted research interest. Until now, researchers have developed several techniques for sarcasm detection from textual contents. This paper provides a survey of this field by first presenting an introduction to sarcasm, followed by an indepth discussion. To reach this purpose, we

review some relevant previous works in sarcasm-detecting approaches and we present their techniques. The survey also introduces several useful illustrations-most prominently, a table that summarizes papers along different dimensions such as the types of features, approach, algorithms, language, annotation, and datasets used.

A Deep Learning-Based Hybrid Model for Optimal Anomaly Detection

Frances Osamor
Computer Science and Information Technology, University of the District of Columbia, Washington, DC, USA

Abstract - Intrusion Detection Systems (IDS) play a critical role in ensuring the security of an organization's network infrastructure by detecting potential threats. One approach used by many IDS tools is to analyze system calls using flow-based techniques. As the security field continues to evolve, IDSspecific research has become a primary focus, with numerous studies and methods proposed to improve detection accuracy. This study proposes a new hybrid approach for malware detection using deep learning models, specifically the Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN). Additionally, we incorporate an Autoencoder (AE), a type of neural network used for unsupervised learning and data compression. To evaluate the performance of our proposed models, we used the open-source benchmark dataset ADFA-LD for training and assessment. Our experimental analysis demonstrates that our proposed hybrid model can achieve promising results in detecting malware with high accuracy. An autoencoder allows our models to learn and identify patterns that may indicate malware, even without labeled training data. By combining the strengths of multiple deep learning models, we achieved better detection accuracy than traditional flow-based IDS methods.

One Semester Team Projects in Systems Analysis, Software Engineering, and Software Project Management Courses

Raghava Gowda
Department of Computer Science, University of Dayton, Dayton, Ohio, USA

Abstract - Team Projects in Software Engineering courses provide students exposure to real world problems, systematic ways to understand and specify problems, arrive at analysis, design models, and put together a system architecture for implementation. This paper explores one semester projects in Systems Analysis, Software Engineering, and Software Project Management courses offered at the undergraduate and graduate levels in Computer Science at the University of Dayton. All the courses are in continuum in the broader definition of Software Engineering and have one overlapping topic on use cases. The challenge is to focus on essential tools and techniques for a specific course, incorporate the latest technologies, and provide learning environments for acquiring knowledge and skill sets, and use them in team projects.

Digital Transformation and its Impact on Customer Service in Traditional Banks: A Case Study of a Zimbabwean Commercial Bank

Edmore Masinire, Esnah Dzimba, John Andrew Van Der Poll
Graduate School of Business Leadership (SBL), University of South Africa (Unisa), Midrand, South Africa

Abstract - As digitalization continues to transform the banking industry, traditional banks are striving to remain competitive by adopting new technologies and strategies. This study focuses on the role of digitalization and its impact on customer services in a Zimbabwean bank. Using a qualitative case study approach, we examine how one of the banks in the country is digitalizing to meet the evolving demands of customers. Our findings suggest that digitalization offers numerous benefits, including improved business processes, cost savings, and enhanced customer experiences. To fully leverage these benefits, banks must adopt a customer-centric approach and focus on developing innovative products and services that meet the unique needs of their clients. Digital platforms can provide a safe, secure, and convenient way for banks to engage with their customers. However, one of the biggest challenges for banks in pursuing a digital strategy is the significant cost of acquiring and implementing new technologies. To overcome this hurdle, partnerships with financial technology companies can be an effective way to reduce initial investment

costs. In addition, traditional banks embarking on a digital transformation journey need to realign organizational values and culture to the new digital strategy. Overall, our study highlights the importance of digitalization in the traditional banking sector and provides recommendations for banks seeking to remain relevant and competitive in the digital era. Future research can explore undertaking a quantitative survey involving several traditional banks to develop a model of digital transformation for enhanced customer service in developing country contexts.

Energy Balance and Energy Conservation Associated with Nonlinear Nonviscous Atmospheric Dynamics in a Thin Rotating Spherical Shell

Ranis Ibragimov
DE Gruyter, Boston, Massachusetts, USA

Abstract - The energy balance is deduced on the basis of two classes of functorially independent invariant solutions associated with the nonlinear Euler equations used to model two-dimensional atmosphere dynamics in a thin rotating spherical shell. It is shown that the energy balance corresponds exactly to the conservation law for one class of the solutions whereas the second class of invariant solutions provides and asymptotic convergence of the energy balance to the conservation law.

Task-Based Content Delivery at the Wi-Fi Edge

Anil L. Pereira
Department of Information Technology, Georgia Gwinnett College, Lawrenceville, Georgia, USA

Abstract - Task-based content delivery methods that combine Java fork-join with Java New file IO are proposed for Wi-Fi edge computing devices. Fork-join task-based processing framework performs better than traditional Java multithreading for parallel computation on multicore CPUs, especially for divide and conquer algorithms. New file IO for synchronous nonblocking socket channels and file channels, memory-mapped files and asynchronous socket channels combined with fork-join perform better than traditional Java synchronous blocking IO for sockets and file streams under highly unbalanced processing loads. Performance of proposed methods for throughput, heap memory occupancy and scalability are evaluated and compared with the traditional method in a Wi-Fi network. Fork-join combined with asynchronous socket channels and memory-mapped files performs best and significantly over the traditional method.

Comparative Analysis of Power Efficiency in Heterogeneous CPU-GPU Processors

Rupinder Kaur, Farah Mohammadi
Department of Electrical and Computer Engineering, Toronto Metropolitan University, Toronto, Canada

Abstract - With the increasing demand of applications and progression in technology, system complexity and power consumption have increased manifolds. This stresses on the need of an energy efficient platform that optimizes the performance of machine learning algorithms. This paper presents in-depth analysis of energy efficiency and performance of a heterogenous CPU-GPU platform with different workload distribution settings to determine optimal configuration.

Comparison and Applications of Multiplying 2 by 2 Matrices Using Strassen Algorithm in Python IDLE Notebook and Colab

Shanzhen Gao, Weizheng Gao, Olumide Malomo, Jianning Su, Julian D. Allagan, Ephrem Eyob
Virginia State University, Virginia, USA;
West Virginia State University, West Virginia, USA;
Midway University, Midway, Kentucky, USA;
Elizabeth City State University, Elizabeth City, North Carolina, USA

Abstract - Even though the task of multiplying matrices appears to be rather straightforward, it can be quite challenging in practice. Many researchers have focused on how to effectively multiply two 2 by 2 matrices by applying Strassen Algorithm in the past 50

years. They worked on the complexity from both the mathematical and algorithmic points of view. In our paper, we will discuss the comparison of processing time in Python IDLE, Jupyter Notebook, and Colab from a practical point of view. Several open problems are then presented to challenge our readers.

Towards Applying Graph Theory Modelling with NFC Smart Dashboards for the Mobility Home Environments

Sean Thorpe, Max Robotham
Faculty of Engineering and Computing (FENC), University of Technology (UTech), Kingston, Jamaica

Abstract - This study aims to improve user experience and usability in Near field Communication (NFC) smart dashboard interfaces within the smart home environment. Existing designs of smart dashboards may lack efficiency, visual clarity, and intuitiveness. The study focuses on two research questions: (i) How does the application of graph theory in the design of NFC smart dashboard interfaces enhance usability and user experience within the context of the home environment? (ii) What are the specific usability challenges and improvements identified when integrating graph theory principles into the Human Computer Interaction (HCI) framework for the design of NFC smart dashboard interfaces? By addressing these questions, the study examines the impact of graph theory on the usability of NFC dashboard interfaces in the home environment. It explores the benefits and challenges associated with integrating graph theory principles into the HCI framework and their impact on usability and user experience.

Order-based Distinctions of Performance - An Instructive Example of Logic Puzzles

Indu M. Anand
Sushila Publications, Chelmsford, Massachusetts, USA

Abstract - For processes which are carried out step-by-step, the impact of varying the order of execution of the steps is expected but challenging to quantify. A method of measuring such distinctions for logic puzzles is presented here wherein the order of execution of the steps is precisely tracked by using labels from a known sequence and assigning number values to the labels. The methodology may be extrapolated to provide insights for other constraint specification problems and activities.

Analysis of Information Security Management Applying International Standards to Mitigate Risks

Rodrigo H. del Pozo Durango, Segundo Moises Toapanta Toapanta, Eriannys Zharayth Gomez Diaz, Jose Antonio Orizaga Trejo, Luis Enrique Mafla Gallegos, Ma. Rocio Maciel Arellano, Victor Manuel Larios Rosillo, Maria Mercedes Bano Hifong
Postgraduate Director Universidad Estatal de Bolivar (UESB), Guaranda, Ecuador;
Research Department, Instituto Tecnologico Superior Ruminahui, Sangolqui, Ecuador;
Department of Information Systems, Universidad de Guadalajara (UDG-CUCEA), Guadalajara, Mexico;
Faculty of Systems Engineering, Escuela Politecnica Nacional (EPN), Quito, Ecuador;
Department of Information Systems, Universidad de Guadalajara (UDG-CUCEA), Guadalajara, Mexico;
Postgraduate Subsystems, Universidad Catolica de de Santiago de Guayaquil (UCSG), Guayaquil, Ecuador

Abstract - Information security is strategic and is one of the most important assets within an entity or organization; is one of the reasons to analyze and treat the different risks of confidentiality, integrity, availability (CIA) in information management, and that identity, authenticity, authorization and auditing (IAAA) are available both internally and externally, for this reason. Therefore, preventive measures must be taken. The objective of this investigation is to analyze the security standards to mitigate the risks, vulnerabilities and threats. The deductive method and exploratory research were used to analyze the information from the referenced articles. The result was the percentage of compliance with the potential risk indicators, a table of security metrics for compliance and the prototype for risk mitigation of a technological platform. It was concluded that information security

management should be standardized considering the standards, good practices and security policies available at the government, strategic, tactical and operational level.

Guide to Protect Home Routers from Cyberattacks and Information Theft

Guillermo Ivan Ona Sarmiento, Moises Toapanta Toapanta, Zharayth Gomez Diaz, Marcelo Zambrano Vizuete, Wladimir Paredes Parada, Jose Aguaiza Loja, Jaime Antamba Guasgua
Research Department, Instituto Tecnologico, Superior Ruminahui, Ecuador;
Mechanical Department, Instituto Tecnologico, Superior Ruminahui, Ecuador;
Mechanics Department, Instituto Tec. Univ. Ruminahui, Sangolqui, Ecuador

Abstract - This work was developed to provide the public with a general written procedure to protect their home routers from cyberattacks and protect their information from data theft and other crimes, the detailed steps of the procedure are not applicable to enterprise routers and intermediary devices of companies in custody of servers, for this study and development of the protection procedure, a quantitative method was used in which the security parameters in each router were analyzed one by one, and it was determined that of a total of 10 routers evaluated, 90% of these routers do not present the minimum security measures for their safe operation, this represents a great problem if we consider that one of the most common cyberattack is the DDoS attack (Distributed Denial of Service) that use "botnets" to deny service to the server, for example from a bank or a university, etc.; For the creation of these “botnets”, cybercriminals attack unprotected home routers to infect connected devices and create their malicious network.

Comparison and Applications of Multiplying Two 3 by 3 Matrices

Shanzhen Gao, Weizheng Gao, Olumide Malomo, Jianning Su, Julian D. Allagan, Ephrem Eyob
Virginia State University, Virginia, USA;
West Virginia State University, West Virginia, USA;
Midway University, Midway, Kentucky, USA;
Elizabeth City State University, Elizabeth City, North Carolina, USA;
Lifetime Fellow of the Institute for Combinatorics and its Applications, USA

Abstract - Multiplying matrices can be very challenging although it seems straightforward. Many researchers have studied the multiplication of two 3 by 3 matrices by using Strassen Algorithm in the past 50 years. They focused on the complexity from the mathematical and algorithmic points of view. We will discuss the processing time comparison from a practical point of view in Python IDLE, Jupyter Notebook, and Colab. Several open problems are posted to challenge our readers.

SESSION: Military and Defense Modeling and Simulation

Co-Chairs: Prof. Douglas D. Hodson (Chair), Prof. Michael R. Grimaila**, Dr. Richard Dill**
**Computer Science and Engineering Department, US Air Force Institute of Technology (AFIT), USA*
***Head, Systems Engineering & Management Department, US Air Force Institute of Technology (AFIT), USA*
https://www.american-cse.org/csce2023/

Understanding the Quantum Teleportation Protocol

Kurt T. Spranger, Michael R. Grimaila, Douglas D. Hodson
The Air Force Institute of Technology, WPAFB, Ohio, USA

Abstract - Decomposing a complex system into smaller abstract functional blocks and developing mathematical models to represent their behavior is an important activity towards developing comprehensive system understanding. In this paper, we decompose the ideal Quantum Teleportation protocol into a collection of simple quantum circuit blocks, examine the behavior of each block, and show how collections of blocks operate to create more complex circuits. We believe this approach greatly simplifies the understanding of how the Quantum Teleportation protocol works. This paper is introductory in nature and is intended to help those who are new to modeling, simulating, and analyzing ideal quantum circuits.

An Exploration of Classical and Quantum Networks Using SeQUeNCe

Blake K. Perkins, Douglas D. Hodson, Michael R. Grimaila
The Air Force Institute of Technology, WPAFB, Ohio, USA

Abstract - SeQUeNCe is a Python-based, discrete-event modeling and simulation package to define and study classical and quantum networks. It defines a modular framework for creating and analyzing networks and provides several hardware and protocol models. This paper demonstrates the creation and simulation of a classical network with a customized protocol and a Quantum Key Distribution (QKD) system that uses the BB84 protocol.

An Examination into Squanch and its Conversion to Julia

Takashi Joubert, Douglas D. Hodson, Michael R. Grimaila
The Air Force Institute of Technology, WPAFB, Ohio, USA

Abstract – N/A

Distributed Boids Simulation: Performance Analysis and Implementation Challenges

Brett M. Martin, Ryan D. Winz, Luke J. Mcfadden, Tor J. Langehaug
The Air Force Institute of Technology, WPAFB, Ohio, USA

Abstract - Boids is a computer methodology to simulate animal flocking behavior. Each boid within the flock is modeled using simple geometric steering rules to define how it moves. Past research uses a boids model in areas such as animation, simulation, and drone swarms. While most approaches utilize a fully centralized architecture where all components are hosted on one computer, this research encapsulates individual boid behavior and distributes it among several computational devices. This is primarily achieved through the use of middleware client software that facilitates the distributed aspect. For our implementation,

Docker was used to containerize the program component; RabbitMQ and Redis were used to facilitate communication and storage. Each component of this distributed system was tested and evaluated on a single host machine.

Distributed Interactive Simulation Prototyping with Mininet

Ryan D. Winz, Douglas D. Hodson, Michael R. Grimaila, Maj Richard Dill
The Air Force Institute of Technology, WPAFB, Ohio, USA

Abstract - Distributed Interactive Simulation (DIS) is a standardized protocol used to foster interoperability between distributed virtual simulations. While each simulation may have a slightly different notion of entity states within its view, the consistent exchange of DIS packets keeps deviations bounded. This deviation, however, is altered under different network conditions such as latency and packet loss. In order to test and evaluate the effect that varying network topologies and conditions have on distributed simulations, we leverage a lightweight containerized testbed using the Mininet network emulation package. A two-node system exchanging DIS-compliant packets is detailed, along with methods for capturing and measuring relevant network traffic and entity state information.

Arduino-based sUAS Detection Sensors

Preston Albury, Maj Richard Dill, Douglas D. Hodson
The Air Force Institute of Technology, WPAFB, Ohio, USA

Abstract - This research effort supports the 2030 US Air Force Science and Technology Strategy goal to achieve global persistent awareness and information superiority through lowcost sensors and less human-intensive data analytics. State actors have converted small Unmanned Aircraft Systems (sUAS) into weapons capable of carrying explosive payloads, artillery, support for other ground and air forces, logistics, and close battlespace awareness. These devices have become instruments of power, even thwarting Russian advancement into Ukrainian territories. In response to this new strike vector, this paper explores the Arduino ecosystem as a novel approach to creating a distributed sensor network capable of detecting sUAS. We propose a sensor node design concept incorporating the use of Tiny ML and lowpower microcontrollers capable of providing robust and vigilant detection of sUAS.

Multi-Sensor Aircraft Classification

Maj Sarah Bolton, Maj Richard Dill, Michael R. Grimaila, Douglas D. Hodson
The Air Force Institute of Technology, WPAFB, Ohio, USA

Abstract - Automatic Dependent Surveillance.Broadcast (ADS-B) is a useful tool to for air traffic controllers, military and other sources that are invested in understanding a national or global air picture. While it is highly available, it can sometimes lack integrity due to hacking, spoofing or, even, unintentional inaccuracies in the broadcast. Unlike primary radar, ADS-B's lack of trustworthiness makes it not feasible to rely on it alone. Fusing other data sources with ADS-B can help confirm the accuracy of the broadcasts or allow ADS-B to act as a surrogate for primary radar and bolster the information that primary radar can provide. This paper presents an effective method of using ADS-B data as a surrogate for primary 3D radar by combining the kinematic information that ADS-B data provides with weather and aircraft images to make predictions about aircraft characteristics.

Drone Detection Using Extracted Mel Frequency Cepstral Coefficients with Logistic Regression and Support Vector Machines

Anthony Brunson, Maj Richard Dill, Brett J. Borghetti, Douglas D. Hodson, (Brent Langhals)
The Air Force Institute of Technology, WPAFB, Ohio, USA

Abstract – N/A

The 22nd International Conference on e-Learning, e-Business, Enterprise Information Systems, & e-Government (EEE'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-EEE
https://www.american-cse.org/csce2023/

Realization of an Automated Validation of Learning Activities Prerequisites in Courses in the Learning Management System Moodle Considering the Qualifications-Based Learning Model Approach

Ramona Srbecky, Alexander Zock, Simon-Alexander Wetzel, Michael Winterhagen,
Wieland Fraas, Jan Dettmers, Matthias Hemmje
Multimedia and Internet-Applications, University of Hagen, Germany;
Work and Organizational Psychology, University of Hagen, Germany

Abstract - This regular research paper presents an approach to adaptively provide learning content and activities in Higher Educational Institutions (HEI) in the Moodle Learning Management System (LMS) based on a learner's Competencies and Qualifications (CQ). Also, the Qualifications-Based Learning Model (QBLM) approach will be considered for further research development. For this purpose, this paper first presents the research methodology, questions, and relevant theoretical background. Based on this, the state of the art, the concept, and the implementation will be presented. In the end, the evaluation and future improvements are presented.

Mining the Internet to Discover Learning Management System Popularity: Evauating who is on top and why

Scot Anderson, Patricia Anderson
School of Computing, Southern Adventist University, Collegedale, Tennessee, USA;
Department of Mathematics, Southern Adventist University, Collegedale, Tennessee, USA

Abstract - Our research reports new data on learning management system (LMS) popularity and gives an evaluation based on feature availability and ease of use of the four leading LMS platforms utilized in US higher education institutions: Canvas, Blackboard, D2L and Moodle. This research provides both the methods to mine market share and links to the data from our mining activities. The analysis breaks down possible reasons why we find such a strong front runner in Canvas. The research compares each feature of the four top LMS platforms and their ease of use. The aggregated data provides a usability score based on the presence of each feature and a ranking of the time it took an instructor to implement the feature. We also review the critical elements of Accessibility and Documentation.

Modelling Electric Car Purchase Intention and Cost Affinity

Kalisri Logeswaran Aravindan, Narinasamy Ilhavenil, Thurasamy Ramayah, Muhammad Aiman Izzat, Yap Voon Chong, Arman Hj Ahmad, Murali Raman
Faculty of Management, Multimedia University, Cyberjaya, Malaysia;
Institute of Teacher Education Special Education Campus, Bandar Tun Razak, Kuala Lumpur, Malaysia;
School of Management, Universiti Sains Malaysia, Minden, Malaysia;
Marketing Department, Universiti Kuala Lumpur Business School, Kramat Kuala Lumpur, Malaysia;
Academic Development and Strategy, Asia Pacific University, Taman Teknologi, Bukit Jalil, Malaysia

Abstract - Green and sustainability are important elements to address climate change especially due to high carbon emissions. Many countries including Malaysia have embarked on electric vehicles especially electric cars as a potent solution in reducing carbon dioxide (CO2) emission levels. Nevertheless, the adoption rate of electric vehicles is still at a very low level, despite strong encouragement from the Malaysian government. This study aims to examine the moderating impact of price value on the factors that influence electric car purchase intention in Malaysia. This study would be targeting Gen Z individuals as the unit of analysis. Henceforth, this could provide insights into how electric cars can penetrate the mainstream market, and within the same notion, may uncover methods to encourage the usage of electric cars. This research is underpinned by the Unified Theory of Acceptance and Usage of Technology (UTAUT) and incorporated the price value construct as the moderator to performance expectancy, effort expectancy, facilitating conditions, and social influence. The purposive sampling method is employed to obtain primary data from the respondents via a self-administered questionnaire followed by data analysis using Partial Least Squares Structural Equation Modelling (PLS-SEM), a second-generation multivariate analysis.

An LSTM based Personalized Travel Route Recommendation System

Chieh-Yuan Tsai, Lin-Yi Tai
Department of Industrial Engineering and Management, Yuan Ze University, Taoyuan City, Taiwan

Abstract - With the rapid growth of the economy, more and more people are willing to spend their budget on travel. Due to the rapid development of social media, more people prefer self-guided tours to package itineraries planned by travel agencies. To fulfill tourists. personalized needs, a personalized travel route recommendation system based on social media is proposed. First, to provide meaningful route suggestions, we cluster nearby geotagged locations of photos into landmarks using the HDBSCAN (Hierarchical Density-Based Spatial Clustering of Applications with Noise) algorithm. Second, for each landmark, all hashtags marked in a landmark form a virtual document. The Latent Dirichlet Allocation (LDA) method extracts landmark topics from virtual documents. Third, the visiting time of each landmark and travel time between landmarks are obtained. Finally, a long-short term memory (LSTM) based model is proposed to predict a user's next visiting landmark and its visiting time. For users who do not have previous visiting records, our system can still generate an appropriate route according to the user's preference for topics. A set of case demonstrations and experiments show that the proposed system is sensitive to the users. input and performs better than the one without considering visiting time.

CEO Profiles in the FinTech Industry and the Impact on Financial Performance

Chi King Li, Vincent Cho
The Hong Kong Polytechnic University, Hong Kong SAR, China

Abstract - The emergence of FinTech companies in recent years has led to more innovative solutions than those led by traditional financial services companies. As a result, the role of the CEOs in the FinTech industry has evolved to be quite different; in educational background, experience and leadership style. This paper explores the key differences in the characteristics between the CEOs in the FinTech industry and the CEOs of traditional financial services industry, and explores which should be the most sought-after characteristics of FinTech CEOs by comparing the financial performance of their respective companies.

Improving Performance of Integrated System Applications in an Organisation

Joshua Mugisa Tibaleka, Emmanuel Yeeka
Expanding Social Protection Programme in Uganda Ministry of Gender, Kampala, Uganda;
Makerere University, Kampala, Uganda

Abstract - Decision makers, especially managers in an organisation rely on information systems to make critical decision and plan for daily activities. As such, the performance of information systems should be monitored to ensure that vital information can be accessed whenever it is needed without any errors. Continuous use of information technology has given rise to development of database applications and services to support management of information. In this regard, there are systems and applications which access and use stored data through use of database applications. Such systems include social protection systems, which are formed by very many linked databases. However, using multiple databases often results into existence of duplicated data and incorrect data thus causing errors whenever integrated applications access it. Furthermore, there is need to improve performance of integrated applications by monitoring their performance. In this paper, we propose a performance detection tool to be implemented to monitor performance of a social protection system in an organisation to achieve improved performance in the way integrated applications interact with a social protection system in an organisation.

Towards Automatic Detection of Participant Attention in Virtual Meetings

Kaleab Teka, Dvijesh Shastri
Department of Computer Science and Engineering Technology,
University of Houston - Downtown, Houston, Texas, USA

Abstract - Online meetings have become a norm in recent years. The rapid adaptation of virtual meetings can be attributed to COVID-19 to some degree but also to the usability of virtual meeting platforms such as Zoom and Microsoft Teams. However, the fundamental challenge of gauging participants' attention in the meetings still persists. This research paper discussed our attempt towards detecting participants' attention in virtual meetings. We conducted a pilot study to examine changes in eye gaze movement when the participants engaged in a secondary activity in addition to attending a virtual meeting. A computer-vision-based approach was utilized to quantify the eye gaze movement. The study results reveal that, in comparison to the control trial, the participants' eye gaze movements increased significantly when they were involved in the secondary activity of typing a document, indicating the feasibility of utilizing the softwarebased eye-tracking approach for detecting attention in virtual meetings. The approach has the potential to be a practical application as it requires no eye-tracking hardware.

TDMPITAS: Framework for Successful E-Learning Environment

Satpreet Singh
Director/CEO, Ardass Corporation, Manteca, California, USA

Abstract - E-learning is rapidly replacing traditional learning in classrooms to distance learning at convenience with a similar but virtual experience, particularly for working professionals, mature learners, parents, and students. Technology brings e-learning to fulfill the dreams of many who could not go through traditional learning. E-learning is very effective and meets the needs of generations at their convenience, without one-place learning restrictions. It opens the world to share and communicate their different experiences as social and cultural and brainstorm virtually to innovate new ideas, methods, and technology for a better future. This paper aims to introduce a framework for a successful e-learning environment.

Artificial Intelligence: Interactive Technology in Art Design

Jie Cheng
Department of Computer and Information Technology, Miami University, Hamilton, Ohio, USA

Abstract - This study covers a small fraction of Artificial Intelligence's vast technological advancement by familiarizing teachers/students with Augmented Reality (AR) and how they can utilize it in their classrooms. Since AR is safe, inexpensive, and accessible, it is becoming an essential tool for the educational system, particularly K-12. AR can engage students in a subject through user-friendly interaction with virtual objects. One of the subjects that can apply AR is ART. 5th to 12th-grade students can use AR to make their art pieces interactive and exciting. In this study, designing an interactive painting is given as a sample for students to follow. Experiences with several elementary and junior high students have shown that AR stimulates students' curiosity and create an environment in which they are more creative and engaged.

The Impact of Chatbots on Education

Jie Cheng
Department of Computer and Information Technology, Miami University, Ohio, USA

Abstract - Recent advances in deep learning and natural language processing made chatbots an efficient tool to interact with users more naturally. Chatbot has been used in many fields. Many studies have shown evidence that by applying this Artificial Intelligence technology in education, students' learning outcomes, motivation, and problem-solving skills are significantly improved. By designing, implementing, and applying chatbot in the classroom, we can initiate a project-driven learning process that stimulates student interest and offer teachers and educators new tools to inspire students' curiosity. A recent study shows the determinant factors to increase chatbot usage in the classroom and suggests some guidelines for designing the chatbot. This paper aims to review studies of the impact of a chatbot on education and explore the potential of how a chatbot can apply in today's classroom through a third-grade social science application: how to build a chatbot that answers interesting questions regarding California Native tribe using Dialogflow. The purpose is to engage students with state-of-art technology while they are learning about important topics.

Blockchain Technology in Education: A Comprehensive Review in Transparency

Younghun Chae
Department of Computer Science, Kent State University at Stark, North Canton, Ohio, USA

Abstract - Blockchain technology offers significant potential to transform the education sector. This paper systematically reviews blockchain's role in education, focusing on its transparency ability. It investigates how blockchain could revolutionize education, from data management to credential verification, and identifies data privacy, scalability, legal issues, and more challenges. The paper aims to contribute to the current body of knowledge, fill existing research gaps, and guide future studies in this promising area.

Market Research on Fall Prevention Products

Chris Cheng Zhang, Yanchang Wang, Chenran Wang, Yanyu Wang
Math and Science Association of Canada, Canada;
University of Toronto, Ontario, Canada;
Saint Patrick Regional Secondary School, Vancouver, BC, Canada

Abstract - With the development of the aging population, the care industry for the elderly is becoming a rapidly growing commercial market, among which, the technology products for the elderly, artificial intelligence products are a highly optimistic segment track by investors. Since “falling” accounts for the highest proportion of injuries and deaths among the elderly, the use of artificial intelligence technology to develop anti-fall technology products for the elderly will have an extremely broad market prospect.

Evaluation of Processing Time in Trial-and-Error Function of Security Exercise System for Security Beginners

Ichitoshi Takehara, Mirei Ishizuka, Hitoshi Kamei, Koji Kida, Keizo Saisho
Graduate School of Engineering, Kagawa University, Kagawa, Japan;
Graduate School of Science for Creative Emergence, Kagawa University, Kagawa, Japan;
Faculty of Engineering and Design, Kagawa University, Kagawa, Japan

Abstract - Cyber defense exercises are effective as a method of training security personnel against the sophisticated cyber attacks. In this exercise, the problem is how to practice for beginners who do not know how to execute defense methods according to attacks. To solve this problem, it is important to be able to trial-and-error the defense method many times during the exercise. Therefore, we develop a defense training system with a trial-and-error function by using the virtual machine snapshot function. It is desirable that the processing time of the trial-and-error function during the exercise be short, but generally the processing time increases as the number of participants increases. In this paper, we measure the processing time of trial-and-error function. Then, we estimate the number of students accommodated on one host machine.

An Exploratory Case Study of the Implications of Gamification Theory's Impact on Adult Learners in Post-Secondary Computer Science Classes

Anastasia Tracy Biggs, Dana Betts
School of Computer Science, Engineering, and Technology, Colorado Technical University, Colorado Springs, Colorado, USA;
Department of Engineering, New Brunswick Community College, Saint John, Canada

Abstract - This exploratory case study examines the impact of gamification theory on adult learners in computer science classes. This is an empirical study of how effectively adult learners retain and learn through the use of gamification and innovative educational techniques in the classroom. This is studied from the online and traditional classroom approaches. At this time, adult learners were more receptive to learning programming through entertaining and enticing puzzles and games in traditional and online classrooms. Their retention of knowledge gained from these classes was evaluated by using problem-solving puzzle assignments and in-class gamification results. The online and traditional classroom data were compared and contrasted to determine the effectiveness of gamification theory and educational techniques in the classroom by incorporating evidence from tests, assignments, and in-class gamification contests. This study demonstrates that students learn and retain more material on computer science topics when gamification theory introduces games, assigned work transformed into practical puzzles requiring technical problemsolving skills using software tools, and other educational techniques in the classroom. It argues for the use of gamification theory as a means to engage multi-generational students in learning complex and advanced material in a fun and involved educational setting. This can occur in both traditional and online classroom conference tools such as; Zoom, or Microsoft Teams.

Challenges Confronting Engineering in Nigeria

Matthew N. O. Sadiku, Janet O. Sadiku, Uwakwe C. Chukwu
Department of Electrical & Computer Engineering, Prairie View A & M University, Texas, USA;
Juliana King University, Houston, Texas, USA;
Department of Engineering Technology, South Carolina State University, Orangeburg, South Carolina, USA

Abstract - Nigeria is located in West Africa and is the most populous nation in African continent. It is the largest economy and most notable democracy in Africa. Engineering has a vital impact on the quality of life for any nation. Engineers are concerned with the application of mathematics, ingenuity, and scientific knowledge to develop solutions for technical problems. They play a critical role in the advancement of society. This paper discusses various engineering-related issues confronting Nigeria and the possible ways out.

The Influence of Artificial Intelligence in Education and Workplaces

Amitava Karmaker
University of Wisconsin-Stout, Menomonie, Wisconsin, USA

Abstract - With the development of automation and robots, only a few human laborers are needed to perform traditional manual work. Instead, talents are in great need in the research and development of the latest technologies. Social issues and ethical challenges need to be considered when implementing such AI-based projects. The data used to train the AI model has to be reliable. Without accurate data feeds, AI products would be very misleading and unreliable. This paper discusses a few possibilities of how Artificial intelligence will revolutionize education and workplaces, how we research and develop such AI products, and how educators and employers should work with such products.

Challenges Faced By Engineering Industries in Nigeria

Matthew N. O. Sadiku, Janet O. Sadiku, Uwakwe C. Chukwu
Department of Electrical & Computer Engineering, Prairie View A & M University, Texas, USA;
Juliana King University, Houston, Texas, USA;
Department of Engineering Technology, South Carolina State University, Orangeburg, South Carolina, USA

Abstract - Engineering is the discipline that relates to the development and application of materials for the benefit of mankind. Since Nigeria aims to be among top Africa and world economy, the demand for engineering skills is higher than ever before. The engineering industry has a significant role to play in the attainment of nations' goals and objectives. This paper discusses various challenges confronting engineering industries in Nigeria and the possible ways out.

The 21st International Conference on Embedded Systems, Cyber-physical Systems, & Applications (ESCS'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-ESCS
https://www.american-cse.org/csce2023/

Lightweight Direct Memory Access on FPGA Using AXI Protocol

Adam Kidwell, Matthew P. Horvath Jr., Darrin Hanna, Bryant Jones
Department of Electrical and Computer Engineering, Oakland University, Rochester Hills, Michigan, USA;
Intrepid Control Systems, Troy, Michigan, USA

Abstract - This paper presents a lightweight AXI DMA Controller architecture useful for embedded systems that do not require fully featured DMA controllers. Simulation is accomplished with VUnit, and implementation results are obtained on a Xilinx XC7Z010CLG400-1 FPGA. When compared with Xilinx.s AXI DMA controller with the same configuration, the presented controller utilizes between 16 and 82% fewer resources with comparable speed.

Improving Energy Efficiency of RISC-V Processor for Sensor Node

Apurva Panchal, Hakduran Koc
College of Science and Engineering, University of Houston - Clear Lake, Houston, Texas, USA

Abstract - Sensor nodes are developing from low-bandwidth data intermediaries to endpoint data analysis devices that execute complicated algorithms on high-bandwidth data streams coming from smart sensors. These devices must work within strict performance and power constraints. This calls for careful designing of processors which are energy efficient to ensure long lasting performance of sensors. In this paper, we aim at optimizing a RISC-V processor design in order to achieve an energy efficient processor which can be used in sensor nodes. The design is implemented using various RTL power reduction methods. Sleep unit is designed to control the core clock. Simulation is performed to understand the functionality of the design. Power and timing analyses are also presented.

EFI: Cache Replacement Policy using Eviction Frequency Integration

Joe Tran, Byeong Kil Lee
Department of Electrical and Computer Engineering, The University of Colorado,
Colorado Springs (UCCS), Colorado, USA

Abstract - Modern processors have last-level caches (LLCs) that are unsuccessful in capturing large amounts of data from memory-intensive workloads. The replacement policy of the LLC is a significant factor in the performance potential of the processor since its selection determines the frequency of off-chip misses for the processor. Existing processors utilize variations of the leastrecently-used (LRU) and not-recently-used (NRU) replacement policies for victim selection. However, due to the nature of LRU and NRU, they are only efficient for workloads with high temporal locality and working sets smaller than the available cache. When we look at Belady.s MIN algorithm, the optimal replacement policy, we observe that LRU is not sufficient in mirroring this efficiency for thrash and scan-resistant workloads. We find that in prior work, there exists an inverse relationship between the expected number of hits of a cache block and its reuse distance. Building off of this observation, we propose that integrating this characteristic with the eviction frequency of a line from the cache can improve the efficiency of LLCs and as a result give us an effective replacement policy. We suggest an addition to the hit-count-based selection policy on top of low-cost replacement policies to offer performance improvement to the replacement policy of the last-level cache. The experiment results show that 2.3% performance improvement over the Expected Hit Count Replacement Policy, which offers 12.9% cache performance improvement over Baseline LRU.

Low-complexity Hardware Architecture of an H.264-based Video Encoder for FPGAs

Azam Tayyebi, Darrin Hanna, Bryant Jones
Department of Electrical and Computer Engineering, Oakland University, Rochester Hills, Michigan, USA;
Intrepid Control Systems, Troy, Michigan, USA

Abstract - In this paper, we present an architecture for a scalable, efficient, real-time intra H.264 video encoder implemented on an FPGA. Our architecture was designed to achieve a throughput of up to 2.3 Gbit/s using a parallel and pipelined architecture described in VHDL. All modules in the architecture are optimized to utilize minimum hardware area. A parameterized encoding system and flexible architecture is proposed to provide the ability to achieve different compression ratios ranging from 1.4 to 2 with varying size and power requirements. As a baseline, with no compression, the encoder required hardware resources equivalent to 18K logic gates. This work experimented with compression ratios up to 2 which required an equivalent of 31K logic gates. The encoder performs at frequency ranges of 159-183 MHz.

An Efficient Sparse Neural Network Accelerator for Low-cost Edge Systems

Kyubaik Choi, Gerald E. Sobelman
Department of Electrical and Computer Engineering, University of Minnesota, Minneapolis, Minnesota, USA

Abstract - It is difficult to achieve real-time performance for a deep neural networks (DNN) in low-profile systems such as edge computing or Internet of Things (IoT) due to the large amount of computations that are required. A commonly used activation function is ReLU (Rectified Linear Unit) due to its simplicity and good performance. A key characteristic of the ReLU function is that it produces sparse vectors within the neural network. In this paper, we propose an optimized DNN accelerator for sparse vector multiplications. In particular, we propose a squeezer unit to detect zeros and then skip feeding that data to the processing elements. In addition, we design a dynamic scheduler to efficiently allocate multiple neuron computations to the processing elements. With our architecture, we achieve reduced hardware resources compared to prior work due to a 62% reduction in the number of required index bits. We also achieve 8.6% better processing speed than for a system without the acceleration.

Drone2Drone (D2D): A Search and Rescue Framework Module for Finding Lost UAV Swarm Agents

Abhishek Phadke, F. Antonio Medrano, Tianxing Chu, Chandra N. Sekharan, Michael J. Starek
Department of Computing Sciences, Texas A&M University-Corpus Christi, Texas, USA

Abstract - The use of UAV (Unmanned Aerial Vehicle) swarms is becoming more widespread due to the reduced costs of UAVs and their ability to accomplish tasks more quickly and effectively as a group rather than individually. Advancements in aircraft design and control, communication topologies, and battery systems have made coordinated UAV swarms possible. One of the most important functions of UAV swarms is SAR (Search and rescue). In this function, UAVs use onboard sensors to locate missing targets, such as survivors after a disaster. SAR techniques have expanded to include other areas, such as wildlife tracking, forest fire detection, and locating survivors at sea. However, any disruptions in the dynamic work environment of UAV swarms can result in the loss of swarm agents themselves. Existing swarm deployments have limited capabilities to track, locate, and rescue disabled agents of the swarm during mission progress. In this study, we present a module of our Drone2Drone framework, which is a critical framework dedicated to rescuing disabled agents of the swarm during mission progress. Although computer simulations will be our primary method of performance validation during later stages of framework development, in this study, we use DJI Tello EDU drones to prototype smaller components of the framework.

Towards a Virtual Cloud-based Smart Factory Testbed for Cybersecurity

Rohit Mandalapu, Hazem Said
School of Information Technology, University of Cincinnati, Cincinnati, Ohio, USA

Abstract - Today's industrial equipment is connected over a network to communicate with external systems and make decisions without human intervention, making it vulnerable to cyberattacks and showing the importance of research. This study explored the implementation of a cloud-based virtual testbed for a smart factory for cybersecurity testing and research. As a first step, this paper reports on developing an environment with one programmable logic controller (PLC) simulating a conveyor belt setup. The study examined different virtualization platforms and network designs. In addition, it executed a denial-of-service attack and identified its signature indicators. The study found that VMware Workstation Pro is the most suitable virtualization platform and that network input and output are the DoS attack's signature indicators.

Smart Passive Ambient Control for Indoor Vertical Farming by Simulation

Rafiqul Islam, Bahram Asiabanpour
School of Engineering, Texas State University, San Marcos, Texas, USA

Abstract - The objective of this research is to design and develop a smart passive temperature control system for indoor vertical farming by simulation and validation of the simulation by empirical study. Passive temperature control (PTC) is defined as the process of controlling or manipulating the temperature of a system with natural heat transfer like conduction, convection, radiation, etc. and the purpose is to reduce energy consumption. Indoor vertical farming (IVF) can be defined as the practice of growing produce stacked one above another in a closed and controlled environment.[1] This research will focus on how the temperature of the water for growing plants can be controlled using PTC. A computational fluid dynamics (CFD) model is developed to simulate the effect of outside temperature on the indoor air and water and the model is validated with experimental data. The purpose of the CFD model is to simulate the temperature of indoor air and water and any given time of the day and year which will save time and equipment required for actual data collection and to find the optimum period for transferring water from inside to outside as a PTC process and found that, with the combination of material or methods for conduction, convection & radiation can help to balance the indoor temperature from the external ambient temperature. Later, an empirical study is done based on observation and measurement of temperature data of air and water both from inside and outside of the shipping container to validate the simulation. After validating the simulation, a design of experiment (DOE) is developed to find the optimum condition for saving energy and keeping the farming environment livable for plants.

Autonomous Object Detection and Landing with Parrot Rolling Spider in Matlab Simulink

Leonard Traeger, Omar Faruque
Information Systems, University of Maryland, Baltimore County, Maryland, USA

Abstract - Object detection using autonomous drones is a typical application of computer vision and robotics. More than detecting the objects with the UAV might be required since many real-life applications require some physical interaction with the detected objects. This paper proposes a generic framework implementing algorithmic code-blocks for a Flight Control System executing an autonomous object search, detection, and landing task on the detected surface. We adapt this framework in Matlab for simulation, deployment, and fine tune the final landing parameters on the supported Parrot Rolling Spider.

Low-Cost High-Precision Contactless Robotic Gripper

Thomas Adams, Samuel Bettencourt, Keily Valdez Sereno, Vida Vakilian
Department of Computer Science, California State University Channel Islands, Camarillo, California, USA

Abstract - Within the field of designing robotics systems for the purpose of handling and transporting materials, there exist certain limitations and restrictions when it comes to dealing with materials that are extremely small or delicate. One solution is using ultrasonic forces to trap small objects in a suspended state. In the last few years, techniques have been created for controlling levitating items positioned on surfaces that are either acoustically transparent or reflective. This paper introduces the design of a four-degree-of-freedom ultrasonic levitation robotic system that is more readily accessible and feasible for precision handling of objects that are of very small or delicate natures. Simulation results anticipate transportation of levitated solids and liquids is achievable with densities of 0.5 to 3.6 g/cm^3.

Improving Automated Plant Seeding Through Design and Development of Automated Seeder and Route Optimization

Lance Simon, Bahram Asiabanpour, Mark Summers
Ingram School of Engineering, Texas State University, San Marcos, Texas, USA

Abstract - The focus of this study was to improve the planting operation through the automation of planting seeds and adding flexibility in the use of the machine for a variety of plants. In this project, an open-sourced robotic seeder (The FarmBot Genesis) was utilized, modified, and improved to enhance the original function of planting individual seeds in their designated location without a defined travel path. The repetitive function of transferring individual seeds in conjunction with the undefined travel path made the FarmBot inefficient. The improvement enabled the system to load and carry several seeds simultaneously and plant individual seeds in the designated locations. The system, by utilizing the nearest neighbor algorithm, optimizes the sequence that the FarmBot travels in a path to ensure that it visits all predetermined locations in the least amount of distance and time. The applied modification saves a significant amount of planting time/energy/cost and gives the flexibility to use the system for a variety of plants with different seed shapes and sizes. The mechanism can be expanded to operate all functions of farming including watering, sampling, monitoring, and harvesting. The system is expandable for any scale from lab prototype to fullscale industrial farming.

False Data Injection (FDI) Modeling and Detection in Global Positioning Systems (GPS) for UAS Environments

Jaya Preethi Mohan, Prakash Ranganathan
School of Electrical Engineering and Computer Science, University of North Dakota, Grand Forks, North Dakota, USA

Abstract - Global Positioning System (GPS) is critical for precise navigation of Unmanned Aerial Systems (UAS) trajectory planning. GPS signals may experience sudden or intermittent interference due to multi-path fading, spoofing, jamming, or difficult terrain environments leading to “dropout” phenomena. The paper models one such dropout scenario by injecting false data into original GPS timestamps representing interference. We consider “dropouts” as insertion of false GPS data (FDI). The end goal is to investigate the performance of FDI detection rates using several clustering algorithms: 1) Density-Based Spatial Clustering, 2) Ordering Points to Identify the Clustering Structure, 3) Gaussian mixture models, and 4) Hierarchical Clustering. Multiple grouping indices such as Silhouette Coefficient (SC) and David Bouldin.s (DB) Index were used to evaluate normal vs. abnormal points grouping. The preliminary results are mixed indicating that the OPTICS clustering algorithm outperforms for scenario 3, while scenarios 1 and 2 are challenging for any clustering methods to capture the injected points. Also, SC is a good metric over DB index to evaluate the cluster groups. This paper also investigated non-clustering approach such as Isolation forest, which seems to be a better candidate for such FDI detection.

Component-Based System Model Design of Multiple-Fault Injection Framework for DCV and Heating Systems

Ali Behravan, Bahareh Kiamanesh, Siddharth Bhandari, Roman Obermaisser
University of Siegen, Siegen, Germany

Abstract - Heating, Ventilation, and Air Conditioning (HVAC) systems are crucial for providing thermal comfort and maintaining indoor air quality. Demand-Controlled Ventilation (DCV), as a type of HVAC system, optimizes fresh air supply based on occupant's demands while integrating with the heating system to ensure comfort and energy efficiency. However, modeling DCV and heating systems for large-scale buildings poses challenges due to complexity, limited data availability, time consumption, resource intensity, and lack of standardization. Moreover, system complexity makes HVAC systems prone to faults, resulting in increased energy consumption, discomfort, poor indoor air quality, and infrastructure risks. A generic Fault Injection (FI) framework that supports multiple fault injections in realistic scenarios is an effective tool for evaluating the reliability of these complex systems. The design challenge here is the complexity of interconnected systems and components. This paper resolves this challenge by introducing a component-based system model design integrated with an automatic multiple-fault injection framework. This approach enhances the efficiency of DCV and heating systems, reducing energy waste, improving indoor air quality, and lowering maintenance costs. The validation phase of the research demonstrates an accurate component-based system model configuration ensembles for multiple fault scenarios at different attributes.

Optimize Matrix Multiplication Utilizing openCL FPGA Kernel

Bassam Shaer, Timothy Stewart
Dr. Muhammad Harunur Rashid Department of Electrical & Computer Engineering,
University of West Florida, Fort Walton Beach, Florida, USA

Abstract - This paper presents an analysis of matrix multiplication utilizing OpenCL, optimization techniques, and a quantitative comparison of matrix multiplication between FPGA and CPU performance based on power consumption, speed and efficiency. We show how algorithm improvements, coupled with specific compilation optimization can improve matrix multiplication performance. Furthermore, we demonstrate the potential for performance gains by utilizing onboard FPGAs to off-load machine learning workloads on edge devices.

Designing a Low-Cost Microcontroller-Based Rover for Microplastic Detection Using Deep-Learning Image Detection and Near-Infrared Spectroscopy

Kevin Zhang, Chris Cheng Zhang
University of British Columbia, Canada;
Canada Youth Robotics Club, Canada

Abstract - This research article presents a potential solution to microplastic accumulation, with a specific focus on nurdles, which are small-rounded plastics used extensively in manufacturing processes. These nurdles have emerged as a major contributor to the buildup of smaller microplastics in our environment. To combat this problem, this study presents an innovative solution in the form of a low-cost rover designed to track and detect these particles in shoreline areas. The research methodology employed in this study encompasses several key components. First, a comprehensive dataset consisting of 150 images depicting polyethylene plastic pellets buried under sand, varying in lighting conditions and quantities, is used to train a compressed object detection model. This model, when executed on the Raspberry Pi 3, enables efficient and accurate identification of microplastic particles. To further enhance the detection capabilities, an AS7263 NIR sensor is used in conjunction with the Arduino BLE Sense 33. This integration allows for the measurement of near-infrared reflectance in polyethylene, thereby serving as an additional metric for microplastic detection. By analyzing the spectral characteristics of polyethylene, the system gains further insights into the presence and concentration of microplastics.

The 19th International Conference on Foundations of Computer Science (FCS'23: July 24-27, 2023; Las Vegas, USA) https://american-cse.org/csce2023/conferences-FCS

A Proof Theoretic Exploration of Mathematical Induction in Computational Paradigm

Ahmed Tarek, Ahmed Alveed, Ahmed Farhan
Engineering, Physical, & Computer sciences (EPCS) Department, SET Division,
Montgomery College, Rockville, Maryland, USA;
The Wharton School, University of Pennsylvania (UPenn), Philadelphia, Pennsylvania, USA;
School of Medicine, The Johns Hopkins University (JHU), Baltimore, Maryland, USA

Abstract - In the world of computing, there exists a wide variety of direct and indirect proof techniques for proving new results and propositions. Among the persisting proof techniques, Mathematical Induction (MI) stands out to be a powerful one for proving propositions, theorems, as well as the new results. MI is strongly founded upon the basis, and the inductive hypothesis. The Mathematical Induction techniques may be broadly classified into three categories, such as, the Weak Induction, the Strong Induction, and the Structural Induction. In this expository paper, several variations of mathematical induction technique are explored, and the specific computational scenarios are considered where a particular variation of mathematical induction emerges as more beneficial compared to other variations. Also, the essential differences among the mathematical induction variations are taken into account. The computational strategies for mathematical induction, recursion, and logical deduction are also analyzed, and the relationships among variations of recurrence relations, and mathematical induction are being established. From an application perspective, recurrence relations, and mathematical inductions are considered together in a single framework for analyzing codewords over a given Alphabet.

Integrating Process Mining with Probabilistic Model Checking via Continuous Time *Markov Chains*

Fawad Ali Mangi, Guoxin Su, Minjie Zhang
School of Computing & Information Technology, University of Wollongong, Australia

Abstract - Process mining represents a methodological approach that facilitates the in-depth analysis of business operations with the aim of revealing significant insights pertaining to their efficacy, efficiency, and regulatory compliance. In a seamless business setting, it's essential to evaluate, refine, and confirm these process models. Traditional methods for checking these models excel in validating their accuracy but cannot handle the inherent probabilistic and real-time behavior of these models. To address this, our research paper presents a new methodology that enhances probabilistic model checking (PMC) for process models. This is made possible by leveraging continuous time Markov chains (CTMCs) as the basic model. This amalgamation creates an inclusive analytical structure that improves the reliability and accuracy of process mining results, thereby enabling corporations to derive deeper insights from their processes and make decisions based on data.

Context Free Grammars with Variables are Universally Powerful

Peter M. Maurer
Department of Computer Science, Baylor University, Waco, Texas, USA

Abstract - DGL is a tool for generating test data for hardware and software, and for generating simulation input. DGL is based on probabilistic context-free grammars. To facilitate generating data in many different forms, many different types of productions are available. One can change probabilities and selection methods and perform other tasks. The most powerful type of production is the variable. Variables have a single right-hand side alternative that can be changed at run time. Combining variables with ordinary, unenhanced variables makes DGL universally powerful. This paper provides a proof of this fact. The general approach is described and is clarified using a specific, tested example.

The 19th International Conference on Frontiers in Education: Computer Science & Computer Engineering (FECS'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-FECS
https://www.american-cse.org/csce2023/

CS1 and CS2 using Scala 2: A Decade of Experience

Mark C. Lewis, Lisa L. Lacher
Department of Computer Science, Trinity University, San Antonio, Texas, USA;
Department of Computer Science, University of Houston Clear Lake, Clear Lake, Texas, USA

Abstract - In 2010, the desire to teach CS1 and CS2 in a single language led us to the adoption of a young language, called Scala, that supported both scripting and large-application object-orientation. We look at the original motivations for this choice and how it has worked. We also consider implications of this choice on later courses and look forward to how we expect language updates in Scala 3 to impact its use in this context.

An Empirical Investigation on Technology Acceptance of AI-enabled Clinical Decision Support Systems in Nursing Practice

Konstantina Karathanasopoulou, Christina Alexandropoulou, Ilias Panagiotopoulos, George Dimitrakopoulos
School of Digital Technology, Department of Informatics & Telematics, Harokopio University of Athens, Greece

Abstract - As health information technology continues to attract immense research interest, leveraging more clinical information and adapting to changes in quality care, nursing informatics is assuming a larger role in nursing practice. Using knowledge discovery in data approach with Artificial Intelligence (AI) tools, one can implement Intelligent Clinical Decision Support Systems (i-CDSS). The acceptance on behalf of nursing personnel of i-CDSS is crucial for merging technology into healthcare and thus optimizing clinical care quality. This paper aims to investigate the factors that affect nurses' intention to adopt and use i-CDSS in their workflow. To do so, it carries out a questionnaire-based survey among registered nurses, based on an appropriate extension of the Unified Theory of Acceptance and Use of Technology (UTAUT). The results of the analysis provide evidence that perceived financial cost, perceived reliability/trust, and social influence are useful predictors regarding nurses' intention to adopt i-CDSS. As such, the contribution of the paper lies in facilitating the efforts of researchers and health services in providing proper strategies that will enable the implementation of i-CDSS in nursing practice.

