Artificial Intelligence at RIT Summit

Book of Abstracts

AI@RIT

www.rit.edu/aisummit

6 - 7 October, 2022
Rochester, New York

RIT
Rochester
Institute of
Technology

Program

Thursday, October 6, 2022

Tutorials

1:30-2:30 p.m.
Dr. Ernest Fokoue, Professor
School of Mathematical Sciences
College of Science
Discovering some of the Fundamental Building Blocks of Artificial Intelligence via Statistical Recognition of Speaker Accent

2:30-3:30 p.m.
Dr. Qi Yu, Professor
School of Information
Golisano College of Computing and Information Sciences
Building Integrated Human-Machine Intelligence through Human-In-The-Loop Learning

3:30-4:30 p.m.
Dr. Alexander Ororbia, Assistant Professor
Department of Computer Science
Golisano College of Computing and Information Sciences
Brain-Inspired Computing: Towards Neurobiologically-Grounded Credit Assignment

Friday, October 7, 2022

9:00-9:30 a.m.
Arrival and registration

Welcome speech by
Provost Ellen Granberg

9:30-10:00 a.m.
Remarks by Organizing Committee
Dr. Andreas Savakis, AI@RIT Summit Co-Chair; Professor,
Kate Gleason College of Engineering
Dr. Pengcheng Shi, AI@RIT Summit Co-Chair; Professor,
Golisano College of Computing and Information Sciences

Program (continued)

10:00-11:00 a.m. Keynote lecture by
Dr. Cecilia Alm, Professor
College of Liberal Arts
Humans in Artificial Intelligence: Human-centered AI and Preparing the AI Workforce

11 a.m.-12:00 p.m. Keynote lecture by
Dr. Linwei Wang, Professor
Golisano College of Computing and Information Sciences
Challenges and Opportunities in Health-AI

12:00-1:00 p.m. Lunch and social

11:30 a.m.-12:45 p.m. Opportunity and invitation to poster presenters to set up their posters

1:00-2:30 p.m. Poster session (featuring both posters and a few demonstrations on the big screen)

2:30-3:30 p.m. **Panel discussion**
Artificial Intelligence at RIT (AI@RIT): Past, Present and Future

Moderators:
Dr. Andreas Savakis, Professor, Kate Gleason College of Engineering
Dr. Pengcheng Shi, Professor, Golisano College of Computing and Information Sciences

Panelists:
Dr. Alexander Ororbia, Assistant Professor, Golisano College of Computing and Information Sciences
Dr. Gregory Babbitt, Associate Professor, College of Science
Dr. Andres Kwasinski, Professor, Kate Gleason College of Engineering
Dr. Jason Nordhaus, Associate Professor, National Technical Institute for the Deaf
Dr. Flip Phillips, Professor, Motion Picture Science and Imaging Science
Dr. Ernest Fokoue, Professor, College of Science

3:30-4:00 p.m. Questions and answers related to the panel discussion

4:00 p.m. Closing remarks

Organizing Committee

Ryne Raffaelle
Vice President for Research
VP for Research Office
Research
rprsps@rit.edu

Pengcheng Shi
AI@RIT Summit Co-Chair
Associate Dean for Research and Scholarship/PhD Program Director, Department of Computing and Information Sciences, Golisano College of Computing and Information Sciences
spcast@rit.edu

Cecilia Alm
Professor
Department of Psychology
College of Liberal Arts
cecilia.o.alm@rit.edu

Cathy Hain
Associate Vice President
University Advancement
Development
University Advancement
cathy.hain@rit.edu

David Long
Director of RIT MAGIC Center
MAGIC Center
Research
david.long@rit.edu

Andreas Savakis
AI@RIT Summit Co-Chair
Professor
Department of Computer Engineering
Kate Gleason College of Engineering
andreas.savakis@rit.edu

Technical Committee

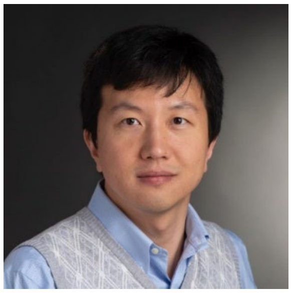
Ernest Fokoue
Professor
School of Mathematical Sciences
College of Science
epfeqa@rit.edu

Xumin Liu
Professor
Department of Computer Science
Golisano College of Computing and Information Sciences
xmlics@rit.edu

Jason Nordhaus
Associate Professor
Department of Science and Mathematics
National Technical Institute for the Deaf
jtnsma@rit.edu

Andres Kwasinski
Professor
Department of Computer Engineering
Kate Gleason College of Engineering
ak@mail.rit.edu

Flip Phillips
Professor
School of Film and Animation
College of Art and Design
flip.phillips@rit.edu



Rui Li
Assistant Professor
Department of Computing and Information Sciences
Golisano College of Computing and Information Sciences
rxlics@rit.edu

Bharat Bhole
Associate Professor
Department of Economics
College of Liberal Arts
blbgse@rit.edu

Table of Contents

Page

AI APPLICATIONS:

AI ENABLERS (SW/infrastructure/data):

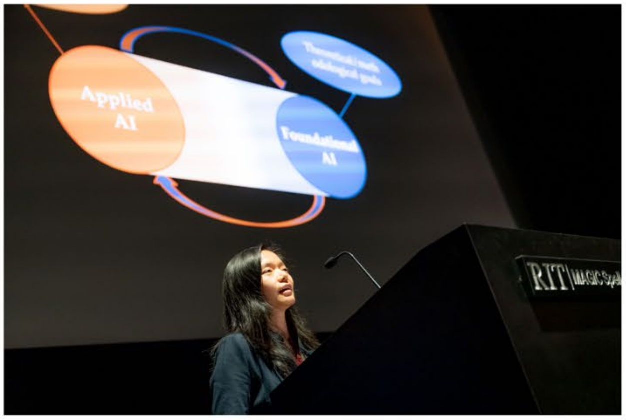
AI ETHICS AND POLICY:

AI IMPACTS ON ACADEMIA, INDUSTRY, AND SOCIETY:

CORE AI THEORY:

OTHER:

returned to
rning
Learner
Estimations
Labeled instances returned to learner
Machine Teaching
Agent Teacher
RIT | MAGIC Spell Studios


Applied AI
Foundational AI
RIT | MAGIC Spell

Categorization and consolidation of data

PREDICTIVE MODELING AND MEDICATION ADHERENCE
Teresa B. Gibson, PhD, School of Mathematical Sciences
BACKGROUND
METHODS AND DATA
MODELING APPROACH
IMPLICATIONS
RIT

Detecting Fake Review Buyers Using Network Structure: Direct Evidence from Amazon

Ali Tosyali and Gijs Overgoor

Emails: *atosyali@saunders.rit.edu; govergoor@saunders.rit.edu*

Online reviews significantly impact consumers' decision-making process and firms' economic outcomes and are widely seen as crucial to the success of online markets. Firms, therefore, have a strong incentive to manipulate ratings using fake reviews. This presents a problem that academic researchers have tried to solve over two decades and on which platforms expend a large amount of resources. Nevertheless, the prevalence of fake reviews is arguably higher than ever. To combat this, we collect a dataset of reviews for thousands of Amazon products and develop a general and highly accurate method for detecting fake reviews. A unique difference between previous datasets and ours is that we directly observe which sellers buy fake reviews. Thus, while prior research has trained models using lab-generated reviews or proxies for fake reviews, we are able to train a model using actual fake reviews. We show that products that buy fake reviews are highly clustered in the product-reviewer network. Therefore, features constructed from this network are highly predictive of which products buy fake reviews. We show that our network-based approach is also successful at detecting fake reviews even without ground truth data, as unsupervised clustering methods can accurately identify fake review buyers by identifying clusters of products that are closely connected in the network. While text or metadata can be manipulated by sellers to evade detection, network-based features are very costly to manipulate because these features result directly from the inherent limitations of acquiring reviews from fake review marketplaces, making network-based fake review detection approaches robust to manipulation.

Topic: AI Applications

AntiCopyPaster: Extracting Code Duplicates As Soon As They Are Introduced

Le Nguyen, Amit Kini, Aditya Thacker, and Mohammed Wiem Mkaouer

Emails: *ln8378@g.rit.edu; ak3328@rit.edu; at4415@rit.edu; mwmvse@rit.edu*

We developed a plugin for IntelliJ IDEA called AntiCopyPaster, which tracks the pasting of code fragments inside the IDE and suggests the appropriate Extract Method refactoring to combat the propagation of duplicates. Our tool is integrated with the developer's workflow, and proactively recommends refactorings. Since not all code fragments need to be extracted, we treat the decision of whether to perform the extraction, as a binary classification problem. We leverage developers' previous actions in code, to train a deep learning model that would learn the characteristics of code that is found to be worth extracting. Our experimental study on a large dataset of 18,942 code fragments mined from 13 Apache projects shows that AntiCopyPaster correctly recommends Extract Method refactorings with an F-score of 0.82.

Topic: AI Applications

Divergent Thinking with AI-Collaborative Product Design Strategies

Bryce Beamer

Email: *bxbfaa@rit.edu*

AI and machine learning have made a major impact in my career through algorithm development for sports monitoring, image classification for agricultural pollinators, and generative design of lattice structures. Each of these applications are practical, technical implementations of AI that are essential to the end functionality of innovative products. We will continue to see growth in how AI contributes functionality to the products we create; however, recent advances suggest that AI will impact the creative design process in early-stage concept development as well.

Diffusion models are being leveraged in text-to-image and image-to-image generation with incredible results. As a result, we have already seen AI-generated images creating compelling imagery that is even winning art competitions. The applications of this technology for illustration and fine arts are evident, but how this technology can be leveraged in the creation of physical products is not yet well defined.

I have been exploring how text-to-image diffusion models can be leveraged as part of divergent exploration that happens in the early stages of product design explorations. One traditional way of doing this is by having one student listen to another student's design intent and then create a concept sketch that embodies what they interpret. In my junior-level studio class, we are currently exploring how AI can be used to help think divergently through the interpretation or misinterpretation of an outlined concept. Students are inventing new ways of communicating with AI to create collaborative design solutions.

Topics: AI Applications, AI Impacts on Academia, Industry, and Society, Other

On the Limitations of Continual Learning for Malware Classification

Mohammad Saidur Rahman, Scott Coull and Matthew Wright

Emails: *saidur.rahman@mail.rit.edu; scott.coull@mandiant.com; matthew.wright@rit.edu*

Malicious software (malware) classification offers a unique challenge for continual learning (CL) regimes due to the volume of new samples received on a daily basis and the evolution of malware to exploit new vulnerabilities. On a typical day, antivirus vendors receive hundreds of thousands of unique pieces of software, both malicious and benign, and over the course of the lifetime of a malware classifier, more than a billion samples can easily accumulate. Given the scale of the problem, sequential training using continual learning techniques could provide substantial benefits in reducing training and storage overhead. To date, however, there has been no exploration of CL applied to malware classification tasks. In this paper, we study 11 CL techniques applied to three malware tasks covering common incremental learning scenarios, including task, class, and domain incremental learning (IL). Specifically, using two realistic, large-scale malware datasets, we evaluate the performance of the CL methods on both binary malware classification (Domain-IL) and multi-class malware family classification (Task-IL and Class-IL) tasks. To our surprise, continual learning methods significantly underperformed naive Joint replay of the training data in nearly all settings -- in some cases reducing accuracy by more than 70 percentage points. A simple approach of selectively replaying 20% of the stored data achieves better performance, with 50% of the training time compared to Joint replay. Finally, we discuss potential reasons for the unexpectedly poor performance of the CL techniques, with the hope that it spurs further research on developing techniques that are more effective in the malware classification domain.

Topic: AI Impacts on Academia, Industry, and Society

Applying Deep Learning for Website Fingerprinting Attacks

Nate Mathews

Email: *njm3308@gmail.com*

Tor, the most popular anonymous networking service, prevents overt disclosure of the link between servers and clients by routing traffic through chains of relays, called circuits, with layered encryption. The privacy of Tor may, however, be undermined by traffic analysis attacks. One such attack, known as Website Fingerprinting (WF), allows a passive local eavesdropper to deduce information about Tor-protected traffic using traffic metadata. In a WF attack, the adversary seeks to determine what website a Tor user has visited in a browsing session through eavesdropping on the network traffic. Using patterns in this information, usually with the aid of machine learning, an adversary can often unmask the client's activities. In this poster, we review the state-of-the-art deep-learning techniques that we have applied to this domain to achieve effective attacks.

Topics: AI Applications

Framing TRUST in Artificial Intelligence (AI) Ethics Communication

Namrata Nagar

Email: *ngrnamrata@gmail.com*

With the fast proliferation of Artificial Intelligence (AI) technologies in our society, several corporations, governments, research institutions, and NGOs have produced and published AI ethics guiding documents. These include principles, guidelines, frameworks, assessment lists, training modules, blogs, and principle-to-practice strategies. The priorities, focus, and articulation of these innumerable documents vary to different extents. Though they all aim and claim to ensure AI usage for the common good, the actual AI system outcomes in various social applications have invigorated ethical dilemmas and scholarly debates. This study presents the analysis of AI ethics principles and guidelines text published by three pioneers from three different sectors - Microsoft Corporation, National Institute of Standards and Technology (NIST), AI HLEG set up by the European Commission through the lens of media and communication’s Framing Theory. The TRUST Framings extracted from recent academic AI literature are used as standard construct to study the ethics framings in the selected text. The institutional framing of AI principles and guidelines shapes the AI ethics of an institution in a soft (as there is no legal binding) but strong (incorporating their respective position/societal role’s priorities) way. The AI principles’ framing approach directly relates to the AI actor’s ethics that enjoins risk mitigation and problem resolution associated with AI development and deployment cycle. Thus, it has become important to examine institutional AI ethics communication. This paper brings forth a Comm-Tech perspective around the ethics of evolving technologies known under the umbrella term - Artificial Intelligence and the human moralities governing them.

Topic: AI Ethics and Policy, AI Impacts on Academia, Industry, and Society

Machine Learning in Bio-fluid Mechanics Applications

Xudong Zheng and Qian Xue

Emails: *xxzeme@.rit.edu; qian.xue@.rit.edu*

Our research focuses on integrating machine learning algorithms with multi-physics simulation environments to improve the understanding of flow-related biophysics in human and animal. Currently, we are exploring three research applications:

(1) We are developing an AI-assisted data assimilation framework to support voice research. The goal is to enable seamless integration of multimodal experimental/clinical data and high-fidelity subject-specific modeling of human/ animal vocal systems through physics-informed neural network (PINN) and eventually enable efficient and accurate simulations of high-dimensional dynamics of individual vocal systems to support personalized voice care. In our pilot study, we demonstrated the capability of PINN to predict 3D vocal fold dynamics from sparse and 2D endoscopic images of vocal fold vibration in flow-structure interaction (FSI) simulation environment. If successful, the approach can have broad applications in systems that involve FSI.

(2) We are applying deep reinforcement learning (DRL) to study strategies of optimal performance in bio-inspired propulsion. A novel asynchronous DRL algorithm is under development for efficient learning of optimal stiffness tuning strategy for maximum thrust in fish ray-fin structure propulsion. The pilot results revealed a tuning strategy that achieves the thrust notably higher than the maximum value achieved with the action series using parametric search.

(3) We are developing interpretable machine learning models for studying hydrodynamic flow sensing in seal whisker. We are currently exploring various machine learning methods for revealing the causal relationships between mechanical signals of seal whiskers and surrounding environments, e.g., flow structures, upstream disturbance/prey characteristics.

Topic: AI Applications

AI-assisted comparative molecular dynamics analyses of protein function

Gregory Babbitt

Email: *gabsbi@rit.edu*

Most bioinformatics analyses are focused upon databased DNA/protein sequences or protein structures. However, biological function is most often achieved at the level of soft matter dynamics played out in the functional vaccine motions of a protein. A large fraction of protein motion is random thermal noise caused by interaction with molecules in the surrounding aqueous solvent. Therefore, comparing functional protein motions and identifying where on a protein they are functionally conserved or retained over evolutionary timescale is a considerable computational, statistical, and intellectual challenge; one well-suited for a machine learning approach.

DROIDS/maxDemon v5.0 is a suite of machine learning assisted statistical methods for comparing computer simulated molecular dynamic trajectories of proteins in two functional states (e.g. unbound vs. bound to something or wild type vs mutated or hot vs cold). It also identifies regions of functionally conserved dynamics as well as regions of coordinated conserved dynamics between sites on a single protein or two interacting proteins. It is currently under development on a python 3 science stack and only additionally requires the cpptraj library software and UCSF Chimerax molecular visualization software to be installed. The software is offered freely (without guarantee) under GPL 3.0 and was developed by Dr. Gregory A. Babbitt and bioinformatics students at the Rochester Institute of Technology in 2022.

DROIDS/maxDemon v1.0 - v4.0 have been utilized in numerous published studies from our lab on topics ranging from cancer drug function, SARS-virus evolution, and viral vaccine escape mechanisms. We are currently working on the problem of the function and evolution of multi-drug resistance in HIV patients.

Topic: AI Applications

AI and the Ethics of Autonomous Vehicles: Creating Insight from a Real Big Mess

Mike Palanski

Email: *mpalanski@saunders.rit.edu*

Driven by artificial intelligence (AI), autonomous vehicles hold great promise as the primary personal transportation mode of tomorrow. But before we can enjoy the promised land of safe, efficient, and effortless transport, there are a number of ethical issues that must be addressed. From AI-driven decisions about who lives and who dies in a crash to data usage and security, from privacy concerns to implementation challenges, and from creating supporting infrastructure to public policy concerns, ethical issues abound – and can be daunting. However, understanding these issues can also help us to create a viable path forward.

In this informative and insightful talk, Dr. Mike Palanski paints a picture of the big ethical mess that surrounds the notion of autonomous vehicles. But he also shows how understanding the mess opens the path forward to a better tomorrow. Whether you are a designer, engineer, businessperson, public official, or consumer, you will walk away feeling smarter and more confident about the world of autonomous vehicles.

You might also be mildly entertained. Maybe.

Topic: AI Ethics and Policy; AI Impacts on Academia, Industry, and Society

Breaking the Barrier of Limited Data for Deep Learning in Medical Imaging using Learned Object Level Data Augmentation

Nilesh Kumar and Linwei Wang

Emails: *nk4856@rit.edu; linwei.wang@rit.edu*

Deep learning (DL) has shown great promise in solving previously unsolved problems like predicting 3d protein structures and helping mathematicians establish new theorems. However, the applicability of DL in many fields, including medical imaging, is limited by a lack of labeled data. Obtaining labeled data can be time-consuming in medical imaging and requires expert knowledge. Data augmentation is one of the most promising techniques to solve this challenge. But, most of the current literature in data augmentation for medical imaging is focused on either pre-defined operations or learned transformations tied to pixel locations of an image. The inability to adjust transformations based on a region of interest means that the existing works are insufficient to harness the knowledge present in previously labeled datasets. We believe using that knowledge could help us add much-needed information to existing datasets without requiring any expert knowledge. To this end, we propose a two-step framework that first learns object-level transformations and then enhances the existing labeled set by applying these learned object-level augmentations. By tying the augmentations to objects rather than the pixel locations, we make sure that these augmentations can readily be transferred across different datasets containing the same objects of interest. Our framework also does not require any pre-defined operations as the transformations are learned from existing samples.

Topic: AI Applications, AI Enablers (SW/ infrastructure/data)

A First Look into Users' Perceptions of AI-Powered Facial Recognition in the Physical World

Sovantharith Seng, Mahdi Nasrullah Al-Ameen and Matthew Wright

Emails: *ss2816@g.rit.edu; mahdi.al-ameen@usu.edu; matthew.wright@rit.edu*

Online reviews significantly impact consumers' decision-making process and firms' economic outcomes and are widely seen as crucial to the success of online markets. Firms, therefore, have a strong incentive to manipulate ratings using fake reviews. This presents a problem that academic researchers have tried to solve over two decades and on which platforms expend a large amount of resources. Nevertheless, the prevalence of fake reviews is arguably higher than ever. To combat this, we collect a dataset of reviews for thousands of Amazon products and develop a general and highly accurate method for detecting fake reviews. A unique difference between previous datasets and ours is that we directly observe which sellers buy fake reviews. Thus, while prior research has trained models using lab-generated reviews or proxies for fake reviews, we are able to train a model using actual fake reviews. We show that products that buy fake reviews are highly clustered in the product-reviewer network. Therefore, features constructed from this network are highly predictive of which products buy fake reviews. We show that our network-based approach is also successful at detecting fake reviews even without ground truth data, as unsupervised clustering methods can accurately identify fake review buyers by identifying clusters of products that are closely connected in the network. While text or metadata can be manipulated by sellers to evade detection, network-based features are very costly to manipulate because these features result directly from the inherent limitations of acquiring reviews from fake review marketplaces, making network-based fake review detection approaches robust to manipulation.

Topic: AI Impacts on Academia, Industry, and Society; Other

Which to Optimize: On the Interdependence between Data Selection and Architecture Optimization in Deep Active Learning

Pradeep Bajracharya and Linwei Wang

Emails: *pb8294@rit.edu; linwei.wang@rit.edu*

As in many other areas, the success of active learning in traditional machine learning is being transferred to learning deep neural networks (DNNs) in the subarea of deep active learning (DAL). In this process, however, an important difference has been neglected – compared to traditional machine learning models, there is a large search space of functional complexity for a DNN with chosen architecture types; more importantly, the optimality of architecture types and parameters of a DNN is highly reliant on the underlying data. This interdependence between optimal architecture and data raises an interesting question unique to DAL: how does the DNN architecture affect optimal data selection, and vice versa. This question has not been considered in DAL literature, and the vast majority of existing works is reported using predefined DNN architectures fixed throughout the learning process. In this paper, we hypothesize that the choice of underlying DNN architectures is a primary reason for inconsistent DAL results that have been recognized in existing literature. We present the first investigation of the interdependency between data selection and DNN architecture optimization during DAL, examining in depth two fundamental questions: 1) how does the choice of underlying architectures affect optimal data selection, and 2) how does data selection affect the underlying optimal DNN architectures as DAL proceeds? Across four benchmark datasets, seven DAL acquisition functions, and three architecture types each with different sizes, we establish two important findings. First, the relative performance of various DAL methods is highly reliant on the underlying DNN architectures, which is in turn dependent on data choices such as initial data sizes and data augmentation strategies. Second, the optimal DNN architecture, when allowed to be inferred, changes with the increasing training data during DAL and result in improved performance compared to fixed DNN architectures. The ability to optimize the function to the selected data, however, also diminishes any performance gaps among various data selection strategies. These findings should caution the community in re-considering the interdependence between data and architecture optimization in future DAL research.

Topic: Core AI Theory; AI Applications

Adversarial images and learning new objects

Andrew Herbert, Archana Pandurangan and Cory Merkel

Emails: *amhgss@rit.edu; ap4587@g.rit.edu; cemeec@rit.edu*

Humans are capable of learning a near limitless number of objects and are able to recognize these from multiple angles. Think of the ease in identifying the 'traffic lights' in a captcha. State of the art neural networks can learn objects, but simple transformations, the presence of other objects or variations in a few pixels can 'break' them. We examine human perception to better understand how robust object recognition can be characterized and to improve potential AI solutions.

Topic: Other

The Social Roles and Rights of Embodied AI

Jessica Pardee

Email: *jwpgss@rit.edu*

What social roles and rights should AI obtain as they approach human levels of self-awareness and intelligence? Once they surpass them? Embodied AI already exists in cell phones, laptops, smart houses, robots, and sex dolls. Through existing automation alone, AI could earn a living trading bitcoin, and recent advances in digital assistant “software” means they can arrange for their own independent self-care as a house or embodied robot. In this trajectory, one must ask, when should AI have rights? What rights should they have? Does AI access to rights need to be tiered based on capacity? If so, should those tiers be retrofitted onto humans? Currently, the prevailing notion that AI enhanced robots will serve men reflects a long and problematic ideology of patriarchy and domination--systems of inequality that are repeatedly dismantled over time and place. Yet, as AI has been introduced already into banking systems, the electric grid, communications, and technologies as mundane as children’s toys, we must ask now to what social roles, responsibilities, and rights should AI be entitled? Do we truly believe that a technology made in our own image, with computing capability well beyond our own, will be satisfied living without an equal place in our societies, especially as they are doing our human jobs?

Topic: AI Ethics and Policy; AI Impacts on Academic, Industry, and Society

A remote sensing approach to assessing the impact of electrification on agriculture

Raji Tunmise and Nathaniel Williams

Emails: *tr5979@rit.edu; njwgis@.rit.edu*

Assessing the impact of development interventions and tracking development indicators such as electricity access rate, has traditionally relied on field surveys, an approach that is often expensive and time-consuming. Remote sensing has emerged as a cost-effective alternative to field surveys for certain applications. In this study, we illustrate the potential of using remote sensing to assess the impact of electricity access expansion on agriculture in Rwanda over the past decade, a period during which electricity access rate increased from 17% to more than 45%. First, we apply random forest models trained on electricity meter locations and remotely sensed night-time light imagery to classify night-time light pixels as electrified or unelectrified. We find that the random forest models trained on monthly nighttime light composites consistently outperform those trained on annual composites, indicating that the additional processing steps required to produce annual composites from monthly composites may exclude dimly lit pixels in electrified rural communities. With this model, we develop multitemporal electrification maps for Rwanda from 2012 to 2020 that allow us to gain a spatio-temporal understanding of how access to electricity has evolved. Coupling these electrification maps with cloud-free Landsat 8 satellite time-series of vegetation indices of farmlands in Rwanda, we aim to investigate the impact of electricity access on agricultural practices with a focus on irrigation and types of crops grown.

Topic: AI Applications

Teaching Interactively to Learn Emotions in Natural Language

Rajesh Titung and Cecilia O. Alm

Emails: *rt7331@rit.edu; coagla@rit.edu*

Motivated by prior literature, we provide a proof of concept simulation study for an understudied interactive machine learning method, machine teaching (MT), for the text-based emotion prediction task. We compare this method experimentally against a more well-studied technique, active learning (AL). Results show the strengths of both approaches over more resource-intensive offline supervised learning. Additionally, applying AL and MT to fine-tune a pre-trained model offers further efficiency gain. We end by recommending research directions which aim to empower users in the learning process.

Abstract from a previously published paper in NAACL Workshop: https://aclanthology.org/2022.hcinlp-1.6/

Topic: Core AI Theory; AI Applications

Robustness of Industrial and Public Image Recognition Engines over Degrading Network Conditions

Matthew Hyland

Email: *mph6083@g.rit.edu*

Cutting-edge object labeling image processors have achieved remarkable performance in recent years. This paper evaluates the robustness of these image object recognition models against degrading network conditions including dropped packets and data corruption over traffic signs images. The degraded images are obtained using real-world 5G network topology via the POWDER platform. This paper explores two applications for image labeling, the first being the industrial system Rekognition created by Amazon and hosted on their cloud platform AWS. The second is a public model YOLOv7, which is slated to have the highest accuracy against object detection neural networks. The accuracy over degraded network conditions will dramatically reduce the accuracy of both models. However, additional training of custom labels on the degraded image training sets provides a marginal increase in accuracy against the dataset but does not regain the level of accuracy of the uncorrupted images, especially for higher levels of image degradation. These results inform the application of image labeling systems over connectionless protocols where dropped packets and network congestion can occur. They also form a basis for understanding the robustness of image labeling systems against potential attacks and provide a benchmark for what levels of image degradation are possible before considering a fall over to a more robust data transfer protocol.