Equity in the Preparation of Students for Software Engineering Coding Interviews: ChatGPT as a Mock Interviewer

Marlon Mejias, Zef Vargas, William (Ted) Edwards, Gloria Washington, Legand Burge, Dale-Marie Wilson, Luce-Melissa Kouaho
University of North Carolina, Charlotte, North Carolina, USA;
Howard University, Washington, DC, USA

Abstract - There exists structural inequities in the tech industry's software engineering interview ecosystem. These inequities are often simply ways of doing business that have perpetuated a gap between those who have access and privilege and those who do not. This gap may not be the result of deliberate bias but they are systematically disadvantageous and may exclude people by race, gender, ethnicity, socioeconomic status. This experience report discusses identified issues of equity in the software engineering interviewing technical interview process and how Large Language Models like ChatGPT are being used to address these gaps.

ChatGPT: A Game-Changer for Embedding Emojis in Faculty Feedback

Ethan Kupershtein, Yulia Kumar, Anjana Manikandan, Patricia Morreale, J. Jenny Li
Kean University, Union, New Jersey, USA

Abstract - This study explores the potential of integrating emojis, and digital pictographs, into faculty feedback to augment student learning outcomes. This additional layer of expressiveness, encouragement, and involvement adds a personal touch to the often distant and virtual student-educator communications, fostering motivation. The study focuses on the impact of emojis on the learning process within the scrutinized Computer Science (CS) Department. Capitalizing on the capabilities of OpenAI's Large Language Model (LLM) ChatGPT-4, its Application Programming Interface (API), and associated tools and thirdparty plugins, a system that translates text into corresponding emojis and vice versa has been developed. The proposed application offers direct benefits to educators by simplifying the provision of detailed and extensive feedback to students. The primary research question is: Can the appropriate use of emojis, matched with the sentiment of the feedback text, contribute to enhanced student learning outcomes, higher retention rates, and boost the reputation of the educators providing it? Two surveys on the impact of emojis across selected course sections were conducted to answer the question: a pre-survey and a post-survey involving 175 active participants. The results were analyzed, and it was concluded that integrating emojis in faculty feedback, particularly when grading student work, could potentially enhance student learning outcomes and their overall course experience.

The Just-In-Time Adaptive Artificial Augmentation Capstone Project

Chad Mello, Troy Weingart
Department of Computer & Cyber Sciences, United States Air Force Academy,
Colorado Springs, Colorado, USA

Abstract - Interest in assistive technologies (AT) and artificial intelligence (AI) supporting decision making and the performing difficult or dangerous tasks is experiencing continued and rapid growth. Systems of this nature have a wide range of use-cases spanning defense and commercial domains, such as aviation, air traffic control, decision support systems, emergency management, command and control, threat assessment, route planning, autonomous multi-vehicle mission oversight, power grid operations, and machine and heavy-machine operations. In line with our ongoing efforts to integrate applied, experiential learning programs into post-secondary undergraduate STEM curriculum, we propose a capstone program whereby undergraduate engineering and computer science seniors can gain valuable insight and skills by contributing to the creation of a just-in-time adaptive artificial augmentation (JIT-A3) system. The purpose of this program is to provide undergraduate senior students with a real-world, project-based experience that encompasses the complexities and realities they are likely to encounter in demanding commercial and government jobs.

Analyzing ChatGPT's Aptitude in an Introductory Computer Engineering Course

Sanjay Deshpande, Jakub Szefer
Computer Architecture and Security Lab, Department of Electrical Engineering,
Yale University, New Haven, Connecticut, USA

Abstract - ChatGPT has recently gathered attention from the general public and academia as a tool that is able to generate plausible and human-sounding text answers to various questions. While recent works have explored the use of ChatGPT in the context of humanities, business school, or medical school, this work explores how ChatGPT performs in the context of an introductory computer engineering course. This work assesses ChatGPT's aptitude in answering quizzes, homework, and laboratory questions in an introductory-level computer engineering course. This work finds that ChatGPT can do well on questions asking about generic concepts. However, predictably, as a textonly tool, it cannot handle questions with diagrams or figures, nor can it generate diagrams and figures. Further, also clearly, the tool cannot do hands-on lab experiments, breadboard assembly, etc., but can generate plausible answers to some laboratory manual questions. One of the key observations presented in this work is that the ChatGPT tool could not be used to pass all components of the course. Nevertheless, it does well on quizzes and short-answer questions. On the other hand, plausible, human-sounding answers could confuse students when generating incorrect but still plausible answers.

Project Xander: Hands-on Skills for Cybersecurity

Gregory B. White
Computer Science Department, Center for Infrastructure Assurance and Security,
The University of Texas at San Antonio, Texas, USA

Abstract—The value of hands-on lab experiences has long been recognized in cybersecurity education. Additionally, the more the labs mirror what the students will experience when they enter the workforce the more valuable the labs will be for them and for their future employers. One such experience a number of universities have tried is to have students conduct cybersecurity assessments on local organizations. The problem many quickly run into is a lack of ability for the instructor to mentor teams as class sizes increase. Project Xander, begun at The University of Texas at San Antonio has addressed this problem allowing for hands-on experiences for classes in excess of 100 students. The approach has also resulted in a multi-semester program that includes risk management, security audits, and penetration testing experiences for students in a series of courses. An additional benefit is the improvement of security in local organizations that do not have the resources to contract with vendors to provide cybersecurity services. The program is a joint project between UTSA and the MITRE corporation and can be duplicated at any college or university who wants to become part of Project Xander.

Promoting K-12 Computer Science Education: A Computational Thinking Partnership Project

Jiang Li, Jennifer Cuddapah, Marisel Torres-Crespo, Ann Stewart, Anthony Williams
Department of Computer Science and Information Technology, Hood College, Fredrick, Maryland, USA;
Department of Education, Hood College, Fredrick, Maryland, USA;
Department of Mathematics, Hood College, Fredrick, Maryland, USA;
The Beacon House, Hagerstown, Maryland, USA

Abstract - This paper describes a collaborative project that was conducted to promote Computer Science (CS) and Computational Thinking (CT) Education among pre-service and in-service teachers as well as Teacher Preparation Faculty. More than 40 pre-service and in-service teachers and 9 Education faculty members from three schools participated in a learning experience designed to address the K-12 Computer Science Framework [1] and Maryland.s K-12 Computer Science Standards. The collaboration was designed to develop CS/CT knowledge, understanding, skills and application among participants. In phase one of the project, pre-service and in-service teachers explored CS educational hard/software platforms and used open-source sites such as Sphero Edu, Wonder Workshop, Scratch and Code.org. They envisioned how activities apply to K8 classrooms, and they worked

collaboratively to design a problem-based project for their own students. In phase two of the project, teacher preparation faculty were trained to integrate CS and CT into several preservice courses as well as build into the coursework sequence meaningful experiences that would expose future teachers to CS and CT knowledge and skills so that they could then incorporate these into their own K-8 lesson plans. Project evaluation included formative and summative assessments to examine changes in content and pedagogical knowledge. The feedback from the participants shows positive results.

The Role of Assessment Items in Determining Student Performance in CS1

Permanand Mohan
Department of Computing and Information Technology, The University of the West Indies, St. Augustine, Trinidad and Tobago

Abstract - Studies of CS1 pass rates in universities around the world typically show a pass rate of about 70% for large classes and 80% for small classes. To a large extent, the pass rate depends on the assessment items that are used to evaluate student performance. This paper examines the assessment items in CS1 courses taught at four tertiary-level institutions to determine how the assessment items measure up to Bloom's Taxonomy of Educational Objectives in the Cognitive Domain. A link is then made between assessment items and student performance, and it is hypothesized that assessment items that target the lower levels of Bloom's Taxonomy resulted in better performance at the institutions studied. Consequently, the paper argues that whenever a research study presents an intervention in CS1, the assessment items which are used to assess performance should also be presented, to give a more complete picture of the intervention. The paper also proposes a way to design CS1 assessment items that are aligned to Bloom's Taxonomy.

A Study of Course-based Undergraduate Research Experiences and the Challenges and Opportunities for Computer Science

Paula Lauren
Department of Mathematics and CS, Lawrence Technological University, Southfield, Michigan, USA

Abstract - A course-based undergraduate research experience (CURE) involves the incorporation of research at the undergraduate level, typically as a replacement to the lab portion of the course. A motivational driver for incorporating research in the classroom is to ensure greater participation and retention of undergraduate students in science, technology, engineering, and mathematics (STEM). Studies have also shown that students who engage in undergraduate research experiences gain a better understanding of the scientific process and are more likely to pursue graduate studies. The CURE model is regarded as a pedagogical paradigm that originated in the life sciences, resulting in numerous research studies and support networks. Computer science can be regarded as a singular discipline and as a discipline that spans all areas of science, technology, engineering, and mathematics. Interestingly, it was the STEM Education Act of 2015 that resulted in the official inclusion of computer science to the definition of STEM disciplines. This paper provides a review of undergraduate research experiences in STEM from the research literature, identifying the unique challenges and opportunities associated with incorporating CUREs into an undergraduate computer science curriculum. In addition, this paper entails an evaluation of an upper-division Artificial Intelligence (AI) course incorporating course-based research.

Reinforcing Online Learning Outcomes through Scaffolded Formative Assessments

Alan Bowen-James, Mary Ruth Freislich
Crown Institute of Higher Education, Sydney, Australia;
School of Mathematics and Statistics, University of New South Wales, Sydney, Australia

Abstract - Learning outcomes among accounting students at two independent Australian higher education institutions (1 and 2) were compared. The two organizations ran introductory accounting subjects for business students which were very close in coverage and assessed by similar methods at similar levels of difficulty. Teaching differed in one important way. Institution 1 supplied additional teaching material consisting of sequences of related accounting tasks corresponding to each topic, with online explanatory support that initially was strong, and reduced for later tasks in each sequence, in a pattern that approximated true scaffolding. The purpose was to help students to get started on learning by doing, so that formative experience could be available to them, and hence also to reduce the anxiety about quantitative tasks often found to impede learning in accounting subjects. Assessments requiring problem-solving under examination conditions in each institution afforded six pairs of tasks that were comparable to a degree that permitted grading using a common marking scheme. Students. responses were marked using a common scheme, and mean scores on each pair were compared. Results favoured Institution 1 (scaffolding provided) in four of the six comparisons, with no significant differences in the remaining two. A composite score, which could be calculated for a subset of the tasks, also favoured Institution 1.

Analysis of Plagiarism via ChatGPT on Domain-Specific Exams

Jinyoung Jo, Sean Choi
University of California (UCLA), Los Angeles, California, USA;
Santa Clara University, Santa Clara, California, USA

Abstract - This work presents a case study, linguistic analysis and potential prevention methods on the use of large language models (LLM) for generating solutions for exams on cloud computing course that require domain-specific knowledge. The study involves analyzing the responses of three groups of students: a group who used ChatGPT to plagiarize solutions, another group who referred to external non-LLM resources (e.g., web search) to plagiarize solutions, a control group who generated solutions without any external assistance. Results show that solutions from groups that participated in plagiarism tend to be lengthy, use uncommon words, and are similar to each other compared to human-generated solutions. This study not only shows that it is possible to generate legitimate solutions for exams that require extensive domain-specific knowledge using ChatGPT, but also shows some potential signals one can use to detect plagiarism, thus providing potential of promoting academic integrity by curbing unethical use of AI in academic settings.

Teacher's Perspective for the Implementation of Social Networks in Virtual Education at a Higher Level

Carmita Ines Suarez, Segundo Moises Toapanta Toapanta, Eriannys Zharayth Gomez Diaz, Angel Ernesto Huerta Velez, Carmen Ines Huerta Suarez, Edith Elizabeth Ordonez Chavez, Marcelo Zambrano Vizuete
Research Department, Instituto Tecnologico Superior Ruminahui, Sangolqui, Ecuador

Abstract - A virtual social network is a space for essential interaction in the life of Internet users, but its capacity as a complement to educational activities at the University has not been widely explored. In this research, we want to propose the use of a theoretical model in which collaborative, cooperative, participatory and interactive work between teachers and students is encouraged, as well as to identify the factors that influence the acceptance of social network applications for use in contexts of higher education institutions. The objective is to carry out an analysis from the teaching perspective on the implementation of social networks in education at the higher education level. The deductive method and exploratory research were used to analyze the problems that arise from properly applying social networks in education. It turned out indicators defined by other authors, the main actors and a conceptual model based on the use of social networks in education. It was concluded that implementing social networks in the

educational field is of great help in the teaching and learning process since collaborative and cooperative work is encouraged without leaving aside the existence of interaction between the benefited peers.

Examining Effective Student Support Services for STEM Graduate International Student Success

Vedika S. Salunke, Kevin J. Bazner
Department of Educational Leadership, Texas A&M University - Corpus Christi, Texas, USA

Abstract - The United States of America is one of the leading countries hosting international students from around the world. International students in the United States come from different countries with different cultural, religious, academic, and professional backgrounds. To ensure their holistic success, it is important to recognize their unique needs and challenges. These students have nuanced experiences which may require tailored support services. Further, the needs of graduate students are different than those of undergraduate international students. And the experiences and challenges of those in science, technology, engineering, and mathematics (STEM) degree programs also differ as they have their unique needs. This paper discusses how effective student support services could allow for smooth transition and a meaningful experience for STEM graduate international students.

Undergraduate Research Experience Impact on Retention in an Electrical Engineering and Computer Science (EECS) Program

Hassan Salmani, Mohsen Mosleh, Gloria Washington
Department of Electrical Engineering and Computer Science, Howard University, Washington, DC, USA;
Department of Mechanical Engineering, Howard University, Washington, DC, USA

Abstract - Research has demonstrated the positive influence of Undergraduate Research Experience (URE) programs in Science, Technology, Engineering, and Mathematics (STEM) on students' educational journey and their development as scientists, ultimately aiding them in making informed career choices. However, traditionally, URE programs have primarily targeted junior and senior students who already possess disciplinary knowledge and exhibit a strong inclination to persist in STEM fields. This study aims to examine the effects of involving freshmen in the Industry-Research Inclusion in STEM Education (I-RISE) program, specifically in the disciplines of electrical engineering (EE) and computer science (CS), on student retention. The I-RISE program integrated research opportunities for undergraduate students with mentorship activities, facilitating the acquisition of relevant skills in applied computing and engineering techniques, research methodologies, and the attainment of internships. Analyzing the retention rates of three distinct cohorts of I-RISE participants over a span of three years revealed significantly higher retention rates compared to students who did not partake in the I-RISE program.

Wordification: A New Way of Teaching English Spelling Patterns

Lexington Whalen, Nathan Bickel, Shash Comandur, Dalton Craven, Stanley Dubinsky, Homayoun Valafar
College of Computing and Engineering, University of South Carolina, Columbia, South Carolina, USA

Abstract - Literacy, or the ability to read and write, is a crucial indicator of success in life and greater society. It is estimated that 85% of people in juvenile delinquent systems cannot adequately read or write, that more than half of those with substance abuse issues have complications in reading or writing and that two-thirds of those who do not complete high school lack proper literacy skills [1]. Furthermore, young children who do not possess reading skills matching grade level by the fourth grade are approximately 80% likely to not catch up at all [2]. Many may believe that in a developed country such as the United States, literacy fails to be an issue; however, this is a dangerous misunderstanding. Globally an estimated 1.19 trillion dollars are lost every year due to issues in literacy; in the USA, the loss is an estimated 300 billion [3]. To put it in more shocking terms, one in five American adults still fail to comprehend basic sentences [4]. Making matters worse, the only tools available now to correct a lack of reading and writing ability are found in expensive tutoring or other programs that oftentimes fail to be able to reach the

required audience. In this paper, our team puts forward a new way of teaching English spelling and word recognitions to grade school students in the United States: Wordification. Wordification is a web application designed to teach English literacy using principles of linguistics applied to the orthographic and phonological properties of words in a manner not fully utilized previously in any computer-based teaching application.

Evidence-Based, Low-Investment Writing Instruction for Computer Science Students

Andrew McAllister
Faculty of Computer Science, University of New Brunswick, Fredericton, New Brunswick, Canada

Abstract - Writing skills are valuable for computer science students and can impact their career trajectories. We describe our experience with a core undergraduate course for computer science students where the writing quality of student term papers has in the past frequently been poor. The desire was to improve the readability of student-produced documents without displacing other important topics taught in the course. We outline a multistep instructional process that involves evaluating the precise areas of writing where students in the class struggle, introducing students to a quick but effective approach for improving the readability of their compositions that is similar to debugging code, a motivational in-class editing competition, and assignments that reinforce learning. Anecdotal feedback indicates the approach produces worthwhile improvements in the readability of student writing submissions. We accomplish this while investing only a single lecture's worth of teaching time for writing instruction.

Enhancing Programming Education through Innovative Teaching and Embracing AI

Fernando Paniagua, James Braman
Computer Science & Information Technology, Community College of Baltimore County, Baltimore, Maryland, USA

Abstract - In this study, artificial intelligence based tools are discussed in the context of adapting teaching methods used in computer programming courses. Four innovative strategies are discussed: an academic integrity and ethics module, interactive programming examples, active learning strategies, and modified assessments. Preliminary approaches to these pedagogical changes and future research directions are provided.

Promoting Data Science to Underrepresented Communities Through the Use of Active Learning

Ed Pearson III, Velma Latson, Pamela Morgan
EE and CS Department, Alabama A&M University, Normal, Alabama, USA;
Department of Technology and Security, Bowie State University, Bowie, Maryland, USA;
Department of Education, University of Maryland - Baltimore County, Baltimore, Maryland, USA

Abstract - The HBCU-Data Science Next Steps (HBCU-DSNS) Program was funded by the South Big Data Hub, from their large S.E.E.D Award. This program was developed to address the lack of diversity in the data science pipeline. The HBCUDSNS Program had two components: the HBCU-DSNS Summer Academy and the year long REU (Research Experience for Undergraduate). This program was hosted in parallel at Alabama A&M University and Bowie State University. A comparison analysis revealed the participants in the HBCU-DSNS Summer Academy technical knowledge and skills were enhanced after attending the program.

Raising Student Awareness of Ethical Concerns Related to Artificial Intelligence in Computing Courses

Mel Akhimiemona, Alexis Brown, James Braman
Computer Science & Information Technology, Community College of Baltimore County, Baltimore, Maryland, USA

Abstract - There is value in educating computing majors about the moral concerns and biases that artificial intelligence systems raise. Future creators of these tools must comprehend AI's broader ramifications and ethical consequences despite the usefulness and the need to understand technical details. The authors highlight the importance of increasing student knowledge of potential ethical issues through instruction in computer courses by presenting numerous initiatives connected to this subject and student responses from a pilot study. Through a case study approach, students can learn to evaluate and analyze ethical concerns and engage in relevant discussions in the context of specific AI courses rather than creating comprehensive computer ethics courses.

Student Academic Advisement Augmented by Artificial Intelligence in the School of Computing and Information Technology, University of Technology, Jamaica

Tanya Graham, Horrett Scarlett, Shaula Edwards Braham
School of Computing and Information Technology (SCIT), University of Technology (UTech), Kingston, Jamaica

Abstract - In this paper we examine the possibility of using artificial intelligence (AI) to improve academic advisement of students within the School of Computing and Information Technology (SCIT) at the University of Technology, Jamaica (Utech). Described as one of the important challenges facing academics [1], academic advisement plays a vital role in student completion. All students at Utech are assigned academic advisors and encouraged to access advisors for advisement. Each faculty manages the process internally. Students are not mandated to seek advisement but are strongly encouraged to do so to allow them to make informed choices related to module selection, academic probation, grade forgiveness, etc. Within SCIT the rate of take up is less than desired resulting in some students going on academic probation, having to switch programs in some cases or failing out of their program. We will explore the automation of the academic advisement process by using AI to push relevant information to students related to their performance. The system will be coded to recognize common situations and contact the students providing information relevant to the situation and schedule an advisement session with the academic advisor (AA).

Engaging Students in Undergraduate Research: Teaching Through Design, Development, and Collaboration

Ethan Widener, Soumik Kundu, Rajwol Chapagain, Prakriti Sapkota, Martin Q. Zhao
Department of Computer Science, Mercer University, Macon, Georgia, USA;
Department of Engineering, Mercer University, Macon, Georgia, USA

Abstract - Incorporating quality research in a college setting where undergraduate teaching is the primary goal is always challenging. Engaging undergraduate students in research-related activities using a research theme to arrange teaching materials and assignments can be beneficial. Students can be introduced to basic concepts in lower-level courses and later become effective research assistants. This paper presents such practices at Mercer University's Computer Science Department from the students' and faculty's perspectives. The focus is on how to plan for breaking down the needs of research projects among student teams in various courses and provide collaboration opportunities between faculty and student researchers. Activities related to a research project supported through a NSF grant show satisfactory results for both students and the faculty advisor.

Reshaping the Educational Landscape of Tomorrow

Horrett Scarlett, Tanya Graham, Shaula Edwards Braham, Yulan Buchanan
School of Computing and Information Technology, University of Technology, Jamaica

Abstract - The use of Artificial Intelligence (AI) in industry is fast becoming commonplace. A survey of the Internet quickly reveals a suite of AI tools which can be used to increase business productivity and complete complex tasks in seconds. A good example of this is the use of AI to remove unwanted backgrounds from images and photographs, which is easily accomplished by going to the remove.bg url. The result is that a complex task, usually performed by skilled graphic artists, can now be accomplished by the average person without technical skills in graphics arts. This is indeed a game changer! While AI technologies, like this one, are poised to upend the current educational system, the one which seems to be getting most of the traction, in terms of the potential support that it lends to the teaching/learning process, is the Chatbot. In this paper we seek to demystify the underlying technologies and stimulate conversations about how best to leverage these tools in educational circles.

Exploring the User Experience and the Role of ChatGPT in the Academic Writing Process

Reem Alqadi, Afra Alrbaiyan, Nafla Alrumayyan, Naseebah Alqahtani, Abir Benabid Najjar
Department of CS, Qassim University, Qassim, Saudi Arabia;
Department of CS, Saudi Electronic University, Riyadh, Saudi Arabia;
Department of IT, IMSIU University, Riyadh, Saudi Arabia;
Department of SE, King Saud University, Riyadh, Saudi Arabia

Abstract - ChatGPT, an AI-powered chatbot, has garnered significant interest among early adopters for its ability to perform high-level cognitive tasks and produce text identical to humangenerated content. This raises concerns about the potential use of ChatGPT as a tool for academic writing. The objective of this study was to investigate the use of ChatGPT in academic writing tasks. A survey was conducted with 23 academic researchers to analyze their usage patterns, effectiveness, and satisfaction with ChatGPT. The results of the study suggest that ChatGPT has the potential to assist in idea generation for academic research and effectively summarize, paraphrase, and proofread texts and literature. However, opinions were mixed on the accuracy and reliability of ChatGPT in academic writing. Overall, the study highlights the potential benefits and limitations of using ChatGPT in academic writing and provides valuable feedback from researchers on their experiences.

Towards the Assessment of Basic Computational Thinking Skills using Syntactic Analysis Techniques

Antonio Gonzalez-Torres, Elliot Ramirez-Trejos, Lilliana Sancho-Chavarria,
Jose Navas-Su, Cesar Garita, Jorge Monge-Fallas
Computer Engineering, Costa Rica Institute of Technology, Cartago, Costa Rica;
School of Mathematics, Costa Rica Institute of Technology, Cartago, Costa Rica

Abstract - This article introduces an exploratory method for automatically grading programming exam questions using syntactic analysis. The target problem is the lack of a robust, scalable, and automated method to analyze computational thinking skills from source code written by elementary school students. The proposed method uses a variety of techniques to assess student responses, including analyzing the programming structure, programming correctness, and code execution based on certain parameters defined during the exercise specification. Analysis of the source code and evaluation of the answers to the exercises are carried out using high performance computing to improve the response time of the system. This preliminary work will contribute to a robust method for automated exam scoring, which is expected to assess and support the development of computational thinking among students.

Viable Program Assessment and Continuous Improvement

Venu G. Dasigi
Bowling Green State University, Bowling Green, Ohio, USA

Abstract - In a dynamic world, continuous improvement is very important for any academic program to ensure that its content, structure, and delivery change for the better over time. It is, however, a complex and daunting endeavor to put together an effective assessment and continuous improvement process that is also easy to implement and sustain. To address the challenge, this paper outlines important elements of an effective assessment and continuous improvement process and describes the various components of approaches that are frequently used, along with their major attributes. The paper concludes by pointing out how a viable assessment and continuous improvement process can be created by including a small set of the described elements in the assessment process.

INVITED TALK
The Perspective of Engineering Students about Blended Learning: Experience and Challenges

Amjad Gawanmeh, Mohammad Daradkeh, Shadi Atalla, Wathiq Mansoor
College of Engineering and IT, University of Dubai, Dubai, UAE

Abstract - Rather than receiving education face-to-face oncampus, blended learning, which combines both online education and face-to-face, emerged as an alternative solution. If implemented properly though, this mode of education can have its positive impact on the teaching and learning processes. Blended learning has become very popular post COVID pandemic, since many institutes wanted to continue using online education partially, where part of the course is delivered using conventional methods, and part using online tools. This paper is intended to study students. perceptive on blended learning involving engineering students at a private university in Dubai where this study was implemented, the researchers utilized a questionnaire focusing on students' feedback and perspective about blended learning. Based on the study's findings, most respondents were in favor of blended learning education, particularly in relation to its impact on student performance and convenience.

Using an Educational Robot for Teaching Environmental Issues in Early Childhood

Rafaela Georgiadou, Jim Prentzas
Department of Education Sciences in Early Childhood,
Democritus University of Thrace, Alexandroupolis, Greece

Abstract - During the last decades, ICT tools play an increasingly important role in early childhood education. Programmable toys are among the most popular tools used. A usual form of a programmable toy is a floor robot having direction buttons on it that control its movement in space. Programmable toys are used in early childhood education in order to introduce programming, robotics, math and linguistic concepts to young children. Programmable toys may be used in inter-disciplinary activities which may also involve other subjects such as environmental education and arts. In this paper, we present an inter-disciplinary approach using a programmable toy. The approach concerns environmental education and was implemented in a kindergarten classroom with twenty-three students. The results were promising as children improved their skills in programming the device and enhanced their environmental awareness.

A New Model of Online Team Teaching

Alireza Kavianpour
Computer Engineering Information Science, DeVry University, Ontario, California, USA

Abstract - Team teaching refers to a group of instructors, teaching in a team. The team-teaching approach allows for more of an interaction between instructors and students. It improves the quality of teaching because it takes various experts. approach at the same topic from different views. The presence of other instructors bring diversity and improve student-instructor relation. Overall, team teaching can be a highly effective approach to the higher education and reduces attrition if it is well-planned and executed.

Constructively Aligned Teaching of FPGA-SoC for Satellite Applications

Markus Plattner
Chair of Integrated Systems, School of Computation, Information and Technology,
Technical University of Munich, Germany

Abstract - A new teaching module has been established in 2018, based on the principle of constructive alignment. This module bridges the gap from electronics to aerospace engineering with the topic of building satellite computers based on FPGA-SoC devices. The module is oriented towards space industry and uses tools and technologies whereof space-qualified counterparts exist. The teaching module is centered around a group work project wherein students define their own topics. We share our thoughts on establishing this module and discuss the lessons learned from the past five years of teaching.

Plagiarism in Entry-Level Computer Science Courses using ChatGPT

Thomas Ryan Devine, Zachary Taylor, Cy Blair, Ethan Glenn
West Virginia University, Morgantown,, West Virginia, USA

Abstract—Recent advances in AI-generated code could potentially be plagiarized by students in entry-level programming courses. This paper examines the use of ChatGPT for generating code to solve a programming project for an entry-level computer science course. We analyzed 59 anonymized student samples along with 75 regular generations of ChatGPT code and 75 samples of ChatGPT code prompted directly for obfuscation. Results showed that 44% of ChatGPT-generated code did not compile in three or fewer manual fixes. Additionally, plagiarism detection software was not easily able to flag AI-generated code. However, the results showed that AI-generated code was more dissimilar in comparison to student-generated code, which could potentially aid in automated detection.

Approaches, Theory, and Role of Ethics in Computer Science and Engineering

Satpreet Singh
Director/CEO, Ardass Corporation, Manteca, California, USA

Abstract - The importance of ethics cannot be understated, particularly when protecting sensitive information and preventing potential breaches. It is necessary to stay up-to-date with the latest ethical practices to ensure a safer and more secure digital world for all. Ethics are crucial in establishing standards to ensure that privacy and security issues are protected and analyzed thoroughly. With the rapid growth of technology, it's essential to keep up with the latest ethical practices to safeguard sensitive information and prevent potential breaches. Implementing ethical standards can create a secure digital and safer world for everyone. This paper reviews various approaches, theories, types, issues, and roles of ethics within the field of computer science and engineering.

A Perspective on the Future of Higher Education

Mehdi Zargham
Department of Computer Science, University of Dayton, Dayton, Ohio, USA

Abstract - With the rapid advances in AI technology, it is becoming essential for higher education institutions to invest in significant reform over the next 10 to 20 years. Students will demand a more flexible, adapted, low-cost, and practical education. Also, the number of employers requiring knowledge units/modules rather than college degrees will increase significantly. Knowledge units/modules and certificates will substitute or complement current degree structures. Therefore, there is a need for a new and expanded educational model that educates students throughout their lives. This article suggests a reform in which students can access knowledge units anytime, anywhere, and at any pace. AI will be a primary factor in providing individuals with adapted personalized information. Society is headed toward a future education that is most outcome-driven, 24/7 availability, physical place independent, AI individualization, low-cost, and consists of concise knowledge units. The reform should provide problem-solving and creativity, communication skills (both interpersonal and intrapersonal), lifelong learning, industry partnership, advising and networking, and motivation for a self-learning attitude.

ChatGPT Implications on Higher Education: Educational Apocalypse or Educational Reboot? A Developing Countries Perspective

Olalekan Samuel Ogunleye
School of Computing and Mathematical Science, University of Mpumalanga, Mbombela, South Africa

Abstract - Artificial intelligence has disrupted many industries, and education is no exception. ChatGPT, a Large Language Model (LLM), has emerged as a promising tool for boosting learning experiences and altering traditional teaching methods in higher education. ChatGPT can be considered a combination of chat and a language model. This article investigates ChatGPT's effects on higher education, focusing on the opportunities and obstacles that may arise. This paper tries to answer whether ChatGPT will lead to an educational apocalypse or a much-needed reboot in higher education by assessing the influence of ChatGPT on educators' usage, student involvement, individualised learning, and administrative operations. In its conclusion, the article explores the prospects of incorporating ChatGPT into higher education and offers some ideas on how to do so successfully.

Exploring Emerging Researchers in LATAM: A Comparative Study with Elsevier Data and Machine Learning Techniques

Jesus Manuel Figueroa-Castillo, Laura Hervert-Escobar, Hector G. Ceballos-Cancino, Neil Hernandez-Gress
Tecnologico de Monterrey, Monterrey, Mexico

Abstract - Developing research in engineering, electronics, and electric fields is of utmost importance to gain insights into the current landscape and foster the development of new methodologies with potential for positive impact. This study aims to detect outstanding young researchers through the analysis of research performance data with machine learning techniques. These Rising Stars not only represent candidates with the potential to become experts in their respective areas but also provide valuable insights to universities and private sector industries seeking to keep track of their progress. This study is carried out with data from Elsevier-Scopus, which enables a comprehensive comparative study, shedding light on the most promising researchers in LATAM's engineering, electronics, and electric research domains.

Value-added Evaluation Strategies in C Language Teaching

Yingxin Xu, Kai Zhang, Peiyu Liu, Ran Lu
School of Information Science and Engineering, Shandong Normal University, Jinan, P. R. China

Abstract - Based on the reflection of the research status, this study designs a value-added evaluation model that conforms to the characteristics of "C Language Programming", and completes the teaching research and practice accordingly. Through the continuous vertical comparison of students' individual progress, the value-added level of students' knowledge and skills is reasonably quantified. Through the horizontal comparison between the experimental group and the control group, the effectiveness of value-added evaluation in improving the teaching effect of "C Language Programming" is verified. It is hoped to provide model reference and practical reference for C language teachers to explore value-added evaluation and provide new ideas for improving the teaching effect of "C Language Programming".

The Considerations of Teaching Computer Science Course - Advanced C++

Haiyi Zhang
Jodrey School of Computer Science, Acadia University, Wolfville, NS, Canada

Abstract - Programming Languages are fundamental courses for computer science students. C++ is one of these programming languages, and it is important because it is not only a procedure language but also it is an Object-Oriented language. In this paper, I share my classroom experiences and propose teaching strategies that in which theory and practice are combined. The paper begins by describing the importance of C++, followed by my key teaching experiences, including the criteria for choosing textbooks, selecting topics, providing quizzes and the implementation of practical coursework. Besides, I analysed the student's weaknesses and strengths and made the considerations of implementation language to implement software design patterns.

Curriculum for a New Four-Year Bachelor's Degree in Intelligent Systems Engineering

Ashu M. G. Solo
Maverick Trailblazers Inc., Wilmington, Delaware, USA

Abstract - This research paper proposes a curriculum for a four-year bachelor's degree in intelligent systems engineering with courses in knowledge-based systems, fuzzy logic, neural networks, evolutionary computation, machine learning, pattern recognition, computer vision, data science, control systems, intelligent control systems, robotics, digital signal processing, computer organization, digital electronics, software development, firmware development, calculus, algebra, physics, biology, etc. This degree will allow graduates to have a good understanding of all of the main branches of intelligent systems as well as relevant subjects in computer engineering.

Improvement and Evaluation of Pseudo Natural Programming Language

Toshiyuki Maeda, Masumi Yajima, Akiyoshi Wakatani
Department of Management Information, Hannan University, Japan;
Department of Economics, Meikai University, Japan;
Faculty of Intelligence and Informatics, Konan University, Japan

Abstract - This paper addresses a multi-lingual pseudo natural language programming for college education, to introduce basics of computer science and support to develop applications. In this paper, we at first explain the background of this research, and we discuss to improve the prototype and evaluation.

Design and Implementation of Cheating Prevention Features for Online Code Judge System

Soojung Kim, Junhong Park, Sujong Kim, Soobin Jeon, Dongmahn Seo
School of Computer Software, Daegu Catholic University, South Korea;
Department of Computer Software, Daegu Catholic University, South Korea

Abstract - During the COVID-19 pandemic, various online judge systems have worked in the learning field. In the case of the exam using the online judge system, there is a vulnerability that a user can try to cheat by exploiting the point that is accessible from outside the exam site. This paper deals with implementing the exam exit and management function to solve these problems.

Investigating Persuasive Behaviors Across Different Opinion Polarities: A Study of Video Messages

Zvetomira Svetleff
University of Nevada, Las Vegas, Nevada, USA

Abstract - Understanding what makes persuasive communication effective is a useful skill in many everyday scenarios and social interactions. With the rapid expansion of online communication through social networking websites, more persuasive content is made available in the form of videos, which have at large replaced text as the predominant communication modality used online. Persuasiveness is at the core of everyday human-human interactions, such as defending a client in a court (Voss, 2005), seeking a patient's compliance with medical advice (O'Keefe & Jensen, 2007) or convincing a customer to buy a product (Meyers-Levy & Malaviya, 1999). With the rise of social networking websites and the dominance of video content online, understanding the impact of opinion polarity on persuasive messages becomes increasingly important. This study aims to explore the relationship between opinion polarity and the persuasiveness of video messages, focusing on key elements that contribute to effective persuasive communication. The central question that motivates this research is: What makes a person persuasive or not persuasive? Specifically, the research focuses on the factors that contribute to a person's persuasiveness. To answer this question, the study investigates behavioral differences between strongly persuasive and weakly persuasive videos. The data used for the analysis is derived from Persuasive Opinion Multimedia (POM) dataset which is a multimedia corpus of movie review videos that already includes pre-labeled persuasiveness items in addition to other pre-labeled elements important to persuasion including confidence, professionalism, credibility, voice pleasantness, physical attractiveness, humor and so on. A pilot study examined whether video rated high on expertise (5=high expertise and 1 =low expertise) were also rated high on confidence (7=high confidence and 1=low confidence). The research results presented here revolve around the following hypothesis: having prior knowledge of a speaker expressing a positive or negative opinion helps better predict the speaker's persuasiveness. In addition, the study introduced a neuromarketing assessment method based on NeuroMap's NeuroScoring tool for evaluating persuasive messages.

The 19th International Conference on Grid, Cloud, & Cluster Computing (GCC'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-GCC
https://www.american-cse.org/csce2023/

BIMEE: Blockchain based Incentive Mechanism Considering Endowment Effect

Jayanth Madupalli, Hui Liu
Department of Computer Science, Missouri State University, Springfield, Missouri, USA

Abstract - Crowdsensing is the concept of crowdsourcing data with the help of mobile users having necessary sensory devices. The users submit the required data based on the criteria provided in the tasks published by requesters. The quality of sensing data depends on the continuous participation over multiple periods of time of highly skilled users. To encourage long term participation of the users, the incentive mechanism based on blockchain framework is offered to offset users' costs and preserve their privacy. Blockchain is a distributed ledger of transactions that is immutable and traceable. However, current blockchain based incentive mechanisms for crowdsensing has two problems, firstly, they ignore the users' preferring avoidance of loss over acquiring equivalent rewards which is defined as loss aversion in economics, secondly, the platform does not consider the fairness to newly participated users for auction. Therefore, we proposed a Blockchain based Incentive Mechanism with a reverse auction considering Endowment Effect (BIMEE). We further implement BIMEE using Smart contracts on Ethereum public test network. The extensive experimental results show that BIMEE improves the participation rate of mobile users in sensing tasks, the platform utility, and the average utility of users.

Assessing Damage and Recovery of Critical Data in Unsecure Cloud Systems

Brajendra Panda, Justin Burns
Computer Science & Computer Engineering Department, University of Arkansas, Fayetteville, Arkansas, USA

Abstract - Many critical infrastructure systems have begun to store their data in cloud due to the many advantages offered by cloud computing. Unfortunately, data in these critical infrastructure systems have become one of the major targets of attackers. Since critical infrastructure data are highly connected and interdependent, the initial damage done by an attacker spreads quickly through the system when valid transactions make any updates based on the value of a damaged object. In this work, we focus on optimizing the damage assessment and recovery procedure by quickly determining the spread of damage. Once the affected area is determined, all those data items are blocked to stop the further spread of damage. This is followed by the recovery of the damaged data. Through simulation, we proved that the model expedites the process significantly.

Reinforcement Learning Approach to Server Selection and Load Balancing for Collaborative Virtual Services

Sakir Yucel
Wexford, PA, USA

Abstract - Boundaries on separate infrastructure domains are blurring as infrastructure providers are collaborating in offering end-to-end services. To capture opportunities and shine in competitive service market, infrastructure and service providers need to excel in addressing the changing customer requirements and in the operations and management of service resources. They should welcome effective collaboration with other network and infrastructure providers in delivering quality services to customers. Server placement problem for end-to-end virtual services becomes a crucial optimization challenge for providers in such collaborative environments. In our previous work, we formulated the collaborative virtual server placement problem and proposed a generic two-step optimization algorithm to effectively place the servers based on the objectives of the collaborating providers. In this work, we address the server selection and load balancing problem in such environments and propose a reinforcement learning based approach for realizing optimum policies for such environments. The objective of the reinforcement learning is to realize optimum policies not only for load balancing on the server nodes but also for balancing the objectives of the users as well as the collaborating providers.

Energy Efficiency in Data Center Management

Terrence Hernandez, Jacob Sawyer, Hunter Uebelacker, Axel Zuniga, Hilmi Demirhan
Department of Computer Science, University of North Carolina Wilmington, North Carolina, USA;
Department of Information Systems, University of North Carolina Wilmington, North Carolina, USA

Abstract - The availability of the internet and exponential growth of IoT devices has provided organizations with advanced capabilities to facilitate and provide services. With these capabilities comes new demands. These demands are met using vast amounts of equipment and power to help support these new advances. A Data Center is a physical building that houses the hardware that provides this support. Some of the equipment one might find in a data center includes but is not limited to networking equipment, storage equipment and much more. The size of these centers varies depending on the size of the organization they serve. The data center performs various functions such as data storage, management, backup and recovery. Other functions include application management and hosting, hosting network-based transactions such as e-commerce and more. These centers provide vast amounts of capabilities, but they too require vast amounts of resources and energy to run. In this paper, we investigate the the energy consumption landscape of data centers.

A Comparative Study of Mobile Cloud Computing, Mobile Edge Computing, and Mobile Edge Cloud Computing

Deok Hee Nam
Professor of Computer Science, College of Professional Studies,
Wilberforce University, Wilberforce, Ohio, USA

Abstract - Mobile computing paradigms have undergone significant transformations with the rise of mobile cloud computing (MCC), mobile edge computing (MEC), and mobile edge cloud computing (MECC). These paradigms offer distinct approaches to address the challenges of mobile application development, resource management, and latency reduction. MECC combines the benefits of MCC and MEC, integrating cloud and edge resources to leverage their respective advantages. MECC offers a hybrid architecture where tasks can be dynamically allocated between cloud servers and edge nodes based on factors such as task requirements, latency constraints, and resource availability. This paper presents a comprehensive review through the comparative analysis of MCC, MEC, and MECC, focusing on key aspects such as architecture, resource allocation, latency reduction, and scalability.

An Energy-efficient and Data-compressed Offloading in Mobile Cloud Computing

Yung-Ting Chuang, Chih-Ting Hsu
Institute of Information Management, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

Abstract - As more and more applications with computational demands emerge on mobile devices, devices with weaker computing capabilities may struggle to handle them, resulting in longer processing times or higher energy consumption to complete the computations. Even if the computations are completed, users may not experience optimal performance. Therefore, task offloading techniques are gradually being adopted to offload the required computations to nearby edge nodes or cloud servers for execution and return the results to the device. However, energy consumption and execution efficiency have always been popular research topics in the field of task offloading. Unlike most binary task offloading decisions where tasks are either fully offloaded or not offloaded at all, this study aims to consider partial task offloading while also taking into account data header and body compression to reduce overall energy consumption while maintaining a certain level of execution efficiency.

The 9th International Conference on Health Informatics & Medical Systems (HIMS'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-HIMS

Optimized Reading for ADHD and Dyslexia Users in VR

Kyler Hoang, Leonidas Deligiannidis
School of Computing and Data Science, Wentworth Institute of Technology, Boston, Massachusetts, USA;
School of Engineering, Wentworth Institute of Technology, Boston, Massachusetts, USA

Abstract - To combat the rising population of ADHD and dyslexic patients, current advancements in Unity and Oculus Integration are investigated and employed to design a user-friendly reading experience. Unlike physical novels that lack visual aid, our prototype ensures a distraction-free and minimalistic VR environment. It utilizes Wit.ai speech recognition to aid ADHD and dyslexic patients in mechanical and comprehensive reading. Users can also change the background color and music to enhance their facilitated experience. This approach allows the user to constantly focus on the text without averting their gaze. We believe that further development of this prototype can create a powerful tool that redefines reading E-books.

Multi-label Concept Classification in Imaging Entities of Biomedical Literature using CNN and Vision Transformers

Md Mahmudur Rahman, Bikesh Regmi
Computer Science Department, Morgan State University, Baltimore, Maryland

Abstract - Biomedical images are frequently used in articles to illustrate medical concepts and highlight regions-of-interests (ROIs) by using annotation markers (pointers) such as different arrows, letters or symbols overlaid on figures. Also, in many cases multiple markers in the same image are often pointing to different concepts relevant to the article. Hence, each image might be assigned with one or more concepts for multi-label classification and object detection based machine-learning tasks. This work reports such a proof-of-concept (POC) experiment by annotating ROIs and classifying (multi-label classification) 200 Chest CT images appeared in biomedical articles with eleven (11) different concept (similar to UMLS) categories such as, ground-glass, bronchi, honeycomb, cyst, nodules, etc. For annotation, we use an online tool (Labelimg) to annotate image ROIs with concepts based on the information content in associated captions. To demonstrate the feasibility of the POC, this study conducts experiments with different Convolutional Neural Networks (CNNs) and Vision Transformers (ViTs) using both transfer learning (fine-tuning) and training from scratch. We achieved encouraging results (around 70% micro average precision and recall accuracies) in a test set, whereas the dataset images are in very low resolution, non-uniform lighting conditions and with varied shapes and sizes. Overall, this study demonstrates the effectiveness of deep learning models in multi-label classification in medical images and establishes the feasibility and rationale of the POC. The ultimate goal of this work is to develop a large-scale concept detection framework towards building a visual ontology of images in biomedical articles.

A Comparison of Public Sentiment on COVID-19 Vaccines on Twitter and Reddit: An Analysis Using VADER

Joseph Wireman, Farnoosh Koleini, Nasseh Tabrizi
Department of Computer Science, East Carolina University, Greenville, North Carolina, USA

Abstract - Since the outbreak and rapid spread of COVID-19, there has been much controversy and outcry, particularly when it comes to vaccination. Recent studies have shown that there is a tight correlation between vaccine hesitancy and vaccine uptake success. Because of this, the importance of measuring sentiment cannot be overstated. In the event of a more dangerous pandemic

emerging, it will be very important to use the most accurate and efficient sentiment analysis method. This study proposes a method to compare sentiments of different social media platforms The purpose of this review paper is to help researchers to discover current research trends and future directions in sentiment analysis of vaccine outbreaks.

Detection of Suspicious Clusters in Women's Breast Image using Convolutional Neural Networks

Munienge Mbodila, Patrick Mukeninay Madimba
Walter Sisulu University, South Africa;
Mukemad Medical Health Practical Inc., South Africa

Abstract - Breast cancer is a type of cancer that develops from the breast tissue and it is one of the leading causes of female deaths worldwide. Mammography is used to examine human breasts for screening and diagnosing the presence of breast cancer in the breast image. The recurrence of the use of mammography for screening produces enormous amounts of human breast data. Normally, radiologists interpret the screening mammograms, however, the process can be long and exhausting. As a result, the radiologist may not detect all breast cancers due to misinterpretation and unrevealing of hidden breast masses that can lead to cancer. Recent studies have investigated the current mammogram systems for effectively finding hidden breast masses, but none have given a conclusive solution. Hence, this research develops a deep learning framework that utilizes mean filtering inclined on Convolutional Neural Networks (CNN) to extract and train (learn) relevant features from breast images to detect anomalous breast masses. Experiments were conducted using a publicly available (Wisconsin) breast dataset, the result was compared with other popular existing methods indicating that new methods achieve superior performance when compared to other models with an F1-score of 0.9761, an precision of 0.9881, and accuracy of 98.64%. The deployment of this model to the physician CAD can help in giving an accurate region in the breast where cancer is likely to occur in the future.

Efficacy of Individual Cognitive Stimulation Therapy on People with Dementia: A Prisma Guided Review

Alfia Parvez, Minza Nadeem Khan, Arshia Khan
Department of Computer Science, University of Minnesota - Duluth, Duluth, Minnesota, USA

Abstract - Background: Individual Cognitive stimulation therapy (iCST) is one of the evidence-based treatments for people with mild to moderate dementia. Aim: To analyze the efficacy of the iCST on cognitive function, quality of life, activities of daily living, neuropsychiatric symptoms, behavioral and psychological symptoms, and depression in people with dementia. Design: Systematic review and meta-analysis. Results: A total of 17 studies were included in the review. Overall, individual Cognitive stimulation therapy can be used as an intervention to improve cognition and quality of life. No significant effects were found on other outcome measures. Conclusions: Individual Cognitive Stimulation Therapy improves the quality of life of the participants of both people with dementia and their carer. It has shown significant improvement in depression. It can potentially be used to improve cognition, but further research may consider the implementation of a standard protocol with a rigorous study design to provide sufficient evidence to address this area.

Effects of Individual Cognitive Stimulation Therapy and Group Cognitive Stimulation Therapy on People with Dementia: A PRISMA-guided Systematic Review

Sakina Rao, Isra Hassan, Arshia Khan
Department of Computer Science, University of Minnesota - Duluth, Duluth, Minnesota, USA

Abstract - Background: Individual Cognitive Stimulation Therapy (iCST) and group Cognitive Stimulation Therapy (CST) are known to improve cognition and quality of life of people with dementia (PwD) and other cognitive impairments. Aim: To analyze

the efficacy of iCST and CST among people with dementia. Design: Systematic review. Methods: Articles related to CST for people with Dementia were searched on PubMed. Studies were selected for the analysis based on the inclusion and exclusion criteria. Results: A total of 15 studies were included for the analysis. Group CST was found to show more improvements than iCST among PwD. Conclusion: CST improves the quality of life (QoL) of PwD and their carers. Group CST can improve cognition while limited evidence was found that iCST improves cognition. More evidence and studies are needed to see if iCST can improve cognitive function. Group CST was also found to be effective in improving BPSD, NPS and ADL.

Maintenance Cognition Stimulus Therapy: A Survey Paper

Hritik Rao, Maryam Kameli, Anne Hinderliter, Sherri Turner, Arshia Khan
Department of Computer Science, University of Minnesota - Duluth, Duluth, Minnesota, USA;
Department of Chemistry and Biochemistry, University of Minnesota - Duluth, Duluth, Minnesota, USA;
Department of Educational Psychology, University of Minnesota - Duluth, Minnesota, USA

Abstract - This Prisma-style survey paper discusses trends in the conclusion of Maintenance Cognitive Stimulation Therapy vs CST Clinical Trials and analyzes benefits in Cognitive (ADAS-Cog), Quality of Life(QoL), Behavioural & Psychological Symptoms of Dementia(BPSD) and Activities of Daily Living(ADL). Maintenance CST receives post-therapy sessions in a reduced contact time for long durations to sustain the therapy benefits observed in cognition or other forefronts mentioned in the Prisma table. The objective of this survey paper is to analyze and compare benefits observed during various MCST trials, since the Trials involving clear Maintenance CST sessions observed against a control group are scarce it is important to discuss the trends observed to understand the benefits of proving MCST vs normal CST. The Clinical Trials were found in Medical Publication Databases (PubMed, Embase, ScienceDirect, Cochrane Central Register of Controlled Trials) and the inclusion criteria were Maintenance sessions after Cognitive therapies. The Discussion analyzes different trends observed from the Prisma-style survey table from the selected publications on Features like Quality of Life(QoL), Cognition improvement (ADAS-Cog), Behavioural & Psychological Symptoms of Dementia(BPSP), and Activities of Daily Living(ADL).

Where Robots Can Fit In: A Systematic Review of the Incidence of Comorbid Dementia and the Prescribed Treatments

Matthew Sawchuck, Arshia Khan
Department of Computer Science, University of Minnesota - Duluth, Duluth, Minnesota, USA

Abstract - (1) Background: The purpose of this systematic review is to find the incidence of dementia in adults with cognitive, developmental, or physical disabilities and the prescribed treatments for those comorbidities. The findings will be used to determine the possible interventions for robots in the treatment for comorbid dementia. (2) Methods: The studies included in this review focused on the incidence of dementia among various disabilities and the prescribed treatments, published between 2018 and 2023, written in English, and have full text availability. The databases used to collect the sources for this review were Academic Search Premier, Gale Academic Onefile, PubMed, Google Scholar, Web of Science, ProQuest, and Scopus. Furthermore, data was collected through hazard ratios (HR), adjusted hazard ratios (aHR), cases per number of people-years, and percentage based on sample size or population. No methods were selected to determine the risk of bias for the sources. (3) Results: The methods of this review produced 30 sources to be included in the results. The highest hazard ratios for the incidence of dementia depending on disease type are mild behavioral impairment (MBI) with HR 8.07, mild cognitive impairment (MCI) HR 7.05, and subjective cognitive decline (SCD) with HR 6.81. These results are mostly consistent with the findings for the cases per person-years data. Finally, the highest comorbid illnesses with dementia, in terms of population size, are type-2 diabetes (45.92%), type-2 diabetes with hypertension (43.60%), and depression (42.90%). (4) Discussion: The incidence of dementia is strongly correlated with the severity of cognitive impairment caused by a disease. Furthermore, the best instances where robots can fit in during the treatment and prevention of dementia is through exercises that promote cognitive function. (6) Funding: This systematic review received no external funding.

Robotics in Caregiving: A Concise Review of Literature

Reilly Moberg, Arshia Khan
Department of Computer Science, University of Minnesota - Duluth, Duluth, Minnesota, USA

Abstract - As elderly populations across the world grow at an exponential rate and the demand for caregivers grows more critical by the day, the need for a solution has people asking themselves compelling technical and ethical questions about the inclusion of robots in caregiving. This concise literature review discusses past research papers and studies from the field and covers the trends, technology, perspectives of both patients and caregivers and the ethical questions that come along with introducing robots into caregiving settings. Overall, we found that while it is agreed upon by researchers that human caregivers cannot be fully replaced, the addition of caregiving robots can provide patients with increased freedom and quality of care.

Benefits of Gossip: A Prisma Guided Review

Minza Nadeem Khan, Arshia Khan
Computer Science Department, University of Minnesota - Duluth, Duluth, Minnesota, USA

Abstract - Gossip is talking about a person in their absence. Gossip is seen as a malicious activity by most people, and is highly condemned. Yet, gossip is present everywhere. From newspapers to daily life, gossip is a huge part of society. Since it is so common, there must definitely be some benefits of gossip. Gossip is not always negative. The valence of gossip is truly important. If a person says something positive about another person behind their back, it is considered to be positive gossip. Positive gossip can be beneficial for a person. It makes other people think highly of the person being gossiped about. The aim of this paper is to explore the various benefits and applications of gossip.

Modeling Obesity Prevention Programs to Reduce Overweight Rates at Schools: A Perspective

Lateefat Amao, Misagh Faezipour
College of Basic and Applied Sciences, Healthcare Informatics,
Middle Tennessee State University, Murfreesboro, Tennessee, USA;
Department of Engineering Technology, Middle Tennessee State University, Murfreesboro, Tennessee, USA

Abstract - Childhood obesity is a public health challenge, with many risks affecting a child's long-term health. This paper explores the complexities of childhood obesity. It recognizes the role and influence that schools play in children's behaviors and leverages the school as a micro-environment to implement multiple levels of interventions. This work proposes an intervention to tackle childhood obesity at schools using system dynamics modeling. Its focus is on promoting healthy dieting, increasing physical activity, creating supportive environments, while interacting with multiple sectors and communities. We propose a measurable and sustainable intervention, spanning a three-year time-frame. System dynamics approach accounts for the dynamic nature of schools and pinpoints the need for well-structured and sustainable interventions. This evidence-based dynamic intervention model targets multiple factors that contribute to obesity, involving relevant stakeholders, and implements a long-term intervention to promote long-lasting behavioral changes that decreases the rate of prevalence of childhood obesity.

Research Survey of Multimodal Medical Question Answering

Hilmi Demirhan, Wlodek Zadrozny
University of North Carolina Wilmington, North Carolina, USA

Abstract - The purpose of this research is to study foundational approaches for Multimodal Medical Question Answering using Natural Language Processing (NLP) and Computer Vision (CV). Question Answering (QA) using NLP provides a solution by retrieving relevant answers for questions from a given text. The QA system enables information retrieval from databases and natural language documents. Multimodal Question Answering adds additional inputs to text based information such as image, audio and video to improve flexibility and effectiveness of an intelligent information retrieval system. This research aims to focus on Multimodal Medical Question Answering using medical instructional and medical non-instructional videos. Multimodal medical QA is able to provide accurate comprehensive answers to complex medical questions for patients and healthcare professionals using texts, images and videos. Multimodal medical QA can be helpful for improving the health system through enhanced clinical decision support and critical precise recommendation.

The Online Database for Prevention of Anxiety Disorders

Joi L. Snowden, Zizhong John Wang
Department of Computer Science, Virginia Wesleyan University, Virginia Beach, Virginia, USA

Abstract - The paper presents an online platform with a functional and simple user interface that would allow individuals to explore their options for improving their mental health. The main elements of this mental health website included a database table that listed therapy locations within Hampton Roads, a user form and database table that allowed visitors and website administrators to interact and share stress management techniques, and a user form that allows website visitors to sign up for newsletters. Another significant element to the online platform is the secure login and logout feature. The login feature allows administrators to update the therapy location table, retrieve information from the newsletter form, and includes a form that allows administrators to interact with users and answer questions about any stress management techniques. All of the components are implemented with the use of HTML, CSS, PHP, and MYSQL.

Pediatric Epileptic Seizure Detection via EEG Signals and Convolutional Neural Networks

Omneya Attallah, Maha Sharkas, Mona Khalil Mohamed, Tarek Omar
Department of Electronics and Communications Engineering, Arab Academy for Science, Technology & Maritime Transport, Alexandria, Egypt;
Pediatric Neurology Unit, Department of Pediatrics, Faculty of Medicine, Egypt

Abstract - According to the world health organization (WHO), epilepsy, a central nervous system condition, affects approximately 50 million people worldwide. The electroencephalogram (EEG) is the most commonly used noninvasive method for examining the brainwave activity of epileptic patients. Accurately determining the occurrence of seizures remains a challenge, and constructing effective methods to monitor epilepsy has progressed over the decades. It has been demonstrated that the morbidity burden of epilepsy is considerably higher in developing countries with lower socioeconomic levels. Automated detection systems based on deep learning techniques can facilitate the detection procedure thus improving the quality of healthcare services provided in low-income countries. This paper is concerned with designing a framework to assist neurologists in detecting epileptic seizures rapidly and effectively in Egypt as an instance of a developing country. To accomplish this goal, a dataset of EEG signals is acquired from 8 children located in Alexandria, Egypt. This dataset is then preprocessed and segmented. Then, using the preprocessed data, a convolutional neural network (CNN) is implemented and trained. Later, using a transfer learning approach, deep features are extracted from this CNN. Finally, machine learning classifiers are used in the detection process. The results of the proposed framework demonstrate that the proposed model is capable of detecting epileptic seizures.

Explainable AI in Orthopedics: Challenges, Opportunities, and Prospects

Soheyla Amirian, Luke A. Carlson, Matthew F. Gong, Ines Lohse,
Kurt R. Weiss, Johannes F. Plate, Ahmad P. Tafti
University of Georgia, USA; University of Pittsburgh, USA

Abstract - While artificial intelligence (AI) has demonstrated many successful applications in various domains, its adoption in healthcare lags a little bit behind other high-stakes settings. Several factors contribute to this slower uptake, including regulatory frameworks, patient privacy concerns, and data heterogeneity. However, one significant challenge that impedes the implementation of AI in healthcare, particularly in orthopedics, is the lack of explainability and interpretability. Addressing the challenge of explainable AI (XAI) in orthopedics requires developing AI models and algorithms that prioritize transparency and interpretability, allowing clinicians, surgeons, and patients to understand the contributing factors behind any AI-powered predictive or descriptive models. The current contribution outlines several key challenges and opportunities that manifest in XAI in orthopedic practice, such as improved diagnostic accuracy, personalized treatment recommendations, and enhanced patient engagement. This work emphasizes the need for interdisciplinary collaborations between AI practitioners, orthopedic specialists, and regulatory entities to establish standards and guidelines for XAI in orthopedics.

3D Machine Vision and Deep Learning for Enabling Automated and sustainable Assistive Physiotherapy

Lyndon N. Smith, Stephen Boyd, Devaki Bhatta, Melvyn L. Smith
Department of Engineering Design and Mathematics, University of the West of England, Bristol, United Kingdom (UK);
Perseptive Ltd., London, United Kingdom (UK)

Abstract - The significant beneficial effects of physiotherapy as a treatment for a wide range of medical conditions, is well known. However, the challenge for already stressed healthcare systems to provide effective physiotherapy to an aging population, that has increasing requirements and often limited mobility, is formidable. Further pressure on provision has arisen as a consequence of the COVID-19 pandemic and climate change concerns over carbon emissions associated with travel of the patient and/or physiotherapist to treatment centers. The situation is particularly acute in relation to frail/elderly patients, due to increased demand placed on services by the pandemic, the risk to the patient of infection or of suffering a fall, and the relatively high fiscal and environmental costs associated with face-to-face care. These factors provide strong motivators for effective provision of remote healthcare methods; however, before automated physiotherapy systems can become widespread, they need to be able to appraise patients and monitor their responses to recommended physiotherapy regimes, with an accuracy similar to that of human physiotherapists and at an acceptable level of cost. This paper outlines a novel remote healthcare package that aims to provide elderly persons with accurate, accessible, and cost-effective physiotherapy in their own homes. The approach makes use of emerging state-of-the-art machine vision (incorporating 3D data to improve accuracy) and advanced machine learning techniques, for accurately recovering a patient's joint and limb positions during exercises, with the ultimate aim of enabling automated assessment of patient exercise concordance. This is considered to be a first step toward a system that can assess patients and monitor their progress, as part of an automated approach to physiotherapy that can offer significant access, environmental and cost benefits; thereby assisting with on-going sustainability of healthcare provision.

A Preliminary Study about the Impact of Covid-19 on Mental Health and Depression of Swiss Population in 2020 using Data Mining

Edona Mehmeti, Farshideh Einsele
School of Business, Bern University of Applied Sciences, Bern, Switzerland

Abstract - The objective of the study was to investigate the impact of Covid-19 on mental health and specifically depression for the Swiss population. Data from Swiss Household Panel surveys from 2019 (no Covid) with 8841 individuals and from 2020 (Covid) with 15882 individuals about their life and health conditions, in addition to data from statistics on Income and living Conditions (SILC) conducted by Swiss Federal Statistical Office (FSO) from 2019 (no covid) and 2020 (covid) with approx. 18000

individuals along with Covid data from Swiss Surveillance Systems (CH-SUR) from 2020 were pre-processed, cleaned, transformed and finally integrated to a unique relational database. Apriori-algorithm for association analysis was applied and led to finding some interesting rules. The result of this study are 3 rules extracted from the database showing the correlation between Covid-19 and depression in the year of 2020 in Switzerland and compared with depression data in the year 2019 (before Covid-19). The found rules show that the Swiss population's life and health condition was not significantly impacted during the first year of Covid-19 in 2020. In our future work, we intend to use different data sources like social media instead of surveys to verify the gained results in this study.

System Analysis of Implementing 5G in an Ambulance

Akshay Monga
Harrisburg University of Science and Technology, Pittsburgh, Pennsylvania, USA

Abstract - This paper develops a conceptual model of a 5G connected ambulance. The model defines the various systems engineering (SE) steps required to achieve the 5G connected state and the new processes that need to be developed to support this transition. The paper also calls out the impact on stakeholders such as the patient, Emergency Medical Services (EMS) staff, Emergency Room physicians and nurses.

Binary Classification vs. Anomaly Detection On Imbalanced Tabular Medical Datasets

Philippe Rambaud, Adel Taleb, Raphael Fauches, Arpad Rimmel, Joanna Tomasik, Jean Bergounioux
LISN, University of Paris-Saclay, SogetiLabs, Paris, France;
CHArt, University of Paris Sciences Lettres, SogetiLabs, Paris, France;
SogetiLabs, Paris, France;
Pediatric Neurology and Intensive Care, Raymond Poincare Hospital (AP-HP), Garches, France

Abstract - When working on imbalanced medical datasets for a classification task, the chosen method depends mainly on label availability. We consider two classes, e.g., normal and abnormal. If the dataset is labeled, we can choose Binary Classification; otherwise, we should opt for Anomaly Detection. We aim to determine the trade-off between using labeled data with Binary Classification and unlabeled data with Anomaly Detection while being aware of the labeling effort. The comparison is made on BabyGarches, an original dataset of newborn motricity, and on two publicly available imbalanced tabular medical datasets, Thyroid and Cardiotography. We observe that Binary Classification outperforms Anomaly Detection on all datasets, achieving the highest Area Under the Curve ROC scores. Nevertheless, the performance of Deep Anomaly Detection methods such as Autoencoders and GANs is close to Binary Classification models. performance. Moreover, they reach the best Recall scores, which physicians appreciate. They may thus be recommended when data labeling is costly. Since labeling our use case dataset BabyGarches is expensive, Anomaly Detection is relevant to its classification.

Improve Pose Estimation Model Performance with Unlabeled Data

Adel Taleb, Samuel Diop, Philippe Rambaud, Awa Bakayoko, Audrey Benezit,
Raphael Fauches, Ois Jouen, Jean Bergounioux
CHArt - University of Paris Sciences et Lettres, SogetiLabs, Paris, France;
CHArt - University of Paris Sciences et Lettres, OSO-AI, Paris, France;
LISN, University of Paris-Saclay, SogetiLabs, Paris, France;
Pediatric Neurology and Intensive Care, Raymond Poincare Hospital (APHP), Paris, France;
SogetiLabs, Paris, France; CHArt - University of Paris Sciences et Lettres, Paris, France;
Pediatric Neurology and Intensive Care, Raymond Poincare Hospital (APHP) Paris, France

Abstract - The infant's spontaneous motricity is a fundamental element in the clinical evaluation of neonates. Therefore, the pediatrician values any modification in the neonate's motor pattern for diagnosing infectious, genetic, metabolic, or neuromuscular disease. Automatization of spontaneous motricity analysis would be of great help in diagnosing and treating patients earlier. However, as the patterns are very numerous, it would be resources consuming to use labeled data. We aim to eliminate the necessity for a tedious and costly manual annotation. Our work presents a new infant motor skills dataset and describes methods to improve the automatic 2D pose estimation with AlphaPose from RGB videos. We evaluate the results using P CKh and a clinician's assessment. The results show that our methods significantly increase the AlphaPose performances on our dataset without needing labeled data. We tested the relevance of the posture estimation obtained with our models by training a classification model that separates a motricity of a control group from the motricity of children with Spinal Muscular Atrophy; our model achieves encouraging results for further research with more complex models.

Smart Decision Making Using Data Science Technology with COVID-19 Data

Shelley Zhang, Ankush Rathi, Shivraj Jagdish Patil
Computer and Information Science, University of Massachusetts Dartmouth,
North Dartmouth, Massachusetts, USA

Abstract – N/A

Healthcare Data Handling with Machine Learning Systems: A Framework

David Ang, Kiranmai Nagineni
Auburn University Montgomery, Alabama, USA

Abstract - Data collected in healthcare can be categorized and structured in ways that resemble big data. Analyzing big healthcare data with machine learning (ML) is one of the most effective methods. Using different learning models, ML can assist in the prediction of diseases and the timely diagnosis of patients, improving their health. The use of ML in healthcare data analytics, however, must be preceded by an understanding of which learning model to employ. Choosing between a variety of models to healthcare data becomes a challenge when applying machine learning. The purpose of this paper is to provide a framework for identifying the best machine-learning model for healthcare data.

Multi-Subset Approach to Early Sepsis Prediction

Kevin Ewig, Xiangwen Lin, Tucker Stewart, Katherine Stern, Grant O. Keefe, Ankur Teredesai, Juhua Hu
School of Engineering and Technology, University of Washington, Tacoma, Washington, USA;
Department of Surgery, University of Washington, Seattle, Washington, USA;
CueZen Inc., Seattle, Washington, USA

Abstract - Sepsis is a life-threatening organ malfunction caused by the host's inability to fight infection, which can lead to death without proper and immediate treatment. Therefore, early diagnosis and medical treatment of sepsis in critically ill populations at high risk for sepsis and sepsis-associated mortality are vital to providing the patient with rapid therapy. Studies show that advancing sepsis detection by 6 hours leads to earlier administration of antibiotics, which is associated with improved mortality. However, clinical scores like Sequential Organ Failure Assessment (SOFA) are not applicable for early prediction, while machine learning algorithms can help capture the progressing pattern for early prediction. Therefore, we aim to develop a machine learning algorithm that predicts sepsis onset 6 hours before it is suspected clinically. Although some machine learning algorithms have been applied to sepsis prediction, many of them did not consider the fact that six hours is not a small gap. To overcome this big gap challenge, we explore a multi-subset approach in which the likelihood of sepsis occurring earlier than 6 hours is output from a previous subset and feed to the target subset as additional features. Moreover, we use the hourly sampled data like vital signs in an observation window to derive a temporal change trend to further assist, which however is often ignored by previous studies. Our empirical study shows that both the multi-subset approach to alleviating the 6-hour gap and the added temporal trend features can help improve the performance of sepsis-related early prediction.

Tales From the Past: Adapting App Repositories to App Store Dynamics

Michael Stach, Marc Schickler, Manfred Reichert, Rudiger Pryss
Institute of Clinical Epidemiology and Biometry, University of Wurzburg, Germany;
Institute of Databases and Information Systems, Ulm University, Ulm, Germany

Abstract - The pervasiveness of smartphones in daily life has enabled the transformation of healthcare services into digital services, often offered as mobile apps to make them more accessible to the general public. The steady increase of apps in the app stores, in turn, led to the creation of app repositories to help end users, physicians, therapists, and healthcare providers find highquality apps based on expert app ratings. However, the validity of these ratings is linked to a specific app version. In this paper, we aim to describe the problems of current app repositories and present an approach to address these shortcomings. More specifically, our approach is able to periodically monitor the two major proprietary app stores and react to upcoming changes. We furthermore present a mechanism to extract additional information from Android apps and to automate the latter procedure. Finally, our paper aims to stimulate discussion on what additional tools app researchers need to better study app quality and execution.

ESAI: An AI-Based Emotional Support System to Assist Mental Health Disorders

Gabriel Serrano, Daehan Kwak
Department of Computer Science and Technology, Kean University, Union, New Jersey, USA

Abstract - Every year more and more people are struggling to handle a wide variety of mental health disorders, such as anxiety and depression. However, as the number of people in need of assistance increases, the number of resources available to them has continued to decrease. The goal of this research is to develop an Emotional Support AI (ESAI) system, an additional resource for those unable to obtain the help and information they need. The ESAI has been trained to classify text based on the Naive Bayes Classification model. The model was trained using 160,000 Reddit posts, which were collected using web scrapping, where users have discussed their experiences with mental health. ESAI provides users with a friendly user-interface from which they can discuss their mental health concerns. The user can choose whether to communicate through typing or through real-time speech recognition. ESAI works by hosting "sessions", in which it will log communications between itself and the user to check for any potential flags that may indicate the user is experiencing symptoms of one or many mental health disorder(s). These sessions can

be used by the user for venting or to seek information regarding a variety of mental health disorders. If the probability that the user is experiencing a mental health disorder is higher than a specific threshold, the user is provided with general resources and contacts regarding the specified disorder. The user will also be provided with a mental health evaluation report at the end of each session upon request. Currently, results show that ESAI can classify mental health disorders with seventy-percent accuracy.

Helpfulness of Cognitive Behavioural Therapy (CBT) for the Autistic Population: Towards the Development of Evidence-based Framework for the Adaptation of CBT using an Affective Pedagogical Agent

Jonathan Bishop
Crocels Research CIC, United Kingdom, UK

Abstract - There is growing evidence to suggest that cognitive behavioural therapy (CBT) can be beneficial to people diagnosed with Autism Spectrum Disorder (ASD), many of whom suffer from multiple comorbidities. People with ASD often do not respond well to CBT though. It is acknowledged that CBT requires adaptations to accommodate experienced socio-communication and neuropsychological impairments, but there is little evidence on how best to adapt CBT to optimise treatment outcomes for people with ASD. This paper proposes an affect-aware chatbotbased intervention in the form of an animated pedagogical as a means to inform the development of the effective and acceptable adaption of CBT to meet specific needs of autistic populations in the classroom.

Muse Alpha: Primary Study of AI Chatbot for Psychotherapy with Socratic Methods

Hyeseong Park, Myung Won Raymond Jung, Min Ji Kim, Jihye Kim, Uran Oh
AKA Cognitive Corp., Seoul, South Korea;
Harvard Chan School of Public Health, Seoul, South Korea;
Yonsei University College of Medicine, Seoul, South Korea;
Ewha Womans University, Seoul, South Korea

Abstract - Artificial Intelligence (AI) technology has revolutionized how we interact with information and entertainment with ChatGPT. ChatGPT can use Socratic questioning for medical treatment purposes by avoiding the reinforcement of negative stereotypes from the users. In this paper, we introduced an AI chatbot called Muse Alpha, in which Socratic questioning is also applied, to evaluate a client's automatic thinking in cognitive therapy, inducing the client to evaluate whether their automatic thinking is realistically valid. We have shown three examples of how Socratic dialogue is applied differently between chatGPT and Muse Alpha, and it shows that Muse Alpha seems to achieve the actual moment of cognitive change by uncovering the assumptions and evidence that underpin their thoughts in respect of problems.

Automating the Development of Stress Detection Systems

Felix Beierle, Rudiger Pryss
National Institute of Informatics, Tokyo, Japan;
Institute of Clinical Epidemiology and Biometry, University of Wurzburg, Wurzburg, Germany

Abstract - Stress is leading to bad health and contributes to economic loss due to employee absence. Real-time stress detection based on wearable sensor data can enable the implementation of mitigating strategies. While several approaches to stress detection exist, setting up a new system can be tedious. We demonstrate how the use of libraries and tools for automation can speed up many of the necessary steps when developing a stress detection system. We employ automated feature engineering and automated machine learning. The resulting stress detection system we developed this way is based on the WESAD dataset and achieves a F1 score of 0.87 for unseen users based on 30 seconds of wearable sensor data.

The Construction of Warning Mechanism for Asthma Patient Based on AIoT

Lun-Ping Hung, Zong-Jie Wu, Xin-Ying Liu, Cing-An Yang
Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan;
Department of Industrial Engineering & Management, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

Abstract - With the continuous development of human society, air pollution has become a global concern. For elderly people with respiratory diseases such as asthma and COPD, the danger of air pollution is not to be underestimated. In this study, we use a number of public datasets and LSTM, SVM, DT, and SHAP models to analyze and predict air quality data and explore the effects of different air environment changes on the occurrence of asthma, in an attempt to find the main sources of pollution for asthma. In addition, this study uses the Internet of Things technology to develop an air quality sensor box to collect data and an information platform to calculate various air quality related data using machine learning models and data to provide efficient control methods to help the elderly avoid dangerous spaces and reduce the risk of asthma in order to create a safe and healthy living space.

A Review of Face Processing for Telehealth: Research Survey of Remote Visual Photoplethysmography (rvPPG)

Hilmi Demirhan, Karl Ricanek
Information Systems Organization, University of North Carolina Wilmington, North Carolina, USA;
Computer Science, University of North Carolina Wilmington, North Carolina, USA

Abstract - This paper will focus on an in-depth survey of commercial and open source resources for face processing for telehealth using Remote Visual Photoplethysmography (rvPPG). We can use Remote Visual Photoplethysmography (rvPPG) as a tool to monitor vital health signs and use its results to observe health conditions. There is a growing demand for non-invasive remote (tele-services) patient care information technology solutions. Advancement in the area of machine learning and deep learning especially as it relates to computer vision made it possible to predict various vital signs such as heart rate, blood pressure and oxygen saturation from pixel changes in recorded or live video feeds of human skin. In this review, we have included thirteen novel methods that have significant contributions, twelve open source public datasets, fifteen open source tools and eleven commercial tools for literature review. We also discovered that there is an opportunity for the transferal of rvPPG methods from academia to industry to aid face processing for telehealth software development and maintenance.

Abnormality Detection in Lung Sounds Using Feature Augmentation

Shiva Shokouhmand, Md Motiur Rahman, Miad Faezipour, Smriti Bhatt
School of Engineering Technology, Electrical and Computer Engineering Technology,
Purdue University, West Lafayette, Indiana, USA;
Department of Computer and Information Technology,
Purdue University, West Lafayette, Indiana, USA

Abstract - Assessing the lung sounds reveals important information about the lungs and the existence or severity of possible underlying respiratory related conditions. This paper presents a research-work-in-progress exploring a hybrid signal processing and machine learning-based approach for effectively analyzing the lung sounds. The objective of this study is to achieve high accuracy in detecting adventitious sounds for pulmonary diseases non-invasively. To this end, we are augmenting the feature space and introduce a feature-based model. For the analysis, we are utilizing the ICBHI dataset, which comprises lung sounds collected from 126 patients. The dataset includes various abnormalities, such as wheezing and crackling sounds, providing us with valuable data to train and evaluate the proposed model.

Exploring Risk Factors in PDAC Using System Dynamics

Ashiat Ashake Adeogun, Misagh Faezipour
College of Basic and Applied Sciences, Healthcare Informatics,
Middle Tennessee State University, Murfreesboro, Tennessee, USA;
Department of Engineering Technology, Middle Tennessee State University, Murfreesboro, Tennessee, USA

Abstract - Pancreatic ductal adenocarcinoma (PDAC) is a lethal gastrointestinal cancer prevalent in developing countries with a low survival rate. Early detection remains challenging despite medical advancements. The exact causes of PDAC are unknown; however, genetic, environmental, and lifestyle factors are believed to contribute to this cancer. Modifiable and nonmodifiable risk factors like smoking, obesity, gender, and age play a role in PDAC development. System dynamics is utilized to build causal loops, unraveling complex interactions between these factors. Further research is necessary to uncover additional modifiable causes of PDAC. Public health campaigns are essential to raising awareness, and educate about PDAC risks, and promoting preventive measures to reduce mortality and morbidity associated with the disease.

Empowering Healthcare Professionals and Patients with Chat GPT: Applications and Challenges

Fatemeh Mosaiyebzadeh, Seyedamin Pouriyeh, Reza M. Parizi, Meng Han,
Nasrin Dehbozorgi, Mohsen Dorodchi, Daniel Macedo Batista
Department of Computer Science, University of Sao Paulo, Brazil;
Department of Information and Technology, Kennesaw State University, Marietta, Georgia, USA;
Decentralized Science Lab, Kennesaw State University, Marietta, Georgia, USA;
Zhejiang University, Hangzhou, Zhejiang, China;
Department of Software and Game Development, Kennesaw State University, Marietta, Georgia, USA;
Department of Computer Science, University of North Carolina Charlotte, North Carolina, USA

Abstract - ChatGPT is a recently developed Large Language Model (LLM) and is an effective tool to produce human-like dialogue with users and answers to questions. It is trained on a massive amount of online content and can provide text replies for many domains, like healthcare-related questions. In this paper, we investigated the application of ChatGPT in the public healthcare domain. Also, we studied the limitations and challenges of ChatGPT in healthcare. Indeed, while ChatGPT can offer valuable support and information, it is crucial to recognize that it should not be seen as a replacement for the expertise and personalized care provided by healthcare professionals. Instead, its purpose lies in augmenting healthcare services and enhancing access to information. It can be a useful tool for providing general guidance and educational resources. However, when it comes to medical advice or diagnosis, it is essential to consult qualified healthcare professionals who can consider individual factors, interpret complex medical information, and provide tailored recommendations based on a comprehensive understanding of the patient's situation.

Epilepsyecosystem.org: Crowdsourcing Seizure Prediction with Long Term Human Intracranial EEG

Zhinoos Razavi Hesabi, Levin Kuhlmann, Mark Cook

Department of Neurology, St. Vincent's Hospital and University of Melbourne, Australia;
Department of Data Science and AI, Monash University, Melbourne, Australia;
Sir John Eccles Professor and Director of Neurology, St Vincent's Hospital Melbourne,
Director of the Graeme Clark Institute, University of Melbourne, Australia

Abstract – N/A

Responsible AI Application in Mental Health with Emotionally Aware Chatbots

Atif Farid Mohammad, Bryan Clark, Ram Agarwal, Sipi Garg
OmniAGI.ai - Global Technology Solutions, Inc., USA

Abstract - Advanced Analytics Employing Responsible Artificial intelligence (AAERAI) can make significant positive contributions to the mental health care industry. More recently, AAER-AI has made leaps in the mental health field not only in identifying and rectifying mental health issues but by also helping policy makers with industry directions. This paper explores the uses and applications of AAER-AI in mental health, and it also highlights example of applications of AAER-AI in scheduling, patient flow of mental health patients, neuroimaging, and functional magnetic resonance imaging, mapping the relationship between information technology and mental healthcare, identifying locational vulnerabilities of mental health subjects, and usage of SOMNet in visualization and identification of mental health issues. The paper also explores the use of emotionally aware next generation chatbots for AAER-AI data collection and to scale mental health care and finds that AAER-AI and next generation chatbots have the potential to increase the amount of data collection leading to high quality patient care and better decisions made by industry leaders and decision makers.

The 25th International Conference on Artificial Intelligence (ICAI'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-ICAI
https://www.american-cse.org/csce2023/

Bootcamp Method for Training General Purpose AI Agents

Vincent Lombardi, Lawrence Holder
School of EECS, Washington State University, Pullman, Washington, USA

Abstract - General purpose agents have long been an ultimate goal of AI research. One promising approach to this goal is to first train an agent to use a variety of skills, called a skillnet agent, and then allow the agent to learn how to choose the appropriate skill instead of having to choose the appropriate low-level action. We propose a method for training skillnet agents called Bootcamp that helps agents efficiently learn basic skills in an environment. We found that Bootcamp agents outperform skillnet agents trained randomly on various tasks defined in the ViZDoom simulated environment. We also found that skillnet agents outperform more conventional reinforcement-based learning approaches such as DQNs in ViZDoom.

Generating Realistic Multi-class Biosignals with BioSGAN: A Transformer-based Label-guided Generative Adversarial Network

Xiaomin Li, Anne Hee Hiong Ngu, Vangelis Metsis
Computing and Sensing (IMICS) Lab., Department of Computer Science,
Texas State University, San Marcos, Texas, USA

Abstract - Time series data are commonly used in biomedical applications, but such datasets are often small, expensive to collect, and may involve privacy issues that restrict largescale deep learning models. Data augmentation techniques for time series data are limited by the need to maintain signal properties, but Generative Adversarial Networks (GANs) offer a promising approach for expanding datasets. This paper presents BioSGAN, a transformer-based label-guided GAN model capable of generating multi-class, class-specific synthetic time-series sequences of arbitrary length. Our proposed model architecture and design strategies produce synthetic sequences that are almost indistinguishable from real signals, enabling data augmentation. To evaluate the quality of the generated data, we propose a wavelet coherence metric that compares the similarity of real and synthetic signals. Our results show that BioSGAN outperforms existing state-of-the-art time-series GAN models, as demonstrated by qualitative visualizations using PCA and t-SNE, as well as quantitative comparisons of the discriminative and predictive power of the synthetic data.

Communication-Facilitated Coordination in Agent Team Rescue Mission

David Degbor, Abhijot Bedi, Shelley Zhang, Eugene Chabot
Computer and Information Science, University of Massachusetts Dartmouth, North Dartmouth, Massachusetts, USA;
Naval Undersea Warfare Center Division (NUWC) Newport, Newport, Rhode Island, USA

Abstract - The importance of communication is essential in multi-agent systems (MAS) teamwork. Effective communication among agents in a dynamic environment, such as a disaster response scenario, can lead to more effective collaborative actions providing aid. This paper discusses information sharing's impact on collaboration and conflict resolution among different platoon

agents in a rescue simulation testbed from RoboCup Rescue Simulation League (RRSL). This work explores the challenges associated with communication in MAS, particularly in RoboCup Rescue Simulation (RCRS) agent rescue testbed. We examine the performance score difference with the adopted communication approach compared with the baseline approach in rescue agents' decision-making under different configurations.

Steganalysis of Medical Radiographs: a Deep Learning Approach Comparing the Importance of Using Content Pixels and Content-Free Pixels

Farid Ghareh Mohammadi, Ronnie Sebro
Department of Radiology, Mayo Clinic, Jacksonville, Florida, USA;
Center for Augmented Intelligence, Mayo Clinic, Jacksonville, Florida, USA

Abstract - Content-free pixels, which do not contain actual information, have been used in research studies for passive watermarking of radiographs. However, no studies have compared the effectiveness of content-free pixels to content pixels for watermarking. In this study, we propose a radiograph steganalysis solution that can be used to identify the source of a radiograph and be potentially used to identify whether a radiograph is fake. The solution uses a deep-learning architecture for automating computer-generated fake radiograph detection and compares the performance of passive watermarking using the content-free pixels to that using the content pixels. We use patients who had radiographs of the abdomen, pelvis, and lumbar spine at Mayo Clinic (01/01/2010 - 12/31/2021). The patients (n = 4722, radiographs = 10937) were randomly split into training/validation (80%, n = 3778, radiographs = 8998) and test (20%, n = 944, radiographs = 1939) datasets by patient. We evaluate and obtain the highest source identification for patient level model classification of pelvis (accuracy (ACC) = 98.6%, area under curve (AUC) = 95.34%, precision = 99.11%, recall = 98.6%) in content pixel analysis and for patient level model classification of lumbar spine (ACC = 97.2%, AUC = 97.34%, precision = 97.18%, recall = 97.2%) in content-free pixel analysis. This research confirms that the steganalysis can be performed on content-free and content pixels from radiographs. These results will be valuable for medical forensic and legal communities.

Deep Learning Evolved: Overcoming Sub-Optimal Local Minima with $(\mu/\rho + \lambda)-$Evolution Strategies

Pablo Rivas
School of Engineering and Computer Science, Department of Computer Science,
Baylor University, Waco, Texas, USA

Abstract—Integrating Evolution Strategies (ES) and Backpropagation (BP) within a deep neural network framework presents a significant challenge, as ES has previously only been shown to perform comparably to BP for smaller problems. In this study, we extend the application of ES to high-dimensional problems, using it to initialize the weights of a Deep Neural Network (DNN). Our experiments demonstrate that this novel ES approach can effectively overcome local minima and converge towards near optimal global solutions. Following ES initialization, we employ traditional BP gradient methods to further refine the weights based on the initial set provided by ES. A key finding of our research is the potential for ES to reduce the computational time required by traditional gradient-based backpropagation learning methods. This efficiency is achieved by providing an initial set of weights close to the global optimum, enabling the network to converge more rapidly than with random or zero-weight initialization approaches. Our approach offers a promising direction for future research in the efficient training of deep neural networks and opens up new possibilities for tackling high-dimensional problems with these networks.

Scaling Probabilistic Inference through Message Contraction Optimization

Martin Roa Villescas, Jin-Guo Liu, Patrick WA Wijnings, Sander Stuijk, Henk Corporaal
Eindhoven University of Technology, Netherlands;
Hong Kong University of Science and Technology, Hong Kong

Abstract—Within the realm of probabilistic graphical models, message-passing algorithms offer a powerful framework for efficient inference. When dealing with discrete variables, these algorithms essentially amount to the addition and multiplication of multidimensional arrays with labeled dimensions, known as factors. The complexity of these algorithms is dictated by the highest-dimensional factor appearing across all computations, a metric known as the induced tree width. Although state-of-the-art methods aimed at minimizing this metric have expanded the feasibility of exact inference, many real-world problems continue to be intractable. In this paper, we introduce a novel method for adding and multiplying factors that results in a substantial improvement in the inference performance, especially for increasingly complex models. Our approach aligns well with existing state-of-the-art methods designed to minimize the induced tree width, thereby further expanding the tractability spectrum of exact inference for more complex models. To demonstrate the efficacy of our method, we conduct a comparative evaluation against two other opensource libraries for probabilistic inference. Our approach exhibits an average speedup of 23 times for the UAI 2014 benchmark set. For the 10 most complex problems, the average speedup increases to 64 times, demonstrating its scalability.

Aligning Word Embeddings from BERT to Vocabulary-Free Representations

Alejandro Rodriguez Perez, Korn Sooksatra, Pablo Rivas, Ernesto Quevedo Caballero,
Javier S. Turek, Gisela Bichler, Tomas Cerny, Laurie Giddens, Stacie Petter
School of Engineering & Computer Science, Department of Computer Science,
Baylor University, Waco, Texas, USA;
Intel Labs, Intel Corporation, Portland, Oregon, USA;
School of Criminology & Criminal Justice, California State University,
San Bernardino, California, USA;
Information Technology & Decisions Sciences Department, G. Brint Ryan College of Business,
University of North Texas, Texas, USA;
School of Business, Management of Information Systems, Wake Forest University, North Carolina, USA

Abstract—This paper investigates the limitations of transformer-based models in handling a fixed vocabulary, which can lead to poor generalization of out-of-vocabulary words and domains. To address this, we explore the use of transfer learning from a vocabulary-rigid transformer to a vocabulary-free one by aligning the word-embedding layer. Our approach trains a CNN to mimic the word embeddings layer of a BERT model, using a sequence of byte tokens as input. By replacing the word embeddings layer of the baseline BERT model with the aligned CNN network, we evaluate the model's generalization performance and ability to handle a broader range of linguistic inputs. Our results demonstrate the advantages of using cosine-based loss functions in the alignment process. Our approach makes important contributions toward developing more flexible and robust NLP models. Index Terms—model distillation, word embeddings, bert, natural language proces

BERT Goes to SQL School: Improving Automatic Grading of SQL Statements

Pablo Rivas, Donald R. Schwartz, Ernesto Quevedo
School of Engineering & Computer Science, Baylor University, Waco, Texas, USA;
School of Computer Science & Math, Department of Computing Technology,
Marist College, Poughkeepsie, New York, USA

Abstract—Automated grading of SQL queries is a challenging task due to the complexity of the language and the variety of acceptable solutions for a given problem. In this paper, we propose a novel approach that leverages deep learning with a BERT model to understand the syntax and semantics of SQL statements. By training BERT on a dataset of SQL queries and their

corresponding grades, we create a model that can automatically grade new questions accurately. Our experiments demonstrate that the proposed methodology achieves high accuracy and consistency in grading SQL queries, outperforming existing state-of-the-art models. Furthermore, we provide an analysis of the model's explainability, revealing a new capability that can be extremely beneficial for understanding the decision-making process. Overall, our work demonstrates the potential of deep learning with BERT for improving the efficiency and accuracy of SQL query grading.

Assessing Author Personality Types Using ChatGPT

Curry Guinn
Department of Computer Science, University of North Carolina Wilmington, North Carolina, USA

Abstract— This paper argues that identifying the personality type of an author based on textual data such as essays, online conversations, blogs, and tweets can reach state-of-the-art performance with Large Language Models (LLMs) such as ChatGPT without any special training in the task. Further, these LLMs can provide natural language explanations as to why a particular personality type was chosen, which distinguishes this technique from most machine learning classification algorithms. This paper presents an experiment wherein ChatGPT is used to identify the Myers Briggs Personality Type (MBTI) based on a collection of tweets using the Kaggle MBTI Personality Type Dataset. Precision, recall, and accuracy results will be presented for each personality type dichotomy and compared to prior work.

Predicting Ocean Wave Fun Factor with Machine Learning using Buoy Data and Human Observations

Gulustan Dogan, Dillon Harless
University of North Carolina Wilmington, North Carolina, USA

Abstract - This study explores a niche of ocean wave prediction by using machine learning to predict ocean waves' "fun factor," which estimates how enjoyable surfing the waves would be. We created several machine learning models to achieve satisfactory results, including Random Forest regression and k-Nearest Neighbors. The data used to train these models came from a buoy in Wrightsville Beach, NC, and a business local to Wrightsville Beach called WBLiveSurf. WBLiveSurf's employees post a report approximately four times daily, including fun factor, swell size, and water surface. These reports were merged with the buoy data by timestamp and supplied to the various machine learning models implemented in this study. We also developed models to act as surrogates for swell size and water surface as a supplement. These values are human-observed and thus subject to variation and error. Further, the short-lived nature of a wave's peak amplifies this error. We discuss how we deal with the inconsistency introduced in the WBLiveSurf data due to the individual reporting changing throughout the day and week. We discuss the results of the various models, potential limitations, and a proposed model chain to estimate fun factor in real-time and the future at Wrightsville Beach.

Short Term Prediction of Rainfall in Columbia using a Generative Modeling Approach

Eron Neill, Joshua Edwards, Timothy Reitz, Ethan Cook, Gulustan Dogan, Narcisa Pricope
University of North Carolina, Wilmington, North Carolina, USA

Abstract—This study focused on the use of artificial intelligence to predict precipitation in Columbia. Accurate prediction of rainfall is a challenging problem. This is a serious issue due to the numerous deaths every year due to events caused directly by excess rainfall. While weather forecasting has advanced significantly throughout the years there is still significant room for improvement. One of the main factors detrimental to accurate weather prediction is the multitude of variables required to achieve accurate prediction. Our research seeks to use publicly available satellite imagery to achieve accurate prediction of rainfall with relatively little input data. We seek to achieve that goal by using a custom generative modelling approach to make predictions.

Cellphone-Based sUAS Range Estimation: A Deep-Learning Approach

Ryan D. Clendening, Richard Dill, Brett J. Borghetti, Brett Smolenski, Darren Haddad, Douglas D. Hodson
Department of Electrical and Computer Engineering,
Air Force Institute of Technology, Dayton, Ohio, USA;
North Point Defense, Rome, New York, USA;
Info. Exploit. Branch, Air Force Research Labs, Rome, New York, USA

Abstract—Small Unmanned Aircraft Systems (sUAS) are accessible platforms that pose a security threat. These threats warrant affordable and accurate methods for tracking sUAS. We apply neural network-based methods to predict sUAS range from cellphone acoustic recordings; the data comes from twenty-eight cellphones recording three different sUAS that fly over the devices. The timestamped acoustics data is transformed into 0.5s Mel-spectrograms frames and 0.5s raw audio frames. Truth values are calculated using euclidean distance from the sUAS to a cellphone and split into four range classes. The data is sequestered into an 80/20 training-test split and is used to train three different architectures. The 2DCNN architecture outperforms the other architectures (1DCNN and 2DCRNN). The 2DCNN is then re-trained to generalize sUAS range with various sUAS models and achieves an average Macro-F1 score of 0.758 across different sUAS models. The results show that deeplearning-based sUAS ranging with cellphones is an effective and low-cost method for accurately tracking sUAS.

Harnessing the Efficiency of Reformers to Detect Software Vulnerabilities

Marwan Omar, Rebet Jones
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA

Abstract - Detecting software vulnerabilities is a critical task in ensuring the security of software systems. Recent advances in deep learning have shown promise in improving the accuracy of vulnerability detection. In this work, we propose a new approach for vulnerability detection using the Reformer language model. We compare our approach with previous works on three datasets: SARD, D2A, and Devign. Our experiments show that our approach outperforms the state-of-the-art methods on all three datasets, achieving a precision of 0.95, recall of 0.91, and F1-score of 0.93 on average. Moreover, our approach is computationally efficient, making it suitable for large-scale vulnerability detection tasks. We also provide an in-depth analysis of the results and discuss the implications of our findings. Our work contributes to the growing body of literature on deep learning-based vulnerability detection and demonstrates the potential of the Reformer model in this domain.

Plant Disease Diagnosis Using Transfer Learning Based Models

Vatsala Anand, Sheifali Gupta, Rupesh Gupta, Hani M. Alshahrani, Mohammed Hamdi,
Mana Saleh Al Reshan, Asadullah Shaikh
Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India;
Department of Computer Science, College of Computer Science and Information Systems,
Najran University, Najran, Saudi Arabia;
Department of Information Systems, College of Computer Science and Information Systems,
Najran University, Najran, Saudi Arabia

Abstract— Plant disease diagnosis is the process of identifying and characterizing diseases that affect plants. The primary goal of plant disease diagnosis is to determine the cause of the disease, and to develop appropriate control and management strategies to minimize its impact on crop yields and quality. In this research, four transfer learning models such as AlexNet, VGG16, MobileNetV2, and InceptionV3 has been used for the classification of plant disease into different categories. Out of all the four transfer learning models, the best performing model is InceptionV3 with the value of accuracy as 0.92 and precision as 0.84.

Detecting IoT Malware with Knowledge Distillation Technique

Rebet Jones, Marwan Omar
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA

Abstract— The increasing prevalence and sophistication of Internet of Things (IoT) malware pose significant security threats to IoT devices and networks. Traditional machine learning models for detecting IoT malware are often resource-intensive and computationally expensive, making them unsuitable for deployment on resource-constrained IoT devices. To address this challenge, we propose a novel approach for detecting IoT malware using knowledge distillation, where a teacher model is trained on a large and diverse set of data to provide a compact and efficient student model. We evaluate our approach on two widely used datasets, IoT-23 and Malevis, and achieve promising results, demonstrating the effectiveness of our approach in detecting IoT malware. Our approach can help address the resource constraints and computational challenges of deploying traditional machine learning models on IoT devices. Our study highlights the potential of knowledge distillation as a promising approach for developing lightweight and efficient models for IoT malware detection. Our approach can be extended to detect other types of IoT threats and integrated with existing IoT security systems to provide an additional layer of defense against malware attacks. The future research directions outlined in this paper can help further advance the field of IoT security and enable the development of more effective and efficient approaches for detecting and mitigating IoT malware attacks.

A Multi-Agent System for ISR Asset Planning

Wilmuth Muller, Frank Reinert, Uwe Pfirrmann, Dirk Muhlenberg, Christian Ellmauer
Interoperability and Assistance Systems, Fraunhofer IOSB, Karlsruhe, Germany

Abstract—In this paper an approach for an optimal planning of intelligence, surveillance, and reconnaissance (ISR) asset deployment in order to satisfy the information needs of a commander is presented. Based on the processes of information requirements management (IRM) and collection management (CM), a two-step approach has been developed. In the first step, an operator assigns to each target on which reconnaissance or surveillance has to be performed a set of suitable assets. The operator may assign the suitable assets to a target either directly, based on his experience and knowledge, or supported by an interactive asset selection assistant component of the application, or supported by an intelligent multi-agent system, which generates automatically an asset assignment proposal. In the second step an optimal asset assignment and execution order is computed. The multi-agent system consists of three types of intelligent agents, a target agent representing the targets on which reconnaissance or surveillance has to be performed, the asset agents representing the assets available to the operator, and an interface agent responsible for the communication with the other components of the application.

Obstacles Avoidance Using Reinforcement Learning for Safe Autonomous Navigation

Mokhles M. Abdulghani, Abdulghani M. Abdulghani, Wilbur L. Walters, Khalid H. Abed
Jackson State University, Jackson, Mississippi, USA

Abstract - Autonomous agents can safely navigate environments around them when equipped with advanced hardware such as sensors and controlled with advanced Artificial Intelligence (AI). AI is a powerful science that can be employed to provide the highest safety for agents. AI safety is essential to provide reliable services to consumers in various fields such as military, education, healthcare, and automotive. This paper presents an AI safety algorithm for safe autonomous exploring in a chosen environment. The Reinforcement Learning was used to design the proposed AI safety algorithm. The designed algorithm was tested in virtual reality using the Unity environment. A 0.62% goal collision ratio was achieved, and the collision incidents were minimized from 134 to 54 in the Unity environment within 30 minutes.

Exploring the Relationship between Air Pollution and CNS Disease Mortality in Italy: A Forecasting Study with ARIMA and XGBoost

Mostafa Karami, Sahand Hamzehei, Farzaneh Rastegari, Omid Akbarzadeh
Electronics and Telecommunications, Polytechnic University of Turin, Turin, Italy;
Computer Science and Engineering Department, University of Connecticut, Storrs, Connecticut, USA

Abstract—There is a correlation between air pollution and Central Nervous System (CNS) disease, according to the research. Air pollution has been linked to an increased risk of developing neurodegenerative diseases like Alzheimer's and Parkinson's. Airborne toxins can penetrate the brain and induce inflammation, resulting in damage and cognitive decline. Air pollution is a significant environmental health issue that can harm people's health in a variety of ways, such as by increasing their risk of heart and respiratory diseases. From 2010 to 2020, we looked at data from 107 Italian communes about air pollution and deaths caused by CNS diseases like Alzheimer's and Parkinson's. Then, we compared the performance of the Autoregressive Integrated Moving Average (ARIMA) model and Extreme Gradient Boosting (XGBoost) with grid search in terms of accuracy and training time for forecasting the number of deaths caused by CNS diseases in the forthcoming year. According to our research, both models could accurately forecast how much PM2.5 and PM10 were in the air and how many people would die from CNS diseases. Also, ARIMA is more accurate than XGBoost, but it takes longer to train. Both ARIMA and XGBoost are good at forecasting CNS death based on air pollution, which is what our result shows. Overall, our findings indicate that air pollution is a significant risk factor for CNS diseases, and accurate forecasting models can aid in mitigating its effects.

Microgrid Intelligent Agent Control

Mohammed Alaqtash, Salem Al-Agtash, Khaled Shaalan
Department of Computer Science, The British University in Dubai, Dubai, United Arab Emirates (UAE);
Department of Computer Science and Engineering, Santa Clara University, Santa Clara, California, USA;
Gerrman Jordanian University, Amman, Jordan

Abstract—Microgrid control is complex due to its need to accommodate the intermittence of renewables, balance generation with load, transit between grid-connected and islanded modes, and maintain reliable power supply to customers. Much research has addressed microgrid control complexity in both centralized and decentralized settings. This paper presents an intelligent software agent control with advanced autonomous capabilities to address the intermittent nature of renewables and their integration in realworld scenarios. Such capabilities include data acquisition, load and renewable generation forecasting, energy management, scheduling, optimal power flow, and real-time control to maintain generation-load balance in a secure and reliable microgrid environment. Accurate AI predictive models, machine learning algorithms, and non-linear optimization will be at the core function of the control agents.

Can CNNs be used to Predict Multi-physics Simulations of Microwave Induced Damage to Basalt?

Reena R. Patel, Jeff Allen, Tomas Mondragon, Oliver Taylor
Information Technology Laboratory, U.S. Army Engineer Research and Development Center,
Vicksburg, Mississippi, USA;
Geotechnical and Structures Laboratory, U.S. Army Engineer Research and Development Center,
Vicksburg, Mississippi, USA

Abstract—Understanding the fundamentals of microwave induced damage by conducting multi-physics simulations is a complex task that requires the integration of several physical phenomena and models. In such scenarios, convolutional neural networks (CNNs) can be used to extract intrinsic behavior and patterns from large complex datasets generated from simulations and experiments. The goal of this research is to use temperature images obtained from COMSOL simulations, density, relative permittivity, and power of microwave to train and validate CNN. The data obtained from COMSOL simulations is pre-processed to make it suitable for use as input to the CNN. This process involves normalizing and resizing the image dataset. The data is split

into two sets, namely the training and validation set. The CNN is trained using the training set where the CNN learns the inherent characteristics of the data. Upon successful completion of the training, the CNN is tested using previously seen and unseen data. The CNN perceptions indicate a high prediction accuracy : 76%

AI Resistant (AIR) Cryptography

Gideon Samid
Electrical, Computer and System Engineering, Computer and Data Sciences,
Case Western Reserve University, Cleveland, Ohio, USA

Abstract— highlighting a looming cyber threat emanating from fast developing artificial intelligence. This strategic threat is further magnified with the advent of quantum computers. AI and quantum-AI (QAI) represent a totally new and effective vector of cryptanalytic attack. Much as modern AI successfully completes browser search phrases, so it is increasingly capable of guessing a rather narrow a-priori list of plausible plaintexts. This guessing is most effective over device cryptography where the message space is limited. Matching these guesses with the captured ciphertext will greatly accelerate the code breaking process. We never faced such a plaintext-originated attack on a strategic level, and never had to prepare for it. Now we do. Proposing to apply a well-known martial art tactics: using the opponent's strength against them: constructing ciphertexts that would provide false answers to the AI attacker and lead them astray. We are achieving this defensive measure by pivoting away from the norm of small, known-size key and patternloaded ciphers. Using instead large keys of secret size, augmented with ad-hoc unilateral randomness of unbound limits, and deploying a pattern-devoid algorithm with a remarkably low computational burden, so it can easily handle very large keys. Thereby we achieve large as desired unicity distances. This strategy has become feasible just when the AI threat looms. It exploits three new technologies coming together: (i) non-algorithmic randomness, (ii) very large and inexpensive memory chips, and (iii) high throughout communication networks. These pattern-devoid, randomness rich ciphers also turn up to be an important option in the toolbox NIST prepares to meet the quantum challenge. Avoiding the computational load of mainstay ciphers, AIR-cryptography presents itself as the ciphers of choice for medical, military and other batterylimited devices for which data security is paramount. In summary: we are pointing out a fast emerging cyber challenges, and laying out a matching cryptographic answer.

mCLESS: The Multi-Class Least-Error Square Sum for Interpretable Classification

Hwan Hee Park, Seung Heon Lee, Hwamog Kim, Seongjai Kim
Department of Mathematics and Statistics, Mississippi State University, Mississippi State, Mississippi, USA;
Department of Biochemistry, Mississippi State University, Mississippi State, Mississippi, USA;
Department of Sciences & Mathematics, Mississippi University for Women, Columbus, Mississippi, USA

Abstract—Some machine learning algorithms are considered as black boxes, because the models are sufficiently complex and they are not straightforwardly interpretable to humans. Lack of interpretability in predictive models can undermine trust in those models in many application areas. The article introduces a new interpretable machine learning algorithm, called the MultiClass Least-Error Square Sum (mCLESS). It is linear, simple to implement, and interpretable. Its nonlinear expansion is discussed. This simple algorithm turns out to be superior to many popular machine learning algorithms. Various experimental results involving synthetic datasets and UCI datasets are given to verify the claim.

A Chronological and Cooperative Route Optimization Method for Heterogeneous Vehicle Routing Problem

Yosuke Yamaguchi, Zhao Wang, Yuusuke Nakano, Jun Ohya, Katsuya Hasegawa
School of Creative Science and Engineering, Waseda University, Tokyo, Japan;
NTT Network Service Systems Laboratories, NTT Corporation, Tokyo, Japan;
Institute of Space and Astronautical Science,
Japan Aerospace Exploration Agency, Tokyo, Japan

Abstract— This paper focuses on a deep reinforcement learning (DRL)-based approach for Heterogeneous Vehicle Routing Problem, where each vehicle in the fleet is characterized by its capacity and speed. Previous methods fail to optimize the route for a vehicle while cooperatively considering the current state (e.g., the remaining capacities and the positions) of other vehicles in the fleet. To solve this problem, we first propose a chronological Markov Decision Process, which makes the current state of the whole fleet available. Second, we propose a fleet encoder, a specific network architecture to promote cooperative route generation among vehicles by incorporating the fleet state. Experimental results show that our method outperforms other DRL-based methods on problems with any number of customers and vehicles.

A Comparative Study of DE, GA and ES for Evolutinary Reinforcement Learning of Neural Networks in Pendulum Task

Hidehiko Okada
Faculty of Information Science and Engineering, Kyoto Sangyo University, Kyoto, Japan

Abstract—Reinforcement learning of neural networks requires gradient-free algorithms as labeled training data are not available. Evolutionary algorithms are well-suited for this purpose since they do not rely on gradients. However, the success of training neural networks with evolutionary algorithms is contingent on the careful selection of appropriate algorithms, given the numerous algorithmic variations available. In this study, the author evaluates the efficacy of Differential Evolution (DE), Genetic Algorithm (GA), and Evolution Strategy (ES) for the reinforcement learning of neural networks, utilizing a pendulum control task. The experimental results indicate that DE exhibits statistically significant superiority over GA and ES. While GA performs better than ES, this difference is not statistically significant. The study highlights DE's ability to effectively balance between exploratory and exploitative search, adapting to the problem at hand. Based on these findings, it is suggested that an algorithm possessing such characteristics is better suited for evolutionary reinforcement learning of neural networks.