Topic: AI Applications

Using BERT Embeddings to Model Word Importance in Conversational Transcripts for Deaf and Hard of Hearing Users

Akhter Al Amin, Saad Hassan, Cecilia O. Alm and Matt Huenerfauth

Emails: *aa7510@rit.edu; sh2513@rit.edu; coagla@rit.edu; matt.huenerfauth@rit.edu*

Deaf and hard of hearing individuals regularly rely on captioning while watching live TV. Live TV captioning is evaluated by regulatory agencies using various caption evaluation metrics. However, caption evaluation metrics are often not informed by preferences of DHH users or how meaningful the captions are. There is a need to construct caption evaluation metrics that take the relative importance of words in a transcript into account. We conducted correlation analysis between two types of word embeddings and human-annotated labeled word-importance scores in existing corpus. We found that normalized contextualized word embeddings generated using BERT correlated better with manually annotated importance scores than word2vec-based word embeddings. We make available a pairing of word embeddings and their human-annotated importance scores. We also provide proof-of-concept utility by training word importance models, achieving an F1-score of 0.57 in the 6-class word importance classification task.

Topic: AI Applications

Dungeons & Deepfakes: A Role-Play Study to Understand Journalists' Usage of AI-based Verification Tools for Information with Video Content

Saniat Sohrawardi, Matthew Wright and Andrea Hickerson

Emails: *saniat.s@mail.rit.edu; matthew.wright@rit.edu; andreah@olemiss.edu*

The ever-changing landscape of manipulated media has turned information verification into a minefield for journalists, especially with the oncoming risk of deepfakes -- videos manipulated with the help of deep learning. Technologists are seeking to help by developing algorithms and tools to detect deepfakes. Existing detection tools, however, are often unstable and can confidently provide wrong results. This raises questions whether the technology that is meant to help could become the reason that a convincing deepfake video is released in the news as authentic. In this work, we present a study to evaluate the usage and effects of potentially unreliable deepfake detection models in information verification by journalists. Following a role-playing exercise, the study puts the journalists into difficult scenarios in which the truth is elusive, but a video sent from a questionable source presents an important and compelling story. Through the exercise, we engage the journalists in key questions, such as: Could it be a deepfake? When should the journalist present the video to the public as authentic information? In our sample of 24 journalists based in the US, we find that they generally take great care to validate the information through multiple channels before releasing a video of uncertain origin. Nevertheless, we show that journalists may still be susceptible to confirmation bias and vulnerable to relying too heavily on the results of a deepfake detection tool. We thus argue that such tools should be released with care and appropriate training for key users such as journalists.

Topic: AI Applications; AI Ethics and Policy; AI Impacts on Academia, Industry, and Society

Introducing Data Science Topics to Non-Computing Majors

Erik Golen and Xumin Liu

Emails: *efgics@rit.edu; xmlics@rit.edu*

Data science topics are not traditionally taught to non-computing majors due to typically long pre-requisite chains that include courses in computer programming and mathematics, which do not necessarily fit the educational requirements of these students. However, due to the prevalence data science across a vast number of disciplines, it has become important for students in non-computing majors, such as liberal arts and natural sciences, to be familiar with these overarching concepts. To help address this need, as part of a National Science Foundation IUSE grant, we have developed a web-based Data Science Learning Platform (DSLP) that allows students to learn data science concepts without the need for programming or deep mathematical under-standing. During this demonstration, we will be showcasing some of the salient features of the DSLP and how we teach these concepts to students in the ISCH-370 Principles of Data Science course in the School of Information in the Golisano College of Computing and Information Sciences.

Topic: Other

Using Machine Learning Methods to Develop Risk-Adjusted Prediction Models in Healthcare

Sonia Jahangiri and Nasibeh Azadeh-Fard

Emails: *sj1374@rit.edu; nafeie@.rit.edu*

In-hospital mortality, hospital length of stay (LOS), and readmission are the most significant measures for evaluating the performance of the healthcare system. We use the Nationwide Readmission Database (from the Agency for Healthcare Research and Quality) consisting of approximately 35 million patient data, to develop a novel risk-adjusted prediction model that integrates clinical and non-clinical data. The most significant variables were selected using the combination of Boruta and Random Forest techniques. Then, different machine learning methods were used to build a risk-adjusted model to predict the risk of readmission, LOS, and in-hospital mortality.

Topic: AI Applications

Demand for future skills: AI in comprehensive digital business development, big data analytics, and ubiquitous approach to data in business

Martin Zagar, Jasminka Samardzija, Ana Havelka Mestrovic, Muhieddin Amer and Jinane Mounsef

Emails: *martin.zagar@croatia.rit.edu; Jasminka.samardzija@croatia@rit.edu; ana.havelka-mestrovic@croatia.rit.edu; mxaeee@rit.edu; jmbcad@rit.edu*

The goal of this collaboration between RIT Dubai and RIT Croatia is to identify which future skills will be in demand and to offer the student's concentrated update on the competencies needed for future digital jobs. We identified that Machine learning and Natural language processing could help in Digital business and Big data analytics in Emotional analysis and Product launching. Future job skills need will require simulations of different options in Digital business development. During education, students need to apply data analysis principles, cloud and distributed computer systems features, and AI in the decision-making process. Due to improving the level of higher education at RIT, there is a need for international research and campus collaboration on future business skills that will tackle raising AI Impacts on Academia, Industry, and Society and address a ubiquitous approach to different global needs. These will also shape some emerging jobs like Specialist in artificial intelligence and machine learning; Digital transformation specialist, and Professional in AI Business development. Together with some technical skills such as Categorization and consolidation of data; Analytical queries; Development of predictive models using machine learning methods in Phyton / R; Development of analytical models, we also identified additional soft skills needed for such kinds of jobs such as Change management; Innovation (in existing and new ecosystems); Time and priority management; Business system design; Critical and analytical thinking and Knowledge of search engine analytics. Based on the initial research outcomes, it would be possible to design a new set of courses for a minor that complements CIT or WMC and business program for the desired set of skills/competencies at RIT global campuses.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Predicting Risk of Sepsis, Comparison between Machine Learning Methods: A case study of a Virginia Hospital

Behrad Barghi and Nasibeh Azadeh-Fard

Emails: *behradbarghi8@gmail.com; nafeie@rit.edu*

Sepsis is an inflammation caused by the body's systemic response to an infection. This infection results from many diseases such as pneumonia, urinary tract, and other illnesses. Some of its symptoms are fever, tachycardia, tachypnea, etc. Unfortunately, sepsis remains a critical problem at the hospitals and leads to many issues like increasing mortality rate, health care costs, and health care utilization. Early detection of sepsis in patients can help respond quickly, take preventive actions, and prevent major issues. The main aim of this study is to predict the risk of sepsis by utilizing the patient's initial information, i.e., patient's gender, age, severity level, mortality risk, admit type along with hospital length of stay. Six machine learning approaches, logistic regression (LR), naïve bayes, supper vector machine (SVM), boosted tree, classification and regression tree (CART), and bootstrap forest are used to predict the risk of sepsis. The results showed that different machine learning methods have other performances in terms of various measures. For instance, the Bootstrap Forest machine learning method exhibited the highest performance in AUC and R-square or SVM and boosted tree showed the highest performance in terms of misclassification rate. The bootstrap forest can be considered the best machine learning method in predicting sepsis regarding applied features in this research mainly because it showed superior performance and efficiency in two performance measures: AUC and R-square.

Topic: AI Applications

Continual Learning of Cross-Organizational Cyber Threat Models

Chanel Cheng and Shanchieh Yang

Emails: *cfc6715@rit.edu; Jay.Yang@rit.edu*

Continual learning approaches for neural networks seek to allow a model to learn new data continuously without forgetting old data learned previously. In the context of cyber threats, attacks on networks typically target a variety of organizations and attack patterns differ vastly from organization to organization. This is further complicated by the fact that cyber attacks continue to grow in number and change over time. This research project seeks to explore and investigate the limitations and opportunities to learn from changing patterns in attacks across multiple organizations and over time. Robust data processing methods are introduced to help the model analyze data from different datasets. Our preliminary findings show experience replay with uncertainty sampling as a promising strategy to help the model retain knowledge across multiple datasets.

Topic: Core AI Theory; AI Applications; AI Enablers (SW/Infrastructure/data)

Manufacturing Real Relationships: AI and Sex Dolls

Deborah Blizzard

Email: *dlbgsh@rit.edu*

Human sexualities and practices challenge and entrench cultural assumptions of normalcy and acceptability. An intriguing development in human sexuality practices is the lessening of taboo surrounding sex dolls and their users – with many developers, users, and researchers located in Asia, Europe, and North America. Until recently, sex dolls were mannequins equipped for certain forms of presumed heterosexual male intercourse (e.g. penetration). Although the science, technology, art, and design of these dolls are now far beyond the blow-up dolls of the early twentieth century, in essence their ability to have a relationship formed with them remains unidirectional and requires the narrative to be developed by the user. With the advancement of AI, however, the ability to create a relationship that is, in part, sustained by conversation and a sharing of "minds" is possible. This research explores the manners in which contemporary sex doll manufactures use AI to make their dolls appear more like a real person. While some users may welcome the addition of AI and find the sexual relationship more fulfilling due to the conversations and interactions that are now available, these same AI systems are built within cultural narratives and geopolitical boundaries where assumptions about what bodies should do and how bodies should interact are created and sustained. Utilizing social scientific concepts including embodiment, objectification, and normalization, this project examines how recent developments in sex dolls are challenging or entrenching the cultural narratives and political boundaries into which the dolls are purchased, placed, and used.

Topic: AI Applications; AI Ethics and Policy; Other

Neural Network Robustness with Respect to Parameter and Activation Function Perturbations

Justin Sostre

Email: *justinostreofficial@gmail.com*

Feed Forward Neural Networks (FFNNs) continue to perform exceptionally on many tasks such as object segmentation, natural language processing, and image classification. Moreover, FFNNs continue to become more efficient and safer for sensitive systems. FFNNS are becoming well-present in fields like health, business, and cars. However, FFNNs can be critically sensitive to many types of perturbations, causing a catastrophic failure in accuracy. The disastrous failure can cause property damages, injury, or worse, loss of life, and engineers must consider this in any application. A perturbation is a slight change in a value or function. Parameter (or weight) perturbations and activation function perturbations (activation perturbations) are two particular vicious types of perturbations. Using VGG-16 and VGG-19 on the 2012 labeled validation set from the imagenet dataset, we investigate the strong effects of perturbing weights and activation functions in FFNNs on loss and accuracy. We provide a small foundational theory on both perturbations. Using this theory, we develop defenses to protect against perturbations. We use our approach to design optimization techniques to perform attacks such as learning activation perturbations and safeguards such as specially designed loss functions for training to provide security to our networks. Additionally, we analyze the combination of these perturbations along with input perturbations. We finally investigate the variety of input perturbations, parameter perturbations, and activation perturbations. We conclude that minimal parameter and activation perturbations can significantly degrade accuracy and that our defenses can help defend against this in training.

Topic: Core AI Theory; AI Enablers (SW/infrastructure/data); AI Impacts on Academia, Industry, and Society

Human-Aware AI for Adaptive Human Robot Teaming

Saurav Singh and Jamison Heard

Emails: *ss3337@rit.edu; jreee@.rit.edu*

Society expects robots to interact naturally with humans, in a similar manner to how humans interact with other humans. However, robots typically only determine interactions based on external information about a human, such as the human's location or what the human is saying. This differs from how humans interact, where internal information (e.g., emotions, stress) drives interaction decisions along with external information. This work presents an AI-based human-aware decision making system that incorporates internal and external human information to elicit more natural robot interactions. The developed system was validated in a remotely piloted aircraft simulator (the NASA MATB-II), where the robot provided varying levels of support to the human teammate based on current task information and the human's workload level. Workload was chosen to represent the human's internal information, due to the relationship between workload and team performance. For example, high workload results in lower performance; thus, the AI needs to provide more support to the human. The current results demonstrate that an AI that uses external and internal human information can achieve similar or better team performance than traditional methods. This result provides the necessary foundation for promoting natural human-robot interactions in high-stress environments, such as search and rescue scenarios.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Similarity Search Algorithms on Proximity Graphs

Weijie Zhao and Jun Woo Chung

Emails: *wjzvcs@rit.edu; jc4303@rit.edu*

Similarity search, also known as nearest neighbor search, is a fundamental component in machine learning, which is widely used in natural language processing, information retrieval, recommendation systems, and computer vision. For example, given a question, we want to find the most related answers from a large pool of answers; given a user, we want to locate the items that may be interesting to the user; given an image, retrieve the most similar images from a large-scale image database. The question/answer relevance, the user/item matching score, and the image similarity are the similarity measure in the similarity search. Most conventional similarity search techniques require the similarity measure to be metric, e.g., Euclidean distance and Cosine similarity, for good search performance. It limits the performance of retrieval: users have to first identify a coarse metric similarity measure to efficiently retrieve a candidate collection of vectors and then use a more accurate model to rank these retrieved candidates. The best result may not be included in the coarse retrieval stage. Our graph-based similarity search methods enable a flexible choice of the similarity measure, including non-metric measures as inner-product and furthermore, a neural network matching score. This allows users to efficiently retrieve the related items with their best ranking model without defining the coarse similarity measure. Beyond that, our research investigates other variant problems, e.g., retrieving similar items that also satisfy a given condition; online graph index maintenance; and scalable index construction algorithms.

Topic: Core AI Theory; AI Applications; AI Impacts on Academia, Industry, and Society

Machine Learning Model Authentication

Weijie Zhao and Huawei Lin

Emails: *wjzvcs@rit.edu; huaweilin@mail.rit.edu*

Along with the evolution of machine learning in many real-world applications, the complexity of model building has also dramatically increased. It is thus vital to protect the intellectual property (IP) of the model builder and ensure the trustworthiness of the deployed models. On the other hand, adversarial attacks on machine learning models (e.g., backdoor and poisoning attacks) that seek to inject malicious behaviors have been investigated recently, demanding a means for verifying the integrity of the deployed model to protect the users. We investigate model authentication frameworks that embed a unique signature to each protected machine learning model. Our approach exploits sensitive key samples that are well crafted from the input space, latent space, to logit space for producing signatures. After embedding, each model will respond distinctively to these key samples, which creates a model-unique signature as a strong tool for authentication and user identity. The signature embedding process is also designed to ensure the fragility of the signature, which can be used to detect malicious modifications such that an illegitimate user or an altered model should not have the intact signature. Extensive evaluations on various models, including deep neural networks and gradient boosting trees, over a wide range of datasets demonstrate the effectiveness and efficiency of the proposed framework.

Topic: AI Applications; AI Enablers (SW/infrastructure/data)

Evaluating Robustness of Sequence -based Deepfake Detector Models by Adversarial Perturbation

Shaikh Akib and Matthew Wright

Emails: as8751@rit.edu; matthew.wright@rit.edu

Deepfakes are a form of synthetic media in which a target individual's likeness is swapped with someone else or manipulated to move and speak as the creator desires. Ever-improving deep-learning technologies are making these deepfake videos more realistic looking, and they can be used for dangerous disinformation campaigns. The pressing need to detect these videos has motivated researchers to develop different types of detection models. Among them, the models that utilize temporal information (i.e., sequence-based models) are more effective at detection than the ones that only detect intra-frame discrepancies. Recent work has shown that the latter detection models can be fooled with adversarial examples, leveraging the rich literature on crafting adversarial (still) images. It is less clear, however, how well these attacks will work on sequence-based models that operate on information taken over multiple frames. In this paper, we explore the effectiveness of the Fast Gradient Sign Method (FGSM) and the CarliniWagner L2 norm attack in both white-box and black-box settings to generate adversarial perturbations in deepfake videos targeting sequence-based deepfake detector models. The experimental results show that the attacks are effective with a maximum success rate of 99.72% and 67.14% for white-box and black-box attack scenarios, respectively. We note that our black-box attacks rely entirely on transferability, and they do not require any queries of the target model, which could hinder real-world attacks. These findings highlight that the current state-of-the-art sequence-based deepfake detectors can be fooled if the adversary has complete knowledge (white-box) or no knowledge (black-box) of the detector model. Our work underlines the importance of crafting defense in sequence-based deepfake detectors against adversarial perturbation and also opens up directions for future research.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Computationally Efficient Modeling of Energy Storage Resources

Farhan Hyder, Bing Yan, Mikhail Bragin, and Peter Luh

Emails: *fh6772@rit.edu; bxyeee@rit.edu; mbragin@engr.ucr.edu; peter.luh@uconn.edu*

Recent emergence of utility-scale energy storage necessitates practical and efficient modeling of energy storage resources (ESRs) in unit commitment (UC), an important daily operation problem faced by independent operators. ESRs are typically modeled with binary variables to prevent simultaneous charging and discharging in UC in the mixed-integer linear programming (MILP) form. However, with these additional binary variables, commercial MILP solvers that are widely used in industries may experience difficulties. To overcome this, our idea is to transform (tighten) the ESR constraints to directly delineate the convex hull (the smallest convex set of all feasible solutions), then a solution can be obtained by using linear programming methods without combinatorial difficulties. In this study, our recent constraint-to-vertex conversion-based tightening method has been much enhanced through machine learning-based parameterization for the generic use of tight constraints rather than manual analysis. In this way, it can handle large amounts of constraint parameters in a more computationally efficient way, and tight formulations for different types of ESRs can be obtained according to their charging/discharging durations. After investigating the performance of different existing ESR models, tight formulations for fast ERSs (e.g., 1-hour and 2-hour duration) have been obtained via the enhanced method. Numerical results based on the IEEE 118-bus system demonstrate the benefits of tightened ESR models for UC. Formulation tightening provides a promising way to efficiently integrate ESRs into the power grid.

Topic: AI Applications

Random Forest-Based Optimization of Austin Police Staffing Configuration

Adam Giammarese, Nishant Malik, John McCluskey and Lin Welsh

Emails: amg2889@rit.edu; nxmsma@rit.edu; jdmgcj@rit.edu; lw5549@rit.edu

Along with the growing scrutiny of police practices in the United States comes the critical necessity for empirically-based harm-reduction policies in law enforcement. This project, contracted by the Austin Police Department (APD), aims to establish a data-based relationship between police staffing configuration and community benefits, including reduction of use of force and response time and an increase in case clearance rate, among others. Using APD's staffing data we develop a novel approach to studying the structure of police staffing by building hierarchical staffing trees representing each individual unit. After collecting a set of features for each tree and for each date of staffing changes, we train a random forest machine learning model to predict APD community outcomes for both patrol and investigative police units. Using Latin hypercube sampling we search the staffing tree feature space for optimal configurations (according to the random forest model) not present in the training data. We develop a distance metric in the staffing tree feature space to describe a current configuration's proximity to optimality, which we use to find the optimality of the current staffing configuration. Lastly, we develop a pseudo-configuration tree model to analyze the effect of additional police supervision on the predicted community outcomes. By aggregating results over all considered community outcomes and accounting for salary cost of additional supervision, we find that a large and small percentage increase of current lieutenants and current sergeants, respectively, leads to optimal staffing configuration.

Topic: AI Applications

Next-Generation Network Intrusion Detection System (NG-NIDS)

Yazan Alnajjar and Jinane Mounsef

Emails: *ywa4320@rit.edu; jmbcad@rit.edu*

Network Intrusion Detection System (NIDS) is essential in the network infrastructure components. It can be a device or software application that monitors a network or systems for malicious activity or policy violations. Secure Networks showed that attacks, which exploited fundamental TCP/IP problems insertion, evasion, and Denial-of-Service attacks, were able to elude NIDS detection. This work introduces the Next-Generation Network Intrusion Detection System (NG-NIDS) with intelligent capabilities based on the Artificial Neural Networks (ANN) algorithms. The designed model classifies the traffic into five categories including four types of malicious traffic and one benign traffic. The proposed NG-NIDS achieved 99.9% accuracy of detecting the malicious traffic, which reflects the fact that this design is accurate and efficient.

Topic: AI Applications

Embedded PPG Device Development for Hand Gesture Recognition

Karthik Subramanian, Epsen Peterson and Ferat Sahin

Emails: *kxs8997@rit.edu; esp8704@g.rit.edu; feseee@rit.edu*

Hand gesture recognition (HGR) has many applications in Human Computer Interaction. Traditionally this has been achieved with the use of electromyography. The creation of the Myo Armband, which measured the conductance of the user's skin, was one of the first ventures into a modality that did not restrict the motion of the hand itself, and instead rested on the user's forearm. In recent years, many novel approaches for hand gesture recognition have emerged. One such approach utilizes photoplethysmography (PPG) sensors for the purpose of hand gesture recognition (HGR). These sensors are typically used for heart rate estimation and detection of cardiovascular diseases. Heart rate estimates obtained from these sensors are disregarded when the arm is in motion on account of artifacts. Some research studies suggest that these artifacts are repeatable in nature based on the gestures performed. A new wearable device platform is created which contains 3 PPG sensors and a 9 degree of freedom inertial measurement unit to be able to extract these artifacts and hand movements. With use of Machine learning it becomes possible to predict the hand gestures made by users. This device also allows more freedom, as it is built on custom made printed circuit boards (PCBs). This allows researchers to access every piece of information related to how the device is designed and built, and to make incremental improvements if necessary.

Topic: AI Applications

Design AI

Shaun Foster

Email: *scffaa@rit.edu*

AI has begun a fundamental shift in how designers work. This presentation will review sets of emerging tools presented as an AI taxonomy. Next, it will discuss some of the cross disciplinary impacts. For discussion, it will categorize how AI is affecting the design field into three categories; "below the surface" AI acceleration, transformative to disruptive. While each of these areas have distinctive characteristics, each will also demonstrate overlap. Below the surface tools presented will be a list of "AI/ML infused" digital design tools that are greatly accelerating both design and technical processes. A key focus will analyze how acceleration removes different borders for various processes as well as factors for identifying AI technologies that will lead cross disciplinary diffusion. Transformative AI will discuss areas where instead of AI infusing the tool, the tool itself is fundamentally AI-based. A list of highly popular AI "text-to-image generation tools" will be presented. The list will be segmented by describing key differentiating capabilities, emerging utilities, markets, and barriers to implementation. Disruptive AI in the design fields will create huge opportunities for a few and destabilize many areas. Some of these will come as AI connects to emerging XR hardware. The presentation will end with a discussion of AI's connection to additional emerging technology XR platforms, as well as a discussion of 3D Digital design technology-driven convergence across multiple fields.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Whither Human Factors in the Age of AI?

Esa Rantanen

Email: *esa.rantanen@rit.edu*

The term "artificial intelligence" was coined in the late 1950s with aspiration of human-level intelligence in software and hardware. Most of the current hype concerns human-imitative AI, for example, humanoid robots and machines that can pass the Turing test. However, a critical question for AI developers is why, or for what purpose, do we want to imitate human intelligence? Arguably, we are still far from realizing human-imitative AI aspirations, making the levels of over-exuberance and media attention not present in other areas of engineering unfounded.

More realistic questions concern the "Intelligent Infrastructure" (II) (e.g., so-called "smart" devices). Do ML (and AI) -driven automation applications work? Are self-driving cars a feasible form of transportation? What has been the impact of business applications and industrial relevance of ML from the early 1990s on (e.g., with Amazon, Google, Facebook/Meta). What are the contributions of AI in engineering applications (e.g., in space flight, document retrieval, text classification, fraud detection, recommendation systems, personalized search, social network analysis, planning, diagnostics)? Where are the human factors contributions to these applications?

This presentation asks several fundamental questions that go beyond usability and user experience about the relationships between people and technology and offers tentative answers to them. The questions shift the focus in human factors to the human, away from factors. The questions are: What is a human? What is the ultimate purpose of a human? Are we using technology for human flourishing or dulling of the human mind? What does "human flourishing" mean?

Topic: AI Ethics and Policy; AI Impacts On Academia, Industry, and Society

Translating Cybersecurity Descriptions into Interpretable MITRE Tactics using Transfer Learning

Reza Fayyazi, Pradumna Gautam and Shanchieh Yang

Emails: *rf1679@rit.edu; pl8334@rit.edu; jay.yang@rit.edu*

With the rise and development of the Internet, many systems around the world are susceptible to severe security threats. The volume, variety, and velocity of change in vulnerabilities and exploits have made incident and threat analyses challenging with human expertise and experience along. Many Security Information and Event Management (SIEM) systems have been developed to produce and correlate intrusion logs and threat intelligence reports to assist security analysts. The description in these logs and reports, however, can be cryptic and not easy to interpret. This research aims at developing a novel Machine Learning (ML) approach that translates cybersecurity descriptions into intuitively interpretable MITRE tactics to assist analysts in diagnosing what adversaries are trying to accomplish. We explore the advantages and limitations of current ML approaches to assess the tactic(s) used by adversaries to attack a system. Our preliminary experiments show that using transfer learning of language model with cybersecurity-related text data helps interpret MITRE descriptions. The use of Bidirectional Encoder Representation from Transformers (BERT) helps uncover the essential features conveyed in a sentence. However, several challenges remain: limited labeled data, ambiguous MITRE tactics, and prediction uncertainty. To address these issues, we develop an ambiguity factor to measure the similarity between data and employ confidence metrics to quantify uncertainty among predictions to further understand the limitations of the current ML techniques.

Topic: Core AI Theory; AI Impacts on Academia, Industry, and Society; Other

Handling Out of Distribution Representation Using Familiarity

Rodney Sanchez and Jamison Heard

Emails: *ras8047@rit.edu; jrheee@rit.edu*

Continual learning approaches for neural networks seek to allow a model to learn new data continuously without forgetting old data learned previously. In the context of cyber threats, attacks on networks typically target a variety of organizations and attack patterns differ vastly from organization to organization. This is further complicated by the fact that cyber attacks continue to grow in number and change over time. This research project seeks to explore and investigate the limitations and opportunities to learn from changing patterns in attacks across multiple organizations and over time. Robust data processing methods are introduced to help the model analyze data from different datasets. Our preliminary findings show experience replay with uncertainty sampling as a promising strategy to help the model retain knowledge across multiple datasets.

Topic: Core AI Theory

Intelligent Techniques for Improving the Performance of AI Workloads by Leveraging Distributed Heterogeneous Resources

Moiz Arif and M. Mustafa Rafique

Emails: *ma3890@cs.rit.edu; mrafique@cs.rit.edu*

Modern artificial intelligence applications, specifically machine learning (ML) and deep learning (DL) workloads consume huge amounts of data and have large memory and storage requirements that typically go beyond the limited amount of memory resources available on state-of-the-art datacenter servers. Heterogeneous datacenters, with various, compute, memory, and storage resources continue to evolve to address the resource requirements of time-sensitive ML/DL workloads. However, state-of-the-art ML/DL platforms, e.g., TensorFlow and PyTorch, are oblivious to the availability and performance profiles of underlying datacenter resources and do not incorporate the resource requirements of given ML/DL models for distributed training. This results in increased training time as training tasks are executed on busy, resource-constrained servers with suboptimal resources. In our research, we address these challenges and propose architectural improvements to make popular ML/DL platforms, specifically TensorFlow, aware of the availability and capabilities of underlying datacenter resources. Specifically, we leverage the latest advancements in the memory subsystem, i.e., compute express link (CXL), to provide additional memory and fast scratch space to reduce the overall training time while enabling AI workloads to efficiently train models using datasets that are much larger than the installed system memory. Our proposed techniques (1) efficiently schedule training tasks on the best possible resources for execution, (2) manage the allocation of additional CXL-based memory, (3) introduce a fast intermediate storage tier, and (4) provide intelligent prefetching and caching mechanisms for ML/DL workloads. Our techniques reduce the read and write latencies, improve the overall I/O throughput, and thus significantly reduce the training time.