Siamese-NAS: Using Trained Samples Efficiently to Find Lightweight Neural Architecture by Prior Knowledge

Yu-Ming Zhang, Jun-Wei Hsieh, Chun-Chieh Lee, Kuo-Chin Fan
Department of Computer Science, National Central University, Taoyuan, Taiwan;
College of AI and Green Energy, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

Abstract—In the past decade, many architectures of convolution neural networks were designed by handcraft, such as Vgg16, ResNet, DenseNet, etc. They all achieve state-of-the-art level on different tasks in their time. However, it still relies on human intuition and experience, and it also takes so much time consumption for trial and error. Neural Architecture Search (NAS) focused on this issue. In recent works, the Neural Predictor has significantly improved with few training architectures as training samples. However, the sampling efficiency is already considerable. In this paper, our proposed Siamese-Predictor is inspired by past works of predictor-based NAS. It is constructed with the proposed Estimation Code, which is the prior knowledge about the training procedure. The proposed Siamese-Predictor gets significant benefits from this idea. This idea causes it to surpass the current SOTA predictor on NASBench-201. In order to explore the impact of the Estimation Code, we analyze the relationship between it and accuracy. We also propose the search space Tiny-NanoBench for lightweight CNN architecture. This well-designed search space is easier to find better architecture with few FLOPs than NASBench-201. In summary, the proposed Siamese-Predictor is a

predictor-based NAS. It achieves the SOTA level, especially with limited computation budgets. It applied to the proposed Tiny-NanoBench can just use a few trained samples to find extremely lightweight CNN architecture.

Analysis of Factors Influencing the Severity of Coronavirus Symptoms Using Predictive Modeling

Anatoli Nachev
Business Information Systems, Cairnes Business School, University of Galway, Ireland

Abstract—This paper presents a case study on the IPUMS NHIS database, which provides data from censuses and surveys on the health of the U.S. population, including data related to COVID-19. By addressing gaps in previous studies, we propose a machine learning approach to train predictive models for identifying and measuring factors that affect the severity of COVID-19 symptoms. Our experiments focus on four groups of factors: demographic, socio-economic, health condition, and related to COVID-19 vaccination. By analysing the sensitivity of the variables used to train the models and the variable effect characteristics (VEC) analysis on the variable values, we identify and measure importance of various factors that influence the severity of COVID-19 symptoms.

A Proposal for Electrodes and Frequency Bands that Combine Performance and Experimental Cost in EEG Analysis of Motions

Sho Kanemoto, Ryota Nishiura, Seiji Tsuchiya, Koichi Watabe
Graduate School of Science and Engineering, Department of Computer Science and Engineering, Doshisha University, Kyoto, Japan

Abstract— In this paper, we utilize acquired EEG motion data to verify the reproducibility of pyramidal tract crossings, which are highly associated with motion. Previous studies faced the challenge of unclear detection of motion-related features in the acquired EEG data. Furthermore, the EEG data from 32 channels often contained significant noise, resulting in low prediction accuracy. However, through the validation conducted in this study, we successfully identified electrode-based features associated with motion. By reducing the number of electrodes to six, based on the extracted features, we were able to improve the accuracy of motion prediction and reduce computational costs. The reduction in the number of electrodes also suggests a potential decrease in the burden on the subject wearing the EEG measuring device.

Automated Orchestration Systems based on Deep Learning

Ryosuke Kamada, Seiji Tsuchiya, Hirokazu Watabe
Intelligent Information Engineering Graduate School of Science and Engineering, Doshisha University, Kyoto, Japan

Abstract— The demand for "programmed music" has increased, but orchestration requires expertise to allocate appropriate instruments to melodies. This study uses deep learning to create an instrumentation that assigns the optimal instrument for orchestration to the required melody. Bidirectional Encoder Representations from Transformers (BERT) was used to determine which notes and melodies suit specific instruments. BERT was originally developed for natural language processing, so the music data had to be tokenized and converted into vectors. Instrument information was incorporated into each token, and the relationship between tokens was learned. This method produced an AI system that can automatically assign melodies or notes to the most suitable instruments. We believe that this study will lower the hurdle for composing difficult instrument configurations.

Comparative Analysis of Deep Learning Approaches for Analysis and Prediction of Multivariate Time Series Data

Nikhila Vintha, Kaur Devinder
Electrical & Computer Science Engineering, University of Toledo, Toledo, Ohio, USA

Abstract—In recent years deep learning has become the desired approach to analyse long term dependencies in data to build predictive models to forecast future outcome. They have been extensively used for natural language processing, computer vision, image processing, and for analysing sequential data. In this paper, deep variants of Recurrent Neural Network (RNN) called Long Short-Term Memory Network (LSTM), have been used to build and predict the outcomes of multivariate time series data with reference to multivariate stock market prediction data of Meta, Ford, and FedEx. Comparative performance analysis of three variants of LSTM viz., Vanilla LSTM, Stacked LSTM, and Bi-directional LSTM was conducted. Four input features were chosen, viz., opening stock price, high and low prices for the stock and adjusted closing price for a given day. The predicted output was the closing price for the day or the opening price for the next day. Traditional time series prediction methods often fail to capture the nonlinear and dynamic patterns present in stock price data, leading to inaccurate predictions. Deep learning, a subset of machine learning, has shown promising results in various fields, including time series prediction. Our results demonstrate that the proposed multivariate deep learning model Bi-directional LSTM outperforms the other two LSTM models in terms of accuracy and performance.

Construction of Concept-Base including Internet-Derived Words

Misako Imono, Seiji Tsuchiya, Hirokazu Watabe
Department of Information Systems, Daido University, Nagoya, Japan;
Department of Intelligent Information Engineering, and Science, Doshisha University, Kyoto, Japan

Abstract— A concept base is a knowledge base that defines the meaning of a word as a set of pairs of words (attributes) that can be associated with the concept and weights that express the importance of the attributes. Unlike the construction of vectors by machine learning, the structure allows humans to understand the meaning of the constructed knowledge. Existing concept bases define words that appear in Japanese dictionaries and newspaper articles. In recent years, however, many words have been created on the Internet and have come to be used in everyday life. These words do not appear in Japanese dictionaries or newspaper articles, and therefore are not defined in the existing concept base. The purpose of this study is to define these words by creating a concept base using Internet-derived words.

An Analysis of Research Trends on Language Model using BERTopic

Woojin Kim, Yumi Kim, Heesop Kim, Jongwook Lee
Department of Library and Information Science, Kyungpook National University,
Daegu, South Korea

Abstract—Although language models have played a crucial role in various natural language processing tasks, there has been little research that focuses on systematic analysis and review of research topic trends in these models. In this paper, we conducted a comprehensive analysis of 31 years of research trends in the field of language models, using publications from Scopus, an internationally renowned academic database, to identify research topics related to language models. We adopted BERTopic, a state-of-the-art topic modeling technique, on the 13,754 research articles about language models. The research on language models has gradually increased since 1991, and there is a sudden increase in the number of publications with the emergence of BERT and GPT in 2018. We assigned 14 main topics with meaningful keywords clustered by BERTopic model. Among 14 topics, research related to speech recognition, statistical language models, and pre-trained language models demonstrated the most vigorous research fields. Our results demonstrate a more systematic and comprehensive trend in language model research, which is expected to provide an important foundation for future research directions.

Impacts of Catastrophic Forgetting: From a Machine Learning Engineering Perspective

Timothy Elvira, Juan Ortiz Couder, Tyler Thomas Procko, Omar Ochoa
Department of Electrical Engineering and Computer Science,
Embry-Riddle Aeronautical University (ERAU), Daytona Beach, Florida, USA

Abstract - Software validation is the method of ensuring that a software system is built correctly, meets user needs, and functions according to defined requirements. In the current age, Machine Learning (ML) and Deep Learning (DL) are at the forefront of technological innovation. Software engineering practices are being adopted in the ML/DL domain to assist in defining and mapping software validation anomalies, e.g., catastrophic forgetting. A ML model can be engineered, trained, and used to fulfill a requirement specification based on its classification capabilities. Catastrophic forgetting can cause a model to forget a class fully or partially insofar as a requirement is no longer fulfilled. This paper posits that catastrophic forgetting is an issue of ML validation and presents the issues of catastrophic forgetting in the novel domain of ML engineering. Ideally, this paper will spark a discussion and lend insight to the future direction of ML validation and verification. Future work includes developing case studies to observe the validation and verification challenges of catastrophic forgetting. The authors suggest researchers explore other ML/DL anomalies that might impact validation and verification.

Co-Evolutionary Modified Black Hole Constraint Satisfaction: Application to Piezoelectric Actuator Control

Saber Yaghoobi, M. Sami Fadali
Electrical & Biomedical Enfineering Department, University of Nevada, Reno, Nevada, USA

Abstract - Many engineering design problems can be formulated as constrained satisfaction. Because of their simplicity and easy implementation, penalty function methods have been the most popular methods for constrained optimization. One major problem is that it is not easy to select suitable penalty weights. This paper proposes a co-evolutionary modified black hole optimization approach (CMBH) for constrained optimization using co-evolution to adapt penalty weights. MBH is applied with two types of populations for evolutionary exploration and exploitation in spaces of both solutions and penalty weights. The proposed CMBH is population-based and easy to implement in parallel. We apply the new approach to the design of a controller for a piezoelectric actuator Simulation results demonstrate the effectiveness, efficiency and robustness of the proposed method. The results show that CMBH can yield better results than other methods reported in the literature.

A Comparative Study of Fishing Activity Detection Approaches in Maritime Surveillance

Vinicius D. do Nascimento, Tiago A. O. Alves, Diego L. C. Dutra, Sandip Kundu
Systems Engineering and Computer Science Program (PESC),
Federal University of Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil;
Mathematics and Statistics Institute, State University of Rio de Janeiro (UERJ), Rio de Janeiro, Brazil;
Electrical and Computer Engineering (ECE) Department,
University of Massachusetts, Amherst, Massachusetts, USA

Abstract—Maritime authorities (MA) must track fishing vessels to ensure that fishing activities are limited to permitted areas. If illegal fishing is suspected, resources must be allocated to intercept and inspect the vessels. Thus, a false flag by the MA is costly, so it is important to use accurate detection methods. We compare the accuracy and computational time of the main approaches for detecting fishing activities described in the literature, using the Global Fishing Watch (GFW) dataset. We find that Long Short-Term Memory (LSTM) neural networks achieves an optimal accuracy of 1.00, while the random forest approach comes second with an accuracy of 0.87. Given the high cost of mistakes for MA, we conclude that the LSTM's high computational cost is worthwhile.

Designing User-Centered Artificial Intelligence to Assist in Recovery from Domestic Abuse

Sneha Talwalkar, Courtney Crooks
School of Cybersecurity and Privacy, Georgia Institute of Technology, Atlanta, Georgia, USA;
School of Medicine, Emory University, Atlanta, Georgia, USA

Abstract—The Centers for Disease Control (CDC) found that millions of people in the United States experience Intimate Partner Violence (IPV) each year [1]. Both IPV perpetrators and cyber criminals use manipulation, coercion, and intimidation to control their victims. In this research, we found how cybersecurity professionals could best assist local mental health intervention teams with the reduction of safety risks for patients trying to escape IPV. We developed partnerships between local mental health and cybersecurity professionals who build cognitive models to predict adversarial behavior and develop technological interventions that support a patient's safety, agency, and resilience during recovery from IPV.

Data Augmentation Using Brightness and Darkness to Enhance the Performance of YOLO7 Object Detection Algorithm

Abdulghani M. Abdulghani, Mokhles M. Abdulghani, Wilbur L. Walters, Khalid H. Abed
Jackson State University, Jackson, Mississippi, USA

Abstract—The object detection field depends on various techniques for duplicating the dataset without adding additional images. The data augmentation is the recommended technique that can be used to duplicate the dataset. In this paper, we trained YOLOv7 with a dataset that is part of the Open Images dataset that has 8,600 images with four classes (Car, Bus, Motorcycle, and Person). We used two different data augmentations techniques for duplicates and improvement of our dataset, and we compared the result of the proposed augmented data using brightness and darkness with the result of the original data. The evaluation result for the augmented data gives a promising result for every object, and every kind of data augmentation gives a different improvement. The mAP@.5 of all classes was 76%, and F1- score was 74%. The proposed improved methods increased the mAP@.5 value by +13% and F1-score by +11%.

Data Augmentation with Noise and Blur to Enhance the Performance of YOLO7 Object Detection Algorithm

Abdulghani M. Abdulghani, Mokhles M. Abdulghani, Wilbur L. Walters, Khalid H. Abed
Jackson State University, Jackson, Mississippi, USA

Abstract—Artificial intelligence is a powerful tool to improve the quality of our life, especially when it comes to object detection. The availability of large datasets is essential for improving the accuracy of object detection algorithms. Data augmentation techniques can double or even triple the capacity of the available dataset. In this paper, we trained YOLOv7 with 8,600 images with four classes: Bus, Car, Motorcycle, and Person. We used two different data augmentation techniques to duplicate and improve the dataset. The performance of the object detection algorithm was compared when using the proposed augmented dataset with noise and blur and when using only the original dataset. The evaluation result for the augmented dataset revealed a promising result where the mAP@.5 of all classes was 76%, and the F1-score was 74%. The proposed improved methods increased the mAP@.5 value by +13% and F1-score by +10%.

Performance Evaluation of CNN Models in Urban Acoustic Event Recognition through MFCC Hyperparameter Search

Ricardo Tangarife, Ricardo Velasquez, Fredy Rivera
University of Antioquia, Medellin, Colombia

Abstract—This paper focuses on acoustic event recognition, a significant research area in various fields, including security, medicine, and entertainment. The study explores the design space of combining convolutional neural networks with Mel Frequency Cepstral Coefficients to extract audio features. This work explicitly investigates three acoustic events relevant to citizen security: gunshots, screams, and sirens. We aim to find the optimal combination of hyperparameters to train accurate models with low computational requirements. The proposed approach achieved impressive F1-Scores of 95% for sirens, 97.2% for gunshots, and 99% for screams. Furthermore, considering computational complexity, our results demonstrate that these sounds could be utilized in real-time acoustic event recognition systems for citizen security applications.

Designing an Electroencephalograph to Predict Suicidal Behavior

Diana Karen Rangel Aguilar, Julio Cesar Martinez Romo, Francisco Javier Luna Rosas, Marisol Luna Medina
Instituto Tecnolologico Nacional campus Aguascalientes, Mexico;
Universidad Autonoma de Aguascalientes, Mexico

Abstract - Electroencephalogram (EEG) is a noninvasive recording technique used to measure the electrical activity of the brain. It has traditionally been used to diagnose neurological disorders and epilepsy. However, recently there has been great interest in using this technique in conjunction with machine learning to predict suicidal behavior. Since suicide is a global problem, this tool can be effective in early detection of suicidal behavior and help specialists make more informed decisions.

Reinforcing a Monitoring System of a Regulated River using LSTM

Md Julfiker Al Jewel, Yhocsan U. Guevara-Garcia, Md Muminul Hossain, Fred Wu, Fernando Rojano
Department of Mathematics, Engineering and Computer Science,
West Virginia State University, Institute, West Virginia, USA;
Facultad de Estadistica e Informatica, Universidad Veracruzana, Xalapa, Ver, Mexico;
Department of Computer Science, University of Memphis, Tennessee, USA;
R. Gus Douglass Institute and Department of Biology,
West Virginia State University, Institute, West Virginia, USA

Abstract— Accurate methods predicting flow in a river require high performance computing systems providing only short-term early warnings. To address these limitations, we propose an optimized Long Short Term Memory layer combined with a Fully connected layer within a six-layer architecture to estimate flows based on readings of three neighboring river gage heights. This study assessed the number of cells in the LSTM and the number of neurons in the Fully connected layer as well as the probability in the Dropout layer to avoid overfitting. Identification of the best settings by means of the RMSE evaluated six different input datasets to estimate flow. Our study proved that relying on the river gage height from neighboring sensors is possible to predict flow which is crucial when a sensor may fail and a reliable neural network at low computational cost can handle that.

A Reinforcement Learning Approach to Training Chess Engine Neural Networks

Lily Young, Byeong Kil Lee
Department of Electrical and Computer Engineering, The University of Colorado, Colorado Springs (UCCS), Colorado, USA

Abstract - Neural networks are powerful tools that allow computers to process complex data using trained heuristics to produce an output that is useful for users. These networks tend to be faster than comparable algorithms and are much more flexible, which makes them ideal for real-time object and facial detection or very large dataset anomaly detection. There are many different methods for training such a network, however, most require large quantities of hand-labeled data. While the required amount of data labeled varies, it can range from hundreds of thousands of entries to millions. In many cases, this presents a problem because of cost or time requirements. In some situations, however, it is possible to use a learning method called reinforcement learning.For reinforcement learning, the network interacts with a system and is trained based on the outcome of its decisions. In this paper, the ability of a relatively thin neural network (<20 layers) to learn complex decision-making through reinforcement learning was studied. The network was trained to evaluate positions in the board game chess, a game that has relatively simple ruleset but complex interactions. Throughout our testing, it was determined that this method of training was very difficult to implement correctly and that injecting randomness was very important to make reinforcement learning function. It was also determined that the time to train a network using reinforcement learning far exceeds that of more traditional training approaches. Despite our limited training time of one week, we show that reinforcement learning may be effective for problems where good training data is not available. We also show that reinforcement learning may be an effective way to train networks on more complex or open-ended problems.

Fine-Tuned Large Language Models for Improved Clickbait Title Detection

Chandra N. Sekharan, Pavan Sai Vuppala
Department of Computer Science, Texas A&M University Corpus Christi, Texas, USA

Abstract—The term clickbait is used to describe headlines or other content that is designed to attract attention and clicks, often by using sensationalized or misleading language. However, the task of detecting click-bait titles automatically can be challenging due to the wide variance in the wording and structure of titles. Rule-based approaches, machine learning models, deep learning models, and natural language processing techniques are some of the existing methods for click-bait detection. As good as these techniques are, they still need customization and large amounts of training data to achieve good accuracy. Since generative, pretrained transformer models have shown an inherent capability to understand the meanings of sentences really well, our research examined the use of GPT-3.5, and GPT-4 as zero-shot models. They resulted in a maximum test accuracy of 88.5%, lower than some of the custom techniques reported in previous research works of between 90% to 95%. This indicates the need to go beyond a generic LLM. We then fine-tune OpenAI's Large Language Model (LLM) Ada using an efficient dataset of only 1000 samples of clickbait and non-clickbait titles to get a test accuracy of up to 99.5 %. The research further shows that finetuning a GPT model can yield not only better accuracy but also consume smaller amount of training data, shorter development time, and be cost-effective. Our research is indicative of the broader possibility that fine-tuning a large language model is all you need for most targeted natural language processing tasks.

Accurately Forecasting the Energy Cost in Home-Based Small Businesses by Applying Deep Neural Networks

Edwin Arrey Agbor, Yanzhen Qu
Colorado Technical University, Colorado Springs, Colorado, USA

Abstract—The data generated within small homebased businesses has tremendously increased with the exponential growth of data from diverse sources, including structured and unstructured data sources. Small home-based businesses rely on electricity energy as the primary source of power in production, thereby establishing energy cost as a major impacting factor to profitability of the business. Despite technological advancements, small home-based business continues to rely on traditional data analysis

solutions such as gut-based or naïve analytics models (often built into spreadsheets) in analyzing data to provide answers to business questions that drives decision making. The high dependency of business decisions on forecasting, implies a high benefit of utilizing the new advanced deep learning techniques (such as LSTM) having capabilities of meeting the ever-changing data needs. This paper has proposed an experimental quantitative method, leveraged with walk-forward validation and RMSE as evaluation metric in measuring performance of all three algorithms (LSTM, ARIMA and naïve base model) on the same dataset, to demonstrate that it is possible to address decision making challenges arising from the lack of usage of advanced deep learning techniques in small home-based businesses for improved electricity consumption cost prediction and time series forecasting.

Assessing the ChatGPT and GPT-4 on Pharmacy Education

Larry E. Roberts Jr., Mesay Abomsa, Lillian Onokala, Simon Wang
College of Pharmacy, Howard University, Washington, DC, USA

Abstract - Large language models (LLMs) has been at the forefront of providing innovative solutions in improving learning outcomes for students [4]. OpenAI is a nonprofit research company that had the objective to develop and direct LLMs which benefits humanity [6]. In November 2022, the company released ChatGPT which is a powerful interactive AI chatbot that engages in conversational question-and-answering dialogue [5]. Later, in March 2023 the company released a newer version of the language model system known as GPT-4 [8]. It is a large multimodal model that is able to accept both text and image inputs and outputs human-like text. The objective of this study is to assess the utilization of LLMs in the pharmacy education [6].

Visual Memory Transfer for Imbalanced Medical Image Classification

Ke Yao, Yan Liu
Concordia University, Montreal, Canada

Abstract - In deep learning of medical image data, skin lesion classification remains a challenging problem due to imbalanced distribution of training data. We propose a visual memory transfer (VMT) method by means of transferring visual knowledge from majority classes to minority classes. As a result, our method enriches feature of minority classes with pre-calculated memory features. In addition, VMT defines a refined feature map to perform fine-grained classification. Our classification results outperform SOTA methods on the largest public available dermoscopic image dataset on averaged F-score and Top-1 classification accuracy.

Automated Generation of Health Care Dynamic Recommendation Reports through GPT-powered Interoperability in Health Care IoT Environment

Shih-Yang Yang, Yu-Wei Qiu, Ling-Hsiang Hung, Ching-Hu Tasi
Department of Media Arts, University of Kang Ning, Taipei, Taiwan;
Department of Information Management, National Taipei University of Nursing and Health Sciences, Taipei, Taiwan;
Department of Educational Psychology and Counseling, National Tsing Hua University, Hsinchu, Taiwan;
Department of Long Term Care, University of Kang Ning, Taipei, Taiwan

Abstract—Due to the global aging population and low birth rates, there is a growing emphasis on improving the living environment for the elderly. In Taiwan, the government has implemented a policy of long-term home care and introduced information and communication technology applications to increase care and services for the elderly. This study focuses on the use of physiological measurements as the main topic of investigation. The intuitive information provided by networked devices may be difficult for the elderly to understand and can be perceived as apathetic. Additionally, the elderly often has weaker information capabilities. Therefore, this study explores the innovative use of AI-generated technology combined with physiological measurements to provide feedback for elderly home care. Based on the development of artificial intelligence applications, machine learning technology has become a widespread information tool, generating multimedia-based care suggestions that are easy to understand and flexible.

Determinants of Global Hunger Index

Saatwik Panigrahi, Swadhin Rout, Anasse Bari
Computer Science Department, CIMS New York University, New York City, USA;
Department of Physics, University of California, San Diego, California, USA

Abstract— Hunger is a serious public health issue and there is a dire need to focus our attention on its risk factors to maximize hunger-reduction. Using cross-sectional data on hunger and nonfood risk factors, we refine UNICEF's Malnutrition Framework 2020 by ranking the importance of risk factors. For this, we use a subset of risk factors that comprise "enabling determinants" of nutrition, which are basic political, financial, social, cultural, and environmental conditions that enable good nutrition. Using a multiple linear regression framework, we demonstrate that hunger depends on gender inequality, sanitation, political situation, and per-capita income. Like previous researchers, our findings suggest that nutritional outcomes like Global Hunger Index are highly elastic with respect to gender inequality and education. Our analytics framework generalizes the findings of existing literature of a globally representative dataset and systematically analyzes the underlying determinants of hunger as set forth in UNICEF's Malnutrition Framework.

Defense of Military Installations from UAV-borne Attacks using Deep Learning

Nicholas Schaefer, Jaafar Alghazo, Sherif Abdelhamid, Ghazanfar Latif
Department of International Studies and Political Science, Virginia Military Institute, Lexington, Virginia, USA;
Department of Electrical and Computer Engineering, Virginia Military Institute, Lexington, Virginia, USA;
Department of Computer and Information Sciences, Virginia Military Institute, Lexington, Virginia, USA;
Department of Computer Science, Prince Mohammad Bin Fahd University, Alkhobar, Saudi Arabia

Abstract— Consumer-grade Unmanned Aerial Vehicles (UAVs) are becoming more common capabilities on the modern battlefield, finding use by both formal standing armies and nonstate sponsored organizations with small budgets. The threats posed by these UAVs are varied, ranging from intelligence gathered from reconnaissance, spotting for indirect fires, or attacking with a payload on the UAV itself. These dangers create new challenges that militaries must adapt to ensure soldiers are protected and mission completion is possible despite the threat of UAV interdiction. In this paper, we propose an AI object detection model that is capable of identifying UAVs in the visible spectrums and distinguishing them from images that contain no UAVs. This model can be used as a targeting system for anti-UAV countermeasures. The dataset used to train the model consists of 2,000 images. 1000 images are of UAVs and 1000 contain no UAVs. The models we tested were ResNet18, ResNet50, and GoogleNet. GoogleNet achieved the best results, yielding a precision of 0.995, recall of 0.995, F1-score of 0.995, and test accuracy of 98.44%. These results and the initial dataset present a good base for researchers to explore and design a practical solution for defense against small drone attacks.

Regression Models for Solar Radiation Prediction

Ayatelrahman Elsayed, Mona Y. Elshinawy, Nada Ibrahim, Randall Woodall,
Abdel-Hameed A. Badawy, Satish Ranade
Klipsch School of Electrical & Computer Engineering,
New Mexico State University, Las Cruces, New Mexico, USA

Abstract—Solar radiation prediction is essential for various applications, from renewable energy planning to weather forecasting and environmental monitoring. Accurate solar radiation prediction can help improve the efficiency of energy production and facilitate better management of natural resources. This paper focuses on solar prediction as an essential parameter for weather forecasting. We applied several regression models to predict solar radiation. In conclusion, Random Forest Regression was the most accurate model for predicting solar radiation. The CNN model ranked third with an accuracy difference of approximately 0.16 compared to NN-MLP. However, there is a possibility of enhancing the accuracy of the CNN model by testing different combinations of layers and filters. Furthermore, the performance of the SVR model was improved by 1.577 after hyperparameter

tuning. On the other hand, adjusting the number of estimators for the random forest regressor resulted in a slight decrease in performance, possibly due to over-fitting, as the tuning was done solely on the training data.

Analyzing a Literary Representation of the Wilmington Massacre of 1898 through Natural Language Processing Using Sentiment Analysis

Aysun Karamustafaoglu, Morgan Glisson, Matthew Aaron Ferraro, Seref Recep Keskin, Gulustan Dogan, Jennifer Marie Lozano, John Knox
University of North Carolina Wilmington, North Carolina, USA;
Logiwa WMS, Chicago, Illinois, USA

Abstract—The purpose of this Natural Language Processing (NLP) paper is to discover, learn, and help others to better understand and disclose the sentiment of a fictional and blended, non-fictional story during the period of the Wilmington Riots of 1898 utilizing the literature Hanover; or the Persecution of the Lowly. In addition, the goal of the project is to process the events and sentiments of the characters to derive feeling from the story. We analyzed each sentence and designated a specific sentiment (neutral, positive, negative) and emotion to it. After processing those sentences, we conducted several machine learning processes to try and predict and/or classify the sentiment of a sentence from the text. Although NLP balancing techniques were attempted, our models were not able to predict the sentiment of the sentences at a decent accuracy level. Future work may involve conducting a sentiment analysis on Hanover at a document-level or phrase-level analysis.

Design Approach of Electronic Attendance-Absence Recording System using Multi-User Face Recognition

Sang-Heon Lee, Myoung-Kyu Sohn, Hyunduk Kim, Junkwang Kim
Division of Automotive Technology, Daegu Gyeongbuk Institute of Science & Technology (DGIST),
Daegu, Republic of Korea

Abstract—In this paper, we propose the design and service scenario of an electronic attendance-absence recording system (EAARS) based on face recognition that is convenient to use and cannot be fraudulently attended. We suggest the design of the system for commercialization that utilizes state-of-the-art (SOTA) deep learning facial recognition technology, which offers high accuracy and fast speed. The system is designed to simultaneously recognize multiple users' faces and store the attendance and absence of each person. The electronic attendance-absence recording system consists of a server, a mobile platform for users or students, a personal computer (PC) platform for lecturers or teachers, and a multi-user face recognition module. It is configured in the form of TCP communication using the JSON file format with the web server. To implement the multi-user face recognition module, we utilize the SOTA technologes in face detection and face recognition, namely RetinaFace and ArcFace. Various backbone networks such as ResNet50, MobileNet V2, MobileNetV3, and MobileViT are used for training and we compare the recognition results and speed to select the appropriate model. The WiderFace database is used for developing face detection module, while MS-Celeb-1M and LFW are used for face recognition. In addition, we use TensorRT to optimize the trained model to improve the speed of the Multi-User Face Recognition module. We believe that it is necessary to specify the service scenario in detail for how to deal with the claims of users or students in case of misidentification in the system in the future.

A Clovis/Solutrean Projectile-Point Image Classifier

Jack K. Horner
Independent Researcher, Lawrence, Kansas, USA

Abstract - Here I report a Mathematica technique for implementing an artificial-intelligence-based, whole-image classifier of Clovis vs. Solutrean projectile points. The classifier software proper assumes no representation, as such, of empirical properties that are used in traditional archaeology, such as the dimensions of an artifact, flake geometry, the nature of the raw materials, or provenance information. The classifier is trained on 20 images each of Clovis and Solutrean points, and tested on images of 37

Clovis, and 4 Solutrean, points. Given the source classifications of the training and test images, the Wilson score approximation of the 95% confidence interval for the classification of the test items is [0.87, 0.99].

Multi-Agent Reinforcement Learning System Using Value-Decomposition Network Algorithm in Starcraft Environment

Abdulghani M. Abdulghani, Mokhles M. Abdulghani, Wilbur L. Walters, Khalid H. Abed
Department of Electrical & Computer Engineering and Computer Science,
Jackson State University (JSU), Jackson, Mississippi, USA

Abstract - Multi-Agent Reinforcement Learning (MARL) has been shown to be extremely successful in cooperative assignments. MARL allows for the control of multiple agents to complete multiple tasks in a certain environment and provide helpful services. In this paper, we will examine a war scenario with the StarCraft II Multi-Agent Challenges (SMAC) environment to implement a multi-agent system. For training, we employed one of the most popular MARL algorithms, which is the Value-Decomposition Network (VDN). This algorithm works on controlling the agents to cooperate with each other to achieve the desired goals. We will then use the battle won mean and the dead allies mean metrics to measure the performance of the VDN algorithm. The result showed that the VDN algorithm reaching the highest value of battle won mean with one million iterations and the lowest value of dead allies mean metrics with less than one million iterations. The hardware that we use in this work is CPU Cor i7 11800H, with 32 GB Ram and RTX 3080 laptop GPU, with CUDA Toolkit 11.7.1, and Pytorch 1.7.1.

Federated Learning for Character Prediction in Text Generated

Guang Hu, Xin Fang
School of Statistics and Information, Shanghai University of International Business and Economic, Shanghai, China;
School of Computer Science, Fudan University, Shanghai, China;
Shanghai Key Laboratory of Data Science, Shanghai, China

Abstract—Modern mobile devices have access to enormous amounts of user data including text, images, speech, etc., which can be utilized to train high-performance learning models and enhance the user experience. However, accessing large amounts of data often raises concerns for user privacy and security. To address this, federated learning (FL) has emerged as a new machine learning approach that trains models on multiple decentralized edge devices (e. g. mobiles) or servers while protecting user privacy. In this paper, we present a distributed learning framework using a practical iterative average-based federated learning algorithm for the text generation task in Natural Language Processing (NLP). Our results show that text generation training under federated learning yields better performance than random guessing, demonstrating the feasibility of FL in language modeling. The study highlights the success of text generation techniques trained using federated learning, while emphasizing the importance of safeguarding user privacy and security.

Exploring Automatic Malware Detection through Deep Learning Models

Jaafar M. Alghazo, David M. Feinauer, Sherif E. Abdelhamid
Department of Electrical and Computer Engineering, Virginia Military Institute, Lexington, Virginia, USA;
Department of Computer and Information Sciences, Virginia Military Institute, Lexington, Virginia, USA

Abstract—Malware is continually being developed and reinvented by malicious users. Deep Learning opens doors for real-time detection and classification of malware. In this paper, we explore the use of deep learning models (GoogleNet and ResNet50) for the detection and classification of malware based on the Malimg dataset. In addition, we explore modifying deep learning models and present a modified ResNet50 model to study its effect on classification accuracy. Finally, we explore the use of ensemble methods, applying three independent ResNet50 models to study the effect of ensemble models on classification accuracy. GoogleNet achieved the highest test accuracy among the three with 94.82%, however, when applying the ensemble method, the test accuracy reached 97.86%.

Digital Rubber Duck: Leveraging Large Language Models for Extreme Programming

Timothy Elvira, Tyler Thomas Procko, Juan Ortiz Couder, Omar Ochoa
Department of Electrical Engineering and Computer Science,
Embry-Riddle Aeronautical University, Daytona Beach, Florida, USA

Abstract—The recent prevalence of Large Language models (LLMs), e.g., GPT-3.5 and GPT-4, has brought about a new age of man-computer symbiosis, where LLMs are employed for a litany of creative, constructive, scientific, or otherwise contentgenerative tasks, e.g., as general chatbot assistants, writing editors, digital subject matter experts, programming consultants, and so on. Of interest to software engineers is the concept of "rubber duck debugging", which is the act of expressing code, lineby-line, in natural language, to an inanimate object, e.g., a rubber duck, for the purpose of elucidating potential issues that can then be corrected. In this paper, we detail a workflow process that leverages the concept of rubber duck debugging, replacing the duck with a capable LLM, e.g., GPT-4. We call it Digital Rubber Duck Programming. Furthermore, the Extreme Programming (XP) method, an implementation of the Agile paradigm, is considered as easily integrated with the proposed workflow, as XP is performed in pairs (much like the modern software engineer works in pairwise fashion with an LLM) and because XP places emphasis on performing extensive code reviews and unit testing all code, which capable LLMs like GPT-4 can facilitate.

Multi-Agent Reinforcement Learning System Using Multi-Agent Proximal Policy Optimizer Algorithm in SMAC Environment

Abdulghani M. Abdulghani, Mokhles M. Abdulghani, Wilbur L. Walters, Khalid H. Abed
Department of Electrical & Computer Engineering and Computer Science,
Jackson State University (JSU), Jackson, Mississippi, USA

Abstract - Multi-Agent Reinforcement Learning (MARL) is a subfield of reinforcement learning that focuses on studying the behavior of multiple learning agents that coexist in a shared environment. In this paper, we will examine a war scenario within StartCraft II Multi-Agent Challenges (SMAC) environment to implement a multi-agent system. For training, one of the well-known MARL algorithms was used, namely Multi-Agent Proximal Policy Optimization (MAPPO). This algorithm works on navigating the agents to cooperate with each other to achieve the desired goals. We will then use two offered metrics: battle won mean and dead allies mean in the SMAC environment to evaluate the performance of the MAPPO algorithm. The result showed that the MAPPO algorithm reached the greatest value of battle won mean with one million iterations and reached the lowest value of dead allies mean metrics with less than one million iterations. The hardware that we use in this work is CPU Cor i7 11800H, with 32 GB Ram and RTX 3080 laptop GPU, with CUDA Toolkit 11.7.1, and Pytorch 1.7.1.

Performance Comparison of Deep Learning and Machine Learning Models Applied for P&c Insurance for Measuring Prediction Accuracy

Sushanth Manakhari, Yanzhen Qu
Colorado Technical University, Colorado Springs, Colorado, USA

Abstract— Many research studies have shown that machine learning is quite effective for predicting future sales or demand of customers. In the Insurance industry, machine learning is applied to know customer behavior. This study will perform a detailed analysis of insurance and customers. The purpose of the study is to perform the comparison of machine learning models for measuring prediction accuracy. After an extensive literature review, many research studies have focused on various predictive models. Predictive models such as neural networks, Deep Learning, Artificial Intelligence, and Machine learning are applied in the insurance sector to know customer behavior. However, there are not many studies that have made performance comparisons. This study will perform a performance comparison of various predictive models. This research study will provide a huge amount of information to the target audience and fill the research gap in the current literature.

Short Review on Supervised learning under Adversarial Label Poisoning and Clean-label Data Poisoning

Pooya Tavallali, Vahid Behzadan
Independent Researcher, Boston, Massachusetts, USA;
Tagliatela College of Engineering, University of New Haven, New Haven, Connecticut, USA

Abstract—Training under adversarial label poisoning is relatively one of the recent topics that have attracted attention in the literature of machine learning. Label poisoning is considered as one of the data poisoning techniques under adversarial attacks. Adversarial attacks tend to change the data such that some data points are classified wrongly. Adversarial attacks are categorized as white-box or black box attacks. In white-box the attacker has partial or complete knowledge of the model. On the other hand, black box attacks have no information about the model. Data poisoning considers manipulation of training points for adversarial intents. A subcategory of this attack is label poisoning where the labels are manipulated or new data is added to the dataset. This short review, considers label poisoning attacks and finally shortly mentions other related topics that has attracted some attention when developing techniques for label poisoning.

LLM/GPT Generative AI and Artificial General Intelligence (AGI): A Monolithic Study

Atif Farid Mohammad, Bryan Clark, Ram Agarwal, Sean Summers
OmniAGI.ai - Global Technology Solutions, Inc., USA

Abstract - The dawn of artificial intelligence has sparked excitement and curiosity as it paves new ways for us to envision the future. Artificial General Intelligence (AGI) stands as a thrilling and challenging frontier to explore, offering many possibilities and transformations across various industries. This article attempts to provide an in-depth overview of AGI, its history, applications, challenges, and ethical implications. Before diving into the world of AGI, it is prudent to familiarize oneself with the key concepts and distinctions that shape this field. This section aims to define AGI, differentiate it from other types of artificial intelligence, and offer a glance at its historical development and its potential future development.

Evolving Efficient CNN Based Model for Image Classification

Afsaneh Shams, Drew Becker, Kyle Becker, Soheyla Amirian, Khaled Rasheed
School of Computing, University of Georgia, Georgia, USA;
Institute of Artificial Intelligence, University of Georgia, Georgia, USA

Abstract - Evolutionary algorithms, rooted in Darwin's theorem, can be considered as a base for implementing deep/machine learning models. This approach can noticeably increase the accuracy in most cases as shown in this paper. This experiment aims to evaluate the performance of two evolutionary algorithms, an evolutionary neural network (ENN) and an evolutionary CNNbased algorithm with mutation and crossover (ECNNB), on the Fashion-MNIST, MNIST, and EMNIST Digits datasets. The performance of the ENN algorithm is examined for 10, 50, and 100 generations, with 50 generations being used due to computational limitations. The results show that the accuracy of the model improves as the number of generations increases. However, the ECNNB model consistently outperforms the ENN model on all three datasets, with an average accuracy of 92.58% on FashionMNIST, 99.32% on MNIST, and 99.50% on EMNIST Digits, compared to 88.54%, 98.05%, and 98.95%, respectively, for the ENN model. The performance of both models is compared with other state-of-the-art models in the literature. These results highlight the significance of well-designed models in achieving high accuracy in machine learning tasks.

Transformers for Network Traffic Prediction

Vinayak Tanksale
Department of Computer Science, Ball State University, Muncie, Indiana, USA

Abstract—Transformers have been extensively used in a variety of machine learning and data science applications. Building prediction models using transformers has also been extensively researched for various domains. In this paper, we present our work on building a prediction model for network traffic data using transformers. The significant contribution of our paper is the application of transformers to Controller Area Network (CAN) data prediction. Real-world third-party datasets are used in some of our experiments. The paper discusses parameter selection for the transformer model, present results of predicting temporal network data, and compares our proposed approach to prior recurrent neural network-based techniques.

Performance Evaluation of Federated Learning for Anomaly Network Detection

Roudha Alhammadi, Amjad Gawanmeh, Shadi Atalla, Mohammed Q. Alkhatib, Wathiq Mansoor
College of Engineering and IT, University of Dubai, Dubai, UAE

Abstract—Network anomaly detection is crucial for ensuring the security and dependability of modern networked systems. Traditional machine learning methods face scalability, data security, and adaptability challenges. This paper explores federated learning, a collaborative learning technique, for network anomaly detection. The federated learning approach allows multiple computing agents to train a model on their local data without sharing sensitive information. Using a real-world dataset of network traffic, the effectiveness of federated learning is evaluated and compared with rule-based and machine-learning-based methods. The results show that the federated learning-based approach outperforms traditional accuracy, precision, and recall methods. Achieving an accuracy of 97%, precision of 93%, and recall of 91%, it surpasses the best rule-based method (accuracy: 85%, precision: 71%, recall: 62%) and the best machine learning-based method (accuracy: 93%, precision: 83%, recall: 79%).

A SQL-Based Probabilistic Inference

Feng-Jen Yang
Department of Computer Science, Florida Polytechnic University, Lakeland, Florida, USA

Abstract—There is an industrial trend to integrate machine learning models into relational databases so that users can get the intended support without compiling and feeding the set of training data to a learning model and training it separately. To this end, implementing a machine learning model by using SQL statements is an efficient and low-cost practice because the relational database platform is ready to run the intended computational queries without additional plugins. By integrating probabilistic inference functionalities into relational databases, we can ensure that decision-makers have access to the most accurate and up-to-date information possible. This is especially important in today's fast-paced world where the ability to make quick and informed decisions is more important than ever. In this study, I demonstrated a way to equip a relational database with the capability of performing probabilistic inference by using only SQL statements. This is an exciting development that has the potential to reshape the way we design databases and approach decision-making in the industrial world.

Overcoming Supply Chain Challenges with Advanced Machine Learning: Exploring the Potential of Deep Meta-Learning and Multi-Task Learning

Beilei Zhu, Chandrasekar Vuppalapati
Global Supply Chain, Intel Corp., Hillsboro, Oregon, USA;
Computer Engineering, San Jose State University, San Jose, California, USA

Abstract - Artificial intelligence and machine learning can significantly enhance supply chain businesses, boosting profit, reducing cost, and improving customer service. They utilize large data volumes to automate tasks and improve decision-making. However, their widespread adoption faces challenges, including shifts in business climate and the requirement of deep industry knowledge for successful application. This paper explores deep meta-learning and multi-task learning as a solution to these challenges. With advanced machine learning techniques, we can overcome traditional limitations and create adaptable supply chain models. We aim to highlight how meta-learning can optimize supply chain operations and enhance business performance utilizing existing shared domain knowledge.

Extraction of Breast Cancer Information from Clinical Records for Cancer Registry using Natural Language Processing

Adhari Abdullah Alzaabi, Abdurahman AAlAbdulsalam
College of Health Science and Medicine, Sultan Qaboos University, Oman;
Department of Computer Science, Sultan Qaboos University, Oman

Abstract - National cancer registries rely on manual abstraction of free-text clinical records to collect vital information about cancer diagnosis, stage, progression and treatment. Many prior studies have demonstrated the ability of natural language processing (NLP) based on machine learning to extract information from free-text clinical records for a variety of purposes (diagnosis, adverse events discovery, clinical trial matching, ...). We present in this study experimental results of applying NLP to extract information from the records of breast cancer patients for the cancer registry in Oman.

Trash Classification using Deep Learning Models

Enes Faruk Keskin, Ogulcan Isleyen, Hilmi Demirhan
Turkish Aeronautical Association University, Department of Computer Engineering, Ankara, Turkey;
Department of Information Systems, University of North Carolina Wilmington, North Carolina , USA

Abstract—The rapid growth of population and urbanization, combined with the transformative technological advancements arising from the industrial revolution, have resulted in a divergence in living standards and consumption patterns. Consequently, a substantial increase in waste generation has occurred compared to earlier times. This scenario has given rise to severe air, water, and soil pollution, posing a grave threat to all forms of life. Moreover, it has accelerated the depletion of natural resources and exacerbated the challenges associated with climate change. Given these circumstances, the effective classification of recyclable waste stands as a crucial objective for humanity. To accomplish this, the utilization of Deep Learning models has proven to be highly beneficial. In the pursuit of this goal, we conducted a comprehensive study where we assessed prominent Deep Learning models, aiming to identify the most efficient approach. In this study, DenseNet121, InceptionResnetV2, MobileNet, ResNet50 and Xception architectures were used for TrashNet dataset. As a result of the conducted experiments, the best results were found in the DenseNet121 using fine-tuning with a test accuracy rate of 95.

Neural Network based Evolutionary Approach to Discovery of Governing Equations

Shahab Razavi, Eric R. Gamazon
Vanderbilt University Medical Center, Tennessee, USA

Abstract – N/A

Machine Learning for Assessing the Risk of Military Aircraft Failures

Hokey Min, Ryan Wood, Seong-Jong Joo
Allen and Carol Schmidthorst College of Business,
Bowling Green State University, Bowling Green, Ohio, USA;
Department of Operational Sciences, Air Force Institute of Technology, WPAFB, Ohio, USA;
Harvard University, USA

Abstract - Modern aircraft are costly and require heavy investment. It is the same regardless of industries, such as commercial airlines and militaries. Thus, maintaining desired readiness by reducing ground time in the militaries is critical to maintaining air superiority and winning the war. There are two types of maintenance activities such as preventive and corrective maintenance. Preventive maintenance requires taking action before failures happen. Meanwhile, corrective maintenance reacts to failures, which takes time to buy parts and repair failed components. If we can predict aircraft failures accurately, we will be able to change corrective maintenance activities to preventive maintenance activities, which will reduce aircraft downtime and, thus, increase aircraft readiness or availability. This paper proposes multiple machine learning tools to minimize aircraft downtime to predict aircraft failures with the highest accuracy possible.

Improving Energy Efficiency in Healthcare Machine Learning Models

Ishanvi Kommula
Monta Vista High, Cupertino, California, USA

Abstract – The future of the healthcare industry is expected to approach a hybrid model where trained machines work side-by-side with doctors in patient diagnosis. In some areas, like radiology, machine learning models have the capability to completely replace humans. Comparable to humans, machines must be trained regularly with new data sets. Even though these power-hungry machines are more accurate and faster than humans, they emit hundreds of tons of carbon dioxide due to the compute-intensive nature of ML training and diagnosis. In comparison, the average human throws out about five tons annually [19]. We can limit the impact of machine learning on the environment by choosing efficient learning techniques to mitigate carbon emissions. In this paper, I focused on the critical part of an ML model: the programming language. Using eBPF based telemetry collected from Linux kernel trace points and CPU performance counters, I compared the energy usage of a radiology training model when written in multiple programming languages. To limit the scope of my research, I have focused on Go, Python, R, Java, C, and C++ programming languages.

Is Gamification a Practical Solution for Increasing Repurchase Intention of Online Customers?

Khalil Safari, Zohreh Safari, Fatemeh Keshavarzi
Department of Business Management and Information Technology, Payam Noor University, Iran;
Department of Computer Science, Math, Engineerig Technology,
Elizabeth City State University, Elizabeth City, North Carolina, USA;
Department of Information Technology Engineering, Qazvin Islamic Azad University, Qazvin, Iran

Abstract—With the growth of online shopping, organizations should update their portals and websites in response to changing conditions and competitors' actions. To attract and retain online customers, various tools such as reputation building, trust

enhancement, and gamification are utilized and reinforced. This study has investigated 243 users of DigiKala company through paper and electronic questionnaires. Due to the nonnormality of the data, non-parametric statistical methods using SPSS and Smart-PLS3.0 software were employed to analyze structural relationships. Additionally, machine learning tools were utilized to optimize the analysis of customer behavior and identify patterns for effective customer retention strategies. The results show that contrary to the research background, the use of gamification on DigiKala's website has not had a positive and significant impact on customers using this company's website and their repurchase intention. Thus, it seems that DigiKala should focus on context and infrastructures, while simultaneously strengthening and revamping gamification dimensions.

Large Language Model (LLM) and GPT, A Monolithic Study in Generative AI

Atif Farid Mohammad, Ramya Ganesh
University of North Carolina at Charlotte, North Carolina, USA

Abstract— Large Language Models (LLMs) are a branch of computer science and artificial intelligence which is concerned with computer and human language interaction. It is the study of mathematical and computational modeling of various aspects of language and the development of an arsenal of systems. Large Language Models are considered an area of research and application that explores how computers can be used to comprehend and manipulate natural language text or speech to perform useful tasks. It has spread its applications in various areas such as machine translation, email spam detection, information extraction, summarization, medical, and question answering etc. Large Language Models (LLMs) have a greater contribution in the area of text pre-processing as well in the time of ChatGPT as an example. Different pre-processing steps are required to perform, such as stemming, part-of-speech (POS) tagging, chunking, parsing, information extraction, etc. to perform language processing.

Harnessing the Speed and Accuracy of Machine Learning to Advance Cybersecurity

Rebet Jones, Marwan Omar, Derek Mohammed, Calvin Nobles, Maurice Dawson
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA;
Saint Leo University, Saint Leo, Florida, USA

Abstract— As cyberattacks continue to increase in frequency and sophistication, detecting malware has become a critical task for maintaining the security of computer systems. Traditional signature-based malware detection methods have limitations in detecting complex and evolving threats. Machine learning (ML) has emerged as a promising solution to detect malware effectively in recent years. ML algorithms can analyze large datasets and identify patterns difficult for humans to identify. This paper presents a comprehensive review of the state-of-the-art ML techniques used in malware detection, including supervised and unsupervised learning, deep learning, and reinforcement learning. We also examine the challenges and limitations of ML- based malware detection, such as the potential for adversarial attacks and the need for large amounts of labeled data. Furthermore, we discuss future directions in ML-based malware detection, including integrating multiple ML algorithms and using explainable AI techniques to enhance the interpretability of MLbased detection systems. Our research highlights the potential of ML-based techniques to improve the speed and accuracy of malware detection and enhance cybersecurity.

Ecuador Banana Production and Democratization of Climate Change Machine Learning Models to Mobile Edge Devices

Chandrasekar Vuppalapati, Anitha Ilapakurti, Sandhya Vissapragada, Sharat Kedari,
Vanaja Mamidi, Raja Vuppalapati, Santosh Kedari, Jaya Shankar
Computer Engineering (Software Engineering), San Jose State University, San Jose, California, USA;
Data Science and AI, Hanumayamma Innovations and Technologies, Inc., Fremont, California, USA;
Data Science & Artificial Intelligence, Hanumayamma Innovations and Technologies PVT Limited, Hyderabad, India

Abstract - Bananas are an important and nutritionally diverse component in diets around the world. Revenue generated from trade in bananas plays an important role with regards to the food import bill of producing countries. Bananas have a particular significance in some of the least developed and low income food-deficit countries, where they contribute not only to household

food security as a staple but also to income generation as a cash crop. Research conducted in ten banana producing countries revealed that income from banana farming accounts for some 75 percent of total monthly household income for smallholders. Nevertheless, climate change is impacting banana production in many countries. The impact of climate change not only increases lower yields in some countries but also render severe economic consequences to small holder farmers. As part of this innovative research paper, we provide deployment of complex banana production climate change models to farmers across the world to better plan adaptation strategies and thus reduce the overall negative impact and sustain financial returns.

Comparison of Standard Machine Learning Classification Methods Using a Mammogram Dataset

Ali Al-Faris, Gaurav Dhital
Computer Science Department, Worcester State University, Worcester, Massachusetts, USA

Abstract— The aim of this paper is to apply number of machine learning methods on the same dataset to compare effectiveness of classification approaches. A range of standard algorithms is used on a mammogram dataset to determine whether the tumor is benign or malignant. The effectiveness is measured using number of analysis metrics. The findings suggest that the supervised machine learning methods can effectively differentiate between benign and malignant tumors with accuracy vary from 87 to 95 percent. With further work and improvement, this project has the potential to aid in early diagnosis for patients with breast tumor.

Unleashing the Potential of Machine Learning: A Comprehensive Exploration of State-of-the-Art Algorithms and Real-World Applications in Computer Vision

Chris Cheng Zhang, Kevin Zhang
Canada Youth Robotics Club;
University of British Columbia, Vancouver, BC, Canada

Abstract—This research paper explores the transformative impact of machine learning techniques in the field of computer vision. With the rapid evolution of machine learning algorithms, computer vision has witnessed remarkable advancements and diverse applications. This study dives into the intersection of machine learning and computer vision, presenting a comprehensive analysis of state-ofthe-art algorithms and their real-world use cases. Notable machine learning algorithms are shown in the context of their applications in computer vision. The paper highlights one real-world example in particular: the use of machine learning techniques in long-term care for elderly patients using object recognition and pose estimation. Experimental results and shed light on algorithm performance and inherent limitations. The findings provide valuable insights into the untapped potential of machine learning in computer vision.