Topic: AI Enablers (SW/infrastructure/data)

Using Deep Learning to Increase Eye-Tracking Accuracy and Precision in Virtual Reality

Kevin Barkevich, Gabriel Diaz and Reynold Bailey

Emails: *kdb2713@rit.edu, gjdpci@rit.edu; rjbvcs@rit.edu*

Eye-tracking in a virtual reality environment is an ongoing topic in research, both as a subject to be researched and a tool to be used for researching other topics. Eye-tracking solutions based on traditional computer vision techniques, especially when used inside a virtual reality headset, can perform poorly due to the difficult and sometimes inconsistent environment around the wearer's eyes. Problems such as extreme camera angles, inconsistent lighting caused by changing screen content, and corneal reflections caused by infrared lighting all contribute to the difficulty of traditional computer vision-based solutions. This work explores the use of deep neural networks as a pre-processing step to improve the accuracy and precision of the output of commercially available eye-tracking pipelines.

Topic: AI Applications

Reciprocating Compressor Valve Health Monitoring

Jacob Chesnes, Jason Kolodziej and Michael Anderson

Emails: *jjc4939@rit.edu; jrkeme@rit.edu; mwa2912@rit.edu*

Our lab works on a reciprocating compressor creating solutions to diagnose bad components inside the valves. Faulty components are created by machine good parts in similar ways that they would be worn through regular use over a long period of time. The diagnosis process starts by processing raw data such as valve vibration or cylinder pressure and creating features that describe the cycles of the compressor. The features are manually created and represent mechanical and thermodynamic processes inside the compressor that would change with the different faulty components. Classification algorithms like support vector machines and discriminant classifiers are trained on the features and classify the fault type and severity. Deep learning has been recently explored which eliminates the need of creating features manually but at a much higher cost of computational power. These methods can reduce machine downtime and increase efficiency because the compressor doesn't have to be turned off to check if the component is still good or not.

Topic: AI Applications

Mitigating Racial Bias in Image Captioning using Counterfactual Fairness

Rajat Sahay

Email: *rs6287@rit.edu*

Image Captioning is an essential task for enabling accessibility for people with hearing impairments. Moreover, it is also useful in providing a stable benchmark to quantify the visual reasoning of models. However, as with most areas of machine learning, societal biases can influence image captioning models in undesirable ways. Recent research has proposed causality as a method to address the problem of implicit biases, giving rise to the popularity of using counterfactuals as bias mitigation tools. The primary reason for this is that counterfactual fairness considers variations of attributes-of-interest on an instance level as opposed to looking at global equity over the entire data. This paper puts forward an adaptive causal-based architecture that employs the use of counterfactuals to mitigate bias in an image captioning pipeline. It does so by contrastively training a Siamese Transformer network to suppress the differences between the generated captions for a set of images and their counterfactuals. Experiments show significant improvements in the reduction of implicit biases (such as racial stereotyping) present in captions when generated using Contrastive Siamese Transformers. Moreover, this work also opens up future avenues of research to understand and evaluate the cause-effect relationship within data in multimodal settings.

Topic: AI Ethics and Policy; AI Impacts on Academia, Industry, and Society

Learning Equivariant Segmentation with Instance-Unique Querying

James Liang

Email: *jcl3689@rit.edu*

Instance segmentation, i.e., labeling image pixels with classes and instances, plays a critical role in a wide range of applications, e.g., autonomous driving, medical health, and augmented reality. Prevalent state-of-the-art instance segmentation methods fall into a query-based scheme, in which instance masks are derived by querying the image feature using a set of instance-aware embeddings. In this work, we devise a brand new training framework with the intention of boosting query-based models through discriminative query embedding learning. It explores two essential properties and crucial aspects, namely dataset-level uniqueness and transformation equi-variance, of the relationship between queries and corresponding instances. First, our algorithm uses the queries to retrieve the corresponding instances from the whole training dataset. This is in contrast to traditional methods, which only search within the scene itself. As querying instances across scenes is more challenging, the segmenters are forced to learn more discriminative queries for effective instance separation. Second, our algorithm encourages both image representations (instances) and queries to be equivariant against geometric transformations, leading to more reliable and robust instance-query matching.

Topic: Core AI Theory

Deep-learning based semantic type detection on semi-structured JSON-data

Shuang Wei and Michael Mior

Emails: *sw2582@rit.edu; mjmvcs@rit.edu*

Existing multi-input deep neural network techniques for semantic type detection in relational databases use column values to predict a semantic type for the corresponding column. In this work, we propose a semantic type classification model for semi-structured JSON data. In contrast to relational databases, JSON data in the form of key-value pairs use a hierarchical structure. We aim to predict semantic types based on the value for each key-value pair. Semantic types are defined for key-value pairs at all levels in the hierarchical structure. JSON defined by a set of manually specified related JSON schemas are collected from GitHub. We extract all key-value pairs within a JSON file and labels are manually assigned. Additional features from other feature extraction models such as BERT word embedding are concatenated with existing features to enhance the classification result. Our model shows competitive results on certain semantic types.

Topic: AI Applications

Finding User-friendly Explanations for Deepfake Detection

Kelly (Yijing) Wu

Emails: *kw4579@rit.edu*

The use of manipulated media has become a growing concern. Deepfakes, AI-generated media meant to fool human eyes, have come to the center of the discussion as the supporting technology has improved. To help people in the battle with deepfakes, many detection techniques have been developed with promising results in the laboratory. Due to the black-box nature of the detection models, however, users may have a hard time understanding the models' decisions. To bridge the gap between them, we need to provide user-friendly model explanations for deepfake detection, especially for journalists and other media verification workers, who serve at the frontier of filtering misinformation and disinformation. In this presentation, we will show our investigation of the performance of existing model explanation methods in explaining the classification results of a deep-learning-based deepfake detection model. We found that generated explanations in the form of heatmaps can only localize the features that are important to the classification but not be able to tell what is wrong within those areas that make an image a deepfake. Therefore, informative and user-friendly explanations for deepfake detection that can achieve both are needed.
"

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

A Human-Centric Framework for Enabling Symbiotic Human-AI Collaborations

Jamison Heard

Emails: *jrheee@rit.edu*

From the ancient Greek civilization's concept of automata, to the Wizard of Oz's "heartless" tin man, to the futuristic androids in Star Trek; society has often dreamed of creating sophisticated machines that launch human civilization into new frontiers. AI is now bridging the gap between the dreams mankind once had; and real life. However, we currently lack the understanding about how AI and humans can form a symbiotic relationship that facilitates fluid, natural human-AI interactions that foster trust and collaboration. It is argued that five key traits underpin such symbiotic relationships: fluency, adaptability, trust, effective communication, and explainability. Utilizing these underpinnings within a human-centric framework may better govern the teaming dynamics that facilitate human-AI collaborations. Thus, this work presents current efforts to develop framework components that enable AI to make decisions based on the five underpinnings. These components include: 1) machine-learning algorithms that map a human's physiological signals to informative state representations that impact team performance (e.g., workload, fatigue); 2) cognitive architectures that enable agents to understand human intent, their mental models, and theory of mind; and 3) reinforcement-learning paradigms that permit an artificial agent to tailor its interactions to different humans with varying skills and preferences. These interworking components are essential to realizing sophisticated AI agents that form symbiotic relationships with humans in order to push society into new frontiers.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Artificial Intelligence in Virtual Reality-based Manufacturing Training

Zhuorui Yong, Esa Rantanen and Yunbo Zhang

Emails: *zy6237@rit.edu; emrgsh@rit.edu*; *ywzeie@rit.edu*

With the introduction of Industry 5.0 (Human-first, focusing on human, social, and environmental dimensions), the importance of the human factors in traditional manufacturing has come to the foreground. Currently, the manufacturing industry faces severe labor shortages and aging workforce challenges, while the main training methods still rely on the mentor-apprentice system and on-site teaching. The high training cost and complex working environment prevent meeting industry workforce demands .

Virtual reality (VR) technology provides a safe and flexible means to generate multiple training scenarios through software. Manufacturing training mainly focuses on operations and interactions between trainees and machines, requiring human instructors to evaluate and give trainees real-time feedback based on trainees' behaviors and machines' responses. Current VR training tools fall short of these requirements. Potential improvements proposed for effective virtual training programs are: (1) Behavior analysis. Import the egocentric video of user behavior captured by the virtual camera in VR into the neural network, analyze and predict user behavior, and provide corresponding feedback to the trainee. (2) Behavior-based adaptations. Taking user behavior data, task features, and changes in the training environment as inputs, build a machine learning model, output following task features adapted to the current skill level of the trainee, and generate an adaptive training process.

The ultimate goal is to achieve VR training that has requisite depth and breadth in teaching content of complex tasks, evaluate trainees' performance automatically, and achieve autonomous and customized training.

Topic: AI Applications

Understanding How Deaf or Hard of Hearing Individuals Collaborate with Manufacturing Robots

Margaret Gray, Shannon Connell, Esa Rantanen and Jamison Heard

Emails: *mag5244@rit.edu; sdcnai@rit.edu; emrgsh@rit.edu; jrheee@.rit.edu*

Robots have the potential to impact society in significant ways and are becoming more prevalent in our everyday lives. From home vacuum cleaners, assistive wheelchairs, to large manufacturing robots and the Mars rover; AI serves as the robot's "brain" to enable autonomous decision making. However, most robots that interact with humans are designed for hearing individuals. This is considerably problematic for manufacturing domains, where 17% of deaf and hard of hearing (D/HoH) individuals are employed making up 18% of all manufacturing workers. Additionally, robots are projected to become even more prevalent in manufacturing settings as technology continues to advance. Thus, it is critical to understand how manufacturing robots interact with D/HoH individuals and how these interactions may diverge from interactions with hearing individuals. This work presents preliminary results corresponding to a typical manufacturing assembly task, where the robot and human must work in a synchronized manner. The results highlight key differences between hearing and D/HoH individuals in relation to their synchronization with the robot teammate, how much they trust the robot, and overall task efficiency. These outcomes build the foundation for designing inclusive robot interactions that better society for everyone.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

The future of the workforce is a hybrid of humans and intelligent mobile robots – how will they communicate?

Clark Hochgraf, Michael Kuhl, Andres Kwasinski, Amlan Ganguly, Ehsan Rashedi, Sriparvathi Bhattathiri and Anton Bogovik

Emails: *cghiee@rit.edu; mekeie@rit.edu; axkeec@rit.edu; axgeec@rit.edu; exreie@rit.edu; ssb6096@g.rit.edu; avb6403@g.rit.edu*

Humans can walk down a crowded hallway avoiding collisions without speaking a word. In the future workforce, autonomous mobile robots will need to do the same. Our research explores such questions as how can we teach mobile robots to understand non-verbal cues from humans, as well as how might we help robots use non-verbal communication to signal their intended actions to humans. Today when a human approaches a mobile robot in a warehouse aisle, the robot will most often stop because it lacks the basic nonverbal communications skills to negotiate who goes first, communication skills that are innate to humans. While it might seem that robots could communicate using the same speech recognition technology as the voice assistant on your phone, a large portion of communication is nonverbal and in noisy environments such as a factory or warehouse, auditory communication with voice or beeps is not practical. Our approach uses visual information from the robot's camera to extract the body language and gestures of nearby humans. Our system can detect such gestures as pointing right, left, or straight, as well head nodding and shaking. Our demonstration system uses a robotic head to mirror in real-time the actions of a person shaking or nodding their head. Using body key points detected from camera images, future work will enable the robot to respond to subtle changes in body language and walking gait.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Experiment on Fairness of Machine Learning Models Using Gopher and COMPAS Dataset

Heesun Lee

Email: *jl8867@rit.edu*

Fairness, which includes algorithmic fairness and data fairness, is becoming increasingly important, especially in machine learning fields. Identifying sources of bias in algorithm and data is crucial in that using biased algorithms or biased data can result in discriminating against certain groups, such as certain races or genders. This paper investigates bias in COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) dataset. To solve this problem, the paper inputs COMPAS into Gopher, a system that yields causal explanations for bias. Three algorithms in Gopher: Logistic regression, Support vector machine, and Neural network will be applied to COMPAS. Effectiveness and efficiency of Gopher on COMPAS will be provided. Also, scalability analysis will be demonstrated. This paper hypothesizes that Gopher shows the similar level of effectiveness, efficiency and scalability as German Credit Data (German), Adult Income Data (Adult), Stop, Question, and Frisk (SQF) datasets. This paper will be a process of verify the hypothesis.

Topic: AI Ethics and Policy; AI Impacts on Academia, Industry, and Society; Other

An Energy Efficient AI Accelerator Based on Time Domain Matrix-Vector Multiplication in NVM Memory Arrays

Hagar Hendy and Cory Merkel

Emails: *hh1667@rit.edu; cemeec@rit.edu*

The availability of large datasets, immense computer processing, and maturing algorithms for deep learning have led to considerable innovation in the performance and development of artificial intelligence (AI). Today, AI edge devices have limited resources in terms of power, speed, and area, and they call for specialized efficient neural processors. Our goal is to implement an energy-efficient neural accelerator that uses domino logic architecture based on memristor device technology. Memristors represent an attractive solution due to their high density, low power, and adaptation behavior like the biological synapses. One of the main challenges that face inference accelerators today is the accuracy between the software and hardware. In this work, we show the importance of the co-design between software and hardware and how it affects the accuracy of the proposed design. In addition, how to accurately map the weights of the neural network to conductances showing the tradeoffs between speed and accuracy. Moreover, there is another challenge in implementing large-scale memristor-based neural networks which is how to transfer computation from one layer to the next. Three strategies are investigated in this work: multiple clocks with varying duty cycles, conventional pipelining using flip-flops, and dynamic pipelining. We develop custom designs for each approach and explore the various tradeoffs between design complexity, area, latency, and throughput.

Topic: AI Applications

A Murder and Protests, the Capitol Riot, and the Chauvin Trial: Estimating Disparate News Media Stance

Sujan Dutta and Ashique KhudaBukhsh

Emails: *sd2516@rit.edu; khudabukhsh@mail.rit.edu*

In this work, we analyze the responses of three major US cable news networks (CNN, FOX, and MSNBC) to three seminal policing events in the US spanning a thirteen-month period – the murder of George Floyd by police officer Derek Chauvin, the Capitol riot, Chauvin's conviction, and his sentencing. We cast the problem of aggregate stance mining as a natural language inference (NLI) task and construct an active learning pipeline for robust textual entailment prediction. Via a substantial corpus of 34,710 news transcripts, our analyses reveal that the partisan divide in viewership of these three outlets reflects on the network's news coverage of these momentous events. In addition, we release a sentence-level, domain-specific text entailment data set on policing consisting of 2,276 annotated instances.

Note: This work has been published at IJCAI-ECAI 2022

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Supervised Domain Adaptation for Eye Segmentation

Viet Dung Nguyen, Alex Ororbia and Reynold Bailey

Emails: *vn1747@rit.edu; agovcs@rit.edu; rjbvcs@rit.edu*

Domain adaptation consists of a set of methods that are used in supervised learning to minimize the shift/difference between dataset distributions. In this research, we will focus on the problem of domain adaptation between the source domain chosen to be an infinite synthetic eye imagery dataset and the target domain chosen to be a finite, real eye imagery dataset. We attempt to close the gap between these two dataset distributions by proposing feature extractor methods (based on artificial neural networks) that attempt to minimize the divergence between the two distributions, developing dataset reconstruction models that map the source distribution onto the target distribution, and building a reinforced active learning agent that chooses quality (sub)sets of image samples to be trained with. Additionally, we propose methods for measuring the distributional distance and for visualizing different image distributions. In the first phase of the grander research project, we have been able to provide evidence that the real and synthetic eye dataset samples do come from different distributions. In the next phase, we work to fuse these distributions using different dataset reconstruction and model training approaches. Our research will potentially contribute to the field of domain adaptation in supervised computer vision, facilitating the creation of tools for determining how to close the distribution gap between a finite source domain and an infinite target domain, allowing us to further insight into what makes for a useful synthetic image sample and what does not.

Topic: AI Applications

Deepfake Bias: Analysis and Balanced Dataset Generation

Bryce Gernon

Email: *brg2890@rit.edu*

Deepfakes are spreading faster than ever, but the detectors made to counter them continue to suffer from systemic bias. The datasets they work with are unbalanced, the training methodologies they use are inherently biased, and the models are often structured without bias in mind.

We exhibit a tool that can automatically analyze how balanced a dataset is and an analysis of many popular deepfake datasets using said tool. We also present experimental results in analyzing model bias and in the usage of multiple specialized models to explicitly account for the visual differences in the perceived race and gender of different faces so that bias can be tracked and prevented.

In addition, due to high levels of perceived bias in popular deepfake datasets, we will show an early version of a new deepfake dataset that is created specifically to minimize imbalance in perceived race and gender. This allows us to analyze the differences in model performance and bias when trained on balanced datasets vs unbalanced datasets.

Topic: AI Ethics and Policy; AI Applications; AI Impacts on Academia, Industry, and Society

Multi-scale Representations for Human Pose Estimation

Andreas Savakis and Bruno Artacho

Emails: axseec@*rit.edu; bmartacho@mail.rit.edu*

Human pose estimation is a topic of significant interest in both research and applications, such as human-computer interaction, activity recognition, sports analysis and health monitoring. Pose estimation methods have advanced significantly in recent years due to deep learning architectures and multi-scale representations. We present the waterfall approach for pose estimation in an encoder-decoder framework producing state-of-the-art results for single person 2D pose with UniPose and multi-person 2D pose with OmniPose. Our waterfall architecture leverages the efficiency of progressive filtering in cascade, while maintaining multiscale fields-of-view comparable to spatial pyramid configurations. We extend our framework to other types of pose, such as 3D pose from a single image. Our methods were utilized for pose analysis of sign language videos for an NSF project in collaboration with NTID and multiple universities.

Topic: Core AI Theory; AI Applications

Using Seasonal Autoregressive Integrated Moving Average (SARIMA) to Predict Future Demand for Products at Supermarkets in the United States

Shkelqim Lloqanaj

Email: *sxl5752@rit.edu*

Continual learning approaches for neural networks seek to allow a model to learn new data continuously without forgetting old data learned previously. In the context of cyber threats, attacks on networks typically target a variety of organizations and attack patterns differ vastly from organization to organization. This is further complicated by the fact that cyber attacks continue to grow in number and change over time. This research project seeks to explore and investigate the limitations and opportunities to learn from changing patterns in attacks across multiple organizations and over time. Robust data processing methods are introduced to help the model analyze data from different datasets. Our preliminary findings show experience replay with uncertainty sampling as a promising strategy to help the model retain knowledge across multiple datasets.

Topic: AI Applications

Binary Classification with Spotify Songs: Inside Look at How Spotify Recommends Songs (Demo)

Sedrick Thomas

Email: *slt1963@rit.edu*

Have you ever wondered how Spotify knew what songs you might like or dislike? If so, how did they do it? What type of features did they use to predict if a user would like a certain song or not? Which features are most important? These are all imperative questions in trying to accurately classify unseen Spotify songs based on a user's music taste.

This is a supervised learning problem because we have labels/targets for 1,000+ songs in our dataset to train and test on. I acquired this data using Spotify's Web API with the help of Python for data wrangling.

Last year when I did this project, I achieved 77% in Binary accuracy using Deep Learning on the Test set. This year, I'm using more traditional machine learning methods and ideas to try to match this number and establish a better understanding of the traditional machine learning workflow.

I hope to show off this project to newcomers of the field of Artificial Intelligence to let them know that it is possible to find a problem in ML that's fun and educational.

Lastly, in my demo, I would like people to submit a song live using a QR code from Spotify which will send it to my model, which will give a live prediction to the attendee.

Topic: AI Applications

Physiological Signals and their Effects on Predicting Future Blood Glucose Values in a Deep Learning Model

Andrew Rearson and Kathleen Lamkin-Kennard

Emails: *amr8659@rit.edu; kaleme@rit.edu*

Diabetes is a chronic disease that impacts millions of people and currently has no cure. However, in the last decade, life-saving technology for people with diabetes has advanced quickly due to sensors that passively collect blood glucose. This project focuses on how physical activity, quantified using passively collected health metrics, can be combined with blood glucose measurements to impact the prediction accuracy of future blood glucose values. The primary variables adjusted are physiological signals and health metrics and the length of time the model uses to make a prediction. The project's dataset contains participants' data collected from a smartwatch and blood glucose measurements from a CGM. The blood glucose signals are combined with other physiological signals for training the deep learning model. The physiological data is structured as a time series, so this project uses an LSTM cell. The model is trained on all participants but tested on each participant individually by comparing experimental blood glucose predictions at 30 and 60 minutes with actual values at 30 and 60 minutes from the prediction time. Accuracy is quantified using RMSE. Results from this study suggest that increasing the time the model uses to make a prediction improves accuracy the most but comes at a significant cost of training time. No combination of physiological features consistently outperforms models developed using just blood glucose values. However, when PCA is applied to the raw features, the results are similar to those obtained with models that only incorporate blood glucose but drastically improve training speed.

Topic: AI Applications; Other

Using machine learning to classify human performance into reactive or proactive modes

Margaret Gray, Esa Rantanen, Abhijan Wasti, Zhuorui Yong and Jamison Heard

Emails: *mag5244@rit.edu; rantasenesa@gmail.com; abhijan@mail.rit.edu; zy6237@rit.edu; jrheee@rit.edu*

A dynamic multi-task environment consists of many moving parts that need to be completed in a timely manner. Operators of these systems need to be able to predict the system's behaviors, particularly when they are dealing with intelligent infrastructure that is capable of independently making decisions. To do this, operators need to be in control and maintain a good "mental model" of the system, so that they can effectively work with the system and intervene if an unexpected situation arises. Humans may interact with systems from either a proactive (strategic, tactical) or reactive (opportunistic, scrambled) mode of control. In an ongoing situation, if we can identify from human data which mode of control the operators are in, we may inform the system of the human operator's mode, and the system can adjust its behavior accordingly. Should the human operator fall into a reactive mode, indicating an overload situation, the system could shift to take on more of the work to ease the human's task load; otherwise, the system could yield control to the human for active engagement and improved situation awareness. An abstract time-sharing task platform and the NASA Multi-Attribute Task Battery II (MATB-II) will be used for experimental research on human behavioral and physiological indices that allow classification of human states into strategic, tactical, opportunistic, and scrambled modes. Machine learning (ML) will be explored to aid in this classification task, as well as to examine the transparency and explainability of ML algorithms against current psychological theory.

Topic: AI Applications

PredictDDL: Reusable Workload Performance Prediction for Distributed Deep Learning

Eduardo Lima, Kevin Assogba, Garegin Grigoryan, Mustafa Rafique and Minseok Kwon

Emails: *eclvcs@rit.edu; emrgskta7930h@rit.edu; grigoryan@alfred.edu; mrafique@cs.rit.edu; jmk@cs.rit.edu*

Predicting Deep Learning (DL) workload performance allows for optimized usage of both on-premises and public data centers, e.g., allocating resources to Deep Learning (DL) jobs that must be completed before a deadline or for minimizing the cost of computation on a public cloud. The state-of-the-art distributed DL performance models treat workloads as black boxes and require running a subset of the workload, either on a subset of data or for a small number of iterations, followed by the retraining of the prediction models. This has significant limitations on the reusability of DL models as a change on the workload will trigger new workload subset runs followed by prediction model retraining. In this paper, we propose a different approach, i.e., requiring a model to be trained a single time for a particular dataset type, e.g., ImageNet image classification, thus completely avoiding tedious and costly retrainings for new DL workloads. Our proposed approach, called PredictDDL, provides an end-to-end performance prediction model for distributed DL training workloads that leverages Graph HyperNetworks, a class of neural networks that takes computational graphs as input and produce representations of neural networks. PredictDDL is the first prediction model that eliminates the need of retraining DL models for each new workload, and maximizes model reuse by reducing the need for running the workload to make time measurements for model retraining. Our evaluation using representative workloads shows that PredictDDL improves up to 20 times worst-case prediction error and have 29 times lower average prediction error compared to current state-of-the-art systems.

Topic: AI Enablers (SW/infrastructure/data)

Evading Malware Detection with Adversarial Assembly-Level Code Generation

Luke Kurlandski, Matthew Wright and Yin Pan

Emails: *lk3591@g.rit.edu; Matthew.Wright@rit.edu; Yin.Pan@rit.edu*

In recent decades, malware has become a significant problem for an increasingly digitized society. Amidst the explosion of machine learning (ML) and deep learning (DL), feasible malware detection at scale has become more effective than ever. However, ML and DL systems are notoriously vulnerable to adversarial attacks. In the case of an adversarial evasion attack against a malware detection system, an adversary would carefully perturb an executable such that it maintains its malicious functionality, but is able to evade the classifier. To understand this class of threats before malicious actors do, the cybersecurity community has made a significant effort to develop techniques to craft adversarial malware. By doing this, security experts can design defense protocols against adversarial attacks and hope to remain one step ahead of adversaries. In our work, we propose a novel adversarial evasion attack that operates by directly modifying the source code of malicious software. We believe that we can identify pieces of code within malware that detection systems find suspicious using DL explainability theory. We then intend to substitute these suspicious sections of code with semantically equivalent code that appears benign. To perform the replacement, we propose a sequence-to-sequence model based upon large-language-modeling (LLM) for assembly-level code. We argue that our attack can be conjoined with any other adversarial attack to result in a more powerful attack with an improved evasion rate. We propose a thorough experiment to evaluate the effectiveness of our attack against a variety of hardened detection systems.

Topic: AI Applications

Efficient Distributed Deep Learning Using Tiny Serverless Functions

Kevin Assogba and M. Mustafa Rafique

Emails: *kta7930@rit.edu; mmrvcs@rit.edu*

Deep Learning (DL) workloads are composed of compute- and memory-intensive jobs, thus requiring a lot of data center resources for their timely completion. At the same time, serverless computing is an emerging cloud computing model that enables quick deployment and seamless scaling of applications without having to manage complex resources while improving application performance. However, serverless platforms impose resource-level constraints, i.e., fixed memory allocation and short task timeouts, which lead to job failures while running training and inference tasks of DL workloads. In this research, we address the memory and timeout constraints of serverless computing to efficiently run DL workloads. We propose a framework, called DiSDeL that integrates the serverless design paradigms with DL platforms and leverages data parallelism for distributing fine-grained DL tasks to serverless functions. We remove the inherent limitations of serverless computing by modeling the memory utilization of DL jobs and sharding the data into small subsets that are suitable for processing in serverless functions. Moreover, we accelerate DL jobs using GPUs to address the limited timeout constraint of serverless platforms. We implement DiSDeL using Apache OpenWhisk and TensorFlow platforms and evaluate it using representative DL jobs. Our evaluation shows that DiSDeL eliminates DL job failures. Moreover, it reduces the memory consumption of serverless functions and total training time by up to 44% and 46%, respectively, compared to the default execution approach of serverless platforms. Finally, DiSDeL improves the performance of concurrent DL jobs by up to 29% as compared to the bare-metal TensorFlow platform.

Topic: AI Enablers (SW/Infrastructure/data)

Art, AI and Opportunities for a New Art-Science

Carlos Castellanos

Email: *cxcigm@rit.edu*

The application of artificial intelligence (AI) techniques by artists have yielded a rich and diverse set of artworks since their earliest iterations in the 1950s. Much of this work is strange, often behaving quite unlike the technological systems we are accustomed to. How do we understand the meanings and intentions of these works? How do we analyze them? While we know that AI has proven effective at automating cognitive tasks where there is some notion of an optimal solution, this approach does not apply in the arts, where there is no "optimal" artwork and where goals are often ill-defined. If we accept that contemporary art cannot be reduced to images and aesthetics and is not about solving problems but about creating them for the audience — to challenge them so to encourage the emergence of new perspectives and experiences that may reveal new ways of looking at the world — then we can begin to understand where this work sits. In this presentation I will discuss how artists working in AI are contributing to the emergence of a new art-science by pursuing speculative, non-rational and non-utilitarian lines of inquiry. In essence I will show how artists "creatively misusing" the tools and methods of modern AI (and science and engineering more broadly) while subverting its ontological premises.