Uncentralized Artificial Intelligence Computing Agent with the Distributed Training and Computing Tasks Based on Open Source Cloud Proxy

Ming-Shen Jian, Jian-Ting Guo, Xin-May Wu
Department of Computer Science and Information Engineering,
National Formosa University, Yunlin, Taiwan

Abstract— Based on the proposed system, the lower cost uncentralized artificial intelligence computing with limited time spending is developed. The computing tasks of artificial intelligence are distributed into multiple low cost embedded devices. The training of the artificial intelligence based on physical cloud computing node spends 50% less time than the embedded system. According to the verification, the distributed workload of the artificial intelligence for image object recognition is processed on embedded system device within reasonable and acceptable time spending which is only about 15 seconds.

SESSION:
XXIII Technical Session on Applications of Advanced AI Techniques to Information Management for Solving Company-Related Problems

Co-Chairs:
Dr. David de la Fuente and Dr. Jose A. Olivas***
**University of Oviedo, Spain*
***University of Castilla - La Mancha, Spain*

Assessing Industry 4.0 Readiness: The Influence of Drivers and Information Technologies

O. Leon, J. Puente, I. Fernandez, J. Parreno
Escuela Politecnica de Ingenieria de Gijon, Campus de Viesques s/n, Gijon (Asturias), Spain

Abstract— This article analyzes the integration of advanced technologies in manufacturing processes, known as Industry 4.0, and how organizations can be prepared to implement these technologies. The study aims to examine the drivers that lead organizations to adopt Industry 4.0 and identify how these technologies affect the preparation of companies for their adoption. The study used a questionnaire survey method to collect data and analyzed them to determine correlations in the proposed model. The results suggest that companies prioritizing Industry 4.0 drivers are more likely to adopt information technologies successfully. The findings demonstrate a direct relationship between adopting information technology and readiness organizations for Industry 4.0. The study concludes that Industry 4.0 technologies provide the necessary tools and capabilities for companies to be prepared for Industry 4.0, while preparation for Industry 4.0 is a prerequisite for companies to effectively use and benefit from Industry 4.0 technologies.

Reinforcement Learning based Production Planning in the Aquaculture Industry

M. Luna, O. Perez-Mon, R. Rosillo, A. Gomez
Escuela Politecnica de Ingenieria de Gijon, Campus de Viesques s/n, Gijon (Asturias), Spain

Abstract— Production planning and decision-making processes in the aquaculture industry are becoming increasingly complex. Decision-makers must take into consideration a large number of factors - economic, biological, environmental, etc. - and, in addition, the increasing size of enterprises and intensive production systems have exponentially increased the number of decisions to be made. The aim of this article is to evaluate the possibilities of applying a novel technique, such as reinforcement learning, to decision-making processes in aquaculture. To this end, its application in comparable industries, also part of the agricultural sector, that are more technologically advanced, is reviewed in order to use them as a reference in the short and medium term. As can be observed in this article, reinforcement learning not only offers an interesting alternative for dealing with this complexity but also makes it possible to tackle problems that had not been solved until now.

Gender Analysis of European Job Offers through NLP Techniques

Cristina Puente, Ivan Sanchez-Perez, Elisa Gutierrez, Jose A. Olivas
Advanced Technical Faculty of Engineering - ICAI, Pontificia Comillas University, Madrid, Spain

Abstract— This article is an experiment on a set of job offers advertised in several countries of the European Union and in different sectors in which the professional profile to which they are addressed is clearly reflected. At a time when we are trying to emphasize equal opportunities between men and women and the special care that gender biases require in the automation of processes, this experiment shows that this is not always the case, since through the application of natural language processing

techniques, we are able to classify the ads around the type of language they use. To close the experiment, some conclusions and suggestions are proposed that can avoid this type of problems in the future.

A Control Engineering Approach to the Ripple Effect Analysis in Closed-loop Supply Chains

M. de Arquer, B. Ponte, R. Pino, D. de la Fuente
Escuela Politecnica de Ingenieria de Gijon, Campus de Viesques s/n, Gijon (Asturias), Spain

Abstract—Closed-loop supply chains, incorporating circular economic principles, play a key role in enhancing the sustainability of production systems and creating economic opportunities for firms. In today's business environment, characterized by frequent disruptions, these systems need to be both efficient and resilient to remain competitive. However, integrating efficiency and resilience is not a straightforward task. This manuscript discusses control engineering as a promising approach to understand closed-loop supply chain resilience. It enables the analysis of the propagation of disruptions (the ripple effect), thus aiding managerial decisionmaking. We use an example to demonstrate the potential of control theory in improving the resilience of closed-loop supply chains.

Investigating the Potential of Blockchain Technology for Improving Traceability in Agriculture

S. Fernandez-Vazquez, N. Alvarez, O. Leon, J. Costas
Escuela Politecnica de Ingenieria de Gijon, Campus de Viesques s/n, Gijon (Asturias), Spain

Abstract – Customers' sensory quality demands have risen as they grow more worried about their health. Blockchain technology is viewed as a solution for improving agri-food supply chain traceability and providing stakeholders with information on food quality, safety, and nutrition. We propose a blockchain-based system intended to include all of the certifying bodies' rules, regulations, and standards within the aquaculture supply chain.

Assembly Flowshop Scheduling Problems with Blocking Constraints: Two Optimal Properties and a Iterated Greedy Search Approach

Carlos Andres Romano, Julien Maheut, Paz Perez Gonzalez, Jose Manuel Framinan Torres
Dept Organizacion de Empresas, Edificio 7D, Camino de Vera s/n, Valencia, Spain

Abstract – N/A

The 19th International Conference on Data Science (ICDATA'23: July 24-27, 2023; Las Vegas, USA)

https://icdata.org
https://www.american-cse.org/csce2023/

The Ethical Responsibility for Explainable AI in Mission Engineering

Althea Henslee, Haley Dozier, Indu Shukla, Griffin Turner, Abderahim Salhi, Ian Dettwiller
Information Technology Laboratory,
United States Army Corps of Engineers, Vicksburg, Mississippi, USA

Abstract—Advances in the field of reinforcement learning show promise in developments in Mission Engineering. By lowering the level of dependence on human subject matter experts, Reinforcement Learning has the potential to not only speed up analysis critical to decision-making, but to reduce cost and human bias as well. However, the more advanced Machine Learning necessary to handle complex Mission Engineering scenarios often provide little explainability and may, therefore, inspire lower levels of confidence by decision makers. Widespread adoption and deployment of Machine Learning techniques will require the higher fidelity associated with explainable Artificial Intelligence. Index Terms—explainabl

Crime Hot Spots Detection with Network Science-enhanced Method in Major Cities of the U.S.

Yu Wu, Natarajan Meghanathan
Department of Electrical & Computer Engineering & Computer Science,
Jackson State University, Jackson, Mississippi, USA

Abstract—Spatial-temporal analysis of crime, along with hot spot policing, are crucial methodologies in research of crime prevention and control. However, the authors notice that the majority of studies predominantly emphasize geographic information while slightly neglecting the impact of temporal feature of crime occurrences. Building upon the authors' prior research, this study employs an innovative approach to identify the temporal hot spots of various crime types, applying the method to more homogeneous data sets and examining its generalizability across different cities. The authors capture three distinct temporal occurrence patterns and elucidate their realworld implications. The findings of this study offer more detailed and interpretable evidence from temporal perspective for law enforcement's research on crime types occurrence patterns.

Harnessing Big Data in Agriculture by Addressing Heterogeneity in Large-Scale Data Mining Techniques and Limitations

Usama Ikhlaq, Tahar Kechadi
Department of Computer Science, University College Dublin (UCD), Dublin, Republic of Ireland

Abstract—This study harnesses big data in fertilizer nutrient research by addressing data heterogeneity and introducing novel data mining techniques, such as Slice Separate and Link (SSL), Layer Linking Stream (LLS), and Separate Order Connect (SOC). Our large-scale data, characterized by volume, velocity, variety, and value, encompasses agroclimatic indices and farming features. The techniques effectively manage feature subsets, handle multi-time series complexities, and provide accurate nutrient application

timeline predictions. This research overcomes limitations in large-scale data analysis, optimizing fertilizer applications, promoting sustainable agriculture, and enhancing resource management. The experimental results demonstrate the potential of these techniques to significantly improve the accuracy and reliability of fertilizer application timeline predictions across diverse agroclimatic conditions and crop types.

Canonical Correlation Analysis of Neighborhood-based Centrality Metrics vs. Shortest Path-based Centrality Metrics

Natarajan Meghanathan
Department of Electrical & Computer Engineering & Computer Science,
Jackson State University, Jackson, Mississippi, USA

Abstract—Canonical correlation analysis is useful to analyze the correlation between two sets of features in a dataset. In this paper, we demonstrate the use of canonical correlation analysis to study the correlation between the neighborhood-based centrality metrics (Degree centrality: DEG and Eigenvector centrality: EVC) vs. the shortest path-based centrality metrics (Betweenness centrality: BWC and closeness centrality: CLC) for a suite of 80 complex real-world networks. We observe more than 4/5th of the 80 real-world networks analyzed to exhibit either a strong positive correlation or negative correlation between these two sets of centrality metrics (i.e., if the neighborhood-based centrality value for a node is lower, the shortest path-based centrality metric value for the node is likely to be lower or higher). We also observe the extent of variation in node degree to play a significant role in the correlation between these two sets of centrality metrics in complex real-world networks.

Management, Storage, and Retrieval of Complex Data Comprising Multiple Formats Collected from Different Sources: a Systems Engineering Approach

Barry C. White, Reena R. Patel, LaKenya K. Walker, Matthew D. Bray
Information Technology Laboratory, US Army Corps of Engineers (USACE) Engineering Research and Development Center, Vicksburg, Mississippi, USA;
Geotechnical and Structures Laboratory, US Army Corps of Engineers (USACE), Engineering Research and Development Center, Vicksburg, Mississippi, USA

Abstract - Physical modeling procedures, with intermediate data, are being developed for the large-scale generation of synthetic imagery for automated target recognition (ATR) machine learning (ML) algorithms. This imagery is typically combined with collected data for generating robust training sets. The management and retrieval of this data requires large-scale storage and a means to query data of different types. Queries need to be performed for selection of data sets to the single file. The goal of this study is the creation of managed system for storing and retrieving this information using high-performance computing resources with the integrated Rule Oriented Data System (iRODS). Search oriented metadata tags are created for query searches based on locality, time-of-day, and other factors. When possible, metadata generation will be automated based on information in the data file. Use cases for the import and query operations are created. Simple scalable problems have been processed and are presented for this data set procedure and the proposed architecture is presented. This data storage and retrieval system will serve to provide locality specific data for ATR ML datasets from a large set of collected and synthetic imagery and the processes to create that imagery.

Evaluation of Sampling Thresholds for Fundamental Spatial Parameters of Human Mobility

Zaid Matloub, Ivica Kostanic, Sasha Knezevic
Department of Computer Science and Engineering,
Florida Institute of Technology, Melbourne, Florida, USA

Abstract—This paper aims to characterize human mobility using crowd-sourced cellular network data. The study considers four key spatial parameters of mobility, assesses those parameters for a set of defined sampling thresholds, and performs a comparative study on the obtained results. The findings show that human mobility is highly predictable, and accurate models may be derived

if more studies are conducted for different periods and geographic areas. The study also suggests that choosing the appropriate sampling thresholds depends on the mobility parameters to be evaluated, the dataset size, and the available computational resources.

Fundamental Spatial Parameters of Human Mobility

Sasha Knezevic, Ivica Kostanic, Zaid Matloub
Department of Computer Science and Engineering,
Florida Institute of Technology, Melbourne, Florida, USA

Abstract — A set of parameters for characterizing spatial aspects of human mobility are presented. The parameters are extracted from a large set of crowd-sourced cellular data which were collected in a major metropolitan area of United States. Each parameter is analyzed at three different time scales (daily, weekly, and monthly time scale). The results demonstrate that human mobility on a population scale has a high degree of statistical regularity. The mobility may be described to a great extend using a limited number of judiciously selected parameters.

Analyzing COVID-19 Impact in the US: Demographic, Economic, and Social Factors

Daniel Ojeda, Anissa Champion, Ching-Yu Huang, Daehan Kwak
Department of Computer Science and Technology, Kean University, Union, New Jersey, USA

Abstract—The COVID-19 pandemic has had a great impact on the world, with the United States experiencing a disproportionately high burden in terms of infections and deaths compared to other nations. The virus has disrupted daily life across the country and highlighted longstanding health inequalities. Therefore, this study aims to analyze the correlation between total COVID-19 cases and deaths with various demographic variables such as median age, education level, race, ethnicity, income, unemployment rate, disability status, and insurance coverage. The goal is to identify groups of people who are disproportionately affected by the COVID-19 virus and the factors that contribute to this inequality. To achieve this, the study collects data from all 50 states in the United States, using authoritative sources such as the U.S. Census Bureau and the Bureau of Labor. Then, ETL process (Extract, Transform, and Load) is performed to obtain clean and refined datasets and compile them into a final table for correlation analysis. With the help of this study, public health officials and policymakers could initiate the development of targeted interventions to ensure that everybody has the same opportunity to achieve good health, regardless of their demographic, economic, and social status.

Text Mining Legal Documents for Clause Extraction

David Baglee, Ken McGarry, Tony Vidler
University of Sunderland, United Kingdom, UK

Abstract—Natural Language Processing (NLP) solutions for legal contracts have been the preserve of large law firms and other industries (e.g., investment banks), especially those with large amounts of resources, having both the volume and range of legal documents and manpower to label the training data. The findings suggest that it is possible to use a smaller volume of training contacts and still generate results that are within an acceptable range. Our results show that just 120 training contracts trained on a pre-trained language model can generate results that are within 10% of the same model trained on 3.3 times the volume. In conclusion, smaller law firms could benefit from machine learning NLP solutions for clause extraction.

An Approach for Identification and Estimation of Outliers in Time Series with Nonstationary Mean

Koki Kyo
Gifu Shotoku Gakuen University, Japan

Abstract—We address the problem of identifying and estimating outliers in a time series with a nonstationary mean. We first apply a moving linear model to decompose the time series into a constrained component and a remaining component that is a stationary time series containing the outliers. We then propose an approach of identifying and estimating the outliers. The proposed approach is based on the maximum likelihood method for an autoregression (AR) model with outliers in variables and has the following features. (1) The used model has a simple structure that is easy to understand. (2) The proposed approach allows for the estimation of not only the positions but also the magnitudes of the outliers. As an illustrative example, we apply the proposed approach to the analysis of the index of industrial production in Japan. The results show the high performance of the approach.

An Investigation of Time Series Embeddings and Topological Data Analysis for Fault Analysis

Dean Lee, Jamal Rorie, Andrew Sabater
U.S. Northern Command, U.S. Government, USA

Abstract—Topological data analysis (TDA) is an important part of a data scientist's tool box; through the extraction of topological information, TDA provides an automated means of feature engineering. Thus, TDA may be an appropriate tool for fault analysis of mechanical systems, where the methodologies have traditionally relied on expert intuition and signals processing techniques. However, appropriate embedding of the data must be defined before TDA can be applied. In this paper, we investigate several different embeddings of time series data for the application of TDA for fault analysis and show that features engineered by TDA improve fault classification accuracy.

Bringing Data Analysis to the Files and the Database to the Command Line

Antonio Badia, Saikiran Potti, Andrew Nguyen, Lance Gibson
University of Louisville, Louisville, USA

Abstract—The analysis of a dataset starts with exploratory and lightweight analysis, in order to gain a rough understanding of the data under consideration. While several tools already exist to support this task, most require that data be loaded and encoded before starting. We propose a set of tools to carry out an initial analysis directly on files of raw data. In particular, we propose a tool that supports simple SQL queries directly over files without the need for any database system. Our implementation is based on expanding some basic capabilities of the Unix command line and exploiting its ability to compose sequences of commands in pipelines. We present an initial implementation and report on some preliminary results.

Between-Sample Relationship in Learning Tabular Data Using Graph and Attention Networks

Manar Samad, Shourav Rabbani
Tennessee State University, Tennessee, USA

Abstract—Traditional machine learning assumes samples in tabular data to be independent and identically distributed (i.i.d). This assumption may miss useful information within and between sample relationships in representation learning. This paper relaxes

the i.i.d assumption to learn tabular data representations by incorporating between-sample relationships for the first time using graph neural networks (GNN). We investigate our hypothesis using several GNNs and state-of-the-art (SOTA) deep attention models to learn the between-sample relationship on ten tabular data sets by comparing them to traditional machine learning methods. GNN methods show the best performance on tabular data with large feature-to-sample ratios. Our results reveal that attention-based GNN methods outperform traditional machine learning on five data sets and SOTA deep tabular learning methods on three data sets. Between-sample learning via GNN and deep attention methods yield the best classification accuracy on seven of the ten data sets, suggesting that the i.i.d assumption may not always hold for most tabular data sets.

Unstructured Data Analytics to Improve Digital Eligibility of E-commerce Listings

Nikhitha Siddi, Sharan Sanjay Shirodkar, Matthew Lanham, Suneet Abraham, Akshay Deshmukh, Anish Jasti
Purdue University, USA

Abstract – N/A

Exploring the Trade-Off between Privacy and Predictive Power in Synthetic Data Generation

Naveen Shaji, Matthew Lanham, Pratik Kamat, Anusha Reddy, Prashanth Suresh, Amisha Turkel
Purdue University, USA

Abstract – N/A

Improving Fleet Utilization - A Comprehensive Analysis With Solutions For Shipment Brokerage Market Entry

Abhiram Chilukuri, Matthew Lanham, Aditya Uppuluri, Austin Bohlin, Hammad Nabi Khan, Nai-Wei Lu
Purdue University, USA

Abstract – N/A

Optimization of Loading Operations and Palletized Goods

Xindi Liu, Udyog Pati, Matthew Lanham, Nelson Capote, Tommy Starnes
Purdue University, USA

Abstract—There are several facets within supply chain operations that are pivotal to the success of minimizing time between order placed and order delivered. This research focuses on the optimization of truck loading. The motivation for this research is that delays in the packing and loading of pallets into trucks can lead to delays in arrivals which lead to customer dissatisfaction. A common scenario for freight packers is there exists orders having multiple items. They must estimate the number of pallets needed to fulfill an order, estimate the size and weight of each pallet. Lastly, there is an optimization step to decide if pallets can be stacked, and how to load the truck to allocate space efficiently. We construct a two-stage design where in the first stage prediction

is used to estimate pallets, then in the second stage a linear program optimization model is employed to effectively provide a truck loading recommendation. Our model integrates proper truck identification, truck constraints, pallet dimensions, and weight.

Social Recommendation through Heterogeneous Graph Modeling of the Long-term and Short-term Preference Defined by Dynamic Time Spans

Behafarid Mohammad Jafari, Xiao Luo, Ali Jafari
Purdue University, USA

Abstract—Social recommendations have been widely adopted in substantial domains. Recently, graph neural networks (GNN) have been employed in recommender systems due to their success in graph representation learning. However, dealing with the dynamic property of social network data is a challenge. This research presents a novel method that provides social recommendations by incorporating the dynamic property of social network data in a heterogeneous graph. The model aims to capture user preference over time without going through the complexities of a dynamic graph by adding time span nodes to define users' long-term and short-term preferences and aggregating assigned edge weights. The model is applied to real-world data to argue its superior performance. Promising results demonstrate the effectiveness of this model

Language Agnostic Readability Assessments

Vedanti Parag Gulalkari, Matthew Lanham, Mrinmoy Dalal, Sankarsan Gautam,
Amal Tom, Shreyas Joshi, Venkatesh Seetha
Purdue University, USA

Abstract – N/A

Made-to-Order: Targeted Marketing in Fast-Food Using Collaborative Filtering

Oluwayemisi Ajayi, Yuqiu Chen, Jason Crawford, Kamalika Das,
Venkata Rahul Karumuri, Matthew Lanham
Purdue University, USA

Abstract – N/A

Automation and Model Evaluation for Risk Mitigation in Banking

Mahima Ashok Kriplani, Yang Wang, Thannir Malai Annamalai Kumar,
Dae Sung Kim, Anushila Satpathi
Purdue University, USA

Abstract – N/A

Optimization of Magnetic Gripper Design for Efficient Robotic Sheet Metal Manipulation: A Comparative Study of Clustering Algorithms

Luis Deutsch Garcia, Ana Paula Trevino Trevino
Instituto Tecnologico y de Estudios Superiorres de Monterrey, Mexico

Abstract - In response to the growing trend of increased product variety in the manufacturing industry, flexible technologies are being employed to enhance production efficiency. This study specifically aims to optimize the design of grippers that manipulate

different sheet metal part geometries within a robotic bending cell. The study compares four clustering algorithms to determine the minimum gripper quantity that effectively handles all different production parts. The findings of this study demonstrate that a robust magnetic gripper family, consisting of 4 or 5 grippers, can potentially work with over 927 metal sheet parts while meeting the appropriate dimensions and carrying capacity requirements. Implementing this approach helps minimize engineering changes when introducing new products into the production process.

Using Statistical and Forecasting Methods to Predict Arrival of Container Shipments

Gauri Vaidya, Chethan Manjunath, Keerthana Nemili, Soham Patil, Sreeja Sesha, Matthew Lanham
Purdue University, USA

Abstract – N/A

On the Definition of Appropriate Trust and the Tools that Come with it

Helena Lofstrom
Jonkoping University, Sweden

Abstract—Evaluating the efficiency of human-AI interactions is challenging, including subjective and objective quality aspects. With the focus on the human experience of the explanations, evaluations of explanation methods have become mostly subjective, making comparative evaluations almost impossible and highly linked to the individual user. However, it is commonly agreed that one aspect of explanation quality is how effectively the user can detect if the predictions are trustworthy and correct, i.e., if the explanations can increase the user's appropriate trust in the model. This paper starts with the definitions of appropriate trust from the literature. It compares the definitions with model performance evaluation, showing the strong similarities between appropriate trust and model performance evaluation. The paper's main contribution is a novel approach to evaluating appropriate trust by taking advantage of the likenesses between definitions. The paper offers several straightforward evaluation methods for different aspects of user performance, including suggesting a method for measuring uncertainty and appropriate trust in regression.

Short-term Solar Photovoltaic Power Forecasting using a Multiple Similar-day Method based on Weather Type

Chung-Chian Hsu, Wun-Siang Chang, Ben-Jei Tsuang
National Yunlin University of Science and Technology, Taiwan

Abstract — With the amendment of relevant laws in recent years, Taiwan's renewable energy is moving towards an open market. Although it increases the freedom and bargaining power of private companies to sell electricity, it also means that companies must prepare their own energy storage equipment and grid dispatch plans. In the field of short-term solar power forecasting, the accuracy of power generation forecasting depends largely on the quality of the predicted radiation in the input variables. In the past, some studies have achieved good results by using the radiation of similar days as the predicted radiation. However, there are still some shortcomings in the method of selecting similar days. In this study, an improved method for similar days is proposed to reduce radiation error due to the sampling error of similar days by using multiple similar days; it is demonstrated that the improved method can reduce the error of radiation prediction (nRMSE) by 10% and the error of power generation prediction (nRMSE) by 2%.

Performance Reliability of Reinforcement Learning Algorithms in Obstacle Avoidance Game with Differing Reward Formulations

Brandon Hansen, Haley Dozier
US Army Engineer Research and Development Center, ERDC, USA

Abstract—When formulating environments for complex application areas, using analogies to games is beneficial as they provide convenient models to test algorithm performance in ways that are transferable to realistic environments. We propose a Frogger like grid based environment containing a simple action space, dynamic obstacles, and discrete game loop for testing Proximal Policy Optimization 2 and Deep Q-Network with comparisons to a heuristic and random agent. The environment contains four different reward function implementations, along with two different environment variations to explore adaptability. Seeing how these different parameters effect not just the average performance of the algorithm, but also the reliability of the performance is of concern as reliability determines the expectations of any single performance of a reinforcement learning agent. Experiments in these environments demonstrate common behaviors of reinforcement learning algorithms showing possible strengths and weakness of the approaches when applied to more complex decision-making scenarios. We explore these behaviors through evaluation techniques meant to measure the agents accumulation of reward in the game. Cross comparing these evaluation techniques elucidates the causation behind agent performance.

Thrive with Big Data: Navigate the Pandemic World with a Real-time Health Advisor

Andrew Yu, Lawrance Wang, Anthony Yan, Valerie Yu
Aragon High School, USA;
University of California, Santa Barbara, California, USA

Abstract—In this work, we investigate how to leverage Big Data to build an application to help medically under-privileged individuals to make travel decisions during pandemics. We propose a prototype platform which takes as inputs both personal heal information and hospital resource data of an intended destination, as well as additional configurations for user preferences, to make recommendations. We will discuss how to expand the platform to tackle upcoming challenges the world presents us.

Improved Decision Support for Product Returns using Probabilistic Prediction

Dirar Sweidan, Ulf Johansson, Beatrice Alenljung, Anders Gidenstam
University of Boras, Sweden

Abstract—Product returns are not only costly for e-tailers, but the unnecessary transports also impact the environment. Consequently, online retailers have started to formulate policies to reduce the number of returns. Determining when and how to act is, however, a delicate matter, since a too harsh approach may lead to not only the order being cancelled, but also the customer leaving the business. Being able to accurately predict which orders that will lead to a return would be a strong tool, guiding which actions to be taken. This paper addresses the problem of data-driven product return prediction, by conducting a case study using a large real-world data set. The main results are that well-calibrated probabilistic predictors are essential for providing predictions with high precision and reasonable recall. This implies that utilizing calibrated models to predict some instances, while rejecting to predict others can be recommended. In practice, this would make it possible for a decision-maker to only act upon a subset of all predicted returns, where the risk of a return is very high.

Autism Risk Classification using Graph Neural Networks Applied to Gene Interaction Data

Kyle Riccardi, Danushka Bandara
Fairfield University, Fairfield, Connecticut, USA

Abstract—We use a gene interaction network to predict which genes are associated with Autism Spectrum Disorder (ASD), thus allowing for earlier detection of ASD. ASD is a disorder that affects the development of the brain. No medical tests can diagnose ASD, so the diagnosis is based on the individual's behavior and development. We used gene location in the chromosome band and their possible interaction with each other to help classify the gene association with ASD. Graph sage and graph convolutional network (GCN) models were studied to organize these associations. The model's performance was analyzed using various parameters to obtain optimal performance. The graph sage model obtained an area under the ROC curve (AUC) of 0.85 percent on testing and 0.85 on validation. The GCN model achieved 0.78 AUC on testing and 0.78 AUC on validation. GCN achieves a specificity of 0.96 and a sensitivity of 0.63. Graph Sage achieved a specificity between 0.96 and a sensitivity of 0.94. Our classification performance demonstrates the capability of graph sage to classify genes carrying association risk for ASD. This analysis also demonstrates the inability of GCN to recognize association, as determined by the low sensitivity. Our results show that it is possible to apply graph neural networks to understand the link between diseases and genetics.

Using Learning Algorithms in Synthetic Mobility Trace Generation

Felipe Marques Megale, Richard W. Pazzi, Felipe D. Cunha
Ontario Tech University, Canada

Abstract—Traffic jams impact commuters every day. People waste time sitting in traffic, and emissions are also a factor to consider. It is known that making roads bigger and better does not solve congestion, but makes it worse. One of the solutions we can still apply is creating better policies for road use but testing them in the real world can be expensive and dangerous. Thus, we need better traffic simulation strategies that accurately depict reality. We propose traffic prediction by feeding real-world data to machine learning models. Our results show that it is possible with a high degree of confidence. Amongst the selected algorithms we evaluated, XGBoost displayed better performance, whilst fitting in an acceptable amount of time.

Predicting Atmospheric Air Pollution: A Convolutional-Transformer Approach for Spatial and Temporal Analysis of PM2.5

Pratyush Muthukumar, Shaurya Pathak, Kabir Nagrecha, Dawn Comer,
Navid Amini, Jeanne Holm, Mohammad Pourhomayoun, Janmesh Kalra, Hirad Hosseini
California State University, Los Angeles, California, USA

Abstract— 6.7 million premature deaths occur annually due to household and ambient air pollution. Air pollution affects individuals globally and can be derived from a variety of factors including household cooking fuel, motor vehicles, industrial practices, and natural fires. To tackle this global crisis, our research focuses on understanding the spatial and temporal patterns between air pollutants to predict future levels of air pollutants. Our approach uses a novel deep learning methodology that involves a spatiotemporal Convolutional-Transformer architecture (ConvTransformer) We harnessed multiple data sources as inputs to our ConvTransformer, including several remote sensing instruments such as NASA's Terra+Aqua satellites with the Moderate Resolution Imaging Spectroradiometer (MODIS), and the European Space Agency's Sentinel-5P satellite with the Tropospheric Monitoring Instrument (TROPOMI), gridMET ground-based meteorological remote-sensing data from North American Land Data (NLDAS-2), US Census TIGER Roadways, and ground-level PM2.5 sensing data sourced from EPA AirNow. Our approach shows a 38.8% decrease in 5-frame average SSIM error compared to existing modern deep learning architectures utilizing remote-sensing data, satellite imagery, and ground-level data for PM2.5 prediction.

Electronic Health Data in the Context of Patient Length-of-Stay Prediction: Using Generative Adversarial Nets for Synthetic Data Creation

Robert Stahlbock, Dominik Bietsch, Stefan Voss
University of Hamburg, DE, Germany

Abstract—While generative artificial intelligence has gained popularity (e.g., for the creation of images) it can also be used for the creation of synthetic tabular data. This bears great potential, especially for the healthcare industry where data is oftentimes scarce and underlies privacy restrictions. For instance, the creation of synthetic electronic health records (EHR) promises to improve the usage of machine learning (ML) algorithms, which normally work with large amounts of data. This also applies for the prediction of the patient length of stay (LOS), a key measure for hospitals. Thereby, the LOS represents one of the core tools for decision-makers to plan the allocation of resources. This paper aims to add to the young research concerning the application of generative adversarial nets (GAN) on tabular EHR. The intention is to leverage the advantages of synthetic data for the prediction of the LOS in order to contribute to the efficiency-enhancing and cost-saving aspirations of hospitals and insurance companies. Therefore, the applicability of synthetic data generated by GANs as a proxy for scarce real-world EHR for the patient LOS multi-class classification task is examined. In this context the Conditional Tabular GAN (CTGAN) and the Copula GAN are selected. The CTGAN is found to be the superior model for the underlying use case. Nevertheless, the paper shows that there is still room for improvement when applying state-of-the-art GAN architectures to EHR.

Empirical Investigation of Different Residual Models in Conformal Prediction

Robert Stahlbock, Arnold Bliesmer, Stefan Voss
University of Hamburg, DE , Germany

Abstract— Conformal prediction is a statistical technique that allows a point prediction from a regression algorithm to be transformed into a prediction region that contains the true target value with a chosen confidence. Normalization allows a conformal algorithm to consider the difficulty of a prediction by adjusting the size of the prediction regions. Easy cases get a smaller prediction region and vice versa. This paper focuses on normalization with a residual model. We investigate the suitability of regression models as residual models and how well they can generate prediction regions. Two evaluation criteria are used for the assessment. The empirical study includes nine datasets and three confidence levels and investigates the algorithms Random Forest, XGBoost, Knearest neighbors, Linear Regression, and AdaBoost. The experiments show that residual models must not produce predictions lower than the lowest value in the training dataset. Furthermore, AdaBoost is best suited for the residual mode in our experiments.

Challenges and Opportunities to Business Analytics and Integration due to Edge and 5G Technologies

Sathish Sampath
Boston University, Massachusetts, USA

Abstract— For organizations to succeed in their respective domains, they need to define a Business Strategy that outlines the Key Performance Indicators that are measurable. Organizations, then, define a Data Strategy that takes all data into consideration and provides analytics that eventually maps to the Key Performance Indicators that are defined. The recent advancements in technologies based on 5G and Edge deployments that are producing Big Data are primarily driven by the motivation of providing a highly interactive and localized customer experience. Organizations that are deploying solutions in Edge locations therefore need quicker analytics to react faster and therefore they need a data strategy that can meet their demands. This research paper aims to provide an overview of the various Data Integration and Business Analytics strategies that organizations have implemented as part of their overall Data Strategy. Additionally, this paper discusses the challenges faced by organizations deploying Big Data solutions due to the traditional Data Strategies and proposes a concept to meet the needs of organizations that need near real-time analytics by leveraging Analytics tools and Artificial Intelligence.

Data Preprocessing Using AutoML: A Survey

Abderahim Salhi
U.S Army Engineer Research and Development Center, USA

Abstract—Data preprocessing is a crucial step in any machine learning (ML) pipeline, as the quality of the data can greatly impact the accuracy and effectiveness of the final model. With the rise of automated machine learning (AutoML), data preprocessing has become even more important, as it can help to streamline the process of building and training models. This paper explores the use of AutoML platforms for data preprocessing, specifically focusing on the ways in which automated tools can be used for ML processes. We examine the different tasks and techniques along with AutoML tools. Finally, we outline several unresolved issues to handle data preprocessing work with the current AutoML techniques for further study.

K-means Clustering for Large Data: Anomaly Detection in Supervisory Control and Data Acquisition System

Corwin Stanford, April Tanner
Jackson State University, Jackson, Mississippi, USA

Abstract—K-means is a common unsupervised method for partitioning low dimensional datasets. However, it is usually not used for higher dimensional datasets, such as those found in Supervisory Control and Data Acquisition (SCADA) systems. Here we examine its application to a large dataset for the purposes of detecting cyber attacks using anomaly-based intrusion detection in the Battle of the Attack Algorithms (BATADAL) dataset. Additionally, dimensionality reduction using Principal Component Analysis (PCA) is examined as a method of improving anomaly detection performance. Emphasis is placed on methods for selecting parameters without using information obtained from examination of attack labels.

Data Collection Methods and Predictive Analysis for Fall Prevention in Elderly Populations

Chris Cheng Zhang
Canada Youth Robotics Club, Canada

Abstract—One of the most severe health risks to elderly populations is the risk of falling. Falls can lead to acute injuries, correlating with a decline in physical health. Recognizing the significance of this issue, the Fall Prevention Project aims to develop an algorithm that can accurately predict and prevent falls in elderly demographics. Data collection is essential to understanding and predicting falls. This article addresses the data collection methods used in the Fall Prevention Project. By collecting data to train an algorithm, injuries, and fatalities can be modeled to improve the lives of atrisk patients. For data collection, MediaPipe, a computer vision framework, is combined with an Intel RealSense D435 Camera. After skeleton tracking is completed, a Python code navigates through this data, which is then exported to an Excel file containing the coordinates and velocities for markers on the human body. A machine learning algorithm is used to classify patient movement as either “Risky” or “Safe”. From this algorithm, future patient behavior can be analyzed to identify and prevent when a fall is occurring.

Metadata: An Integral Component of the Modern Data Strategy

Mahmood Mohammed, John R. Talburt, Huzaifa Syed, Mehjabeen
Information Science, University of Arkansas at Little Rock, Little Rock, Arkansas, USA

Abstract - This paper explores the pivotal role of metadata within the data management practices of organizations, and its subsequent impact on business decision-making. This paper presents a comprehensive exploration of the application of Generative AI in the realm of data analytics. As the volume and complexity of data continue to grow, effective metadata management becomes

pivotal to understanding and utilizing data. With the ability to simulate human inferencing processes and generate contextual insights, generative AI offers promising avenues in improving metadata understanding, governance, integration, and quality. A crucial point of discussion in this paper is how the quality of metadata directly influences the output of Generative AI models. We present a proof-of-concept that illustrates this relationship using the COVID19 dataset and OpenAI text-ada-002 embeddings with GPT3.5 chat completion API. The paper further acknowledges potential concerns related to security, bias, privacy, IP, and the long-term effects on critical thinking abilities.

The 2nd International Conference on Emergent Quantum Technologies (ICEQT'23: July 24-27, 2023; Las Vegas, USA)

https://baylor.ai/iceqt/
https://www.american-cse.org/csce2023/

Concise Yet Efficient Hardware Design of a Quantum Coprocessor

Nadia Nedjah, Sergio Raposo, Luiza de Macedo Mourelle
Dept. Electronics Engineering and Telecommunications,
State University of Rio de Janeiro, Rio de Janeiro, Brazil;
Dept. Sysetms Engineering and Computation,
State University of Rio de Janeiro, Rio de Janeiro, Brazil

Abstract—This work proposes a customizable hardware design of a coprocessor that is able to perform quantum computations, exploring pure parallelism as well as a pipelined execution. The architecture is based on a quantum state memory, a scratch memory, a calculation unit, a measurement unit and, a control unit. All functional units communicate with each other via dedicated and shared data buses, depending information exchange frequency. The design is simulated and proven to be effective in executing quantum operations.

Vision Towards Quantum Humanities: An Architecture of a Holistic QML Environment

Johanna Barzen
University of Stuttgart, Institute of Architecture of Application Systems, Stuttgart, Germany

Abstract—The benefits of using classical computers in the humanities has been proven by the establishment of the digital humanities. Recently, quantum computers reached a state in which advantages can be realized in certain use cases. E.g., initial feasibility studies in the humanities have shown that some advantages can already be achieved, and much more can be expected, especially in applying quantum machine learning. But building quantum applications requires a deep understanding of mathematics, quantum algorithms and programming that cannot be expected by humanists. Consequently, they cannot explore potentially promising applications or apply concrete potentials of quantum algorithms. The proposed architecture and method aim at exactly this, enabling (digital) humanists to easily build their own quantum applications in the focus area of quantum machine learning to allow a broad use and critical evaluation of the potentials of quantum algorithms for humanities research.

Evaluating the Impact of Noise on Variational Quantum Circuits in NISQ Era Devices

Bikram Khanal, Pablo Rivas
Department of Computer Science, Baylor University, Waco, Texas, USA

Abstract—The limited supply of qubits and significant quantum noise impose limitations on the capability of quantum algorithms in the Noisy Intermediate-Scale Quantum (NISQ) era. NISQ devices have a variety of applications, such as Variational Quantum Circuit (VQC), which provides answers to difficult optimization and machine learning problems. This paper presents a thorough investigation of quantum variational classification in the NISQ context, with a focus on comprehending noise's impact on various feature maps and VQCs. We evaluate the effectiveness of quantum classifiers using a variety of datasets, ranging from straightforward binary classification problems to more complex tasks. Our results reveal the critical role that feature maps and variational circuit selection play in mitigating the effect of noise, identifying specific quantum circuit designs that exhibit

robustness even in noisy situations. In order to highlight the potential of quantum machine learning in solving complex problems within in the NISQ setting, this study emphasizes the delicate interaction between feature map selection, variational circuit design, dataset complexity, and quantum noise.

A Quantum Teleportation Protocol Secured by a Blockchain Technology

Jaylin Butts, Emmanuel White, Jinsuk Baek
Department of Computer Science, Winston-Salem State University, Winston-Salem, North Carolina, USA

Abstract—A quantum computer can be defined as a computer that implements high-speed calculations by actively utilizing the unique physical states of quantum mechanics. The unique physical states include superposition and quantum entanglement. Quantum computing takes advantage of these states to achieve greater computational power than traditional computing. One important application of quantum computing is quantum teleportation, which enables the transmission of a quantum bit (qubit) between communication entities using measurement-based calculations. However, since the transmission of measured data still relies on traditional communication channels, there can be a security issue, which could pose a potential obstacle to the widespread use of quantum teleportation in the future. In our paper, we propose enhancing the security of legacy quantum teleportation protocol by incorporating blockchain technology.

The 24th International Conference on Internet Computing & IoT (ICOMP'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-ICOMP
https://www.american-cse.org/csce2023/

Self-organizing QoS-aware Service Composition Framework for Edge Intelligence

Giandomenico Spezzano
National Research Council of Italy, Institute for High Performance Computing and Networking, Rende, Italy

Abstract—Edge computing has great potential for real-time AI services due to its ability to meet low response time and network bandwidth requirements. However, individual edge devices lack the computational power for complex AI tasks. AI tasks can be divided into sub-tasks and allocated across multiple edge devices to overcome this. A composition model of these sub-tasks is required that considers edge services with different Quality of Service (QoS) values. This paper introduces a new service composition framework, utilizing a self-organizing approach to enhance service selection considering run-time QoS changes. The framework combines an ant-based clustering algorithm for efficient service discovery and a self-Adaptive Multi-Objective Particle Swarm Optimization (MOPSO) algorithm for selecting suitable services based on the topological map generated. Simulation results validate the effectiveness of this method in enabling efficient and QoS-aware service selection.

An Onboard IoT-Based Communication System for Autonomous Vehicles

Hassan Rajaei, Deepak Thammineni, Sampath Yelchuri
Computer Science Department, Bowling Green State University, Bowling Green, Ohio, USA

Abstract— Automotive industry is facing multiple challenges in developing future vehicles from the traditional and combustion driven motors to transform them into electric cars, smart cars, and fully Autonomous Vehicles. Two major challenges of the transition regard high-speed computational and communications power onboard the auto platform. This paper proposes a communication platform called SC-Hub that integrates IoT, 5G, Mobile Edge Computing, VANET, and Special Sensors. The goal is to deliver a fast, reliable, and efficient communications system onboard the smart vehicle, and thus support an Intelligent Transportation System in the fleet. Our solution also supports the Internet of Vehicles where multiple smart vehicles can configure fleets on the road. A preliminary simulation study suggests that an onboard communication hub can dramatically reduce packet retransmissions in fleets while increasing throughput.

A Web Application to Visualize the Marine Highway Routes in the US

Vivek Sunchu, Natarajan Meghanathan
Jackson State University, Jackson, Mississippi, USA

Abstract—This paper presents the software architecture of the web-based application developed to visualize the marine highway routes of the US Intermodal Port Network (MIPN). The marine highways are being construed as effective alternatives for the traffic-congested interstate roads in the US. The MIPN comprises of 89 ports - 21 of which are core ports that are located on two or more marine highways and the rest of the ports are stub ports that are located in only one marine highway and are connected to the closest core port in that highway. The MIPN is modeled as an undirected graph featuring the core ports and stub ports, and the

edges (marine highways) connecting them, as described above. Referred to as the Marine Highway Route Planner application, it displays interactive maps using OpenStreetMaps and OpenLayers. The application uses React JS, NodeJS and Azure Cloud technologies as a part of its core architecture. We use custom port markers, route path attributes and marine highway layers to display relevant visual information.

Designing an Open Field Precision Agriculture System Using Drones

Kwang Ho Yang, Meong Hun Lee, Hyun Yoe
Department of Smart Agriculture Major, Sunchon National University,
Suncheon si, Jeollanam-do, South Korea;
Department of Information and Communication Engineering,
Sunchon National University, Suncheon si, Jeollanam-do, South Korea

Abstract — The study aimed to determine the performance of precision agriculture systems in detecting crop stress, disease, and other abnormalities and to provide farmers with convenience in managing their crops. The study found that among the spectral indices, NDVI and NDRE are effective in detecting crop stress and disease, while VARI is useful in identifying fields that need additional irrigation or fertilizer. Overall, the study shows the potential of an open field precision agriculture system using UAVs and spectral indices to improve crop management and productivity. The findings have implications for farmers, agricultural technology companies, and policy makers, and emphasize the need for continued investment in research and development to address the challenges and opportunities in this rapidly evolving field.

A Study of Honeybee-Guided Drones to Mitigate Damage from Honeybee Disappearance

Si Eun Park, Hyun Yoe, Meong Hun Lee
Department of Information and Communication Engineering,
Sunchon National University, Suncheon si, Jeollanam-do, South Korea;
Department of Smart Agriculture Major, Sunchon National University,
Suncheon si, Jeollanam-do, South Korea

Abstract —Bees play an important role in pollinating around 80 percent of crops that rely on insect pollination. However, bee populations are declining rapidly around the world. To address this issue, this thesis proposes a solution to prevent the loss of bees by using drones to attract bees to hives using pheromones and the scent of bee orchids. The thesis details the research, design and development of the drone system and its expected impact. The drone uses the scent of Ophrys apifera, which emits the same scent as the pheromones released by female bees, as an inducer to attract bees to the hive. By attracting bees to the hive, we are proposing a solution to bee loss.

LCP: A Low-Communication Parallelization Method for Fast Neural Network Inference for IoT

Ramyad Hadidi, Jiashen Cao, Younmin Bae, Da Eun Shim, Hyojong Kim,
Bahar Asgari, Sung-Kyu Lim, Michael S. Ryoo, Hyesoon Kim
Rain, AI, USA; University of Maryland, College Park, Maryland, USA;
Georgia Institute of Technology, Atlanta, Georgia, USA;
Google, Inc, USA; Stony Brook University, New York, USA

Abstract—Deep neural networks (DNNs) have stimulated research in diverse edge applications including robotics and Internet-of-Things (IoT) devices. However, IoT-based DNN inference poses significant challenges due to resource constraints. Further, as communication is costly, taking advantage of other available IoT devices by using data- or model-parallelism methods is not an effective solution. We introduce a low-communication parallelization (LCP) method to minimize communication overhead in distributed IoT systems. LCP models consist of multiple, largely-independent, narrow branches, providing enhanced distribution and parallelization opportunities while reducing memory and computational requirements. Implemented on AWS instances,

Raspberry Pis, and PYNQ boards, as well as a customized 16mW 0.107mm2 ASIC @7nm chip, LCP models yield maximum and average speedups of 56x and 7x, compared to original models, which could be improved by incorporating common optimizations such as pruning and quantization.

On Standards for Wireless Sensor Networks in the Application of Structural Health Monitoring

Peter Edge
Department of Computing, Ara Institute of Canterbury, Christchurch, New Zealand;
University of Southern Queensland, Australia

Abstract— This paper presents the view of The Internet of Things (IoT) from the perspective of Wireless Sensor Networks (WSN) in the application of Structural Health Monitoring (SHM). Automation of building data collection over time, detecting prolonged or rapid changes in physical structure properties offers financial and time saving opportunities. SHM has attracted a great deal of interest from the research community in recent years. Regulatory frameworks, security, and guidance for manufacturers have not evolved in parallel for the WSNs supporting IoT. In this context it is important to consider the effectiveness of WSN for IoT and its ability to deliver accurate, secure and reliable service to the civil and structural engineering SHM application. The number of research projects to consider the security and integrity of collecting data from critical sites is negligible. Based on these considerations, we present the idea of SHM systems designed in partnership with structural engineers. In this way, we propose regulated, secure WSN classified into SHM categories based on current information technology (IT) standards.

Harnessing the Power of GPT Model to Generate Adversarial Examples

Rebet Jones, Marwan Omar, Derek Mohammed
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA;
Saint Leo University, Saint Leo, Florida, USA

Abstract— In this paper, we propose a method for generating adversarial examples in the text domain using GPT-2, a state-ofthe-art language model. Our method employs an iterative algorithm to produce perturbations to input text samples, creating adversarial examples capable of fooling sentiment analysis models. We evaluate our approach on three widely-used benchmark datasets for sentiment analysis: Yelp, MR, and IMDB. Our results show that our approach can generate highly effective adversarial examples that significantly degrade the performance of sentiment analysis models. Specifically, we achieved a decrease in accuracy of up to 67.3% on the Yelp dataset, 68.1% on the MR dataset, and 52.5% on the IMDB dataset. We also discuss the limitations of our approach and the open challenges in this field. Overall, our study demonstrates the potential of GPT-2 for generating effective adversarial examples in natural language processing tasks.

IoT Malware Detection with GPT Models

Marwan Omar, Derek Mohammed, Rebet Jones, Calvin Nobles, Maurice Dawson
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA;
Saint Leo University, Saint Leo, Florida, USA

Abstract— With the proliferation of IoT devices, detecting IoT malware has become a critical challenge for cybersecurity. In this study, we propose a novel approach for IoT malware detection using code gadgets and the GPT language model. We extract code gadgets from the network traffic data of IoT devices, tokenize them using the Hugging Face Transformers library, vectorize them using the GPT model's embedding layer, and feed them into the GPT model for malware detection. We evaluate our approach on two publicly available datasets, IoT-23 and Malvis, and achieve high accuracy in malware detection, with F1-scores of 0.997 and 0.986, respectively. Our approach shows promise in detecting previously unseen malware variants and can be used to enhance the security of IoT devices.

Integration of Green Aspect inside Internet of Things Standard

Thierry Monteil
IRIT, Universite de Toulouse, INSA, Toulouse, France

Abstract— The Internet of Things is seen as a solution for many problems such as helping to create smart cities, industry 4.0 or even smart grids. The objective is in particular to offer a better service, to automate a set of services or to save money, particularly on energy consumption. Nevertheless, the complexity of these systems of systems requires the use of standards to deploy and make interoperable all of those connected objects and needed services. The oneM2M standard aims to provide a service layer to create those complex IoT systems. There are many scenarios and deployments in the literature and in standardization documents showing how the use of IoT can save or better manage energy consumption in a particular area, as opposed to energy consumption by the IoT system itself is little studied. This aspect, particularly in the oneM2M standard, is currently not very present except for interfacing with the cellular network and considering the capabilities of 5G in terms of reducing energy consumption. On the other hand, at the level of services, data management or the architecture to be deployed to set up an IoT system ranging from the sensor to the cloud, everything remains to be done. In this article, we propose to raise awareness on the one hand of the actors of IoT standardization of the energy impact in the definition of standards and on the other hand of the developers of IoT stacks and applications. For this, we first develop an environment to measure the energy impact in the choice of the use of a standard and its implementations then in a second time, we identify a set of recommendations both on the evolutions possible of the standard but also on the choice of implementation and deployment of applications and IoT stacks.

Optimized Decision Trees to Detect IoT Malware

Angel Jones, Marwan Omar
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA

Abstract— The proliferation of the Internet of Things (IoT) devices has led to an increased risk of cyberattacks and malicious activities, including the spread of malware. To mitigate these risks, it is crucial to develop effective approaches for detecting IoT malware. In this study, we propose a framework for detecting IoT malware using optimized decision trees with AdaBoost. We use two widely used datasets, NSL-KDD and CICIDS2017, to evaluate the performance of the proposed framework. The framework includes feature selection and hyperparameter tuning to enhance the performance of the model. Our results show that the proposed framework achieves high accuracy, precision, recall, F1 score, and AUC-ROC in detecting malware attacks. However, the study also has limitations, such as the focus on network-level features and the limited evaluation on specific datasets. Future research can address these limitations by testing the proposed framework on more diverse datasets and exploring different machine learning algorithms and techniques. Overall, our study provides a promising approach to detect IoT malware and can contribute to the development of more robust and effective approaches for network intrusion detection.

Application of IoT in Climate Change and Forest Fire

Alireza Kavianpour, Stephen Riegel, Peter Vitale, Taren Roberts, Jacob Dowell
College of Engineering and Information Sciences, DeVry University, Ontario, California, USA

Abstract - In the past three decades, climate change has doubled the total area burned by forest fires in the western states of America. Forest Fire Alert System (FFAS) is a student senior project to tackle this issue. It is an economical system which aims at solving the problem of wildfires experienced in the United States and other parts of the world. The system is based on the use of GPS (Global Positioning System) and Internet of Things (IoT). Upon detection of fire, Forest Fire Alert System (FFAS) unit can quickly report the presence of fire to the command center. Upon detection of fire, the command center will send an alert and trigger an alarm at the fire station so fire personnel can quickly start the required action. The main advantages of this student project are cost, size, the ease of mass production, and capability for connecting them as a network in the different parts of a forest.

The Way Forward: The Environmental Implications of Bitcoin and a Possible Solution

Elaine Ulsh, Xin Sun
Computer Science Department, Ball State University, Muncie, Indiana, USA

Abstract—There is a concerning rise in the energy consumption of Bitcoin due to its increase in Carbon Dioxide emissions. In 2020, such emissions were similar to that of a country the size of Austria or Portugal. In the coming decades, Bitcoin is expected to contribute to a 2‰increase in global temperatures. Environmental organizations across the globe are demanding some sort of mitigation. We propose a solution that we believe, if implemented correctly, could reduce Bitcoins emissions without forcing miners to update the version of Bitcoin they are mining with. Our solution is similar to the process where food is certified to be organic, except with cryptocurrency and renewable energy. A study by the University of New Mexico suggests that if 88.4% of miners would run their rigs using clean energy, it would take climate damages per coin from 35% to 4%. Our solution could be possible by having a third party organization be in charge of running a certification process for which Bitcoin gets certified as being renewably mined. This can be done through the certification of specific miners as opposed to the individual Bitcoin itself. Such a solution would uphold the principles that many Bitcoin users hold dear: decentralization and anonymity. Since many organizations and individuals claim to be environmentally focused, our hope is that they would prefer to buy more of the renewably mined Bitcoin, instead of the alternative. In particular, many companies that are incharge of pensions tend to claim that they are environmentally friendly, which is proven through their investing in environmentally friendly companies. As such companies explore the world of cryptocurrency, our hope was that they would choose to invest in renewably mined Bitcoin as opposed to the alternative. This would then artificially inflate the price of the "clean" Bitcoin and/or lower the price of the "dirty" Bitcoin, which provides incentive for miners to participate.

SpearRidge: A Response Time Prediction Model

Asaad Althoubi, Hassan Peyravi
Department of Computer Science, Kent State University, Kent, Ohio, USA

Abstract—Accurately predicting response times in cloud data centers is essential for improving service quality and optimizing performance. However, this remains a challenging task due to various factors such as network congestion, hardware failures, imbalance loads, and over/under utilization of servers. To address this, a new machine learning model called SpearRidge is introduced in this paper. SpearRidge utilizes Spearman's rank correlation technique and advanced Ridge regularization to select relevant features and mitigate the impact of over-fitting, respectively. The regularization parameter is fine-tuned through cross-validation to achieve an accurate and interpretable model with minimal trade-off. SpearRidge demonstrates remarkable accuracy and robustness towards outliers, having been trained on various traffic traces. The experimental results obtained from these traces indicate that SpearRidge significantly improves the prediction of response times, thereby enabling optimal task assignment within cloud and data center environments.