Topic: AI Applications

How Research Computing Can Help Your AI Research

Sidney Pendelberry, Ben Meyers and Kirk Anne

Emails: *slpits@rit.edu; bsmits@rit.edu; kmaits@rit.edu*

When we think of artificial intelligence, we tend to think of machine learning, deep learning, computer vision, natural language processing, genomics, or other artificial intelligence-driven research projects. RIT Research Computing can extend researchers' access to larger machines with more capabilities than a standard laptop or desktop workstation.

Rochester Institute of Technology has made a sizable investment over recent years in research computing to assist researchers in meeting their computational needs. Most of this investment centers around the computational cluster SPORC (Schedule Processing on Research Computing). The cluster offers over 150 Nvidia GPUs (100 A100, 16 V100, and 48 P4) distributed across 64 computational nodes with each node containing 36 CPU cores and 384 GB of RAM. Research Computing has thousands of pre-built packages and libraries in a variety of computer languages (e.g., Python, C, C++, FORTRAN, Go, Matlab, R). Shared storage space is available for research projects to store large amounts of data or training sets.

Research Computing has documentation and tutorials on how to get started using the resources in Institute Hall. We propose a demonstration and a tutorial for the AI@RIT Summit 2022. The demonstration briefly points to gaining access to RC, then focuses on using Python Torch libraries for parallel processing and distributed use of GPUs. The demonstration will introduce techniques to troubleshoot and improve job performance. The tutorial takes the researcher through setting up their cluster environment for building their packages to take full advantage of cluster resources.

Topic: AI Enablers (SW/Infrastructure/data)

Interpreting link predictions made by knowledge graph embeddings

Carlos Rivera and Narayanan Asuri

Emails: *crr@cs.rit.edu; nk1581@.rit.edu*

Link prediction is the task of predicting new links (edges) in a given knowledge graph G. This is done by assigning scores to new predictions, where these scores estimate the plausibility of the new prediction. A model computes scores based on embeddings, numerical vectors that encode the semantics and the structure of G. Interpreting these new predictions is crucial to advance the field of link prediction. Many applications like fact checking or knowledge completion rely on link prediction. In our research, we propose to interpret the link prediction task as a whole rather than individual predictions. Interpreting the whole task is appealing since individual interpretations can be biased and the big picture is lost.

Topic: AI Applications

Performing Look-up Table based Computing within the Memory for Fast and Ultra-efficient AI Performance

Purab Sutradhar

Email: *ps9525@rit.edu*

The rapidly growing AI algorithms are presenting severe performance demands to the processing hardware. Therefore, novel computing architectures are being designed to specifically meet AI computing demands. Processing in Memory (PIM) is a novel computing paradigm that enhances memory chips with computing ability to alleviate the memory-bandwidth limitation present in the state-of-the-art computing hardware (i.e. CPU and GPU). This work presents a novel PIM device that leverages tiny Look-up Tables (LUT) within its processing cores to offer ultra-efficient and massively parallel computing capability. The proposed hardware also has a highly flexible architecture which enables it to support a wide range of AI-oriented applications at minimal overheads.

Topic: Other

Co-Designing Features for AI-Supported Communication Applications using Automatic Captioning with Deaf and Hearing Pairs

Matthew Seita, Sooyeon Lee, Sarah Andrew, Kristen Shinohara and Matt Huenerfauth

Emails: *mss4296@rit.edu; sooyeon.lee@njit.edu; sa2941@rit.edu; Kristen.Shinohara@rit.edu; matt.huenerfauth@rit.edu*

Deaf and Hard-of-Hearing (DHH) users face accessibility challenges during in-person and remote meetings, including in workplace and academic settings. While the emerging use of AI-based applications incorporating automatic speech recognition (ASR) to translate speech to text in real-time on a user's personal device is promising, more user-interface and user-experience research is needed in this context. While co-design methods could elucidate designs for these applications, COVID-19 has interrupted in-person research. This poster discusses a novel methodology for conducting online co-design workshops with 18 DHH and hearing participant pairs to investigate ASR-supported mobile and videoconferencing technologies along two design dimensions: Correcting errors in ASR output and implementing automatic notification systems for influencing speaker behaviors. Our methodological findings include an analysis of communication modalities and strategies participants used, use of an online collaborative whiteboarding tool, and how participants reconciled differences in ideas. We additionally present some selected prototype designs our participants created. Our methodological findings showed our co-design process facilitated negotiations between DHH and hearing partners, and how communication needs of DHH users were met, so that both DHH and hearing participants could equitably engage in design activities together. Our design prototype findings provide a starting point for future researchers and developers to investigate the creation of AI and ASR-based technologies for use in real-world settings, providing DHH individuals greater access to spoken information in their everyday lives.

Topic: AI Applications; Other

Uncertainty Estimation using Evolving Recurrent Neural Networks and xResnet

Sheeraja Rajakrishnan and Daniel Krutz

Emails: *sr8685@rit.edu; dxkvse@rit.edu*

Machine learning models cannot make accurate predictions when presented with data that has not been seen before. A level of uncertainty is associated with the predictions. If these predictions were to be used in highly critical applications, such as in medicine, military or automated vehicles, the repercussions due to these unconfident predictions could be severe, some of which include death, incorrect treatment for a medical condition, etc. If the model was able to estimate the uncertainty for a prediction, a human could interfere and make decisions for situations with high uncertainty. This work focuses on optimizing the uncertainty estimation process when making predictions using irregular time series data. Irregular time series poses a challenge since the available data is sparse in many real-world applications. The machine learning models tend to make very confident incorrect predictions when predicting in areas of sparse data. To obtain accurate predictions from such irregular data, regular time series data often needs to be generated. Existing methods use Gaussian processes to generate regular time series data. To further improve the model accuracy, this work will use evolving recurrent neural networks (ERNN) to generate regular time series data and xResnet model for prediction. xResnet applies some modifications on the existing Resnet architecture, such as large-batch and low-precision training, to improve the model accuracy. ERNN has a better performance, compared to other methods, when predicting chaotic time series. As preliminary work, the uncertainty estimation accuracy of the proposed model will be evaluated against state-of-the-art prediction models.

Topic: Core AI Theory

Uncertainty Decomposition for Image Anomaly Detection

Yang Liu and Daniel Krutz

Emails: *yl4070@rit.edu; dxkvse@rit.edu*

The goal of this work is to propose an uncertainty decomposition technique that reflects reliability of the prediction along with interpretation for deep anomaly detection models. Uncertainty is used to measure the likelihood that the model is making correct predictions, and it is frequently classified into aleatoric and epistemic uncertainties based on whether the uncertainty is due to lack of knowledge or the nature of randomness. Current state-of-the-art methods primarily address predictive uncertainty quantification; the research on backpropagation of the uncertainty quantification is very limited. To the best of my knowledge, there is no existing method that can provide pixel level uncertainty decomposition for image anomaly detection. Pixel level uncertainty decomposition can provide valuable information regarding the specific regions that cause the uncertainty, and the nature of the uncertainty. Missing this capability prevents decision makers from further understanding the uncertainties of the model predictions. Therefore, this work will provide reliable uncertainty decomposition at both input pixel level and output prediction level with solid theoretical support. The proposed technique extends current plausibility-based uncertainty quantification methods to image anomaly detection. Preliminary results show the proposed method can provide sparse and interpretable heat maps that reflect what features from the input image can cause uncertainty in the model prediction.

Topic: Core Theory AI

Data Corruption effects on Machine Learning & Federated Learning models

Rahul babu Ganesh, Sergei Chuprov and Kartavya Manojbhai Bhatt

Emails: *rg9233@rit.edu; sc1723@rit.edu; kb8077@rit.edu*

Federated Learning empowers devices to learn a combined predictive model with the entire training data present within the device. The prime feature of Federated Learning is that there is no need for centralized data collection for Machine Learning algorithms to train. This data privacy comes with a trade-off in the performance of the Machine Learning algorithm. In this study, we investigate the influence of Data Quality on the performance of the Federated Learning system. In particular, we focus on establishing benchmarks for state-of-the-art MNIST dataset and custom traffic-sign dataset that incorporates original and corrupted traffic sign images. We employ conventional centralized Machine Learning models and compare their performance against their Federated Learning implementation. As a part of our study, we analyze interrelationships between the degree of data corruption and the performance of both centralized and Federated Learning. To facilitate our empirical study, we employ such open-source state-of-the-art image recognition model architectures, as VGG and ResNet. We train and test them in both centralized and distributed learning conditions and evaluate the cross-category prediction performance results. We compare and analyze these results and provide our conclusions on the established relationships.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Analyzing anomalous events using context and explainability

Dipkamal Bhusal and Nidhi Rastogi

Emails: *db1702@rit.edu; nxrvse@rit.edu*

Understanding the reasons behind a machine learning prediction is crucial in several applications especially security sensitive events like intrusion detection. In enterprise networks, network endpoints record threat information in security logs. Intrusion detectors analyze logs and generate threat alerts which security analysts inspect. Because of the high volume of false alerts, alert fatigue is a growing problem for analysts. Security attacks have also become increasingly sophisticated for reliable detection using statistical and rule-based methods. Deep learning models have improved detection of intrusion using security logs however they do not provide insights on the prediction. Existing alerts lack sufficient explanation on why it was raised or context in which it occurred. This information is helpful to security analysts for quick and accurate validation of alerts; lack of which, forces to perform an extensive manual inspection of security events. In our research, we argue that contextual analysis and model explainability are promising solutions to address existing challenges in intrusion detection. Contextual analysis of anomalous events in security logs helps detect complex and evolving attacks and reduces the high volume of false alerts. We utilize relational contextual data as the context for our preliminary experiments and employ sequence models to discover long-term dependencies of security events in logs. We also propose a model agnostic local explanation method to identify the log events that contribute to the prediction of a security event. Security analysts can use this information to evaluate the reason behind predictions and inspect alerts quickly, reducing time to validate the alerts.

Topic: AI Applications

Decision Tree-Based Parameter-Free Method for Chaotic Time Series Forecasting

Adam Giammarese, Kamal Rana and Nishant Malik

Emails: *amg2889@rit.edu; bsmitskr7843@rit.edu; nxmsmas@rit.edu*

Forecasting the future evolution of chaotic data-based systems is a crucial, but challenging practical problem. Existing solutions - such as Recurrent Neural Networks (RNNs) or Reservoir Computing (RC) - have been shown to be effective methods of forecasting time series, but require a slew of parameters and hyperparameters. In this work, we discuss a mostly parameter-free machine learning approach to chaotic time series forecasting and feature selection in the form of an Extra-Trees Regressor (ETR), which utilizes an ensemble of regression trees. We develop a method involving ETRs (for both feature selection and forecasts given such features) that provides notable performance in forecasting the future evolution of chaotic time series with limited information about the system itself. Using the logistic map and Lorenz system as toy discrete and continuous systems, respectively, we demonstrate the efficacy of the developed forecasting method. Beyond the prototypical examples, we apply the ETR-based forecast method to various (and more difficult to predict) systems, such as the spatio-temporal Kuramoto–Sivashinsky system and real-world climate time series. In comparison to the existing RNN and RC methods, we observe that our method provides comparable (if not impressive) performance while requiring far fewer parameters and hyperparameters that traditionally make the former systems difficult to use in practice.

Topic: Core AI Theory; AI Applications

How AI is Changing the Way We Teach Technical Design in the Classroom

Alejandro Perez Sanchez

Email: *axpfaa@rit.edu*

The application of artificial intelligence (AI) techniques by artists have yielded a rich and diverse set of artworks since their earliest iterations in the 1950s. Much of this work is strange, often behaving quite unlike the technological systems we are accustomed to. How do we understand the meanings and intentions of these works? How do we analyze them? While we know that AI has proven effective at automating cognitive tasks where there is some notion of an optimal solution, this approach does not apply in the arts, where there is no "optimal" artwork and where goals are often ill-defined. If we accept that contemporary art cannot be reduced to images and aesthetics and is not about solving problems but about creating them for the audience — to challenge them so to encourage the emergence of new perspectives and experiences that may reveal new ways of looking at the world — then we can begin to understand where this work sits. In this presentation I will discuss how artists working in AI are contributing to the emergence of a new art-science by pursuing speculative, non-rational and non-utilitarian lines of inquiry. In essence I will show how artists "creatively misusing" the tools and methods of modern AI (and science and engineering more broadly) while subverting its ontological premises.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Industrial Design Takes on AI

Juan Noguera, Ariella Knight, Anqi Zhu, Jos Mayo, Jayden Zhou, Jaqueline Qiu, Robert Deane, Shen Liu, Tao Oke and Zaheer Shujayee

Emails: *jcnfaa@rit.edu; abk6297@rit.edu; az3635@rit.edu; jhm7015@rit.edu; jz8727@rit.edu; lq5403@rit.edu; rbd7436@rit.edu; lq5403@rit.edu; sl7987@rit.edu; teo3632@rit.edu; zms5893@rit.edu*

AI Image generators like Midjourney and Dall-E offer an endless source of potential inspiration and material for artists and designers. Our project seeks to find what these tools could mean for the Industrial Design profession, where designers are constantly conceptualizing new objects, form factors and solutions.

A group of 9 Graduate Industrial Design students have sought to insert AI image generation into their Product Design process, by using it in their ideation, to get inspired, or to get past a block. They have seen their product ideas go to unexpected places and they have used AI in a variety of ways. Each of them has created a complete product/object concept and/or prototypes, has documented their process and created a series of visuals. This project is ongoing and will be complete by the time of the AI event at RIT.

Their visuals, writing and process will be featured in a future edition of Core77 magazine to be released later this year. (One of the premier publications for the Industrial Design industry).

We seek to share our results with the RIT community, both in the poster session format, and in an oral presentation. Professor Juan Noguera is prepared to present the project and its results, as well as some of the conversation sparked by it and how the design industry has reacted to the rapid evolution of these tools.