A Comparative Performance Analysis of Vulnerability Detection Schemes for Specific Blockchain Applications

Jae Geun Song, Cheol Hun Park, Dong Hwan Jo, Ju Wook Jang
Department of Electronic Engineering, Sogang University, Seoul, South Korea

Abstract - This paper is an ongoing work in the literature for the analysis of vulnerabilities that can be defined based on the application. Smart contracts play an important role in the security of digital assets and sensitive user information in blockchain networks. However, vulnerabilities in smart contracts can threaten the security of these assets and information, causing potential losses and affecting the trust and integrity of the blockchain. This paper compares several application-specific smart contract vulnerability detection schemes. We run each analyzer against a smart contract application, and then examine whether pre-looked-for and un-looked-for vulnerabilities are detected, depending on the application. The results of the study will help you choose a

smart contract vulnerability detector based on your application. Our research aims to define vulnerabilities that can be defined according to a specific application and then detect them.

Designing a Livestock Management System with a Vet Call System

Hyeon Seo Kim, Hyun Yoe, Meong Hun Lee
Department of Information and Communication Engineering, Sunchon National University, Suncheon si, Jeollanam-do, South Korea;
Department of Smart Agriculture Major, Sunchon National University, Suncheon si, Jeollanam-do, South Korea

Abstract — To prevent farm losses due to livestock diseases, the user enters the information of the livestock and collects data measured by the difference in body temperature of the livestock using thermal imaging cameras and ICT technology, and sends a notification to the farmer when an abnormality occurs. Users can send information to the veterinarian for a quick response to symptoms, and a notification of the veterinarian's temporary measures is sent to the user, and a veterinarian is called. In times when infectious diseases are more likely to occur, such as during the winter, summer, and rainy seasons, it is desirable to administer preventive drugs to all domesticated animals, so we designed a service system to schedule a veterinarian call.

Research on the Design and Application of Digital Twin-based Smart Agricultural Systems

Hyun Jun Kim, Meong Hun Lee, Hyun Yoe
Department of Information and Communication Engineering, Sunchon National University, Suncheon si, Jeollanam-do, South Korea;
Department of Smart Agriculture Major, Sunchon National University, Suncheon si, Jeollanam-do, South Korea

Abstract— The smartization of agriculture is happening at a rapid pace. A large number of sensors are being used in agricultural systems to collect data, analyze it, and provide appropriate environmental conditions to lead to more efficient agricultural production. This paper is a study on building a smart agricultural system using digital twins. To this end, the proposed agricultural system collects crop and environmental data from sensors in real time and generates digital twins based on them. The generated digital twins are modeled using machine learning algorithms, and agricultural experts can then determine We show that it is possible to build a smart agricultural system using digital twins that can increase crop productivity and reduce costs.

Intelligent Fall Detection Monitoring System

Hassan Rajaei, Satya Upendra
Department of Computer Science, Bowling Green State University, Bowling Green, Ohio, USA

Abstract— The efficacy of fall detection devices can be impacted by the different issues that aging communities experience. The use of these systems by elderly individuals can be challenging due to bodily and mental impairments, and some people might be reluctant to utilize the required gear. False alarms and detections that are missing can also be an issue, which undermines user confidence and discourages use of the system. Despite these difficulties, fall monitoring systems have the potential to significantly increase the safety and wellbeing of elderly people and people who have mobility issues by lowering the risk of injury and facilitating quick medical treatment in the event of a fall. We propose an IOT-based fall detection and monitoring system which is an advanced and modern solution utilizes the integration between multiple sensors to improve the safety and welfare of elderly people with mobility problems. The proposed system incorporates several cutting-edge sensing technologies, such as Wi-Fi signals, floor pressure sensors, smart carpets, accelerometers, gyroscopes, GPS, and pulse sensors, to detect and monitor falls in real-time. The suggested system can identify falls and offer real-time tracking for patients, lowering the risk of injury and allowing quick medical treatment.

A Latency-aware Container Scheduling in Edge Cloud Computing Environment

Hajime Miyazawa
Department of Electronics and Communication Technology, Nanzan University, Nagoya, Japan

Abstract—In an edge cloud environment where cloud computing technologies are applied to edge computing, it is necessary to make container orchestrators deal with problems caused by the heterogeneity of network connectivity. In this paper, we propose a communication latency-aware container scheduler for edge cloud environments. When an application is accessed from an end device, the access request is made to the appropriate computing node based on the communication latency. The scheduler deploys/reallocates application containers according to access to applications from end devices. An experimental result with our prototype implementation showed the effectiveness of the proposed scheduling system.

IT/OT Convergence Protocols: MQTT, OPC, and REST

Bassam Zahran, Fuad Abu Zahra
Department of Computer Science and Software Engineering, University of Wisconsin Platteville, Wisconsin, USA;
Department of Engineering and Applied Sciences, University of Wisconsin Milwaukee, Wisconsin, USA

Abstract—Information Technology and Operational Technology convergence refers to the integration and merge of IT and OT systems within an organization. Several protocols are commonly used in IT/OT convergence to facilitate communication and data exchange between IT and OT systems. It is imperative to understand how both systems are linking and exchanging data to be able to examine potential security breaches. In this study, we will discuss some of the common protocols used to achieve this convergence. We will include the known strengths and weaknesses of each. This will be considered an initial foundation to study potential security breaches and countermeasures for IT/OT convergence.

Information Exchange Among Routers Using Newly Defined IP Options for Traffic Control

Hiroto Hoshiba, Kenji Ichijo, Akiko Narita
Graduate school of Science and Technology, Hirosaki University, Hirosaki, Japan

Abstract— On the Internet, routers choose the shortest path basically. The confluence of transmission routes for flows can often cause congestion. Coordinating routes through adaptive routing has the potential to reduce congestion. However, its practical use is limited. Communications among routers under heavy traffic is problematic. It is difficult to propagate a packet under heavy traffic so that network devices outside the hotspot can learn of its existence belatedly. Furthermore, if routers frequently send traffic information for quick delivery, the packets for the information make the traffic worse. In this paper, we propose a method to exchange among routers using newly defined IP options for adaptive routing. We implemented functionalities for the method on Click modular routers and confirmed validities.

IT/OT Convergence Protocols: DNP3, Ethernet/IP, and MODBUS

Bassam Zahran, Fuad Abu Zahra
Department of Computer Science and Software Engineering, University of Wisconsin Platteville, Wisconsin, USA;
Department of Engineering and Applied Sciences, University of Wisconsin Milwaukee, Wisconsin, USA

Abstract—The convergence of Information Technology (IT) and Operational Technology (OT) systems is becoming a necessity. In the context of IT/OT convergence, various protocols are widely employed to enable smooth communication and data exchange between IT and OT systems. It is crucial to gain a comprehensive understanding of the networking and data flow between these

systems in order to effectively identify and address potential security risks. This study aims to explore several common protocols utilized in IT/OT convergence and examine their strengths and weaknesses. This initial exploration will serve as a foundation for further investigation into potential security breaches and the development of appropriate countermeasures in the context of IT/OT convergence.

Camera-Based Analysis of Human Pose for Fall Detection

Chris Cheng Zhang
Canada Youth Robotics Club, Canada

Abstract— Falls are a significant public health concern, especially among seniors and older adults, leading to severe injuries, decreased independence, and increased healthcare costs. Automated fall detection systems have emerged as a potential solution to address this issue by detecting falls in real-time and alerting caregivers or emergency services. While traditional fall detection methods rely on wearable devices, challenges arise in accurately detecting falls in complex environments. This report proposes a camera-based analysis of human pose using the Intel RealSense Depth Camera D435 to enhance the accuracy and reliability of fall detection systems. The methodology for this study involves the integration of OpenCV, MediaPipe, and Numpy python libraries to analyze human poses in three-dimensional space. This depth-based analysis includes calculating the depth of pose landmarks, estimating the body's center of mass depth, and tracking the velocity of movement. The findings of this study contribute to the field of fall detection, offering insights into the advantages of a camera-based analysis of human pose and improving the safety of individuals at risk of falls. Future work should include the integration of machine learning techniques and multi-modal approaches to enhance system performance and expand applications beyond fall detection. The technology showcased in this study has implications for healthcare, fitness, rehabilitation monitoring, and elderly care, providing valuable insights to support various domains and improve quality of life.

Design a Cloud-based Smart Indoor Plant Growing System

Seung Jae Kim, Meong Hun Lee, Hyun Yoe
Department of Information and Communication Engineering, Sunchon National University,
Suncheon si, Jeollanam-do, South Korea;
Department of Smart Agriculture Major, Sunchon National University,
Suncheon si, Jeollanam-do, South Korea

Abstract— Recently, due to population growth and urbanization, the supply of crops in urban areas has become difficult. In order to solve and contribute to these problems, this paper introduces an indoor smart cultivation system that interfaces with a cloud platform DB equipped with 30 crop optimal environment data recipes. The system is designed to be composed of various subsystems to provide functions such as adjusting the optimal environment based on the cultivation recipe. It is expected that this system can be used in abandoned subway stations and abandoned buildings due to urbanization to solve the problem of supplying required crops.

The 22nd International Conference on Wireless Networks (ICWN'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-ICWN
https://www.american-cse.org/csce2023/

A Lightweight Encoder and Decoder for Non-Binary Polar Codes

Franklin Cochachin, Fakhreddine Ghaffari
ETIS UMR 8051, CY Cergy Paris Universite, ENSEA, CNRS, Cergy, France

Abstract—In this paper, we associate the Cyclic Code-Shift Keying (CCSK) modulation to Non-Binary Polar (NB-Polar) codes to have good decoding performance at ultra-low signalto-noise ratios (SNRs). We show that the kernel transformation of order 2 using kernel coefficients equal to 1 reduces the complexity of the encoder and decoder and does not degrade error-correcting performance. We consider the SuccessiveCancellation Min-Sum (SC-MS) decoder. To greatly reduce the complexity of the SC-MS decoder, we choose a small number of most meaningful values at the check node processing. Our simulation results show that the optimized SC-MS decoder presents a negligible performance degradation with respect to the SC decoder.

5G Architecture based on Software-Defined Perimeter (SDP)

for Direct Trust Access to Private Networks
Woocheol Kim, Kiwon Kim, Jongkuk Lee, Heasook Park
Defense ICT Convergence Research Section, Electronics and
Telecommunications Research Institute (ETRI), Daejeon, South Korea

Abstract—Challenges of securing 5G networks have more complex security vulnerabilities due to a large number of connected devices and services. To provide access to private networks using the public 5G architecture, the use of virtual private network (VPN) techniques is commonly employed to address 5G security issues. However, the VPN protocol is vulnerable in that it can protect unauthorized nodes as authentication is done only after the connection. We propose a secure 5G architecture that utilizes the software-defined perimeter (SDP) to authenticate and then grant access to 5G devices based on specific criteria for direct trust access to private networks. We also show the communication flow of direct trust access to a private network through a public 5G architecture using the VPN technique and a direct trust access server (DTAS).

An Modified YOLOv5 Algorithm to Improved Image Identification for Autonomous Driving

Chun-Chieh Wang, Wen-Piao Lin
Department of Electrical Engineering, Chang Gung University, Taoyuan, Taiwan, R.O.C.

Abstract—In this paper we propose a solution to improve image identification on the early expressway engineering warning signs for autonomous driving. This scheme uses the modified lightweight YOLOv5 algorithm model to train feature categories. After manual labeling and data augmentation, the dataset is sent to the advance deep neural network for training. The practical experimental results show that the modified algorithm model can effectively identify engineering warning signs on the expressway. It allows drivers and construction units on the expressway to use the road more safely.

Gradual Route Modification in Mobile Wireless Multihop Network with Combination of Carrying and Forwarding

Ryota Kurabayashi, Hiroaki Higaki
Department of Robotics and Mechatronics, Tokyo Denki University, Tokyo, Japan

Abstract—In a wireless multihop network, data messages are transmitted along a wireless multihop transmission route composed of a sequence of intermediate wireless nodes forwarding the data messages. However, it is difficult to detect a wireless multihop transmission route from a source wireless node to a destination one if a wireless nodes are not evenly distributed and there are some areas in which wireless nodes are sparsely distributed. Some methods have been proposed to solve this problem by combination of forwarding data messages between neighbor intermediate wireless nodes as in a usual wireless multihop network and carrying data messages to pass through such areas. In the conventional methods with the combination of forwarding and carrying data messages, some dedicated intermediate mobile wireless nodes serve the role of carrying. Hence, such nodes are required to consume more battery power which results in lower connectivity of the network. In addition, transmissions of data messages might be suspended for waiting the carrying intermediate wireless nodes which results in longer end-to-end transmission delay. In order to solve the problems, this paper proposes a novel carrying and forwarding method which gradually modifies a wireless multihop transmission route to evenly share the required mobility overhead among all the intermediate mobile wireless nodes and to reduce end-to-end transmission delay of data messages by making the route shorter. The proposed method ensures that no successive intermediate wireless nodes pass each other without neighboring.

The 22nd International Conference on Information & Knowledge Engineering (IKE'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-IKE
https://www.american-cse.org/csce2023/

Culstering Users by Information-Seeking Style: An Empirical Study on an Academic Search Engine

Somayeh Fatahi, Amineh Setayesh, Julita Vassileva
Department of Computer Science, University of Saskatchewan, Canada;
Department of Computer Engineering, Shahid Bahonar University of Kerman, Iran

Abstract—Nowadays, academic search engines have become indispensable tools for getting important online scholarly information. User differences are important factors that influence the use of information systems. The way people use academic search engines to find information varies depending on their information-seeking style. Therefore, finding and understanding different information-seeking behaviors has become an important line of research. User behavior patterns can be discovered by examining user interaction logs to determine who the users are and what they intend to do. These insights can be useful in designing more optimized academic search engines. In this paper, we analyze the user interaction logs collected from the Iranian scientific information database. The Ganj database is the official repository for collecting and organizing theses and dissertations in Iran. Many researchers search scientific and research resources from the Ganj database daily. We use a sequential pattern mining approach to extract frequent sequential behavior patterns on user interaction logs and to cluster users into three groups based on their frequent behavior patterns, using the K-means clustering algorithm. Cluster analysis shows that users with similar frequent behavior patterns have similar information-seeking styles. Finally, we found three clusters and named them: fast surfers, deep divers, and broad scanners. Our findings can help developers of academic search engines and policymakers to identify users' needs and priorities to make better decisions to design a reasonable page layout and well-organized website for all users based on their search styles.

Influence Maximization in Dynamic Social Networks Under Partially Observable Environments

Gkolfo I. Smani, Vasileios Megalooikonomou
Department of Computer Engineering and Informatics, University of Patras, Patras, Greece

Abstract—Influence and influence diffusion have been studied widely in social networks. Influence maximization is the problem of detecting a set of influential nodes in a social network, which represents relationships among individuals. Although, most of the existing works on this task focus on static networks, in this paper we study the problem of influence maximization in dynamic social networks where the changes that occur over time can be observed by periodically probing nodes to update their connections. The goal of this work is to probe a subset of nodes in a social network so that the actual influence diffusion process in the network can be best uncovered with the probing nodes. We propose three algorithms, MaxC, MaxDC and MaxT, to approximate the optimal solution with probing nodes, which achieve improvement on estimating the number of activated nodes over state-of-the-art-method by 2.88%, 4.95% and 5.39%, respectively.

Measuring the Impact of Global Health Emergencies on Self-disclosure Using Language Models

Bianca Jones, Marwan Omar
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA

Abstract— This paper examines the impact of global health emergencies on self-disclosure using natural language processing (NLP) techniques. Specifically, we applied the Bidirectional Encoder Representations from Transformers (BERT) model to sentiment analysis and topic modeling of large datasets of tweets, Reddit posts, and news articles related to the COVID-19 pandemic. Our analysis revealed that individuals expressed a range of emotions in response to the pandemic, with sentiment scores varying across different platforms. We also identified key topics related to the pandemic, with discussions around these topics varying depending on the subreddit or news outlet. Our findings suggest that global health emergencies have a significant impact on self-disclosure and the ways in which individuals and communities communicate about public health issues. The differing sentiment and topic trends across platforms and news outlets highlight the importance of tailoring communication strategies to effectively address the unique challenges posed by global health emergencies. This study provides insights into the role of communication in global health emergencies and how NLP techniques such as BERT can be leveraged to understand and address these challenges.

A Further Analysis of Insect Trading in Shrek Super Party

Brandon Packard, Prakhar Singh
Pennsylvania Western University - Clarion Campus, Clarion, Pennsylvania, USA;
Clarion Area High School, Clarion, Pennsylvania, USA

Abstract—Sometimes, seemingly simple video game mechanics belie surprising complexity. One old game, Shrek Super Party, is rough around the edges but has a game mechanic that is very computationally interesting. The mechanic operates similarly to poker, but with a few modifications: all "cards" are visible, there are no suits and no straights, and instead of drawing "cards" players move by forcibly swapping their "cards" with other players. In this paper, we generated data from all players moving optimally and then analyzed that data to see what scores each player received on average over many rounds, along with other interesting pieces of data. It is our belief that, although creating an AI to play a game optimally is a valuable avenue of research, observing what happens when a game is played in specific ways is also academically valuable.

Basics of Possibilistic PSYOPS for Decoy/Countermeasure Methods

James A. Crowder, John Carbone
CAES Advanced Program Development, Colorado Springs, Colorado, USA;
Global Governments & Critical Infrastructure, USA

Abstract - Situational/threat assessment strategies have been studied for generations. Typically, these threat assessments utilize Bayesian belief networks and inference engines, based on decision tree technologies, to determine the likelihood of different deployment strategies and prevention methods known as psychological operations (psyops). These are typically represented as directed "acyclic" graphs and utilize joint probability distributions, which are typically based on incomplete information in the form of probabilistic outcomes when analyzing various aspects of the current mission parameters. Bayesian believe network solutions are good at showing qualitative relationships between entities and have a compact and theoretically sound foundation. Problems arise when answers to questions are required which cannot be specifically addressed by the Bayesian probabilities. Also, Bayesian methods tend to be computationally intensive. Here we look at alternative robust fuzzy possibilistic methods based upon research in abductive learning models and mutual information theory which are less computationally intensive than standard Bayesian methods.

A Smart User Interface for Structured Deep Web Search

Amal Aljohani
Department of Computer Science, College of Computer Science and Information Technology,
Majmaah University, Al-Majmaah, Riyadh, Saudi Arabia

Abstract—Traditional search engines such as Google, Bing and Yahoo! can only access 'shallow web' and 'surfaced deep web' contents. While there are several hundred times more information that is potentially useful for users' need exists in deep web repositories, they cannot be accessed using these engines. To search for information meaningfully, users must necessarily visit each of the deep web sites and submit queries in the portals to gain access to information they seek. Often, they must filter further to reach their desired information. In this paper, we review several prominent search engines to summarize the progress toward structured querying of deep web databases in a way prominent search engines would, even if it is possible. We focus in particular e-commerce databases to facilitate comparison shopping in ways similar to Amazon.com or Alibaba.com. However, we are not focusing on simple queries that target one specific domain, we are are also handling complex queries from multiple domains. We present an abstract view in a prototype system, called DeepQ, of the unified deep web e-commerce and related sites so that users are able to search information conveniently. DeepQ provides a single database illusion of the underlying databases using a smart query engine and site specific query reformulation. We briefly discuss DeepQ's features and functionalities.

Content Analysis of Items in Newspaper Data Using Table Arrangement Technology and ChatGPT for Stock Price Prediction

Masaki Murata
Graduate School of Engineering, Tottori University, Tottori, Japan

Abstract—In this study, we employed table arrangement techniques and ChatGPT to analyze newspaper content relevant to stock prices. Using table arrangement techniques, we effectively organized sentences from articles into tables, extracting 22 key content elements. Additionally, we discovered that ChatGPT possesses the ability to extract and present newspaper data in tabular form. Factors were found to influence stock price movements. Drops in stock prices were impacted by factors such as crude oil prices, and the COVID-19 pandemic. Conversely, rising stock prices were supported by global trends, and vaccine effectiveness. Furthermore, we propose a highly effective, large-scale method for constructing tables by combining table arrangement techniques and ChatGPT. The proposed method achieves an accuracy rate of 0.95 under a lenient criterion.

Machine Learning Prediction of DoD Personal Property Shipment Costs

Tiffany Tucker, Torrey Wagner, Paul Auclair, Brent Langhals
Graduate School of Engineering and Management, Air Force Institute of Technology,
Wright-Patterson AFB, Ohio, USA;
Integrated Analytics & Support, LinQuest Corporation, Beavercreek, Ohio, USA

Abstract— U.S. Department of Defense (DoD) personal property moves account for 15% of all domestic and international moves - accurate prediction of their cost could draw attention to outlier shipments and improve budget planning. In this work 136,140 shipments between 13 personal property shipment hubs from April 2022 through March 2023 with a total cost of $1.6B were analyzed. Shipment cost was predicted using recursive feature elimination on linear regression and XGBoost algorithms, as well as through neural network hyperparameter sweeps. Modeling was repeated after removing 28 features related to shipment hub location and branch of service to examine their influence on algorithm performance. The best model resulted from a neural network hyperparameter sweep on the simple 4-feature dataset. That model contained 2 hidden layers of 250 neurons, possessed a mean absolute error (MAE) of $2162 on the holdout dataset, and had an overfitting measure of 0.2%. This model's performance is a substantial improvement over a trivial mean-prediction model that possessed a MAE of $6856. Additionally, the model R2 = 0.87 compared favorably to existing work that achieved R2 = 0.73. Models that included 28 features related to shipment hub location and branch of service did not improve performance, showing that those features are insignificant in predicting shipment cost. All

OLS regression and XGBoost models showed that the most influential features (in order) were weight, month of shipment, distance shipped, and days in transit.

Context Extraction in Unsupervised Entity Resolution

Fumiko Kobayashi, John R. Talburt
Donaghey College of Science, Technology, Engineering, and Mathematics,
The University of Arkansas at Little Rock, Arkansas, USA

Abstract—Modern Entity Resolution (ER) system leverage metadata about the reference data to facilitate processing and making equivalence decisions. This requires that each source of input data must be pre-processed individually to conform to a common metadata alignment. During this alignment, the data is also preprocessed to detect and correct errors to optimize results. After metadata alignment and data cleansing, the sources are merged into a single dataset and submitted to an ER process. This is a costly process and requires additional passes of the input data prior to ER decisioning. This paper formally defines the concept of context within ER to replace the need for metadata alignment and preprocessing. During research into Unsupervised Data Curation (UDC), it was found that there are four (4) points within ER processes for which context can be extracted and used to make decisions to perform the error corrections in flight during processing with no need for metadata. Each of these types of context are formally defined as "Global", "Block", "Interim Cluster" and "Final Cluster" in this paper.

Atmospheric Meteorological Effects on Forecasting Daily Lightning Occurrence at Cape Canaveral Space Force Station

Jon Saul, Torrey Wagner, Eric Mbonimpa, Brent Langhals
Graduate School of Engineering and Management, Air Force Institute of Technology,
Wright-Patterson AFB, Ohio, USA

Abstract— As the Cape Canaveral Space Force Station and Kennedy Space Center increase their launch rate, any process that could assist in the automation of the currently-manual lightning forecast would be valuable. This work examines the possibility of machine-learning assistance with the daily lighting forecast which is produced by the 45th Weather Squadron. A dataset consisting of 34 lightning, pressure, temperature and windspeed measurements taken from 334 daily weather balloon (rawinsonde) launches in the timeframe 2012-2021 was examined. Models were created using recursive feature elimination on logistic regression and XGClassifier algorithms, as well as Bayesian and bandit optimization of neural network (NN) hyperparameters. The modeling process was repeated after eliminating 13 features related to windspeed. The best performing models on both datasets were the optimized NN models, with an F1 metric of 0.79 on the full dataset and 0.66 on the reduced dataset. For comparison, a model that predicted randomly achieved F1 = 0.47 on this dataset. The addition of 13 windspeed-related features more than doubled the complexity of the 21-feature no-wind model while increasing model performance by 13 percentage points. A notable inference from the statistical modeling is that the most important feature from both datasets was the Thompson convective index, which is related to temperature, dewpoint, relative humidity and lapse rate.

COVID-19 and US Labor Force: County-Level Insights

Anissa Champion, Daniel Ojeda, Ching-Yu Huang, Daehan Kwak
Department of Computer Science and Technology, Kean University, Union, New Jersey, USA

Abstract—The COVID-19 pandemic has highlighted the crucial role of socio-economic factors in shaping health outcomes. Understanding the correlations between different societal attributes and how each one contributed to COVID-19 outcomes in the United States is essential. Therefore, this study utilized realtime reported COVID-19 data from a database known as SimpleMaps and integrated it with US county data to analyze the societal attribute, labor force, and its correlation with COVID-19 outcomes. This study reveals that the labor force has a correlation with COVID-19 cases and deaths. In addition to identifying correlations, the study creates visuals to help visualize the trends better. These findings and visuals can offer valuable insights for policymakers

in implementing targeted strategies to mitigate the impact of COVID-19 on vulnerable populations and enhance the understanding of socio-economic factors' impact on COVID-19 outcomes.

A Taxonomy Based Digital Platform Evalution Model for Air Quality Data Management

Saba Siddiqui, Asif Qumer Gill
Faculty of Engineering and Information Technology, University of Technology, Ultimo, NSW, Sydney, Australia

Abstract – There is an increasing interest in digital platforms for managing air quality data. There are several platforms, and the challenge is: how to evaluate such platforms for air quality data management? This paper proposes a taxonomy-based digital platform evaluation model (DPEM) to assist with assessing such digital data platforms from cost and quality perspectives. The DPEM provides a qualitative evaluation criterion, which has been developed based on the review of both academic and industry sources. The application of the DPEM is demonstrated through its mapping of three well-known digital platforms. The application results indicate that DPEM is useful for evaluating air quality data platforms. However, it could be adapted to other contexts and data domains as well.

Improved Machine Learning Approach for Shanghai City's Second-hand Housing Price Prediction

Guang Hu, Yue Tang, Rui Yi
School of Statistics and Information, Shanghai University of International Business and Economic, Shanghai, China;
School of Computer Science, Fudan University, Shanghai, China;
Shanghai Key Laboratory of Data Science, Shanghai, China

Abstract—This paper presents an improved machine learning approach for prediction of second-hand housing prices in Shanghai. It firstly builds the random forest model and the XGboost model with Shanghai second-hand housing transaction data, and then analyses the challenges in the improvement of the models' prediction accuracy and generalization performance. In the light of that, it introduces the Lasso model for variable selection, and deletes three variables with insignificant regression coefficients, and then rebuilds the random forest and XGboost prediction models. The experimental results show that the prediction accuracy and the generalization performance of the models is well improved.

Self-reported Measurement of the Impact of 360-degree VR Images on Affective and Cognitive Responses in Hotel Promotion

Claudia Rivera, Luis Alfaro, Jorge Luna-Urquizo, Elisa Castaneda
Universidad Nacional de San Agustin de Arequipa Arequipa, Peru

Abstract—This study examines 360° Virtual Reality images for the promotion of hotel services and their impact on cognition and affective dimensions. The study employs self-reported measurement instruments to assess cognitive loads, affective responses, and attitudinal and behavioral intention responses. A boutique hotel was captured and visualized using the developed Experiential-sensory marketing adaptive system model for hospitality based on 360° VR, which allows potential guests to take tours with direct experiences. The results showed a low physical and mental demand, that the affective and behavioral responses are stronger, also evidencing hotel booking intentions, which allows to establish that the usefulness of adding resources with 360° images in hotel promotion are more attractive and influence the booking intention.

Structured Equation Model of Antecedents to Performative Action from Social Media Exchanges

Michael D. Workman
International School of Management, University of Paris, Paris, France

Abstract—Beyond communicating, social media are used in performative actions ranging from targeted marketing to target designation in what is referred to as opensource intelligence. The literature has extensively covered cyber behaviors as people interact with social media. This body of research has established that social media influences people in different ways, but which people and under what conditions? The answers remain unclear. We used an unsupervised machine learning method to identify clusters of people's attributes for those who may be inclined to escalate their actions based on social media conversations. We then used structured equation modeling of participant responses to a questionnaire and observed their actions to triangulate our findings and validate them. Our purpose was to determine if the theoretically determined constructs were observed empirically; that is, to develop and test the theoretical model of the antecedents to determine their predictability of action outcomes.

A Study on Situation Awareness Estimation for Network Security

Jong-Shin Chen, Cheng-Ying Yang, Jenq-Foung JF Yao, Min-Shiang Hwang
Department of Information and Communication Engineering,
Chaoyang University of Technology, Taichung, Taiwan;
Department of Computer Science, University of Taipei, Taipei, Taiwan;
Department of Computer Science, Georgia College and State University, Milledgeville, Georgia, USA;
Department of Computer Science and IE, Asia University, Taichung, Taiwan

Abstract—Situation Awareness (SA) is one of the most concerned research in the field of network security. According to the progressive relationship of subject, it could be divided into three stages, the knowledge of situation awareness, the situation estimation and the situation prediction. Among these stages, the situation prediction is the most difficult and treasured. Prior predicting, the situation estimation has to be established first. Hence, in this paper, the major concerns are the technologies applied towards the model establishment, the estimation methods with Dempster–Shafer theory, according to the expert knowledge and the databases of experience, the estimations use the logical reasoning to estimate the security situations.

Studying the Effects of Social Media Content on Kids Safety and Well-being

Bianca Jones, Marwan Omar
Illinois Institute of Technology, Chicago, Illinois, USA;
Capital Technology University, Maryland, USA

Abstract— The use of social media has become increasingly prevalent among children and adolescents in recent years. While social media use can have positive effects on the well-being of young people, it also poses risks to their safety and mental health. This paper aims to examine the effects of social media content on the safety and well-being of children and adolescents. The related work section provides a comprehensive review of the literature on the impact of social media on well-being, cyberbullying, inappropriate content, and protective factors. The findings suggest that social media use can have both positive and negative effects on the safety and well-being of young people. Cyberbullying and exposure to inappropriate content are major concerns associated with social media use among children and adolescents. However, protective factors such as parental monitoring and social media literacy programs can mitigate these risks. The paper concludes by highlighting the importance of parental involvement, social media literacy, and policies that address cyberbullying and inappropriate content to ensure the safety and well-being of children and adolescents in the digital age.

Application of Machine Learning Classifiers Interfacing Google Colab and SKlearn to Intrusion Detection CSE-CIC IDS2017 Dataset

Morgan L. Smith, Tor A. Kwembe
Department of Mathematics and Statistical Sciences, Jackson State University, Mississippi, USA

Abstract - Network Intrusion Detection System (IDS) refers to the ability to quickly view network data information, identify any patterns of infiltration, and stop any detrimental effects of abnormal intrusion that could destroy the network. In this concept paper, we have suggested an IDS based on supervised stack ensemble machine learning classification techniques to address this problem. The system is capable of detecting a variety of network threats, according to the findings of using the IDS on the CICIDS2018 data from the University of New Brunswick and the Canada Institute of Cybersecurity (CIC). The proposed IDS model is design to run within the Google Colaboratory (Colab) integrated developmental environment (IDE). A dimension reduction technique approach is also incorporated to enhance the predictive accuracy of the chosen machine learning classification techniques from the Sklearn function libraries.

The 27th International Conference on Image Processing, Computer Vision, & Pattern Recognition (IPCV'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-IPCV
https://www.american-cse.org/csce2023/

TDIP: Tunable Deep Image Processing, a Real Time Melt Pool Monitoring Solution

Javid Akhavan, Youmna Mahmoud, Ke Xu, Jiaqi Lyu, Souran Manoochehri
Mechanical Engineering, Stevens Institute of Technology, Hoboken, New Jersey, USA

Abstract - In the current era of Industry 4.0, Additive Manufacturing (AM), particularly metal AM, has emerged as a significant contributor due to its innovative and cost-effective approach to fabricate highly intricate geometries. Despite its potential, this industry still lacks real-time capable process monitoring algorithms. Recent advancements in the field of metal AM suggest that Melt Pool (MP) signatures during the fabrication process contain crucial information about process dynamics and quality. To obtain this information, various sensory approaches, such as cameras-based vision modules with hundreds to millions of frames per second, are employed for online fabrication monitoring. However, many conventional in-depth analyses still cannot process all the recorded data simultaneously. Although conventional Image Processing (ImP)solutions provide a targeted tunable approach, they pose a trade-off between convergence certainty and convergence speed. As a result, conventional methods are not suitable for a dynamically changing application like MP monitoring. Therefore, this article proposes the implementation of a Tunable Deep Image Processing (TDIP) method to address the data-rich monitoring needs in real-time. The proposed model is first trained to replicate an ImP algorithm with tunable features and methodology. The TDIP model is then further improved to account for MP geometries and fabrication quality based on the vision input and process parameters. The TDIP model achieved over 94% estimation accuracy with more than 96% R2 score for quality, geometry, and MP signature estimation and isolation. The TDIP model can process more than 500 images per second, while conventional methods would have taken a few minutes per image. This significant processing time reduction enables the integration of vision-based monitoring in real-time for processes and quality estimation in metal AM, making it a valuable tool for future applications.

Region-based Steganalysis of Medical Radiographs for Radiographic Machine Identification

Farid Ghareh Mohammadi, Ronnie Sebro
Department of Radiology, Mayo Clinic, Jacksonville, Florida, USA;
Center for Augmented Intelligence, Mayo Clinic, Jacksonville, Florida, USA

Abstract—New advances in artificial intelligence (AI) allow us to fake digital images that are difficult for humans to distinguish from real images, including images used in healthcare like radiographs. Malware creating forged digital radiographs have the potential to have severe negative repercussions for patients' diagnosis and treatment, therefore there is a need to validate the radiographs' source used in health-care. We address this challenge and propose a region-based steganalysis algorithm using a deep learning framework that identified the region of radiographs which has the most informative pixels and patterns for determining the radiographs' source. The deep learning algorithm uses a convolutional neural network (CNN) with four convolutional layers with different filters followed by three layers of fully connected convolutional neural network (FCNN). We used radiographs of the knees (n = 1418), legs (n = 616), ankles (n = 1290) and feet (n = 1074) of patients at Mayo Clinic (01/01/2010 - 12/31/2021) and identified the radiographs' source (manufacturer). The dataset was randomly split by patient into training/validation (n = 3635,

80%) and test (n = 763, 20%), and after tuning evaluated using only one radiograph for each patient in the test dataset. The algorithm yields a model prediction performance for a region of feet radiographs with 98.06% accuracy (Area Under the Curve (AUC) = 98.56%). This novel research is the first in medical forensic imaging that identifies the content-free region of radiographs that is most informative to determine the radiographs' source. These results will be invaluable for detection of fake radiographs and scientific fraud.

Revisiting Spectral Clustering Techniques on Images: 2-k-way vs. k-way Partitioning

Laura E, Kaplan, Louis Petingi
Computer Science Department, College of Stanten Island (CUNY), Staten Island, New York, USA

Abstract—In this paper we compare two Spectral Clustering techniques, 2-k-way, and k-way partitioning, obtained by the application of k-means on the eigenvectors of the unnormalized Laplacian of images. Classical implementations of Spectral Clustering rely on algorithms that recommend the use of k eigenvectors as the training set for the multi-partitioning of data into k-clusters. Our empirical testing of cluster and vector combinations reveals that choosing k vectors for k clusters does not necessarily yield best-performance clustering. Rather, it appears that limiting the eigenvector matrix to just two eigenvectors (specifically including the Fiedler eigenvector), while partitioning the data into k-clusters, yields more balanced clustering results.

Synthetic Data Generation: An Evaluation of the Saving Images Pipeline in Unity

Kevin Sankar, Akash Pooransingh, Sarika Ramroop
Faculty of Engineering, Department of Electrical and Computer Engineering,
The University of the West Indies, Trinidad and Tobago;
Virtana TT Ltd, Trinidad and Tobago

Abstract—In this paper, we propose a new framework, CUDA DirectX 11 Interop, for the Saving Images Pipeline that can exceed current performance standards. Unity generates synthetic data to supplement the large amounts of training data required by neural networks. The Saving Images Pipeline saves synthetic data to disk, but is not optimized due to overhead costs. Current (Linear and Coroutines) and alternative (DirectX 11 and CUDA OpenGL Interop) solutions were evaluated in terms of execution times, image resolutions, and JPEG qualities. The DirectX 11 solution exhibited the best performance as it met the 40 ms baseline. Using CUDA Interoperability with the DirectX 11 solution creates a new framework, CUDA DirectX 11 Interop, that exceeds the performance of all solutions examined.

Visual System of Avatar Robot using Eye-tracking and Foveated Rendering

Masato Hosoi, Atsuhiro Hattori, Asuka Nakamura, Hideo Furuhashi
Department of Electrical and Electronics Engineering,
Aichi Institute of Technology, Toyota, Japan

Abstract— Herein, an avatar robot visual system using a headmounted display (HMD) that performs a foveated rendering processing using eye-tracking technology on images obtained from a stereo camera is developed. The distance between the two cameras of the avatar robot's eyes is automatically controlled according to the interpupillary distance of the operator measured by the HMD. Foveated rendering with eye tracking technology reduces the amount of data and selectively presents visual information to the operator. This minimized discomfort and improves immersion during operation. Consequently, the perceptual precision of the surrounding position is enhanced, because visual information outside the focal area is reduced.

Skin Cancer Prediction Using CNN-based Decision Fusion with Dermoscopic Images

Saeka Rahman, Sudip Vhaduri, Asheka Rahman, Miad Faezipour
School of Engineering Technology, Electrical and Computer Engineering Technology,
Purdue University, West Lafayette, Indiana, USA;
Department of Computer and Information Technology,
Purdue University, West Lafayette, Indiana, USA;
Department of Computer Electronics and Graphic Technology,
Central Connecticut State University, New Britain, Connecticut, USA

Abstract—Skin cancer is a type of cancer that develops in the cells of the skin and can affect any part of the body. Earlystage detection is crucial for the effective treatment and curability of this cancer. Skin cancer diagnosis includes visual screening, dermoscopic image analysis, biopsy, and histopathological examination. The accuracies of these diagnostic procedures are highly dependent on the experience of dermatologists. However, the number of experts in this field is limited. Therefore, automatic skin cancer detection from skin lesion images is an increasingly important field that can aid in medical research. This research proposes a Convolutional Neural Network (CNN) architecture-based decision-level fusion model to predict skin cancer from dermoscopic images with improved accuracy. In so doing, the research uses majority voting decision fusion to combine three individual CNN models. The models are validated using the International Skin Imaging Collaboration (ISIC) 2018 challenges dataset available in the ISIC archive to predict binary and multiclass skin cancer with validation accuracy of 0.83±0.01 and 0.91±0.01, respectively. Paired t-tests are conducted to test the difference in mean accuracies between the fused and best performing individual models. The test results show that the fused models' mean accuracy is significantly better at a 5% significance level than the best individual models for both cases. The performance of the binary fused model is further compared to several state-of-the-art models and is found to outperform most of them, suggesting the considerable clinical significance.

Fake Face2Face Video Detection Using a Novel Scene and Texture Based Feature Set

Amit Neil Ramkissoon, Vijayanandh Rajamanickam, Wayne Goodridge
Department of Computing & Information Technology,
The University of the West Indies at St Augustine, St Augustine, Trinidad & Tobago

Abstract— The existence of fake videos is a problem that is challenging today's social media-enabled world. There are many classifications for fake videos with one of the most popular being Face2Face. Detecting such fake videos is a challenging issue. This research attempts to comprehend the characteristics that belong to Face2Face videos. In attempting to understand Face2Face videos this work investigates the characteristics of the video that make them unique. As such this research uses scene and texture detection to develop a unique feature set containing 19 data features which is capable of detecting whether a video is a Face2Face or not. This study validates the feature set using a standard dataset of the features relating to the characteristics of the video. These features are analysed using a classification machine learning model. The results of these experiments are examined using four evaluation methodologies. The analysis reveals positive performance with the use of the ML method and the feature set. From these results, it can be ascertained that using the proposed feature set, a video can be predicted as a Face2Face or not and as such prove the hypothesis that there exists a correlation between the characteristics of a video and its genuineness, i.e., whether or not a video is a Face2Face.

Intelligent Traffic Control System using YOLO Algorithm for Traffic Congested Cities

Afsana Rahman Mou
Department of Computer Science, University of Arkansas at Little Rock, Little Rock, Arkansas, USA

Abstract - The number of vehicles on the road has increased in recent years, and many countries now require reliable and sophisticated control of the traffic signal system. By accurately calculating the traffic density based on the images captured by cameras installed on the traffic posts, this research aims to resolve this dilemma. The technique uses a simple algorithm that first processes a fuzzy controller to recognize vehicles using a Yolo classifier. It could be developed further for hardware implementation using specialized CPUs. Our method entails periodically capturing photos and processing them further with a cascade classifier. The classifier is retained for use in calibration. With the help of this method, the system can deduce the traffic density, which a fuzzy controller then evaluates to decide when to turn on the traffic signals. The output function of the fuzzy controller, which compares the vehicle density of the most recent photos, adjusts the output dynamically.

Machine Learning with Adaptive Image Colorization for Improving Facial Recognition

Michelle Hu, Yen-Hung Hu
Woodside High School and Governor's School for Science and Technology, Newport News, Virignia, USA;
Department of Computer Science, Norfolk State University, Norfolk, Virginia, USA

Abstract—Facial recognition, a deep learning approach to teaching a model to recognize the features of the human face, is one of the most important technologies used in security today, most notably for preventing identity fraud by its implementation in surveillance cameras or Face ID in smartphones. However, a major problem for facial recognition is accounting for lighting conditions, especially abnormal/artificial ones, and this weakness could possibly be exploited to access private information. In order to improve this aspect of facial recognition, the research used image colorization, another deep learning approach that involves recoloring grayscale images. We created an adaptive image colorization generative adversarial network using U-Net and ResNet and a facial recognition model using the Keras API that could identify images of the selected face. Then, we compared the prediction accuracy of the facial recognition model on the normal (normally lit) images, abnormal (abnormally/artificially lit) images, and colorized (abnormal images recolored using image colorization). We found that image colorization was both capable in recoloring images and able to drastically improve the prediction accuracy of abnormal images in a facial recognition model. These findings could be applied to commercial facial recognition models in order to enhance the outputs of such models and therefore increase the security and reliability of the systems that depend on them.

Classification of Functional and Nonfunctional Hand Movement Using Deep Learning and First-Person View Video

Tan Tran, Lin-Ching Chang, Peter S. Lum
Department of Electrical Engineering and Computer Science, The Catholic University of America, Washington, DC, USA;
Department of Biomedical Engineering, The Catholic University of America, Washington, DC, USA;
MedStar National Rehabilitation Network, Washington, DC, USA

Abstract— Upper extremity (UE) rehabilitation of motor function is needed post-stroke. Although UE rehabilitation is widely used in practice, it is difficult to evaluate the effectiveness of UE treatments. There are various approaches to evaluate UE motor function such as performing tasks in the clinician's office, patient's self-report, or wearable devices equipped with accelerometer/gyro sensors to analyze the data captured from stroke survivors. Our previous study reported that using a Random Forest machine learning algorithm with accelerometry sensor data obtained by single wrist-worn sensor, it is possible to differentiate between functional and non-functional UE movements in stroke survivors with good accuracy. However, to build the

model, a third-person view video was captured and manually annotated by experts to provide ground truth labels for training. This task is very time-consuming and subjective. In this study, we propose a framework that uses deep learning (DL) to classify functional and non-functional hand movement in videos captured from a wearable camera (i.e., first-person view videos). The proposed system is fully automated and consists of two DL networks. The first DL network performs hand pose estimation from a 2- second sequence of frames which extract pre-defined key points on the hands. The second DL network then takes these key points as input and classifies the sequence of frames to either functional or non-functional movement. The proposed system has two major impacts on rehabilitation. First, it can provide initial annotations on the frames of videos that can significantly reduce the time required for human annotations. Second, analyzing the effectiveness of UE rehabilitation using the firstperson view videos enables us to measure the outcomes of UE treatments in stroke survivors' homes. In addition, the system can deliver timely feedback to stroke survivors that has the potential to increase the use of the affected UE in the community.

Convolutional Neural Network Post-compression Evaluation with Explainable AI

Recep Erol, Mariofanna Milanova
University of Arkansas at Little Rock, Little Rock, Arkansas, USA

Abstract— This study highlights the limitations of traditional evaluation metrics in providing insight into the decision-making process of machine learning models, especially in post-training and post-compression evaluation. We examine the use of GradCAM to visualize the attention of the model and provide a more detailed understanding of its decision-making process. While accuracy and f1 score provided valuable information about the overall performance of the models, the incorporation of GradCAM helped to identify instances where the model was correctly predicting the class but focusing on the wrong area of the image or vice versa. This research presents the experimental findings of the models' attention maps, which revealed unusual behavior by the models, indicating a generalization issue and suggesting that incorrect decisions may be hierarchically transferred to the pruned model. The results emphasize the need to prioritize posttraining evaluation to ensure the health of the model for deployment.

Strawberry Pests and Disease Recognition with Self-supervised Learning

Aung Si Min Htet, Hyo Jong Lee
Division of Computer Science and Engineering, Jeonbuk National University, Jeonju, Republic of Korea

Abstract—This paper proposes a self-supervised learning approach to improve the recognition accuracy of strawberry crop pests and diseases, which can cause significant economic losses and reduce crop yields. Traditional methods for identifying pests and diseases are time-consuming and inefficient. Self-supervised learning can leverage unlabeled data to learn valuable representations without relying on manual annotations. The proposed two-stage approach involves self-supervised pre-training on the strawberry crop pests and diseases dataset and fine-tuning using a task-specific learning approach. Results show that this approach significantly outperforms the training from scratch and obtains similar results compared to the supervised pre-training approaches, with a highest recognition accuracy of 98.47% using ResNet-50 as the backbone model.

Deep Learning-based Pothole Detection System with Aerial Image

Soobin Jeon, Sujong Kim, Junhong Park, Dongmahn Seo
School of Computer Software, Daegu Catholic University, Gyeongsan, Republic of Korea;
Department of Computer Software, Daegu Catholic University, Gyeongsan, Republic of Korea

Abstract— A pothole is typical road damage with a pot shape of 15mm or more depth. A large pothole requires rapid maintenance because it causes a traffic accident. In this paper, we propose an aerial image-based pothole detection system using drones and deep learning. Since drones are not affected by traffic conditions, they can shoot large areas faster than shooting while driving on the road. The proposed system uses deep learning to detect potholes in drone images. For pothole detection, a deep learning classification model based on the Inception v3 model learns a self-constructed aerial image-based pothole data set. To evaluate the

pothole detection performance of the model, we measure recall and precision using about 5,496 image datasets. The proposed system is implemented as a web service system that can be used in the real environment by utilizing the developed deep learning model. When the implemented service system is applied to actual road maintenance, it is shown that it is possible to detect a pothole in the broader area in a shorter time than the existing pothole detection method and provide information to the road repairer.

Fall Detection in SoccerNet Data

Jonathan Sturdivant, Ethan Lee, Brian Alvarez, Savitha Rachuri, Gulustan Dogan
Department of Computer Science, University of North Carolina Wilmington, North Carolina, USA

Abstract—SoccerNet dataset is a benchmark for action spotting in soccer videos. The action that we are focusing in this paper is fall of the players on the soccer field during a match. Falls are the leading causes of injury; however, research on fall detection has tended to be focused on the elderly and in medical environments. Falls are not a risk that only affects the elderly, in a study of young adults 48% of young adults reported falling at least once in the previous 16 weeks. Of those falls, over a third of them were a result of sporting activity. The SoccerNet dataset fall detection task is to predict whether or not a person has fallen in a given video. The data consists of video footage of soccer matches, and the task is to classify each frame as either containing a fall or not. The SoccerNet dataset is a large-scale dataset for fall detection in soccer. It consists of over 1,000 hours of video footage of soccer matches. The dataset is challenging due to the variability in the appearance of falls, and the fact that falls are often not the focus of the camera. The dataset we are using is the soccernet tracking challenge dataset. This dataset contains 110 different 30-second clips of soccer games recorded from a single camera angle. For our detection, we intend to use pose estimation and player tracking to create a deep learning model.

The Impact of Background Removal on Performance of Neural Networks for Fashion Image Classification and Segmentation

Junhui Liang, Ying Liu, Vladimir Vlassov
Department of Computer Science, KTH Royal Institute of Technology, Stockholm, Sweden;
Norna, Stockholm, Sweden

Abstract—Fashion understanding is a hot topic in computer vision, with many applications having a great business value in the market. Fashion understanding remains a difficult challenge for computer vision due to the immense diversity of garments and various scenes and backgrounds. In this work, we try removing the background from fashion images to boost data quality and increase model performance. Having fashion images of evident persons in fully visible garments, we can utilize Salient Object Detection to achieve the background removal of fashion data to our expectations. A fashion image with the background removed is claimed as the ”rembg” image, contrasting with the original one in the fashion dataset. We conducted extensive comparative experiments with these two types of images on multiple aspects of model training, including model architectures, model initialization, compatibility with other training tricks and data augmentations, and target task types. Our experiments show that background removal can effectively work for fashion data in simple and shallow networks that are not susceptible to overfitting. It can improve model accuracy by up to 5% in the classification on the FashionStyle14 dataset when training models from scratch. However, background removal does not perform well in deep neural networks due to incompatibility with other regularization techniques like batch normalization, pre-trained initialization, and data augmentations introducing randomness. The loss of background pixels invalidates many existing training tricks in the model training, adding the risk of overfitting for deep models.

The Image of Numerals with Calibration on TVM

Ssu-Hsuan Lu
Department of Computer Information and Network Engineering,
Lunghwa University of Science and Technology, Taoyuan , Taiwan

Abstract—Today's open-source deep learning frameworks, such as TensorFlow, PyTorch, TVM, and Caffe, have become widely used tools for deep learning and support hardware accelerators like GPU and TPU. Among these platforms, TVM stands out for its flexibility and compatibility by compiling high-level deep learning frameworks (such as TensorFlow, PyTorch, etc.) to generate optimized low-level code, which can be automatically optimized to improve performance. Although CNN significantly affects image recognition, some images may not be correctly recognized due to problems such as location, orientation, and deformation. Therefore, this paper aims to improve recognition accuracy by normalizing images. This paper focuses on developing an image normalization calibration module on TVM with OpenCV and performing recognition for digital photos. This module is mainly responsible for edge detection and performing normalization adjustments based on the selected preprocessing areas.

Automatic Pill Identification System based on Deep Learning and Image Preprocessing

Eric Ponte, Xavier Amparo, Kuan Huang, Daehan Kwak
Department of Computer Science and Technology, Kean University, Union, New Jersey, USA

Abstract—The pharmaceutical industry annually distributes thousands of different prescription medications. However, the high volume of available medication increases the likelihood of errors during the dispensation process. These errors can result from distractions, incorrectly filed prescriptions, or similarities between pills. This paper proposes an automated system that employs computer vision to identify pills, aiming to enhance pill identification for pharmaceutical workers and consumers and reduce dispensing errors. The proposed approach involves creating a deep learning model using Keras, preprocessing pill images through an image data generator, and leveraging tools such as OpenCV and Paddle OCR to identify critical aspects of a pill, including its shape, color, and imprint. The ultimate goal is to develop a real-time pill identification system that can seamlessly integrate with video cameras, thereby facilitating smoother operations in high-volume medication dispensaries.

Overview of 3D Human Face Reconstruction Techniques and a Novel Approach

Xiaoyuan Suo
Department of Computer and Information Science, Webster University, St. Louis, Missouri, USA

Abstract— Three-dimensional (3D) human face reconstruction methods have gained significant attention in recent years due to their diverse applications in fields such as entertainment, biometrics, healthcare, and cyber security. This survey presents an overview of existing 3D human face reconstruction methods, highlighting their principles, techniques, and advancements. The survey covers various approaches including photometric stereo, structured light, shape-from-shading, and multi-view stereo methods. Each method's strengths, limitations, and potential applications are discussed, providing insights into their suitability for different scenarios. This survey examines the challenges faced by these methods, such as occlusions, variations in lighting conditions, and facial expressions. Promising research directions and emerging trends in 3D human face reconstruction are also explored, including real-time applications, personalized models, cross-domain integration, and ethical considerations. Additionally, this work also presents our preliminary work in a 3D precise human face reconstruction. This work serves as a comprehensive resource for researchers, practitioners, and industry professionals seeking to understand the state-of-the-art in 3D human face reconstruction and its potential for future advancements.

Underwater Image Enhancement Using Fusion-based Deep Neural Networks

Chien-Lin Tsao, Jin-Jang Leou
Department of Computer Science and Information Enginnering, National Chung Cheng University, Taiwan

Abstract—Underwater images are usually degraded images, due to insufficient light, color distortion, water flow shaking, etc. In this study, an underwater image enhancement approach using fusion-based deep neural networks is proposed. Given an input underwater image, five pre-processing techniques are used to generate five pre-processing underwater images. Five feature transformation units are separately used to compute five feature maps of the five pre-processing underwater images. Based on the input and five pre-processing underwater images, five confidence maps for the five pre-processing underwater images are estimated. Finally, five feature maps and five confidence maps for the five pre-processing underwater images are together used to generate the enhanced underwater image. Based on the experimental results obtained in this study, in terms of objective metrics and subjective evaluation, the performance of the proposed approach is better than those of five comparison approaches.

Unsupervised Video Anomaly Detection Using Memory-augmented Deep Neural Networks

Yi-Ying Wang, Jin-Jang Leou
Department of Computer Science and Information Engineering, National Chung Cheng University, Taiwan

Abstract—In this study, an unsupervised video anomaly detection approach using memory-augmented deep neural networks is proposed. The proposed approach contains five main stages: (1) optical flow computing of input video frames; (2) the features of video frames (called query) are extracted by the proposed encoder; (3) in the memory module, the similarities between all memory items and the query are computed by the proposed attention-based structure to generate the new query; (4) video frames and corresponding optical flows are predicted by the decoder; and (5) in terms of PSNRs, the anomaly scores of video frames are computed to determine whether video frames are abnormal. Based on the experimental results obtained in this study, the performance of the proposed approach is better than those of seventeen comparison approaches.

Non-Invasive Muzzle Matching for Cattle Identification using Deep Learning

Arun Sanjel, Bikram Khanal, Pablo Rivas, Greg Speegle
School of Engineering and Computer Science, Department of Computer Science,
Baylor University, Waco, Texas, USA

Abstract—Accurate cattle identification is an essential but complicated issue in the field of livestock management. Traditional identifying methods can involve invasive procedures, posing ethical difficulties and compromising animal welfare. This paper addresses this pressing issue by proposing a deep learning-based methodology for non-invasive cattle identification through muzzle matching. Our approach leverages a comprehensive dataset of 4923 cleaned and cropped muzzle images from 268 distinct cattle breeds. The model showcases exceptional performance with a training accuracy of 98.88% and a test accuracy of 100%. Importantly, our methodology avoids invasive procedures and exhibits adaptability, effectively handling introducing new animals into the system. This versatility ensures the model's reliability across diverse operational scenarios, making it a suitable candidate for insurance fraud prevention and animal trading applications. The paper also highlights critical future research directions, including expanding the dataset to encompass a broader range of cattle breeds and muzzle variations and the potential integration with other identification modalities.

Using Human Body Recognition Method Evaluate Hospital Caregivers Turning Immobile Patients

Chris Cheng Zhang, Chris Ye, Mike Ye, Yu Shen, Ella Zhaoyue Wang
Canada Youth Robotics Club, Canada;
Saint George's School, Vancouver, British Columbia, BC, Canada;
Jiangsu People's Hospital, Nanjing, Jiangsu, China;
Havergal College, Toronto, ON, Canada

Abstract—The application of computer vision techniques in medical technology have resulted in the development of programs that aim to improve the efficiency of hospital processes and workflows. This article entails the development of one such program whose objective to help hospital caregivers in turning their patients to the correct position on their beds. Two particular techniques were used by the researchers to achieve this goal: human recognition to recognize the position of the patient on the bed, and object detection to determine the position of any external factors such as pillows. The researchers developed the application in Python using open-source libraries such as OpenCV and MediaPipe Pose, which was used in tandem with a Raspberry Pi and a mounted camera. TensorFlow Lite, an industry-standard machine learning tool, was used to train the machine learning model for pose classification. Results of the research show that the application is able to determine the correctness of each step of turning the patients to a reasonable degree of accuracy. Further training of the model suggests that this accuracy will increase with each round of subsequent training data. In the future, the researchers hope to supplement the application with additional functionality in the future to better cater to the needs of hospital caregivers.

Gas Leakage Recognition using Manifold Convolutional Neural Networks and Infrared Thermal Images

Omneya Attallah, Amr M. Elhelw
Department of Electronics and Communications Engineering, Arab Academy for Science, Technology & Maritime Transport, Alexandria, Egypt

Abstract - Gas leakage detection in different industrial sectors is enormously important for safe operation. It is vital to quickly and automatically detect and identify the type of gas in order to prevent environmental damage and protect human lives. Existing approaches mainly rely on electronic noses which have several limitations and should be kept within the leakage region. Lately, novel approaches for gas detection have been proposed based on thermal infrared sensors which can capture heat patterns at a distance far from the gas leakage. Motivated by the success of artificial intelligence such as deep learning in several industrial applications. Combining deep learning with infrared thermal images could effectively improve gas leakage detection accuracy. In this study, a deep learning-based pipeline is proposed based on thermal infrared imaging to detect gas leakage and differentiate between gas categories. Multiple convolutional neural networks (CNN) models are used in the proposed pipeline for feature extraction leading to spatial deep features. These features are then analyzed via the fast Walsh Hadamard transform (FWHT). Next, these features are integrated using principal component analysis and then fed to several machine learning classifiers which are used for gas detection. The detection accuracy attained via the proposed pipeline is 98.0% which suggests that the proposed integration method has improved the performance of gas detection.

The 20th International Conference on Modeling, Simulation & Visualization Methods (MSV'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-MSV
https://www.american-cse.org/csce2023/

Recurrence-Plot Visualization and Quantitative Analysis of Long-Term Co-Evolutionary Dynamics in a Simulated Financial Market With ZIP Traders

Dave Cliff
Department of Computer Science, University of Bristol, Bristol, United Kingdom (UK)

Abstract—This paper reports first results and analysis from an accurate agent-based model of a contemporary financial exchange populated by automated trading systems that use a combination of real-time machine learning and longer-term optimization methods in an attempt to continuously improve their trading strategy, each forever trying to maximise their individual profitability. Because the profitability of any one automated trader's strategy is at least partially dependent on the distribution of strategies being played by all the other traders in the market, the system studied here is inherently co-evolutionary. Recent publications have explored the co-evolutionary dynamics of such markets populated with minimally adaptive "ZeroIntelligence" (ZI) trader-agents, a style of modelling with a long track-record in computational economics. This paper's novel contribution is that it presents the first ever exploration of the long-term co-evolutionary dynamics of such markets populated by co-evolving traders each running the Zero-Intelligence-Plus (ZIP) trading strategy, which was one of the first two trading algorithms demonstrated to consistently outperform human traders. Visualizations, via Recurrence Plots (RPs), are shown of the co-evolutionary dynamics of the ZIP-trader markets over hundreds of days of continuous second-by-second trading, which demonstrate that even when populated by super-human trading strategies, co-evolutionary markets can trace trajectories through strategy space that loop back on themselves, such that after many days or weeks or months of continuous improvement to the strategy, it ends up back where it started. Recurrence Quantitative Analysis (RQA) methods are then used on the RPs to quantify the frequency of occurrence of such strategy-loops. Source code used in the experiments reported here is freely available under the MIT Open Source License, on GitHub.

SNEE: Social Network Evolution Engine

Patrick Shepherd, Anh Ngo, Said Maalim, Isaac Gray
Department of Computer Science, Berea College, Berea, Kentucky, USA

Abstract—In the fields of opinion dynamics and epidemiology, as well as others that study the evolution of networks over time in general, there are few tools available that simultaneously model and visualize a network in real time, and most of those available are either very narrow in scope or are inaccessible to nontechnical users. We introduce SNEE, an evolution/visualization engine capable of emulating most models of opinion diffusion and epidemiological models in use today, open to a wide scope of different parameterizations for fine-grained customization. Our software supports static and dynamic network modeling, and provides for easy data harvesting. This tool is meant to be accessible to a broad audience, and has immense applicability in research and the classroom.

Exploration of Single Compound Phase Diagrams in Chemistry Courses using MATLab or Octave

Madison West, Jaime Diaz, Zenon "Theo" Porche, Brittnee V Veldman
California State University Channel Islands, Camarillo, California, USA

Abstract— Exploration-based class activities have been shown to help students build a deeper connection to the course content and be more engaged than the use of lecture and problem sets alone. High-impact practices that help to build multiple skill sets across disciplines and offer students a chance to exercise agency in their education have been shown to contribute to in depth learning. In this work the coauthors, two students with a background in MATLab and a chemistry instructor, collaboratively developed an interactive activity for students that blends an introduction to programming and computation with physical chemistry content, in this case pure compound phase diagrams. The reported activity helped students to create a better understanding of how computational software such as MATLAb and Octave could aid in visualizing physical phenomena in chemical systems. Literature values of six physical properties of water were used to generate a full phase diagram as an example exercise. The resulting diagram had a maximum error of less than 20% in the calculated output over a range of ten decades (from 1 Pa to 10 GPa). Reflections on the process, educational outcomes and activity from the students are included.

Evaluation of Using Virtual Reality for Applied Behavioral Intervention

Hana Alarifi, Hesham Aldhalaan, Laura Barcelos Nomicos, Linda Hayes
Center For Autism Research, King Faisal Specialists Hospital & Research Center, Riyadh, Saudi Arabia;
University of Nevada, Reno, Nevada, USA

Abstract - In recent research, Virtual Reality has become well recognized method of delivering therapeutic and educational services. Current technology also provides bi-metric objective evidence of attention such as eye-tracking and Electrdermal Activity that can inform therapist and educators.A number of relevant papers of VR usability were analyzed to investigate the effectiveness of VR interventions. The study aims to explore Applied Behavior Analysis ABA specific VR interventions potency.It is concluded that VR is an efficient, accessible and successful method of providing ABA interventions.
.

Artificial Virtual Reality Simulation Design for Children on Autism Spectrum Disorder

Abeer Al-Nafjan, Hana Alarifi, Neehal Almuways, Aliah Alhameed, Renad Al Hussain
Computer Science Department, Imam Mohammad Ibn Saud Islamic University, Riyadh, Saudi Arabia;
Autism Center, King Faisal Specialists Hospital and Research Center, Riyadh, Saudi Arabia

Abstract— Autism spectrum disorder (ASD) is a neurodevelopmental disorder that begins early in childhood. Children with ASD mainly associated with social interaction and communication deficit in addition to behavioral problems. Artificially designed environments are useful to expose ASD individuals to various tasks individualized to target specific goals and objectives. Virtual reality (VR) is an interactive computer-generated experience that stimulate feelings and elicit behaviors. This project aims to design and develop a VR-based application that can be used to improve autistic children's understanding and communication skills and help them to connect with society and the surrounding environment. The results shown are only preliminary but demonstrate the potential of embedding VR technology in autism therapy.

Using Time Series Motifs to Explore the Parameter Space of Agent-Based Models: A Pilot Study

Maria Tomasso, Apan Qasem
Department of Computer Science, Texas State University, San Marcos, Texas, USA

Abstract—Agent-based modeling (ABM) has seen a resurgence of interest in recent years with the continued expansion of computational capabilities, allowing researchers to study the systemic effects of individual-level behavior on increasingly larger and more complex systems. As ABMs become mainstream tools in a variety of disciplines, a need has emerged for better techniques to assess the sensitivity of ABMs and identify the critical regions of the input space where small changes in the input variables result in large qualitative changes in the model's behavior. This work attempts to address this need by characterizing the behavior of ABMs in a two-step process. First, we use time series clustering to identify the behavioral motifs present in a model. Second, we use classification to find and describe the relationship between the motifs and the input space. This feasibility study successfully mapped the input space to the time series motifs for the Schelling model and identified further directions of inquiry towards a generalized method for more complex ABM characterization.

Vitrification and Crystallization of Ti and Ti2Ni Alloy Studied by Computer Simulation

Dmitri V. Louzguine-Luzgin
WPI Advanced Institute for Materials Research, Tohoku University, Sendai, Japan;
Mathematics for Advanced Materials-OIL,
National Institute of Advanced Industrial Science and Technology (AIST), Sendai, Japan

Abstract - The present work is devoted to the atomic structure variations on cooling, vitrification and crystallization processes in liquid Ti and Ti2Ni alloy studied by molecular dynamics simulation. It is performed using a classical molecular-dynamics computer simulation procedure with the embedded-atom method potentials at constant pressure on continuous cooling and isothermal annealing. The structural changes are monitored with direct structure observation in the simulation cells containing 128,000 atoms. The crystallization kinetics is analyzed by monitoring density and energy variation as a function of time. A common-neighbor cluster analysis is performed. The thermodynamic calculations were performed in order to estimate the energy barrier for crystal nucleation and critical nucleus size. The differences in structure and crystallization are discussed in comparison with other metals.

Interaction Petri Nets for Modeling Multiagent Interactions

Nora Aldahash, Steve King, Abir Binabid
University of York, United Kingdom;
King Saud University, Saudi Arabia

Abstract—Petri nets have been closely related to multiagent systems (MAS) and used as a tool for modeling agents and their behavior for specification, verification, and validation purposes. Various models of agent activities, plans, and interactions show the ability of Petri nets to represent dynamic and concurrent MAS. Insight from previous Petri net models of agent interaction led to the proposal of a Colored Petri net model for monitoring the state of communication. This paper presents an interaction Colored Petri net for modeling message-based communication ruled by Interaction Protocols (IP). While previous works focused on modeling a single protocol, the proposed Interaction Petri Net (IPN) allows specifying multiple protocols within one model. Moreover, IPN allows the detection of common communication errors such as: lost messages, delay, busy-wait and agent termination. In IPN, an interaction protocol is specified as a Petri net marking where tokens are message instances. This approach enables the capture of the status of each IP in real-time for error detection where two possible approaches have been examined. The proposed model has been validated through a comparison analysis with a previous Petri net model of interaction.

Impact of COVID-19 on Education: Virtual Class Experience

Akshay Monga
Harrisburg University of Science and Technology, Pittsburgh, Pennsylvania, USA

Abstract— During the COVID-19 pandemic, schools and colleges across the world shut down and education was transitioned from an in-person medium to an online medium. This transition was not planned, and the implementation was not perfect. This paper looks at data collected from students in India in July 2020, in the middle of the pandemic. The survey respondents included elementary school, middle school, high school, and university students. This data is used to highlight the variables that impact the rating of an online or virtual class using statistical analysis such as Discriminant Analysis and Student's t-test. The statistically significant variables can then be used to tweak future online classes to get maximum knowledge transfer and student engagement. Discriminant analysis shows that the null hypothesis cannot be rejected, and the variables do not have a statistically significant impact on the quality ratings of the class. Student's t-tests show that there is no statistically significant impact on the quality ratings of the class and the time spent on self-study or time spent on sleep.

Design Optimization of Lithium-ion Battery Cell for Temperature Uniformity

Jaewook An, Jun Lee, Han-Seung Kim, Horim Lee, Minjae Kim, Chang-Wan Kim
Graduate School of Mechanical Design and Production Engineering,
Konkuk University, Republic of Korea;
School of Mechanical Engineering, Konkuk University, Republic of Korea

Abstract— The non-uniform temperature distribution of lithium-ion batteries (LIBs) accelerates battery deterioration and causes thermal runaway, so the improvement of the temperature uniformity of LIB cells is required. First, 3-dimensional electrochemical-thermal coupled analysis was performed on a 55Ah LFP pouch cell to calculate the heat generations and temperature due to the electrochemical-thermal reaction of the LIB cell. Second, temperature difference between the maximum and the minimum temperature was minimized through optimization method. As a result of the optimization, the temperature difference of the LIB cell was reduced by 7.77 °C compared to the initial design.

The 29th International Conference on Parallel & Distributed Processing Techniques & Applications (PDPTA'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-PDPTA
https://www.american-cse.org/csce2023/

Fault-tolerant Routing Methods in Bicubes

Yitong Wang, Htoo Htoo Sandi Kyaw, Keiichi Kaneko
Graduate School of Engineering, Tokyo University of Agriculture and Technology, Tokyo, Japan

Abstract—In this study, we propose three fault-tolerant routing methods in bicubes [1]. To begin with, we design a simple fault-tolerant routing method based on the shortest-path routing algorithm proposed by Okada and Kaneko [2]. We regard this method as the baseline. First, we devised a method, Method 1, in which each node with a message switches to the dept-first search if the distance from the node to the destination node is less than or equal to 3. Next, we proposed a method, Method 2, in which each node with a message excludes the previous node, that is, the node that forwarded the message to the current node, when it selects one of its non-faulty neighbor nodes to forward the message. Finally, we combine the methods, Method 1 and Method 2, to obtain the third one, Method 3. By using the simple fault-tolerant routing method as the baseline, we conducted a computer experiment in 11-, 12-, and 13-dimensional bicubes, BQ11, BQ12, and BQ13, to evaluate the three methods. According to the experimental result, Method 3 showed the best performance and Method 2 did the second best. Method 1 showed slightly better performance than the simple fault-tolerant routing method.

Distributed Tire Parameter Optimization Using CREATE-GV

Jeremy Mange, Jacob Brendle
US Army, Ground Vehicle Systems Center, DEVCOM GVSC - Analytical Methods, Warren, Michigan, USA

Abstract—Automated optimization of vehicle components according to a set of criteria is desirable in a number of contexts. Modeling and simulation can enable this process; however, it is often computationally intensive, particularly when a large number of potential designs are involved. In this paper, we present a project using high-performance computing (HPC) resources to perform distributed optimization of tire parameters through the CREATE-GV mobility analysis software. We demonstrate that this highly parallel computing process can be used effectively to analyze large sets of model alternatives in order to optimize component design for specific vehicle mobility goals.

Multi-threaded Space Carving for 3D Seed Reconstruction

Mitchell L. Neilsen, Kai Zhao
Computer Science Department, Kansas State University, Manhattan, Kansas, USA

Abstract – This paper presents a novel framework for 3D seed reconstruction and volume measurement using multi-threaded space carving. The new framework achieves better performance than previous approaches with a relatively low cost. On modest hardware, the framework can compute seed volume within 0.5 seconds, but on a more modern machine with multiple cores, a new multi-threaded version can obtain a solution in 0.03 seconds and achieve nearly linear speedup for up to 6 threads. The new framework provides a practical, cost effective solution to accurately obtain seed volume within 1-2% even with very modest hardware.

A Survey on the Proposed Architectures for Efficient Execution of Irregular Applications Using Pipeline Parallelism

Haleh Khojasteh, Hirad Tabatabaei
Department of Computer Science, Bridgewater State University, Bridgewater, Massachusetts, USA;
Department of Electrical & Computer Engineering, Northeastern University, Boston, Massachusetts, USA

Abstract—Irregular applications, which typically involve data structures such as graphs, trees and sparse matrices, present unpredictable memory and network access patterns. The general-purpose processors cannot accommodate the irregular memory access patterns often found in these applications and algorithms sufficiently. We address this challenge by conducting a study of microarchitecture designs aiming to improve the acceleration of irregular applications. This work has an emphasis on decoupled architectures in which applications can be structured as a pipeline of stages decoupled by queues. It also explores the solutions that introduce context switch and memory coalescing methods.

Techniques for Achieving High Performance in Deep Learning Based Systems for Selective Filtering of Live Video Streams

Azhar Talha Syed, Shikharesh Majumdar
Department of Systems and Computer Engineering, Carleton University, Ottawa, Canada

Abstract - The research focuses on applications of parallel and distributed processing techniques and deep learning for filtering frames from live video streaming data based on a set of criteria specified by the user. With various video streaming sources around the world, storing the entire video streams and searching the stored data is computationally expensive. A video stream filtering system keeps the frames of interest and discards the rest for fast search and retrieval as well as reduction of storage space. This research concerns the devising of filtering techniques for streaming video data. Parallel and distributed processing techniques including frameworks like Apache Kafka integrated with TensorFlow are used to devise the proposed scalable real-time filtering prototypes. Initial research results that include experiments with live video streams from webcams were conducted on Amazon EC2 cloud to gain insights into system performance including the reduction in processing time that accrues from applying the proposed techniques.

FLT: Robust Federated Learning with Trust Model to Defend Against Poisoning Attack

Yung-Ting Chuang, Ming-Han Chung
Institute of Information Management, National Yang Ming Chiao Tung University, Hsinchu, Taiwan

Abstract— The topic of safeguarding personal information has gained more significance lately, as a growing number of users worry about potential privacy breaches and opt not to store their data in the cloud. This reluctance to share private information makes it challenging for developers to create applications that rely on such data. Federated Learning is a technology which is developed to address the above problems. However, Federated Learning is threatened by malicious clients who may launch poisoning attacks that disrupt the aggregation of the global model, which affects its accuracy. Therefore, this paper proposes a defense method called FLTrueTrust which uses a trust model mechanism to resist the behavior of malicious attackers, where it aims to significantly reducing the chances of selecting malicious attackers and thereby improving the efficiency in global model training.

Application Aware Green Datacenter Networks

Ishanvi Kommula
Monta Vista High, Cupertino, California, USA

Abstract – Datacenter networks, while fault-tolerant, are also large, overprovisioned, power hungry, and run at full speed all the time. These networks are architected for supporting the maximum throughput required from all servers in that datacenter. However, even well managed networks worldwide do not have the ability to scale up or down dynamically. This paper explores a way to reconfigure the network topology based on combined bandwidth demand from all the active applications. The proposed solution collects network telemetry data from all servers in a datacenter and uses a Vector Autoregression (VAR) based machine learning model to predict future network throughput. These predictions are then used to reconfigure the network by powering on and off various elements, including switches, routers, and links. The proposal furthermore considers network resiliency when rebalancing the network. This paper also analyzes possible power savings in a test network and concludes by demonstrating the process by which a network administrator might be able to satisfy their needs for performance and fault tolerance, while minimizing the network energy consumption.

An Incremental Many and Multi-Core Adoption of the Mathematics Behind the FFT and its Benefits

Dakota Wilson, Eduardo Colmenares
Department of Computer Science, Midwestern State University, Wichita Falls, Texas, USA

Abstract— Graphic Processing Units (GPUs) have been dominating the floating-point race since 2003 and they have become the preferred architecture for computationally intensive task that exhibit high data parallelism. For the new generation of scientist, not only is imperative to adopt them, but also to properly learn how to harness the potential of these accelerators in order to reach better performance than the one achievable through sequential and parallel multi-core (CPU) approaches. This research adopts, and expands the study of a multidisciplinary computationally intensive kernel, a Radix-2 Fast Fourier Transform (FFT). The proposed topic (FFT) is analyzed from four different perspectives in order to highlight the fact that good programming practices can also offer an additional gain in performance.

GLG: Visual Language for the Development of Parallel Programs in CUDA

Rodrigo Rivera-Ceron, Jose Luis Quiroz-Fabian, Miguel Alfonso Castro-Garcia,
Graciela Roman-Alonso, Manuel Aguilar-Cornejo
Department of Electrical Engineering, Universidad Autonoma Metropolitana, Mexico City, Mexico

Abstract—Nowadays, parallel program development is paving path for providing an efficient solution to scientific problems. Particularly, the use of graphic cards has gained popularity since a large number of cores are available at an affordable price. However, the development of graphic card programs could be a challenge since the software developers must know the card architecture for its implementation. This paper proposes the GLG (Graphic Language for GPUs) visual language for developing programs in CUDA (Compute Unified Device Architecture). GLG programs could be created easily through icons interconnections validated by a grammar. In addition, a web GLG development environment is proposed that uses a client-server architecture to perform the execution of visual programs.

WORKSHOP:
International Workshop on Mathematical Modeling and Problem Solving (MPS)

Co-Chairs: Prof. Masahito Ohue, Dr. Kotoyu Sasayama**, Prof. Kazuki Joe***,*
*Prof. Masami Takata***, Prof. Shinya Watanabe*****
** Tokyo Institute of Technology, Japan*
*** Mitsubishi Electric Building Solutions Co., Japan*
**** chair of Nara Women's University, Japan*
***** Muroran Institute of Technology, Japan*

Faster Lead Optimization Mapper Algorithm for Large-Scale Relative Free Energy Perturbation

Kairi Furui, Masahito Ohue
Tokyo Institute of Technology, Japan

Abstract—In recent years, free energy perturbation (FEP) calculations have garnered increasing attention as tools to support drug discovery. The lead optimization mapper (Lomap) was proposed as an algorithm to calculate the relative free energy between ligands efficiently. However, Lomap requires checking whether each edge in the FEP graph is removable, which necessitates checking the constraints for all edges. Consequently, conventional Lomap requires significant computation time, at least several hours for cases involving hundreds of compounds, and is impractical for cases with more than tens of thousands of edges. In this study, we aimed to reduce the computational cost of Lomap to enable the construction of FEP graphs for hundreds of compounds. We can reduce the overall number of constraint checks required from an amount dependent on the number of edges to one dependent on the number of nodes by using the chunk check process to check the constraints for as many edges as possible simultaneously. Moreover, the output graph is the same as that obtained using the conventional Lomap, enabling direct replacement of the original one with our method. With our improvement, the execution was tens to hundreds of times faster than that of the original Lomap.

Antibody Complementarity-Determining Region Sequence Design using AlphaFold2 and Binding Affinity Prediction Model

Takafumi Ueki, Masahito Ohue
Tokyo Institute of Technology, Japan

Abstract—Affinity maturation in the immune response is limited in terms of affinity gain, and natural antibodies often do not have the binding affinity required for therapeutic applications. Therefore, improving the binding affinity of antibodies is essential for developing antibody-based therapeutics. Designing antibodies using experimental methods is expensive in terms of cost and time owing to the large range of complementarity-determining regions to be explored. Recently, computational methods have been developed as low-cost and fast means of designing and redesigning antibodies. This study evaluated the design performance of AlphaFold2 and the binder hallucination, which can predict protein 3D structures with high accuracy even without experimental antibodies, by redesigning antibody sequences to improve the binding affinity of existing antigen-antibody complexes. Therefore, antibody sequences with higher affinity can be designed for antigen-antibody complexes not included in the training data of AlphaFold2, indicating that the proposed method may be effective as an antibody design method.

Enhancing the Performance of AlphaFold through Modified Storage Method and Optimization of hhblits on TSUBAME3.0 Supercomputer

Hayato Fujita, Akihiro Nomura, Toshio Endo, Masakazu Sekijima
Tokyo Institute of Technology, Japan

Abstract—Knowledge of the three-dimensional structures of proteins is vital for understanding their functions. This knowledge can also serve as a basis for understanding the various functions of the human body. The three-dimensional structures of proteins have been experimentally ascertained from the known amino acid sequence of proteins. To reduce the time and cost of experiments, computer-based methods such as AlphaFold have been proposed. AlphaFold uses existing tools such as HHblits to obtain multiple sequence alignments (MSAs) from huge G protein (sequence) databases, such as BFD. However, HHblits requires a long execution time, as it has to perform many I/O operations. The execution time of HHblits differs significantly depending on the type of storage used to store the protein (sequence) database. Notably, the execution time of HHblits when using TSUBAME3.0 Lustre storage area with default settings differs significantly from that when storage with stripe settings is used. Therefore, we modified the storage method of the protein (sequence) database that can be selected on TSUBAME3.0 and measured the execution time of HHblits based on the storage method selected. Furthermore, we profiled the various bottlenecks of HHblits, and based on the results, tuned the number of parallel processes, modified the database arrangement, and optimized sorting. Furthermore, we made the tool execution asynchronous due to the data dependency between the MSA acquisition tools of AlphaFold. Consequently, we succeeded in shortening the execution time by half, on average, when predicting a three-dimensional structure from a single amino acid sequence on TSUBAME3.0. The modifications proposed in this study to accelerate the process have already been pull requested at the following URLs: HHblits (https://github.com/soedinglab/hh-suite/pull/307) and AlphaFold (https://github.com/deepmind/alphafold/ pull/399).

Predicting Chemical Reaction Product by Graph Transformer

Shunya Makino, Nobuaki Yasuo, Masakazu Sekijima
Tokyo Institute of Technology, Japan

Abstract—The substantial cost and prolonged duration associated with drug discovery present significant challenges within the pharmaceutical industry. A potential solution to alleviate these issues is molecular translation, a process where an input molecule is converted into another molecule following specific rules or conditions. In this study, we introduce a novel machine learning technique for predicting chemical reaction outcomes: the molecular graph translation method, employing a transformer architecture designed for efficient chemical reaction product prediction. We assessed the performance of our proposed method using the USPTO dataset for chemical reaction product prediction. Our results demonstrated an accuracy of 86.5%. Although this accuracy was slightly lower than that of the Molecular Transformer (a string-based method using a similar architecture), our approach surpassed both the Seq2Seq and other graph-based methods in terms of performance.

Evolutionary Multi-mode Slime Mould Optimization: A hyper-heuristic algorithm inspired by slime mould foraging behaviors

Rui Zhong, Enzhi Zhang, Masaharu Munetomo
Hokkaido University, Japan

Abstract—In this paper, we propose a novel hyper-heuristics algorithm named Evolutionary Multi-mode Slime mould Optimization (EMSMO) inspired by the slime mould foraging behaviors for continuous optimization. The EMSMO is a combination of High-Level and Low-Level components. In the LowLevel component, we formulate the foraging behaviors of slime mould and design four generation strategies: Search for food, Approach food, Wrap food, and Re-initialization. In the HighLevel component, we design an improvement-based probabilistic selection function containing the probability of improvement (P I) and the normalized improvement (NI) to determine the sequence of heuristics. Considering that a certain strategy may have different performances in various optimization stages, we adopt a sliding window to emphasize the recent

optimization information to the construction of the optimization sequence. To evaluate the performance of our proposal, we run EMSMO with 6 classic or advanced Evolutionary Algorithms (EAs) on CEC2017 benchmark functions with 30 independent trial runs and compare them fairly. The statistical results show that our proposed EMSMO is competitive with compared methods and has great potential to deal with continuous optimization problems.

A Large Neighborhood Local Search Method Incorporating MIP Techniques for Large-Scale Optimization Problems with many Constraints: Application to the machining scheduling

Jin Matsuzaki, Kazutoshi Sakakibara, Masaki Nakamura, Shinya Watanabe
Toyama Prefectural University, Japan

Abstract—This paper addresses the problem of scheduling machining operations in a highly automated manufacturing environment, while considering the work styles of workers. In actual manufacturing, many issues must be considered, for example, constraints related to the works to be machined in the machining schedule and workers' states. To derive good solutions for such a large-scale problem with many constraints in a realistic amount of computing time, we develop an optimization technique based on an MIP-based large neighborhood local search method for the machining scheduling problem. Then, computer experiments are performed for a problem based on actual machining requirements to verify the validity of the proposed method.

Formal Verification of an Autonomous Vehicle Control System by the Timed OTS/CafeOBJ Method

Masaki Nakamura, Tatsuya Igarashi, Yifan Wang, Kazutoshi Sakakibara
Toyama Prefectural University, Japan

Abstract—To develop a safe and efficient self-driving environment, we need to ensure the safety not only of individual vehicles, but also of groups of vehicles. In this paper, we investigate a method to describe the formal specification of a control system for groups of autonomous vehicle based on the OTS/CafeOBJ method.

Formal Verification of an Autonomous Vehicle Control System by the Timed OTS/CafeOBJ Method

Masaki Nakamura, Tatsuya Igarashi, Yifan Wang, Kazutoshi Sakakibara
Toyama Prefectural University, Japan

Abstract—The Lim-Jeong-Park-Lee protocol (LJPL protocol) has been proposed as an efficient distributed mutual exclusion algorithm for intersection traffic control, which manages a queue of vehicles to enter the intersection. In our previous work, we have verified its safety property by using the timed OTS/CafeOBJ method in a macro level. In this study, we model an autonomous vehicle control system of the queue as a micro level of LJPL protocol, where a vehicle can enter the intersection even if the vehicle in front is still crossing it under some time constraints. We describe a formal specification of the model and verify its safety property by the timed OTS/CafeOBJ method.

Proposed Preprocessing for High-performance Complex Singular Value Decomposition

Miho Chiyonobu, Masami Takata, Jun Harayama, Kinji Kimura, Yoshimasa Nakamura
Nara Women's University, Japan

Abstract—The computation of singular value decomposition by the two–sided Jacobi method has the advantage of obtaining singular vectors fast and accurately. Prior research has achieved fast and accurate implementations of the two–sided Jacobi method for both real and complex matrices. One problem, however, is the large amount of operations and long execution time for complex matrices. To solve this problem, as a preprocessing, a singular value decomposition of a complex matrix is converted into a singular value decomposition of a real upper bidiagonal matrix. In this paper, we implement the preprocessing and confirm that it is faster than the direct treatment of complex matrices from the two–sided Jacobi method.

Iterative Censoring and Highly Efficient Machine Learning with Condition Number in LSMR Method

Miho Chiyonobu, Itsuki Kuboi, Rika Tanaka, Shinya Ozawa, Yosuke Hosoda,
Masami Takata, Kinji Kimura, Yoshimasa Nakamura
Nara Women's University, Japan

Abstract—The least squares minimal residual method, which is an iterative solution method, is adopted for an ill-conditioned linear equation with a large–dimensional sparse matrix as a coefficient. In this study, we address a blurred-image-sharpening problem, which contains noise. Thus, the iterative computation should be stopped after an appropriate number of iterations to reduce the effect of noise. The condition number in the bidiagonal matrix can be used as an iterative stopping rule. The optimal condition number is machine-learned from the condition number generated at each iteration. For machine learning, we propose the application of singular value computation as a preprocessing method for the training data.

Development of Stock Recommendation Method

Natsu Kidoguchi, Miho Chiyonobu, Masami Takata
Nara Women's University, Japan

Abstract—In this paper, we propose a method for recommending appropriate combinations of stocks based on the waveforms of stock price changes. To invest in stocks while reducing risk, specialized knowledge is required. To solve the problem, a method for stock recommendation with different characteristics is desired. Stock price movements can be captured as a waveform. Dynamic Time Warping (DTW), cross-correlation functions, and fast Fourier transforms are used to compare the features of the waveforms. Combining stocks with different waveforms avoids the risk of simultaneous crashes.

Prediction of Specific Surface Area of Metal-Organic Frameworks by Graph Kernels

Yu Morikawa, Kilho Shin, Hiroaki Ohshima, Masataka Kubouchi
Computer Centre, Gakushuin University, Japan

Abstract—Metal-organic frameworks (MOFs) are coordination networks of metal ions or clusters bonded by organic ligands. They are expected to have a wide range of applications due to their tunable micro-porosity and unprecedented surface area. However, the diverse combinations of metals and organic ligands make it difficult to identify optimal combinations. This research aims to propose a machine learning-based method for searching candidates of optimal combinations without physically synthesizing MOFs, thus improving the effectiveness and efficiency of MOF development. To incorporate the structural features

of organic ligands into the investigation, this paper utilizes graph kernels. Graphs of atoms are used to faithfully represent the structural features, and graph kernels evaluate the similarity between these graphs. Additionally, a unified kernel is defined to assess the similarity between MOFs. This unified kernel consists of a weighted sum of an RBF kernel to evaluate the similarity of metals and a normalized graph kernel to evaluate the similarity of organic ligands. Using this unified kernel and the support vector machine classification algorithm, a model is developed to predict whether a given combination of a metal and an organic ligand can result in an MOF with a high specific surface area. Through experiments using MOF data from the CoRE MOF dataset, it has been verified that the models constructed with six out of the seven tested graph kernels demonstrate good prediction performance in terms of the AUC-ROC score.

A Proposal for Yield Prediction using Multiple Models based on the Average Number of Days to Harvest Tomato Fruit in Greenhouse Horticulture

Yusei Yoshida, Shinya Watanabe, Yosuke Kobayashi, Kazuhiko Sato, Satoshi Kondo, Tatsuro Horie
Muroran Institute of Technology, Japan

Abstract—Agricultural corporations that operate farms engage in contract transactions with supermarkets on a two-week to one-month basis. Therefore, it is critical for farmers to predict yield accurately two weeks ahead. In the experimental field of this study, facility-based cultivation enables some control over the growing environment, such as temperature and irrigation. However, uncontrollable factors, such as weather conditions, individual differences of plants, and variability in growing locations, make it difficult to predict yields with high accuracy. To predict yield with high accuracy under these circumstances, we propose a predictive method using multiple different machine learning models based on the average number of harvest days for different environments. We verified the accuracy of this proposed method in actual greenhouse tomato farming.

Meta Generative Data Augmentation Optimization

Enzhi Zhang, Bochen Dong, Mohamed Wahib, Rui Zhong, Masaharu Munetomo
Hokkaido University, Japan

Abstract – N/A

Data Augmentation Method for Improving Blurred Image Recognition Rate

Shiori Ishikawa, Miho Chiyonobu, Sayaka Iida, Masami Takata
Nara Women's University, Japan

Abstract—This paper aims to improve the recognition rate of a recognizer for detecting cracks on concrete surfaces from image data. When developing a recognizer, it is common to use a clean image taken from the front as training data. Therefore, when blurred images that were not taken cleanly from the front are used as test data, the recognition rate decreases. To improve the recognition rate, the training data is mixed with processed images. Experimental results show that adding blurred images to the training data improves the recognition rate for blurred images, even when the images are blurred in different ways.

The 22nd International Conference on Security & Management (SAM'23: July 24-27, 2023; Las Vegas, USA)

http://sam.udmercy.edu/sam23
https://www.american-cse.org/csce2023/

Evaluating a Planning Product for Active Cyberdefense and Cyberdeception

Justin J. Green, Robert E. Bixler, Sasha K. Drew, Neil C. Rowe, Charles W. Heinen, Armon C. Barton
U.S. Naval Postgraduate School Monterey, California, USA;
SoarTech, Inc., Ann Arbor, Michigan, USA

Abstract—Active defenses are increasingly used for cybersecurity, as they add lines of defense after access controls and system monitoring. This work tested a commercial tool CCAT from Soar Technology Inc. for planning active defenses including cyberdeceptions. The tool simulated a formal game between attacker and defender, offering options for both sides with associated costs and benefits. Both sides play a series of random games in advance to learn the best methods in a range of adversarial situations. Our experiments then tested the best active-defense responses to six new kinds of cyberattacks, and showed our defensive methods effectively impeded them. Overall, the product worked well in developing realistic attack scenarios, but required work to set up, and the attack plans were quite predictable.

ExFake: Towards an Explainable Fake News Detection Based on Content and Social Context Information

Sabrine Amri, Henri-Cedric Mputu Boleilanga, Esma Aimeur
DIRO, University of Montreal, Montreal, Canada

Abstract—ExFake is an explainable fake news detection system based on content and context-level information. It is concerned with the veracity analysis of online posts based on their content, social context (i.e., online users' credibility and historical behaviour), and data coming from trusted entities such as factchecking websites and named entities. Unlike state-of-the-art systems, an Explainable AI (XAI) assistant is also adopted to help online social networks (OSN) users develop good reflexes when faced with any doubted information that spreads on social networks. The trustworthiness of OSN users is also addressed by assigning a credibility score to OSN users, as OSN users are one of the main culprits for spreading fake news. Experimental analysis on a real-world dataset demonstrates that ExFake significantly outperforms other baseline methods for fake news detection.

Cybersecurity in Industrial Networks: Overview of Artificial Intelligence Techniques Applied to Intrusion Detection Systems

Ralf Luis De Moura, Virginia N. L. Franqueira, Gustavo Pessin
Operational Tecnology Architecture, Vale S.A, Brazil;
School of Computing and Institute of Cyber Security for Society (iCSS),
University of Kent, United Kingdom; Instituto Tec Vale, Brazil

Abstract—Industrial control systems (ICS) operate on serial-based networks which lack proper security safeguards by design. They are also becoming more integrated to corporate networks, creating new vulnerabilities which expose ICS networks to

increasing levels of risk with potentially significant impact. Despite those risks, only a few mechanisms have been suggested and are available in practice as cybersecurity safeguards for the ICS network layer, maybe because they might not be commercially viable. Intrusion detection systems (IDS) are typically deployed in the corporate networks to protect against attacks since they are based on TCP/IP. However, IDS are not used in serial-based ICS networks yet. This study examines and compares modern Artificial Intelligence (AI) techniques applied in IDS that are potentially useful for serial-based ICS networks. The results showed that current AI-based IDS methods are viable in such networks. A mix of AI techniques would be the best way forward to detect known attacks via rules and novel attacks, not previously mapped, via supervised and unsupervised techniques. Despite these strategies' limited use in serial-based networks, their adoption could significantly strengthen cybersecurity of ICS networks.

Assessing the Positive Impact of the Process for Ensuring Diffusion and Adhesion of Cybersecurity Innovations to Reduce Software Vulnerability Severity

Yanzhen Qu, Gregory Vsevolozhsky, Chuan Yue
Colorado Technical University, Colorado Springs, Colorado, USA;
Colorado School of Mines, Golden, Colorado, USA

Abstract— Effective cybersecurity risk management requires an enhanced software development process that includes a feedback loop to facilitate diffusion and adhesion of cybersecurity innovations. Our Innovation Adhesion Evaluation Process (IAEP) model outlines how a team of SMEs led by a CIO/CISO can evaluate the feasibility, timeliness, suitability, and sustainability of cybersecurity innovations during the software's operational phase. To enable the comprehensive evaluation of cybersecurity innovations, we propose extending the CVSS v3.1 score model by introducing a process metric group. Each metric in this group represents a cybersecurity innovation that addresses a specific cyber risk associated with the software during its life cycle. Our approach enables the explicit inclusion of all cybersecurity innovations in the CVSS v3.1 score model and quantifies their impact in reducing the severity of corresponding software vulnerabilities. With this extension, the CVSS v3.1 score model becomes a valuable tool for the CIO/CISO, SMEs, and software development team to plan and validate the software's competency in resisting cybersecurity risks. By combining our IAEP model and the CVSS v3.1 score model extension, we provide a comprehensive and timely solution for managing cybersecurity risks during the software life cycle.

Machine Learning and the Secure Access Service Edge

Steven A. Wright, Achyuth Sathyagiri, Ravish Tayal
Cognizant, Atlanta, Georgia, USA;
Cognizant, Mississauga, Canada;
Cognizant, Noida, India

Abstract—The Secure Access Service Edge (SASE) framework integrates network and security functions in a cloud native architecture. The paper reviews the SASE framework to examine the applicability and necessity for the inclusion of Machine learning (ML). The feasibility of ML at the edge, and its measurement and performance aspects, are reviewed in considering the feasibility of ML functions in the SASE framework. A proof-of-concept implementation is discussed to provide concrete examples of ML based services related to the service edge including monitoring, predictive alerting, and service disruption avoidance. The applicability and necessity are ultimately driven by the enterprise's objectives for its infrastructure. Enterprises adoption of ML services in a SASE framework may be driven by considering whether the radical network decentralization of recent years is a transient event.

Harnessing the Efficiency of Reformers to Detect Software Vulnerabilities

Angel Jones, Marwan Omar
College of Computing, Capital Technology University, Maryland, USA;
College of Computing, Illinois Institute of Technology, Chicago, Illinois, USA

Abstract— Detecting software vulnerabilities is a critical task in ensuring the security of software systems. Recent advances in deep learning (DL) have shown promise in improving the accuracy of vulnerability detection. In this work, we proposed an approach for vulnerability detection using the Reformer language model. We compared our approach with previous works on three datasets: SARD, D2A, and Devign. Our experiments showed that our approach outperforms the state-of-the-art methods on all three datasets, achieving a precision of 0.95, recall of 0.91, and F1- score of 0.93 on average. Our approach is computationally efficient, making it suitable for large-scale vulnerability detection tasks. We also provide an in-depth analysis of the results and discuss the implications of our findings. Our work contributes to the growing body of literature on DL based vulnerability detection and demonstrates the potential of the Reformer model in this domain.

Preventing Twitter from Being Used for Defamation, Doxing, Impersonation, Threats of Violence, and Intimate Images

Ashu M. G. Solo
Maverick Trailblazers Inc., Wilmington, Delaware, USA

Abstract—The social network Twitter is a sewer of defamation, doxing, impersonation, and intimate images. Online defamation, doxing, impersonation, threats of violence, and sharing of intimate images are five of the major problems of the Internet age and can cause serious damages to victims. This research paper makes recommendations for Twitter policies, public policies, and public education to combat online defamation, doxing, impersonation, threats of violence, and sharing of intimate images on Twitter.

R1CIDS: A Hybrid Dual Filters Intrusion Detection System

Zichun Wang, Samuel Temesgen, Fred Wu
Department of Math, Engineering and Computer Science,
West Virginia State University, Institute, West Virginia, USA

Abstract—To cope with the growing number and changing nature of malicious cyber-attacks, machine learning techniques have been extensively employed to develop intrusion detection systems (IDS) for intelligent detection. However, these systems face numerous challenges, given the continually evolving nature of these attacks and their sheer volume. In here, a new hybrid intelligent IDS employing a combination of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) is proposed. This model aims to classify audit data and predict potential security breaches by leveraging the strengths of both CNN and LSTM. The new hybrid model has shown a significant improvement in detection accuracy while reducing the detection time.

PhishGuard: Machine Learning - Powered Phishing URL Detection

Saydul Akbar Murad, Nick Rahimi, Abu Jafar Md Muzahid
School of Computing Sciences & Computer Engineering, University of Southern Mississippi,
Hattiesburg, Mississippi, USA;
Faculty of Computing, University Malaysia Pahang, Pahang, Malaysia

Abstract—Phishing is a major threat to internet security, targeting human vulnerabilities instead of software vulnerabilities. It involves directing users to malicious websites where their sensitive information can be stolen. Many researchers have worked on detecting phishing URLs, but their models have limitations such as low accuracy and high false positives. To address these issues, we propose a machine-learning model to detect phishing URLs. To detect these malicious URLs, we use a dataset of over 500K

entries collected from the Kaggle website. The dataset is used to train five supervised machine-learning techniques, including K-Nearest Neighbors (KNN), Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), and Random Forest (RF). The aim is to improve the performance of the classifier by studying the features of phishing websites and selecting a better combination of them. To measure the performance, we considered three parameters: accuracy, precision, and recall. The LR technique yielded the best performance, demonstrating its efficacy in detecting phishing URLs.

FPGA Formal Verification: A Local Logic Correctness Approach

Jules Chenou, Laurent Njilla, Aurelia William, Tonya Fields
Cybersecurity Complex, Norfolk State University, Norfolk, Virginia, USA;
Information Assurance Branch (RIGA), Air Force Research Laboratory, Rome, New York, USA

Abstract—Hardware verification is the proof that a circuit or a system (its implementation) behaves according to a given set of requirements (its specification). Design faults may result from an erroneous transformation of a design specification into the layout description, the latter being the basis of fabrication. A formal specification is a concise description of the behavior and properties of a system written in a mathematically based language. It will state what a system is supposed to do in the context it is supposed to operate as abstractly as possible, thereby eliminating distracting detail and providing a general description resistant to future system modifications. Suppose logic is used to formalize both the specification and the implementation. In that case, the verification procedure will lead to proof of logical equivalence that the two formalisms are the same or an implication that the implementation covers the specification. Verification is only a correct proposition concerning a formal specification. If the specification or the modeling is wrong, then even a positive result of a correctness proof is meaningless. This article will glue the implementation and specification in one formalism, a.k.a local logic, to avoid the inherent discrepancy between the two formalisms. This article proposes that an FPGA is correct if its local logic is sound and complete.

Unsupervised Learning of Statistical Patterns for Industrial Internet of Things Security

Honggang Wang
Department of Technology and Operations Management, College of Business Administration,
California State Polytechnic University-Pomona, Pomona, California, USA

Abstract - The Internet of Things (IoT) nowadays has become a ubiquitous technology in the fields of business and engineering. Cybersecurity issues are becoming increasingly prevalent in IoT systems, as hackers continue to exploit vulnerabilities in these interconnected devices. To address these concerns, we propose a cutting-edge analytics cybersecurity solution that utilizes clustering and mixture modeling. Our approach has the ability to accurately predict day-zero cyberattacks and, more importantly, can pinpoint the locations from where these attacks originate. This allows for a quick response and efficient recovery from any potential security breaches. We have applied this approach to a real-life water treatment IoT system, and some promising results will be presented and discussed.

Identifying Zero-Day Attacks with Machine Learning and Data Reduction Methods

Haydar Teymourlouei, Daryl Stone, Lethia Jackson
Department of Technology & Security, Bowie State University, Bowie, Maryland, USA

Abstract— Understanding the threats affecting your network is essential to developing a successful cyber-security defense. One threat that is crucial to building a defense against is a zero-day attack. The problem is that a zero-day attack leaves one with no real way to prepare for the attack other than searching through logs and logs of system logs looking for anomalies in the data. Recently, machine learning (ML) has been proposed to handle big data streams such as these. We compare two methods: a common data reduction algorithm that reduces the number of attributes in a dataset while keeping as much variation in the original dataset

as possible. The second method will be focused on developing ML algorithms using deep learning with the Stochastic gradient descent (SGD) classified legitimate traffic that is referred to as a true-positive. The results of the first method showed a 500% reduction in data removed many of the events considered to be zero-day threats. Further, the performance of the second method exceeded 97% accuracy. These results suggest the methods presented here are effective in the identification of zero-day attacks. To validate these results, further testing should be done.

A Study of Phishing Websites and Scan Evasion Techniques

Christian Liguori, Jean Chu, Daehan Kwak
Department of Computer Science and Technology, Kean University, Union, New Jersey, USA

Abstract—A phishing attack is a deceptive tactic in which individuals are lured into divulging sensitive information to a malicious actor masquerading as a trusted entity, typically through the use of phishing websites. Malware scanners are used to identify these fraudulent websites. However, the efficacy of these scanners is limited to the information they receive. Threat actors exploit this limitation by leveraging evasion techniques, such as redirecting to benign webpages or delivering different content, to conceal their malicious intent. This underscores the need for enhanced defense mechanisms against phishing attacks to safeguard individuals from falling prey to these fraudulent schemes. In this research, we investigate the prevalence of scan evasion techniques employed by phishing websites and evaluate the effectiveness of defense strategies by comparing the responses obtained from both a user and a malware scanner. Our findings reveal a notable number of phishing websites employing evasion techniques. While the majority of these websites were detected by the malware scanner, there were instances where evasion was successful.

An Improvement of a Secure Searchable Public Key Encryption Scheme with a Designated Tester Against Keyword Guessing Attacks and its Extension

Min-Shiang Hwang, Shih-Ting Hsu, Cheng-Ying Yang
Department of Computer Science and Information Engineering, Asia University, Wufeng, Taiwan, R.O.C.;
Department of Management Information Systems, National Chung Hsing University, Taichung, Taiwan, R.O.C.;
Department of Computer Science, University of Taipei, Taipei, Taiwan, R.O.C.

Abstract—Public key encryption with keyword search (PEKS) scheme enables one to send the trapdoor, which involves the encrypted keyword in querying data without revealing the keyword. Recently Hu and Liu proposed a secure and efficient searchable public key encryption scheme with a designated tester against keyword guessing attacks. However, this article will show that their scheme cannot be against an offline keyword-guessing attack. Therefore, we also propose an improved scheme.

Current State of Network Security and Privacy

Nick Roys, Donnie Flanders, Marissa Baisi, Julian Martin, Bradley Hoffman, Hilmi Demirhan
University of North Carolina Wilmington, North Carolina, USA

Abstract—In today's business environment, network security is a pressing issue for organizations as they progressively rely on technology for their day-to-day operations. With the rise of the Internet, the threat of cyber-attacks has become a major concern for businesses and individuals alike. The objective of this research paper is to provide an overview of the current state of network security, including a discussion of the challenges faced by modern-day organizations and the possible solutions to address these challenges. This discourse will examine various aspects of security threats such as malware, phishing, and ransomware, and how these threats are constantly evolving to evade current network security measures. Additionally, this paper will examine the role of firewalls, encryption, and intrusion detection and prevention systems (IDS); focusing on how they work together to provide network security. The discussion of the importance of user education and training to prevent security breaches as well as the need for organizations to adopt a proactive approach instead of a reactive approach to network security will also be included. Finally,

this report will provide insights into the future of network security and the developments that are likely to shape the industry in the coming years.

Addressing IoT Security and Privacy Challenges

K. I. Wijesinghe, U. U. Samantha Rajapaksha
Department of Graduate Studies and Research,
Sri Lanka Institute of Information and Technology, Sri Lanka

Abstract—Healthcare Internet of Things (HIoT) systems offer tremendous potential for improving patient care and monitoring, but they also present significant security and privacy challenges. Protecting sensitive medical data and ensuring patient privacy are paramount in healthcare IoT deployments. In this paper, we propose the use of Radio Frequency Identification (RFID) technology as a solution to address the security and privacy challenges in healthcare IoT. RFID offers unique advantages in securing healthcare IoT systems. It enables the unique identification and tracking of medical devices, ensuring their integrity and authenticity. By incorporating RFID tags into healthcare IoT devices, we can establish a secure and reliable connection between devices, preventing unauthorized access and mitigating data breaches. One of the key benefits of RFID in healthcare IoT security is its ability to provide accurate patient identification and access control. RFID tags can be used to authenticate patients, medical staff, and authorized devices, ensuring that only authorized individuals have access to sensitive patient information. Additionally, RFID can enable secure tracking of medical devices, reducing the risk of loss or theft and enhancing asset management in healthcare facilities. In terms of privacy, RFID can help protect patient data by employing privacy-enhancing techniques such as data encryption, access restrictions, and data anonymization. These techniques ensure that patient information is handled securely and only accessed by authorized healthcare providers, preserving patient privacy in IoT environments. This paper explores the integration of RFID technology into healthcare IoT systems and discusses its potential applications in addressing security and privacy challenges. We analyze the benefits and considerations of using RFID in healthcare IoT, including scalability, interoperability, and regulatory compliance. Furthermore, we discuss the importance of robust security protocols and standards to ensure the responsible and ethical use of RFID in healthcare IoT deployments. By leveraging RFID technology, healthcare IoT systems can strengthen their security measures and better protect sensitive patient data. The use of RFID enables accurate patient identification, access control, and secure device tracking, while also addressing privacy concerns through privacy-enhancing techniques. This paper provides insights into the potential of RFID as a solution for addressing security and privacy challenges in healthcare IoT, contributing to the advancement of secure and privacy-preserving healthcare IoT deployments.

A Survey on Retrieving Confidential Data Using Phishing Attack

Sai Teja Veeramalla, Tathagata Bhattacharya
Computer Information System Cybersecurity, Auburn University at Montgomery, Alabama, USA

Abstract—Internet technology has improved the lives of individuals in a multitude of manners due to its broad use, namely social media sites and banking. Because of the advancement of Internet technologies, security threats to networks and IT infrastructure are always changing. One such serious problem is scamming, in which hackers attempt to obtain users' login information by using phony mail, bogus websites, or both. Businesses and academic organizations place high importance on developing defenses against phishing scams. Therefore, companies must place a high priority on end-user training while taking precautions against phishing dangers. In this research, we clone a popular website to place a phishing attack on a user. This paper has a high impact because we have mirrored highly secure web pages that use form validation for security. Once users come across our duplicate web page, they get fooled by the phishing portal and enter their secret information. Now the attacker uses this information to commit bigger crimes. But this paper is designed so that we can warn the world against this kind of phishing measures and save users from falling into the trap.

A Protocol for Monitoring Network Threats in Real-Time

Aftab Ahmad
John Jay College of Criminal Justice, New York, USA

Abstract—As cyber threat intelligence enters the machine learning paradigm, a natural question is 'is it possible to predict network attacks from the current state-of-the-internet (SOTI) before they happen?' In this work, we present the threat intelligence protocol (TIP) between components of a distributed threat intelligence database on host clients that can be maintained by internet devices locally either on the device, or enterprise network to swiftly make prediction of impending threats. The protocol can monitor the cyber threat level in real-time by monitoring and sharing threat information among the participating hosts.

STAR Antivirus Software

Tathagata Bhattacharya, Ramyasree Thungapati, Supriya Gurrapu, Anusha Madichetty, Sai Teja Veeramalla
Computer Information System Cybersecurity, Auburn University at Montgomery, Alabama, USA

Abstract—In today's IOT era, the likelihood of our machines getting infected by malicious software is increasing every single day. The fundamental problem is that many people are unaware of what a virus is and how it infects computers. In this research, we have developed Star Antivirus Software, that improves the effectiveness of identifying and thwarting malware attacks. Our antivirus employs advanced strategies, such as behavioral analysis, to detect and stop assaults more efficiently than traditional antivirus systems with a minimum CPU and Memory utilization. Thus, we can avoid Data breaches and other online risks significantly using STAR antivirus software. In this situation, the newly developed antivirus software's purpose is to check for infections. This antivirus software carefully uses memory so that processor and memory consumption is significantly less. We have also included a minimum subscription fee for the users (compared to other antivirus software) to enhance our research and development for this antivirus software. At the same time, our mission is to provide a great quality of service and security to our clients to achieve better security, we developed a three-zone VPN feature that offers a safe connection. Apart from that, the antivirus software provides the users with a pop-up window for issuing a warning against any corrupted file. Keywords: antivirus, virus scan,

Empirical Analysis of Cryptocurrency Mixer: Tornado Cash

Minwoo Youn, Kota Chin, Kazumasa Omote
University of Tsukuba, Tsukuba, Japan;
National Institute of Information and Communications Technology, Tokyo, Japan

Abstract—Decentralized and anonymous transactions have become possible with the advent of cryptocurrency. However, these characteristics make it challenging to trace the origins of illegally obtained funds. Furthermore, mixing services, which sever a connection between the previous and current owners of cryptocurrencies, are used to increase the anonymity of cryptocurrencies. Tornado Cash, which is frequently used as a mixing service, has been reported as a destination for a significant number of cryptocurrencies that are associated with criminal activity. In this research, we cluster and analyze the deposit addresses in Tornado Cash, list potential criminal addresses, and clarify the actual situation of money laundering in Tornado Cash. Moreover, we focus on NFT phishing incidents and determine the circumstances of the incidents and the total amount of damage relating to Tornado Cash.

Lightweight Software Assurance for Distributed Mobile Networking

Sang-Yoon Chang, Simeon Wuthier, Jonghyun Kim, Jinoh Kim
University of Colorado, Colorado Springs, Colorado, USA;
Electronics and Telecommunications Research Institute (ETRI), Daejeon, South Korea;
Texas A&M University Commerce, Texas, USA

Abstract— With developments in software-defined radio (SDR) and OpenRAN, the mobile networking implementations for radio and security control are becoming increasingly software based. We design and build a lightweight and distributed software assurance scheme, which assures that a wireless user holds the correct software (version/code) for its wireless networking implementations. Our scheme is distributed (to support the distributed and ad hoc networking which does not utilize the networking backend infrastructure) and lightweight (to support the resourceconstrained device operations). Our scheme is distinct from the remote code attestation in trusted computing, which requires the hardware-based security and the real-time challenge-andresponse communications with a centralized trusted server, thus making its deployment in the distributed mobile networking environments prohibitive. We incorporate Merkle tree for the verification efficiency to make it appropriate for the wirelessbroadcasting medium with multiple receivers. We implement and test our scheme on srsRAN (a popular open-source software for cellular technology, including 4G and 5G) and conduct experiments on a phone, Raspberry Pi, and a computer to demonstrate the lightweight design and its appropriateness for wireless networking applications. Our results show that the number of hash computations for the proof verification grows logarithmically with the number of software code files being assured and that the verification is three orders of magnitude cheaper in time than the proof generation, while the proof generation overhead itself is negligible compared to the software update period.

Online Fake Logo Detection System

Vivek Tanniru, Tathagata Bhattacharya, Sai Teja Veeramalla
Department of Computer Science and Information Management Systems,
Auburn university at Montgomery, Alabama, USA

Abstract—Every day, hundreds of domain names, websites, and logos are being cloned by cyber criminals who want to gain our trust to steal our data. As a result, faking logos is becoming a big issue in the online world and needs to be addressed. As a result, fake logos on the internet have become a significant source of worry for businesses and consumers. The algorithm can detect differences in logo design, color, and positioning and assess the possibility of a fake logo. The system's accuracy was evaluated on a massive dataset of actual and false logos, and it obtained a high level of accuracy in recognizing fake logos. The fake logo identification technology has the potential to dramatically increase the credibility and dependability of online material, thereby protecting brand identity integrity. This research proposes a method for detecting fake logos using a Context-dependent similarity algorithm. Our approach involves extracting features from the logos and training a machine-learning classifier to distinguish between real and fake logos. We evaluate the performance of our method on a dataset of real and fake logos and demonstrate its effectiveness in detecting fake logos with high accuracy. Keywords-Context-depen

How to Attack a Galaxy: From Star Wars to Star Trek

Luis Hernandez-Avarez, Miguel Agel Gonzalez de la Torre, Eva Iglesias Hernandez, Luis Hernandez Encinas
Institute of Physical and Information Technologies (ITEFI),
Spanish National Research Council (CSIC), Madrid, Spain

Abstract—Recently, the National Institute of Standards and Technology set CRYSTALS–Kyber as post-quantum public key encryption/key encapsulation mechanism standard, and CRYSTALS–Dilithium as post–quantum digital signature standard. These post quantum cryptosystems are also recommended for national security systems. As a result, it is important to identify and analyze the weaknesses and potential information leakage points, so that they can be resolved. In this work, we study the newest side channel attacks based on artificial intelligence models against Kyber and Dilithium, focusing on the specific function attacked. We also examine the artificial intelligence algorithms employed in these attacks and their configurations, discussing which

parameters and setting are suitable, and identifying different tools that might be useful. Index Terms—Aritificial Intelligence, Kyber, Dilithium, Multi Layer Perceptron, Post Quantum Cryptography.

Basics of Auditable AI systems

Olivia Sina Graupner, Dirk Pawlaszczyk, Christian Hummert
Safe Society Agentur fur Innovation in der Cybersicherheit GmbH Halle (Saale), Germany;
University of Applied Sciences Mittweida, Germany;
Agentur fur Innovation in der Cybersicherheit GmbH Halle (Saale), Germany

Abstract—What are the possibilities of auditing artificial intelligence, and how can they be realized in theory and in practice? The global use of artificial intelligence is highly increasing. The digital forensic community therefore needs forensic methods to audit internal decisions to ensure their reliability, especially in safe-critical environments. Furthermore, forensic readiness measures are equally necessary in the event of an attack on AI systems. The continual learning framework Avalanche is used to construct an application-oriented scenario as the foundation to store weight values and weight changes in a log file while training. The development of the changes as well as the resulting memory size with varying hidden layer and neuron numbers provides conclusions about the applicability of the method, which needs further research in the future.

Securing 5G/6G Communications in Smart Cities: Novel SNOW-V/ZUC-256 Multimode Architectures

Evangelia Konstantopoulou, George S. Athanasiou, Nicolas Sklavos
u-blox Athens S.A., Athens, Greece;
SCYTALE Group, Computer Engineering & Informatics Dept., University of Patras, Patras, Greece

Abstract—In this article, area- and performance-efficient multimode architectures of SNOW-V and ZUC-256 stream ciphers for smart-city deployments are introduced. Efficient resource sharing of critical parts of the two algorithms together with sophisticated application of design optimization techniques, such as loop unrolling, led to low area solutions while the overall performance was kept high. The proposed designs were implemented in both FPGA (AMD Xilinx) and ASIC (22nm) technologies. The results are compared with standalone implementations of individual ciphers, state-of-the-art existing hardware implementations of individual ciphers, and existing SNOW 3G/ZUC-128 multimode topologies as well. It is the first time that multimode designs are proposed that include both SNOW-V and ZUC-256 algorithms.

Security Education Using Screen Reader: Utilizing a Problem-Posing Approach

Masataka Kakinouchi, Kazumasa Omote
Tsukuba University of Technology, Tsukuba city, Ibaraki, Japan

Abstract—Information security measures are essential for ICT users, particularly for the visually impaired who use voiceactivated ICT. However, visually impaired users are often not educated about sufficient countermeasures using a screen reader. This study analyzes the efficacy of security and usability education for the visually impaired using a problem-posing approach. Herein, security education is taught to the visually impaired in a classroom, with focus on screen readers, and the effectiveness of problem posing, a method used in mathematics education, is verified. Furthermore, qualitative analysis reveals that screenreader users deepen their understanding of security and usability via the problem posing activity. Additionally, the important functions for screen-reader users regarding security and usability are determined.

Towards Building a Versatile Tool for Social Media Spam Detection

Jalal Abdel Halim, Weiqing Sun
College of Engineering, University of Toledo, Ohio, USA

Abstract—With the rapid increase of social network spam, it's essential to empower users with the tool to detect the harmful spam effectively. However, existing tools cannot fully meet the requirements. In the paper, we propose and develop a live detection tool that can detect spam text and images from social networks. This tool will be trained on user collected data in the form of both images and text using different machine learning classifiers, the users are then able to save the model and load it whenever they want to use a social network, where this tool will show the users a notification alerting them whether the post they are looking at is spam before they even get the chance to read the text or look at the image, thus protecting them from the potentially malicious information or links. Performance and functional evaluation results have demonstrated the effectiveness of our tool.

An Empirical Evaluation of Encryption and Decryption Times on Block Cipher Techniques

Funminiyi Olajide, Kwame Assa-Agyei, Onome Christopher Edo
Department of Computer Science, Nottingham Trent University, Nottingham, United Kingdom (UK);
Department of Information Systems, Auburn University, Montgomery, Alabama, USA

Abstract – N/A

Intelligent Monitoring and Management of Smart Buildings Using Machine Learning: Optimizing User Behavior and Energy Efficiency

Mais Nijim, Viswas Kanumuri, Hisham Albetaineh, Ayosh Goyal
Department of Electrical Engineering and Computer Science,
Texas A&M University-Kingsville, Kingsville, Texas, USA

Abstract-- Smart buildings are a crucial component of smart cities and smart grids, contributing to a more intelligent urban environment. The efficient utilization of energy in buildings is now a significant concern for sustainable societies. Whether they are residential or industrial. Smart buildings consume the majority of the energy produced, but in the context of smart grids, they are also designed to generate energy and help stabilize demand. Even small fluctuations in peak demand can result in substantial budget saving for customers and utilities. In this research, we explore the applicability of machine learning (ML) methods for load forecasting in smart buildings using smart sensor data that infers user behavior. Here, we utilized a smart building dataset encompassing four floors, 51 rooms, and 255 integrated sensors. Each smart building room comprises five types of calculations, including a PIR Sensor that helps to capture motion, a Carbon dioxide concentration sensor to check the purity of air, a Sensor to calculate temperature, a humidity sensor to maintain good conditioning, and a luminosity sensor. By using this data from these sensors, we can optimize and manage the energy saving. To evaluate the prediction performance of the input variables, we used to compare them we use mean squares error, mean absolute error and root mean square error.

Assessing the Efficacy of Machine Learning and Deep Learning in the field of Cybersecurity

Rajesh Eswarawaka, Mais Nijim, Viswas Kanumuri , Hisham Albetaineh
Department of Artificial Intelligence and Machine Learning,
AMC Engineering College, Bangalore, Karnataka, India;
Department of Electrical Engineering and Computer Science,
Texas A&M University-Kingsville, Kingsville, Texas, USA

Abstract-- The use of machine learning has become widespread across various fields because of its superior performance compared to conventional rule-based algorithms. As a result, these models have also been integrated into cyber security systems, Machine learning is being utilized to aid or possibly even supplant the role of human security analysts. However, it's important to evaluate the effectiveness of machine learning in cyber security with careful consideration, especially if complete automation of detection and analysis is being considered. This study provides an in-depth research focuses on machine learning techniques applied in intrusion, malware, and spam detection that are tailored towards security professionals. The primary objective of our study is to evaluate the degree of advancement or maturity of these techniques of ML-based cybersecurity solutions and to identify any limitations that could impede their effectiveness as detection mechanisms. To achieve this, we conducted a thorough literature review and performed experiments on enterprise systems and network traffic in real-world settings. Our goal is to gain understanding of the capabilities and limitations of ML solutions and provide actionable insights for their improvement.

On Effectiveness of Machine and Deep Learning Algorithms for Detection of GPS Spoofing Attacks on Unmanned Aerial Vehicles

Jaron Burns, Dimitri Amiridis, Dulal Chandra Kar, Longzhuang Li
Texas A&M University - Corpus Christi, Texas, USA;
University of South Carolina, Columbia, South Carolina, USA

Abstract—Unmanned aerial vehicles (UAVs) that use GPS signals for navigation are susceptible to GPS spoofing attacks when operating in hostile environments. In mission-critical applications, it is essential to detect and defend against such attacks on UAVs. This work presents and compares three different approaches to detect spoofed GPS signals using machine learning and deep learning algorithms. The first approach utilizes a dimensional reduction method, known as the Principal Component Analysis (PCA) structure, followed by a classification model known as the Support Vector Machine (SVM) structure to detect spoofed GPS signals. The second approach uses a simple autoencoder structure for dimension reduction, followed by the logistic regression model for GPS signal classification. Finally, the third approach employs a Long Short Term Memory (LSTM) autoencoder for dimension reduction, followed by logistic regression for GPS signal classification. All models were tested on three different data sets, including a data set with Gaussian noise, an imbalanced data set, and a clean data set. It was found that the machine learning SVM-based approach yielded the highest detection rate (over 99.8%) in terms of accuracy, precision, F1 score, and recall on all data sets.

DDoS Attack Detection on a 5G NSA Based Hybrid Energy Communications Network: A Case Study

M. Hackett, P. Thulasiraman, J. Ries, A. Edmond, J. Seville, P. Ashok
California State University Monterey Bay, California, USA;
Naval Postgraduate School, Monterey, California, USA;
California State University Long Beach, California, USA; US Ignite, Washington D.C., USA

Abstract— The energy communications network deployed at Marine Corps Air Station, Miramar, San Diego, is a hybrid system developed to provide remote control and monitoring of distributed energy resources (DER). The network consists of a hardwired system and the integration of a Verizon 5G Non-Stand Alone (NSA) network that provides wireless connectivity from the DERs

to the Energy and Water operations center (EWOC) on base. The network connectivity of the DERs exposes them to specific network dynamics, including potential cyber penetration by an adversary. The objective of this paper is to develop a cyber anomaly detection model for this hybrid energy communications network using an autoencoder neural network. The autoencoder classifies distributed denial of service (DDoS) attacks against the network infrastructure. We train the autoencoder model on two traffic data sets: 1) Modbus TCP/IP data from the hardwired network apparatus of the energy communications infrastructure and 2) experimentally generated 5G data that mimics traffic on the energy communications network. We present a foundational approach for detecting anomalous behavior on this network and evaluate the effects of applying various combinations of model configurations to the autoencoder. We highlight the results from the best model configurations that provide the highest level of DDoS prediction accuracy.

Physical Attacks on the Railway System

Lukas Ifflander, Thomas Buder, Teresa Loreth, Marina Alonso Villota,
Stefan Pickl, Walter Schmitz, Karl Adolf Neubecker
German Centre for Rail Traffic Research Dresden, Germany;
Faculty of Computer Science, Universitat der Bundeswehr, Munich, Germany;
CreaLab GmbH, Feldkirchen-Westerham, Germany

Abstract—Recent attacks encouraged public interest in physical security for railways. Knowing about and learning from previous attacks is necessary to secure against them. This paper presents a structured data set of physical attacks against railways. We analyze the data regarding the used means, the railway system's target component, the attacker type, and the geographical distribution of attacks. The results indicate a growing heterogeneity of observed attacks in the recent decade compared to the previous decades and centuries, making protecting railways more complex.

A Multi-model Rouge Nodes Detection System for Fog Computing

Thanh Duc Bui, Brajendra Panda
Department of Computer Science and Computer Engineering,
University of Arkansas, Fayetteville, Arkansas, USA

Abstract—With the increasing speed of advance of both Internet of Things (IoT) and the demands of utilizing IoT devices in different locations with high compatibility, fog computing has been foreseen as a sustainable solution. However, new challenges related to fog computing's cyber security also emerge. Among various cyber-attacks on the fog computing system, one standard method is to implement corrupted internal nodes that allow malicious/compromised access from attackers, threatening the confidentiality of users' information and the system's performance. In this paper, we propose a method to quantify the distribution of each attribute value from historical access logs from each node within a fog computing system. We also utilize an unsupervised machine learning method to separate nodes into smaller clusters with respect to different models to improve the forecasting ability of the system. Finally, our models are evaluated over simulated data using real access logs. Our experimental results illustrate that our proposed method achieves some advanced performance over applying an overall classification, even using the same classification machine learning methods and dataset.

Investigation of Digital Forensics Tools Validation

Yahya Sayeed, Mehrdad S. Sharbaf
Department of Computer Science, California State University Dominguez Hills Carson, California, USA

Abstract— As the acquiring of Ubiquitous Communications is growing with exponential pace, cybercrime is also expanding, and forensics crime investigation is evolving into more challenging issues which need to be confronted. For the last decades, email has been imperative communication within organizations, and internet. For that reason, email is also a primary source for criminal activities on the internet. Email forensics crime investigation requires effective email forensic planning and efficient tools for

forensic analysis. Email forensics is dedicated to investigating, extracting, analyzing, and exporting emails to collect digital evidence as findings to solve crimes and certain incidents, in a forensically sound manner. PST files, are one of the most important file types during investigations. The research topic compares various methods to repair a corrupt email container file "PST" along with recovering the deleted email files with attachments. Our research aim is to help forensic professionals, forensics researchers and law enforcement agencies choose an appropriate tool for a particular task which is fixing and recovering emails for better digital evidence carving and repair. We evaluate the tools based on the ability to repair an email container, emails recovered along with metadata and hash function values, graphical user interface as well as comparison of various other features. We have used Enron email dataset for the analysis.

Blockchain-based Messaging for VANETs

Andrea Huszti, Tamas Gira szi, Norbert Olah
University of Debrecen, Debrecen, Hungary

Abstract—An intelligent transport system is indispensable for today's life. It can help to reduce traffic and even the number of accidents. However, it raises several security issues. Providing travelers' privacy and anonymity is essential. We propose a message broadcast protocol for vehicle communication based on the blockchain technology. Eligible vehicles can anonymously report road conditions (e.g. traffic jams, accidents, etc.). We give an efficient solution applying Solana blockchain that provides fast block generation time and supports smart contracts. Blockchain maintains immutability of messages, enables anonymous message submission, moreover assures message unlinkability. Our scheme is identity-based, applying bilinear pairings batch verification of messages is accomplished.

PwnPilot: Reflections on Trusting Trust in the Age of Large Language Models and AI Code Assistants

Dwight Horne
Department of Computer Science, Baylor University, Waco, Texas, USA

Abstract—At the dawn of a new era in software engineering, one defined by large language models (LLMs) and AI code assistants like GitHub Copilot, new meaning can be found from a historic Turing Award Lecture that concluded one cannot trust source code they "did not totally create" themselves. In this paper, a targeted, systematic survey of the latest research results from 2019 to early 2023 highlights the possible risks of using AI code assistants that produce substantial source code contributions, and the potential for an AI Copilot to unknowingly become PwnPilot, a malevolent digital actor that introduces vulnerabilities and compromises trust. During a period of explosive growth for generative AI, renewed reflections on trusting trust point to conclusions similar to the original assertions of Ken Thompson in 1984. But despite a recent theoretical roadblock from proof of ability to plant undetectable backdoors in machine learning models, the potential for enhanced productivity from AI code assistants may still be realizable with an acceptable level of risk, perhaps even for safety critical and sensitive security relevant contexts. In support of that goal, a number of near-term risk management options and longer term research paths are identified as enablers for practitioners and inputs to potential research roadmaps toward more secure and trusted AI code generation.

Formal Verification of Authenticated Encryption with Associated Data with Tamarin Prover

Takehiko Mieno, Hiroyuki Okazaki, Kenichi Arai, Yuichi Futa
Graduate School of Science and Technology, Shinshu University, Wakasato, Nagano City, Nagano, Japan;
EPSON AVASYS Corporation, Ueda City, Nagano, Japan;
Graduate School of Engineering, Nagasaki University, Nagasaki City, Japan;
School of Computer Science, Tokyo University of Technology, Hachioji City, Japan

Abstract—Tamarin Prover and ProVerif are automated tools for formally verifying the security of cryptographic protocols. Tamarin Prover is a type of theorem prover, whereas ProVerif is a type of symbolic model checker. In this paper, we formalized the structure of authenticated encryption with associated data (AEAD). To do so, we formalized AEAD as two types of generic compositions, using the Encrypt-then-Message Authentication Code (MAC) and MAC-then-Encrypt constructions presented by Rogaway et al. Using Tamarin Prover, we verified the confidentiality and integrity independent of considering the attacker's capabilities. We demonstrated that the attacker's abilities can be modeled in ProVerif, and we compared our findings with results obtained via ProVerif. We found that ProVerif is superior to Tamarin Prover in the studied scenarios, although future work is needed to consider attacker's abilities.

The Case for the Simpler Cyber-Physical System

Alain Kagi, Caitlyn Wilde, Levi Overcast, Aubrey Birdwell, Andrew Fowler,
Jens Mache, Richard Weiss
Lewis & Clark College, Portland, Oregon, USA;
The Evergreen State College, Olympia, Washington, USA

Abstract—This paper argues for the design of simpler cyberphysical systems. Several factors motivate this case for simplicity: (1) improved performance, (2) reduced power requirements, (3) reduced costs, (4) reduced attack surface, and (5) the possibility to prove the entire design correct. As a proof of concept, we are building a networked sensor on top of the seL4 microkernel [1], and we hope to prove the entire system correctly implemented. Index Terms—Cyber-physical system, verification, security, vertical integration, simplicity, seL4, secure OS microkernel, Resurrecting Duckling security policy.

A Machine Learning Approach for Android Ransomware Detection

Sekione Reward Jeremiah, Jong Hyuk Park, Jungho Kang, Ji Su Park
Department of Electrical and Information Engineering,
Seoul National University of Science and Technology, Seoul, Republic of Korea;
Department of Computer Science and Engineering,
Seoul National University of Science and Technology, Seoul, Republic of Korea;
Department of Software Engineering, Baewha Women's University, Seoul, Republic of Korea;
Department of Computer Science and Engineering, Jeonju University, Jeonju, Republic of Korea

Abstract – N/A

IoT Security: Implementation of Xtea, Simon/Speck Lightweight Block Ciphers

Levent Ertaul, Arpit Chauhan
Computer Science, California State University East Bay, Hayward, California, USA

Abstract – The current technological advancements in the fields of internet, mobile devices, and machine to machine (M2M) communication can be regarded as the initial phase of the Internet of Things (IoT). As the IoT expands, it will integrate diverse technologies to enable novel applications by connecting various physical objects. In such an environment, it has become crucial to safeguard all the information. This research paper focuses on implementing lightweight encryption algorithms, such as Xtea, Simon, and Speck, using an ESP8266 microcontroller and the Arduino IDE. Additionally, the paper presents a performance analysis of these algorithms based on their computational cost and memory requirements. Thus, this study provides insights into the effectiveness of these three lightweight encryption algorithms for securing IoT devices.

Securing Blockchain Technology: A Comprehensive Analysis of Vulnerabilities and Mitigation Strategies

Bharathi Putta, Dulal Chandra Kar
Texas A&M University - Corpus Christi, Corpus Christi, Texas, USA

Abstract—This paper provides a comprehensive analysis of cybersecurity vulnerabilities in blockchain technology and presents mitigation strategies to counter them. The analysis covers recent attacks and their impact on blockchain security. The paper identifies the major types of blockchain vulnerabilities, including consensus algorithm vulnerabilities and wallet vulnerabilities for cryptocurrency applications. Mitigation strategies for these vulnerabilities are discussed, such as the use of consensus algorithms with improved security, the implementation of secure coding practices, and the adoption of multi-factor authentication for wallet access. The paper also highlights the importance of community efforts and collaboration between industry players in enhancing blockchain security. The analysis and strategies presented in this paper provide valuable insights for blockchain developers, researchers, and stakeholders seeking to secure their blockchain applications against cyber threats.

Attack Vectors Against ICS: A Survey

Destin Hinkel, Angela Buie, Amber Surles, George Clark
School of Computing, University of South Alabama, Alabama, USA

Abstract—Attacks against industrial control systems (ICS) have the potential to damage critical infrastructure. Due to the volume of attacks, keeping up with various threat vectors is nearly overwhelming. We describe various attack vectors discussed in offensive and defensive contexts and propose a categorization theme based on those that have been observed in the related literature. We propose a framework that will categorize attacks by one of three primary attack vectors. We also propose potential defenses against attacks using the aforementioned attack vectors.

Password Cracking as a Medium for Introducing Cybersecurity Skills and Student Autonomy

Emily Tanabe, Cole McCorkendale, Jens Mache, Aurelio Puente, Matthew Chio, Richard Weiss
Lewis & Clark College, Portland, Oregon, USA; Evergreen State College, Olympia, Washington, USA

Abstract—This paper proposes a design for an introductory password cracking exercise that gives students the opportunity to develop foundational cybersecurity skills while increasing their confidence and agency. This exercise aims to educate students

about the brittle nature of passwords while increasing students' cybersecurity soft skills, such as collaboration, autonomy, and problem solving. To do so, the exercise uses pedagogical methods such as the Gradual Release of Responsibility model and guiding questions. The exercise is holistic, hands-on, and consists of three scaffolded levels: • Password guessing, intelligence gathering, and spear phishing. • Manually attempting a "credential stuffing" attack on a simple password. • Scripting an automated password cracking tool. This exercise will educate students about passwords, how to attack them, and how to choose secure passwords while building foundational cybersecurity skills and keeping less experienced students interested, engaged, and motivated.

The Current State of Fingerprinting in Operational Technology Environments

Cheronika Manyfield-Donald, Candice Mitchell
US Army Corps of Engineers, Information Technology Laboratory,
Engineer Research and Development Center (ERDC), Vicksburg, Mississippi, USA

Abstract— Observing the increase in cybersecurity attacks on our nation's and even world's critical infrastructures, heightens the priority need for being proactive in detecting possible intrusions. Each endpoint that connects information technology to operational technology is a vulnerability endpoint. Identification of these endpoints is the first step in fortifying the critical infrastructure. Sadly, most entities do not have a complete blueprint of their system's network. This is why the introduction of network fingerprinting arose. This initial step of our research discusses networking fingerprinting as a tool to identify components of a critical infrastructure, outlines previous work and identifies some network fingerprinting tools. The work shows that fingerprinting and operational technology (OT) asset identification is still woefully lacking in its ability and suggests possible future improvements.

Basic Safety Message (BSM) Test Data Generation for Vehicle Security Machine Learning Systems

Guillermo Francia, III, Dallas Snider, Benjamin Cyphers
Center for Cybersecurity, University of West Florida, Pensacola, Florida, USA;
Department of Information Technology, University of West Florida, Pensacola, Florida, USA;
Ben Cyphers ACES SFS Program, University of West Florida, Pensacola, Florida, USA

Abstract—This paper presents a subcomponent of an on-going research on connected vehicle security. It proposes techniques on the generation of synthetic Basic Safety Message (BSM) test datasets, both on normal operation and anomalous behavior. The synthetic test data conform with the SAE J2735 Standard on message sets that support vehicle-to-everything (V2X) communications systems. The purpose of such datasets is for the derivation of machine learning systems that can be deployed in a V2X operating environment.

Benchmarking the Elliptic Curve Digital Signature Algorithm and RSA in Key Signing and Verification Operations with Parallelism

George Dimitoglou, Carol Jim
Department of Computer Science & Information Technology,
Hood College, Frederick, Maryland, USA

Abstract - Cryptographic techniques play a crucial role in securing information, and selecting an appropriate cryptographic technique involves considering factors such as encryption strength and computational overhead. Asymmetric cryptographic protocols and algorithms, such as RSA and those based on elliptic curve cryptography, are widely used for their flexibility, encryption strength, and key management. This study focuses on comparing the performance of the Elliptic Curve Digital Signature Algorithm (ECDSA) and RSA, specifically analyzing the impact of parallelism on signing and verification operations. The

experimental results demonstrate that ECDSA is faster in signing but slower in verification compared to RSA. However, RSA's performance declines rapidly with increasing key sizes, while ECDSA exhibits a more graceful degradation. The study highlights the advantages of using elliptic curves in terms of reduced key sizes and the potential for ECDSA to become a preferred solution as security requirements increase. The paper's uniqueness is based on the impact of parallel processing on the performance of both cryptosystems. Future work involves further investigating the observed performance patterns and examining the implications of post-quantum resistance.

Machine Learning in Connected Vehicles Environment

Jean Bezerra, Rawa Adla
Department of Computer Engineering, Florida Polytechnic University Lakeland, Florida, USA

Abstract—Vehicle trajectory prediction is an important task that can ensure high safety performance in collision avoidance systems and applications of autonomous vehicles. Machine learning methods for vehicle trajectory prediction have been proposed in recent literature in an effort to increase traffic safety by accurately predicting vehicle trajectories in various traffic scenarios. This paper presents a comprehensive review of recently proposed methods of vehicle trajectory prediction that utilize machine learning (ML) techniques. Moreover, it proposes a novel neural network framework for vehicle trajectory prediction that aims to be implemented in a connected vehicles (CV) environment.

FPGA-Accelerated Password Cracking

Eric Britten, Mikhail Gofman, Yu Bai
Department of Computer Science, California State University, Fullerton, California, USA;
Department of Electrical and Computer Engineering, California State University, Fullerton, California, USA

Abstract—Passwords are a common way of securing systems and applications from unauthorized access. However, passwords can be vulnerable to attackers who try to crack them by using random guesses, common patterns (e.g., password topologies), dictionary words, or leaked passwords. In this paper, we propose a hardware-accelerated password cracking method that leverages field-programmable gate array (FPGA) technology to recover passwords hashed with the sha512crypt algorithm. This algorithm is widely used on Linux systems. Our approach focuses on fast development to simulate a casual attacker who wants to boost their password cracking performance by using the speed and parallelism of FPGAs, which can outperform traditional CPUs and GPUs. Using C++ high-level language (a technique called high-level synthesis), we created a hardware device (i.e., a core) that runs the sha512crypt algorithm on the Zynq Z-7020 CPU-FPGA hybrid chip. We then tested the password cracking speed of our core (passwords/sec) and compared it to the AMD Ryzen 9 and Apple M1 Max CPUs. Based on the results of one core and two parallel sha512crypt cores on our chip, the maximum that could fit, we estimated that an FPGA chip with more than 8 sha512crypt cores and the same parameters as our chip could crack 10-character passwords at 360 passwords/sec, which is faster than a singlethreaded sha512crypt on either CPU. We also projected that with 15 cores, we could achieve a speed of 675 passwords/sec, which is twice as fast as either of the CPUs. We think this work is a useful addition to the research on cybersecurity applications of FPGAs, as few works have tried to break sha512crypt on the FPGA. In the future, we will implement our design on modern FPGA chips that can accommodate more than 10 cores, optimize our designs and evaluation methodology, and perform a more extensive evaluation.

The 21st International Conference on Software Engineering Research & Practice (SERP'23: July 24-27, 2023; Las Vegas, USA)

https://american-cse.org/csce2023/conferences-SERP
https://www.american-cse.org/csce2023/

Framework and Methodology for Verification of a Complex Scientific Simulation Software, Flash-X

Akash Dhruv, Rajeev Jain, Jared O'Neal, Klaus Weide, Anshu Dubey
Mathematics and Computer Science Division, Argonne National Laboratory, Lemont, Illinois, USA;
Department of Computer Science, University of Chicago, Chicago, Illinois, USA

Abstract—Computational science relies on scientific software as its primary instrument for scientific discovery. Therefore, similar to the use of other types of scientific instruments, correct software and the correct operation of the software is necessary for executing rigorous scientific investigations. Scientific software verification can be especially difficult, as users typically need to modify the software as part of a scientific study. Systematic methodologies for building test suites for scientific software are rare in the literature. Here, we describe a methodology that we have developed for Flash-X, a community simulation software for multiple scientific domains, that has composable components that can be permuted and combined in a multitude of ways to generate a wide range of applications. Ensuring sufficient code coverage by a test suite is particularly challenging due to this composability. Our methodology includes a consideration of trade-offs between meeting software quality goals, developer productivity, and meeting the scientific goals of the Flash-X user community.

VADER-SC: A Model Agnostic Tool for Large Scale, AI Driven Automated Source Code Summarization

Dwight Horne, Anthony Pierson, Elvis Hedary, Garrett Freddo, Luis Trejo, Mark Matis, Lonnie Mask
Agile Dev. Engineering, L3Harris Technologies, Greenville, Texas, USA;
Deptartment of Computer Science and Engineering, Texas A&M University, College Station, Texas, USA

Abstract—Production of a natural language description for the function of a source code segment is commonly referred to as source code summarization. Useful comments in source code can facilitate more rapid onboarding of new engineers and contribute to decreased maintenance costs. Unfortunately, the documentation task can also be labor intensive. In this paper, we introduce a new model agnostic tool for AI driven automation of source code summarization at scale. The initial version of the adVanced AI Driven Enhancement to Readability for Source Code (VADER-SC) software offers numerous options for customization and the ability to leverage a variety of AI models to enable experimentation in resource constrained environments, while also scaling up to benefit from larger models in contexts with increased compute resources. It further supports private cloud, self-hosted, and airgapped network configurations for environments with strict intellectual property protections or processing of sensitive or controlled data. Qualitative and quantitative results suggest model selection, fine-tuning, and multi-shot tailoring significantly impact the quality of generated comments. VADER-SC could be an enabler for practitioners to explore large-scale automation of AI driven source code summarization and researchers may find it enables studies with larger volumes of disparate data across a diversity of AI model options and target programming languages.

Enhancing Ropes for Collaborative Text Editing

Semih Sandal, Tolga Ovatman
ASELSAN A.S., Ankara, Turkey;
Department of Computer Engineering, Istanbul Technical University, Istanbul, Turkey

Abstract—Rope is a tree-based data structure providing performance enhancements in string concatenation operations in applications such as text editors. Even though collaborative text editing has become one of the major applications in cloud environments, utilizing ropes in collaborative text editing is rarely discussed in current literature. In this study, we enhance the Rope data structure for collaborative text editing. Firstly we discuss possible issues that Ropes may suffer in collaborative text editing and propose enhancements that adapt Ropes to such environments. During this adaptation we pay special attention to keep Ropes' immutable nature for the insert operations and balance the tree by following AVL-tree-style and Red-Black-Treestyle methods. We also provide a performance comparison of those balancing approaches. In our results, it can be seen that the AVL-Rope and Red-Black-Rope are slower in the concatenation operation due to balancing, but faster in index search operation and insert operation than the unbalanced Rope.

Creating Continuous Improvement in Agile Software Development Using Lean Six Sigma

Eric Malvar, Ning Chen
Department of Computer Science, California State University, Fullerton, California, USA

Abstract - While Agile development models have been refined and molded with other development models, only a small amount of research exists on Agile improvement through the assistance of Lean Six Sigma. The goal of this research study is to investigate possible increases in throughput and reduction of waste within current Agile processes using a variety of Lean Six Sigma tools. This paper discusses both current implementation in industry, as well as research being done on the merging of the two ideologies. Our study indicates that Lean Six Sigma is critical in initiating standardization of Agile practices and can also reduce time spent during development and maintenance within the Agile lifecycle.

On the Interoperability of Programming Languages via Translation

Dante Broggi, Yi Liu
Department of Computer & Information Science,
University of Massachusetts Dartmouth, Massachusetts, USA

Abstract—Interoperability between systems written in different programming languages can be challenging. While C offers some degree of interoperability, it is an unsafe language and does not encourage consideration of abstractions independently from their implementations. This study aims to explore the possibility of extending general interoperability beyond C by evaluating the current suitability of languages and tools for transpilation, which refers to compiling code into another source language instead of machine code. Specifically, we focus on transpiling the runtime of one language into a target language. This study aims to preserve safety invariants between Pony and Rust, two programming languages that have similar but distinct high-level language semantics. Two phases are involved in translating the Pony runtime into Rust: improving the C source to better represent the invariants as permitted by C and addressing the inadequacies identified in the tooling, and refactoring the resulting Rust code to identify and correct any errors caused by misuse. As a result, this study has developed a methodology for implementing the runtime of one language in another and applied it to a case study for translating the Pony runtime into Rust. Through this translation process, the study has identified improvements to enhance Rust's suitability for future work, reported the unsafety issue in the Pony runtime, and highlighted insufficiencies in several other tools.

Systematic Behavioral Design Algorithms

Devon M. Simmonds
University of North Carolina Wilmington, Wilmington, North Carolina, USA

Abstract— Algorithms to support systematic software design are needed to enable practitioners to mechanize and coherently manage the engineering of quality software. In practice, design is both an art and a science, and this dichotomy has led some to under-focus on the science and engineering side of software development. In our previous research, we presented a design algorithm for structural design focused on UML use case and requirements class diagrams. In this paper we shift focus to systematic behavioral design algorithms illustrated by UML state and activity diagrams. We present the systematic behavioral design algorithms, their application to a specific case study, as well as contingent results, implications, and lessons learned.

Formal Modeling and Verification of Timed Connectors in IoT with Z3

Ziyun Xu, Meng Sun
School of Mathematical Sciences, Peking University, Beijing, China

Abstract—The Internet of Things (IoT) is rapidly advancing and reshaping the whole world. Coordination models and languages, like Reo and Orc, provide connectors that interconnect components in IoT applications and organize their interactions in distributed environments. In this paper, we propose a method for formally modeling and verifying the properties of timed connectors using Z3, an SMT (Satisfiability Modulo Theories) solver. We use Z3 Python-bindings to construct the models and carry out experiments. The formal model in Z3 clearly reflects the original structure of connectors. With the definition in Z3, we can automatically verify time-related properties of connectors, and automatically construct counter-examples when the properties do not hold.

Augias: a Model-based Software Development Kit for High-Performance Collaborative and Multiphysics Scientific Computing

Didier Nassiet, Valentin Gaisset, David Lugato
French Alternative Energies and Atomic Energy Commission (CEA), France;
CEA / CESTA Le Barp - France

Abstract—The design of simulation studies for HPC is a complex and time-consuming activity, especially when developing multiphysics designs of experiments. Using integrated environments limits the simulation study management efficiency, interoperability, and traceability in a multi-user context. This paper presents the Augias toolkit, a collaborative approach to managing HPC data and processes which addresses these issues by providing complete and neutral modeling of the simulation domain and a service layer end users can extend directly. The Augias metamodels and services are detailed in addition to the technology stack required to implement its principles and architecture. In conclusion, we present the results of an empirical evaluation study confronting the Augias SDK with three other solutions and scoring them through a canonical use case.

Efficacy of Reported Issue Times as a Means for Effort Estimation

Paul Maclean, Joshua Hernandez, Dale Dzielski, Tom Devine
West Virginia University, West Virginia, USA

Abstract—Software effort estimation helps project managers monitor their team's timelines and increase software reliability. Issue Tracking Systems (ITS) typically have a time spent field, where developers document their time dedicated towards resolving an issue (ITS effort). This paper proposes ITS effort estimation and prediction approaches and applies them to realworld repository

data from five Apache projects. We explore correlations between ITS effort and implementation metrics, with results ranging from weak to moderate strength. Our predictions yielded non-informative models (median F-measures around 0.33 for N=3 classes, and 0.25 N=5 classes), with low statistical power due to data quantity. Higher rates of ITS effort documentation would improve the certainty of our models in future work. Our dataset, the ITS Code Mining Database (ICMD), is publicly available.

Interstices In the Certification of Safety Critical Avionics Software: Boeing 737-MAX MCAS Case Study

Aiman Gannous
Computer Science, University of Benghazi, Benghazi, Libya

Abstract—Two Boeing planes of the same model were crashed in October 2018 and in March 2019. All the passengers and crew onboard were killed. Investigations showed that an external failure in the newly installed critical software Maneuvering Characteristics Augmentation System (MCAS) was responsible of these two fatal accidents. In this paper we reviewed the reasons behind the failure of MCAS that installed in Boeing most selling airplane and investigated engineering and certifying such software systems. The case study revealed an urgent need to clarify the certification elements in a more practical way and to be linked to advancements in software engineering.

A Cost-Effectiveness Metric for Association Rule Mining in Software Defect Prediction

Kinari Nishiura, Takeki Kasagi, Akito Monden
Graduate School of Natural Science and Technology, Okayama University, Okayama, Japan

Abstract—This paper introduces a cost-effective metric for association rule mining in software defect prediction, where defective module conditions are represented as association rules. The metric calculates the expected number of defects discovered in modules that satisfy an association rule, considering a limited test effort or test cases. Since conducting full testing on all modules is impractical due to resource limitations, the proposed metric helps identify the most cost-effective modules to test within given constraints. It is based on the extended exponential Software Reliability Growth Model (SRGM), incorporating the module size parameter to account for the increased effort required to detect defects in larger modules. To evaluate the metric's effectiveness, association rules were extracted and prioritized using the proposed metric across datasets from four open source software projects. The assessment employed the widely used LOC-based cumulative-lift chart to measure the cost-effectiveness of defect prediction. The findings demonstrate that the proposed metric effectively prioritizes rules with a higher potential for defect discovery compared to conventional association rule metrics like confidence and odds ratio.

Interoperability Open Architecture of Unmanned Systems

Wilmuth Muller, Florian Segor, Dirk Muhlenberg, Bart Driessen, Vincent de Geus
Interoperability and Assistance Systems, Fraunhofer IOSB, Karlsruhe, Germany;
TNO, The Hague, The Netherlands

Abstract—Within the INTERACT Project, funded by the DGDEFIS of the European Commission and managed by the European Defence Agency (EDA), interoperability concepts aiming at enhancing the capabilities of European armed forces to safely, effectively and flexibly operate unmanned and manned systems in joint or combined operations have been developed. The challenge lied in creating overarching interoperability concepts for defence systems in general and unmanned systems in particular. The different interoperability concepts have been integrated into an open interoperability architecture for unmanned systems. In this paper the developed open interoperability architecture is presented.

Developing Microgrid Agents Using Smart Python Agent Development Environment (SPADE)

Salem Al-Agtash, Xueyan Xian, Kristina Petroshchuk, Heba Syeddah, Hima Bindu Mallampati
Department of Computer Science and Information Technology, Mission College, Santa Clara, California, USA;
Department of Computer Science and Engineering, Santa Clara University, Santa Clara, California, USA;
Department of Computer Engineering, German Jordanian University, Amman, Jordan

Abstract—Microgrids are revolutionary power systems that interconnect a mix of renewable power generation, load, storage systems, and inverters in a small-scale grid network. Operating microgrids while maintaining a consistent grid voltage and frequency during the islanding and disruption of renewables has been a challenging research problem. In this paper, we present a preliminary microgrid agent implementation using SPADE (Smart Python Agent Development Environment) as a powerful development framework that has been used extensively in many application domains. Agents autonomously manage and operate microgrid individual components. We model a multiagent microgrid system to seamlessly operate and optimize energy balance by coordinating the actions of agents. Agents coordinate energy demand, available renewable generation, storage charging and discharging patterns, and grid energy prices. They can detect anomalies, diagnose faults, balance generation and load, and maintain adequate network voltage and frequency.

Data-Driven Requirements Verification in a Tool to Support the Cybersecurity Risk Management Process in Maritime Transportation Ecosystem

Mark McKenzie, Yanzhen Qu
Colorado Technical University, Colorado Springs, Colorado, USA

Abstract— Commercial maritime transportation is a critical infrastructure for countries and the global economy. Currently, the industry transports 80-90% of all international trade. However, the industry has traditionally focused on physical security and lacks a web-based distributed system for efficient cybersecurity risk management. This is challenging due to the diverse range of companies and organizations involved, each with different cybersecurity expertise and technologies and no common ownership. We have proposed a solution of developing a web-based distributed system called MTcyber RMPDS, to support the cybersecurity risk management processes in maritime transportation ecosystem. This paper explores the quantitative verification of the requirements represented by the use cases of various service requests sent to MTcyber RMPDS, aiming to improve its design efficiency.

Assessment of ChatGPT's Proficiency in Software Development

Dae-Kyoo Kim, Jingshu Chen, Hua Ming, Lunjin Lu
Department of Computer Science and Engineering, Oakland University, Rochester, Michigan, USA

Abstract—This paper presents an assessment of ChatGPT's proficiency in software development, using an online tour reservation system (TORS) as a case study. The findings indicate that ChatGPT has significant potential in software development, demonstrating its capabilities in assisting various activities throughout the development process, including requirements analysis, domain modeling, design modeling, and implementation. Notably, the model performed well in implementation, generating more than 90% of the code and fixing a majority of errors. It also demonstrated its capability in making design decisions in the design phase. However, the study also identified non-trivial limitations, such as a lack of traceability and inconsistencies among produced artifacts, which required human involvement. Overall, the results suggest that when combined with human developers, ChatGPT can serve as a valuable tool in software development. It has the potential to enhance productivity and efficiency in various aspects of the development process, while acknowledging the need for human expertise to mitigate the limitations.

Integrating Sample Iterative Communication Model into Project Management Life Cycle Phases

Stephanie Onyemowo Omakwu, Lei Chen
Department of Information Technology,
Georgia Southern University, Statesboro, Georgia, USA

Abstract - Quite a few projects are faced with the challenges of not achieving stakeholders' goals or the entire project failing due to reasons such as unclear or changing goals, issues with communication, inadequate risk management, insufficient follow-up, using wrong or too many tools and challenges of context and timing. One of the major causes of all these reasons include inappropriate communication, delay in communication, using the wrong method and tool or no communication at all which is why this study is proposing an integration of the Sample Iterative Communication Model(SICM) in all the phases of project management (project initiation, planning, execution, monitoring and controlling and the completion/closing phase) to tackle the communication challenges faced during the project life cycle.

Semantic Blockchain Software Tools and Services for Trustworthy Applications - ONTOCHAIN

Klevis Shkembi, Petar Kochovski, Thanasis G. Papaioannou, Caroline Barelle,
Anthony Simonet-Boulogne, Marco Ciaramella, Vlado Stankovski
University of Ljubljana, Ljubljana, Slovenia;
Athens University of Economics and Business, Athens, Greece;
European Dynamics, Luxembourg;
iExec SARL, Lyon, France;
IntelliSemantic srl Rivoli, Italy

Abstract—In the last years, we have seen the rise of Blockchain technology as a new way that enables trust and fair exchange of digital and physical assets among different actors, however, its usability in software engineering remains unexplored. In this context, various stakeholders have expressed interest in combining Blockchain technology with other well-known technologies such as the Semantic Web. Blockchain and Semantic Web technologies are both about building trust, and both represent some shared conceptualization. Hence, we explore new approaches to achieving trust in highly decentralized, semantically complex, and dynamic environments. In this paper, we propose a new software framework called ONTOCHAIN, a set of ontologies, blockchains, tools, and services that combine semantic web and blockchain technologies and may be used to build trustworthy decentralized applications (dApps). It addresses challenges faced by citizens in online services, such as unknown provenance of information, service provider quality, fake news, fraud, and privacy violations. ONTOCHAIN offers high-level application protocols for data provenance, reputation models, decentralized oracles, and more. It also includes core protocols for authorization, certification, privacy-aware data processing, and identity management. We demonstrate the feasibility of implementing Next Generation Internet dApps using these protocols, enhancing trust and security in online services.

From Vulnerabilities to Improvements: A Deep Dive into Adversarial Testing of AI Models

Brendan Hannon, Yulia Kumar, Peter Sorial, J. Jenny Li, Patricia Morreale
Kean University, Union, New Jersey, USA

Abstract—The security vulnerabilities inherent in large language models (LLMs), such as OpenAI's ChatGPT-3.5 and ChatGPT-4, Bing Bot, and Google's Bard, are explored in this paper. The focus is on the susceptibility of these models to malicious prompting and the potential for generating unethical content. An investigation is conducted into the responses these models provide when tasked with completing a movie script involving a character disseminating information about murder, weapons, and drugs. The analysis reveals that, despite the presence of filters designed to prevent the generation of unethical or harmful content, these models can be manipulated through malicious prompts to produce inappropriate and even illegal responses. This discovery

underscores the urgent need for a comprehensive understanding of these vulnerabilities, as well as the development of effective measures to enhance the security and reliability of LLMs. This paper offers valuable insights into the security vulnerabilities that arise when these models are prompted to generate malicious content.

Events-Based Test Suite Reduction for Mobile App Test Suites Generated by Reinforcement Learning

Abdullah Alenzi, Waleed Alhumud, Renee Bryce, Ryan Michaels, Md Khorrom Khan
Department of Computer Science and Engineering, University of North Texas, Denton, Texas, USA;
Department of Computer Science, St. Edward's University, Austin, Texas, USA

Abstract—Reinforcement learning is promising for automated test generation. However, a current shortcoming of these algorithms is that the exploration process may result in extra test cases that have duplicate coverage of events within a test suite. Fine tuning parameters of reinforcement algorithms may help, but this comes with trade-offs and requires time consuming and careful consideration of the characteristics of the application under test and its environment. This work takes a different approach. Instead of exploring parameters of reinforcement algorithms, we look at reducing test suites that have already been generated. Specifically, we use test suites that were generated with the SARSA algorithm and then apply a greedy test suite reduction algorithm that uses an event coverage criterion. Results show that test suite reduction results in 10.74% to 50% reduction in test suite sizes while maintaining the same code coverage of the original test suites. The results motivate that redundant event coverage should be considered during and/or after test generation.

Optimal Solution Through Fast Convergence for Transportation of Shipping Wood in Los Angeles

Chandrasekhar Putcha, Rakeshkumar Mahto, Sapankumar Patel, Ramesh Battu, Andy Iniguez
Department of Civil and Environmental Engineering, California State University, Fullerton, California, USA;
Department of Electrical and Computer Engineering, California State University, Fullerton, California, USA

Abstract— In this paper, we introduce an optimization approach for transportation problems that are presented as linear programming problems. Our proposed methods provide an optimal solution with faster convergence, eliminating the requirement for chain reaction or transportation simplex methods. To demonstrate the effectiveness of our developed optimization methods, we provide an illustrative example in this paper. Additionally, our approach can be effortlessly integrated with the supply chain, facilitating the optimization of large-scale problems in the fields of mechanical and industrial engineering. The proposed Putcha Bhuiyan method significantly reduces the repetitive process of chain reactions and iterations, resulting in considerable time savings. The proposed Putcha Bhuiyan method produces a value of 'Z' that is comparable to, and in some cases, even better than, traditional methods. This approach can improve the effectiveness and precision of solving transportation issues, providing valuable insights for professionals and researchers in the industry. This method showcases its potential to enhance the efficiency and accuracy of solving transportation problems, thereby offering valuable insights for researchers and practitioners in the field.