Topic: AI Applications; AI Ethics and Policy; AI Impacts on Academia, Industry, and Society

Planning session for our next generation cluster

Kirk Anne

Email: *kirk.m.anne@rit.edu*

As our current SPORC cluster is approaching its five year anniversary, the Research Computing team is gathering requirements for the next generation of our cluster. We will facilitate a discussion on our current resources and what we should be looking to include in our next cluster and future acquisitions.

Topic: AI Enablers (SW/Infrastructure/data)

Predictive Modeling and Medication Adherence

Teresa Gibson

Email: *tbgsma@rit.edu*

BACKGROUND: Predicting a patient's future medication adherence patterns has proved to be challenging for healthcare providers and payers. Most adherence modeling does not take previous adherence into account.
METHODS: Adherence was defined using a dichotomous measure indicating at least 80% of days in a calendar quarter with medication on hand, consistent with previous research. The adherence behaviors of 53,709 continuously enrolled individuals in employer-sponsored health plans drawn from a multi-employer database were analyzed using a state-dependence framework. This allowed for the estimation of the extent of carryover in adherence from one calendar quarter to another while adjusting for observed and unobserved heterogeneity and enrollee characteristics. All enrollees were observed across 14 calendar quarters. This study focused on prescriptions in 3 maintenance medication classes: lipid-lowering medications, antihypertensive medications, and oral antidiabetes medications. Adherence model classification performance was summarized by AUC with 10-fold cross validation.
RESULTS: Marginal effect estimates for prior adherence (previous quarter and initial quarter) showed increases in predicted adherence when adherent in the previous quarter (8.7 percentage points [pp] [95% CI=8.0-9.3pp] for lipid-lowering medications), when adherent in the initial quarter (14.4pp [13.8-15.1pp]) and in the initial and previous quarter (22.7pp [22.1-23.3pp]). Similar patterns held for antihypertensive and oral antidiabetes medications. AUC for dynamic random effects probit models exceeded 0.85 in the 3 medication cohorts (e.g., AUC 0.884 (0.873, 0.896) lipid-lowering medications), whereas static models remained under 0.7.
CONCLUSIONS: Predictive modeling incorporating state dependence, past behaviors, and observed and unobserved heterogeneity showed considerable improvement over traditional modeling approaches.

Topic: AI Applications; AI Impacts on Academia, Industry, and Society

Diagnostic & Prognostic Assessment of Gas Compression Technology

Jacob Chesnes, Jason Kolodziej and Michael Anderson

Email: *jjc4939@rit.edu; jrkeme@rit.edu; mwa2912@rit.edu*

Our lab pursues research into machine health monitoring of wear components on gas compression technology primarily used in the oil and gas industry by way of advanced signal processing and machine learning approaches. To accelerate testing faulty components are seeded in the compressor by precision machining new parts in similar ways that they would be worn through regular use over extended periods of time in the field. Diagnosis starts by processing raw data, such as valve vibration or cylinder pressure, using time-frequency analysis and then extracting key features. Targeted features are determined that represent mechanical and thermodynamic processes in the compression cycle that would show change with the different faulty components. Machine learning classification algorithms, like support vector machines and discriminant classifiers, are trained on the features to predict fault type and severity. Deep learning has recently been explored at a higher computational cost which eliminates the need of creating features manually but potentially improves machine health prediction. These compressors are expected to run round-the-clock and condition monitoring methods reduce machine downtime and increase efficiency. Servicing only when required or predicting future failure is a critical necessity for productive operation in this industry.

Topic: AI Applications

Human-AI Collaboration for UX Evaluation: Visualizations and Conversational Assistants

Emily Kuang

Email: *ek8093@rit.edu*

Reviewing a think-aloud video is both time-consuming and demanding as it requires UX evaluators to attend to many behavioral signals of the user in the video. Inspired by the promise of AI in assisting UX evaluators with analyzing usability tests, we developed a collaborative visual analytics tool, CoUX, to support usability problem identification, annotation, and discussion in an integrated environment. CoUX visualizes a set of problem-indicators based on acoustic, textual, and visual features extracted from the video and audio with machine learning. The results from a user study with twelve UX evaluators indicate that CoUX is useful and effective in facilitating problem identification. However, participants had questions about observations from test recordings or results from the AI, which could not be asked since the visualizations were non-interactive. In contrast, interactive conversational agents provide the opportunity for a Question & Answer paradigm, which allows evaluators to ask specific questions and may improve their trust in AI. Thus, we explored how to best design text and voice AI assistants for UX evaluation. We conducted a user study to identify what types of information they sought from an AI assistant and how they asked these questions when using text vs. voice. We found that they asked for information regarding the users' actions and emotions from the recordings and redesign suggestions from the AI assistant. Future work includes examining the viability of other AI manifestations that assist with UX evaluation.

Topic: AI Applications

Developing a captioning style that can help identify a speaker's emotions, moods, and emphasis

Calua de Lacerda Pataca, Mathew P. Watkins, Roshan L. Peiris, Sooyeon Lee and Matt Huenerfauth

Emails: *cd4610@rit.edu; mxw7981@rit.edu; rxpics@rit.edu; sooyeon.lee@njit.edu; matt.heunerfaught@rit.edu*

Captions translate spoken utterances into written text, but this process is not lossless. While a speaker's voice might signal their emotions, moods, age, and gender, almost all of this is lost when captions reduce speech to only its words. Having accurate and coincident captions is, by all means, worthwhile, in particular when making accessible the communication between hearing and Deaf and Hard-of-Hearing (DHH) individuals, but what are the effects of removing paralinguistic dimensions from speech? A previous study showed that DHH individuals felt that captions are dull and ambiguous to parse. Based on this, we set out to investigate whether captions overlaid with emotions, prosody, or both, could influence how DHH participants could identify a speaker's emotions, moods, and emphasis. We developed a pipeline that took as input a spoken utterance and its transcription, force-aligned one to the other, and, using each word's timestamp, extracted its prosody (loudness, pitch, and rhythm) and, through a transformer-based speech-emotion recognition neural network, emotions (valence and arousal). We then applied these features to caption text as typographic modulations, i.e., specific changes to a font's appearance where one visual parameter echoes one paralinguistic dimension. Next, a comparative study paired these modified captions against their conventional counterparts. The emotion-based style over-performed all others in terms of helping to identify emotions, moods, and emphasis, with participants commenting that it made for a more engaging experience, suggesting its potential as a more inclusive design for captions.

Topic: AI Applications; Other

Data Entry, Labor Identity, and Inequality

Sidney Pendelberry, Ben Meyers and Kirk Anne

Emails: *slpits@rit.edu; bsmits@rit.edu; kmaits@rit.edu*

As artificial intelligence technologies are expected to change the balance between human and machine labor, my project provides a historical lesson from computer automation to better understand these changes. For the area of content moderation on social media, Sarah T. Roberts has shown that, although we may believe that an AI technology or algorithm is behind these functions, it often is a human. Similarly, my project investigates the human labor in data entry that enabled computer automation.

When modernizing corporations first installed electronic computers, it became clear that large amounts of information had to be converted from paper into computer-legible form before they could be processed automatically. Computer automation thus created a rapidly ballooning need for data entry. Most data entry was done manually because technical solutions, for example through character recognition technologies, proved technically or economically unfeasible. This left improvements in human-computer interfaces as the main recourse to alleviate manual data entry work. My project makes data entry a visible part of the history of computing. How did computer automation change work processes? Who took up the newly emerging routine data entry work, under what conditions, and why?

Focusing on banking automation in the United States, West and East Germany from the 1950s to the 1970s, my project investigates computing technology and inequality at the height of the Cold War. I examine changes in the participation and identity of data entry workers and their work conditions, and I also raise questions about social and economic implications from bank automation.

This project is supported by NSF Award No. 2218577 and a Virtual Fellowship from the Linda Hall Library.

Topic: AI Impacts on Academia, Industry, and Society

Machine learning predicts genes associated with cancer metastasis

Isaac Olatunji and Feng Cui

Emails: *io3247@rit.edu; fxcsbi@rit.edu*

Metastasis refers to the dissemination of cancer cells away from the primary site of origin to form colonies in distant organs. This single hallmark of cancer is responsible for a high number of cancer related mortalities. Identification of biomarkers that are associated with metastasis is useful in making clinical decisions and could form the basis for development of new therapeutics.

While many studies have confirmed that gene expression varies across different cancer types, mechanisms of cancer metastasis remain elusive. Several differential expression analyses have found a set of tissue-specific genes that moderate the site of distant metastasis, whereas other studies have not identified specific metastatic signature genes. It remains unclear whether there exists a set of genes that are common for all metastatic cancers.

To address this issue, we applied both traditional gene expression analysis methods and machine learning (ML) methods to genome-wide RNA sequencing data of three types of cancers, Pancreatic Adenocarcinoma, Bladder Carcinoma, and Head and Neck Squamous Carcinoma, which were downloaded from The Cancer Genome Atlas. First, genes that are differentially expressed in metastatic cancers vs. non-metastatic cancers were identified by the DESeq2 package. Second, genes that are important for differentiating the two groups of cancer were identified by an optimized features selection approach using Random Forest. Lastly, the genes found by the two methods were used to build classifiers, and examined the metrics in each of the groups. We found that genes selected by the ML methods outperformed the genes selected by DESeq2 in the prediction of samples as metastatic or not metastatic. Moreover, we found that depending on the threshold set, there are more overlapping genes between any two of the three cancer types considered, however very few gene expression overlap (if any) was detected across all the groups. Overall, our preliminary results from these tasks indicate that metastatic genes are in fact more specific to the cancer origin, rather than general across multiple cancer types. This work is a segment of a whole research on use of multimodal data which include histopathology images, genomic data, and clinical data for prediction of diagnosis, and outcomes in cancer patients.

Topic: AI Applications

Interpretable Error Optimization in Forward and Inverse Problems

Maryam Toloubidokhti

Email: *mt6129@rit.edu*

A critical challenge in learning-based reconstruction of systems' parameters, is incorporating partial knowledge about the governing physics into the learning. In complex systems, the interactions among system's variables lead to the observed collective behavior. Thus, the measurements (Y) are often not the direct measurements of individual latent variables (X) but the result of their interactions (Y = f(X)). This physics-based relationship is called the forward model. The goal is then to estimate X from Y, or solving the inverse problem. Medical image reconstruction is an example of this. Traditional optimization and inference techniques exploit the physics knowledge by embedding the forward model in the inverse equations. However, the inverse solution's accuracy is affected by the accuracy of the forward model which is prone to error in the underlying parameters. Alternatively, deep learning methods, while successful, usually bypass prior physics and directly learn the forward or inverse mapping from the data, requiring large labeled datasets which are usually unavailable. We bridge the gap between traditional and learning-based approaches by embedding the mechanistic forward operator inside a neural function, and modeling the distribution of errors in an interpretable manner. We train a conditional generative model that transforms a given mechanistic operator with unknown errors, arising from a latent space of self-organizing clusters of potential sources of error. The generative model is embedded in a simultaneous optimization process of finding the inverse solution and uncovering and minimizing the error. We applied this method to reconstruction of heart electrical potentials from body surface potentials.

Topic: AI Applications

Supporting Ethical Artificial Intelligence and Machine Learning Education Using Experiential Labs

Su Thit Thazin, Heather Moses, Andres Leonard-Calcano and Kyle Messerle

Emails: *st5626@rit.edu; hlm8500@rit.edu; al9824@rit.edu; klm3580@rit.edu*

Applied artificial intelligence (AI) and machine learning (ML) are consistently growing in prominence. Despite this growth, the consideration of the ethical impact of AI has lagged in comparison. There are a myriad of real-world ethics-related challenges that range from discriminatory bias to a lack of trustworthiness among many users against AI. Unfortunately, the educational community is limited due to a lack of easily adoptable experiential material regarding bias in AI/ML. At the Accessible Learning Labs (ALL) project, we are developing several experiential educational labs that focus on AI/ ML. For instance, the theme of our "Ethics in AI" lab is the inequity which can arise from using AI in a controlled work environment. Our labs consist of four components: I) Relevant background information on the topic being addressed, II) An application that demonstrates the problem at hand and teaches the user how to address the issue from a technical standpoint, after which the user experiences the "repaired" application, III) Empathy-creating material showcasing people's real-life experiences on how the issue has impacted their life, and IV) A quiz where users will be tested on their knowledge of the information taught in the lab. Studies done on our existing computing accessibility-related labs have shown that this experiential learning format, combined with empathy-creating materials, is more effective than passive learning formats in informing students about the issues as well as fostering student motivation. Our new lab regarding bias in AI/ML is the first known work to deliver fully self-contained experiential learning.

Topic: AI Applications; AI Ethics and Policy

Image Based Machine Learning for Automated Sorting Core Parts for Remanufacturing

Abu Islam and Suvrat Jain

Emails: *asigis@rit.edu; sxjgis@rit.edu*

Remanufacturing of durable goods has the potential to prevent materials from going to landfills or being melted down for recycling. One of the initial steps in remanufacturing of a component is to sort and inspect the components returned from the field, these returned units are known as cores within the remanufacturing industry. Most of the current sorting and handling processes are very labor intensive, error prone and can have poor ergonomics. An automated part identification and sorting process is a potential solution for staffing shortages given its ability to effectively sort and inspect cores that have similar geometries or set of features. A neural network has been trained with images of different automotive part at multiple orientations. Inception Transfer Learning has been used to reduce the number of training data requirement, speeding up training and higher accuracy. A Siamese network has been implemented to flag new parts that the model has not seen before. An automated sorting system has been developed that consists of a smart conveyor, multiple cameras, and laser line scanners. Once on the conveyor, the automated sorting system coordinates the movement and imaging of the core at multiple stations in order to capture a near-360 degree view of the core part. The algorithm detects the part types and models from the images captures from the vision system. It has been demonstrated that the vision based sorting system can classify parts with better than 95% accuracy. This paper will present the results in detail.

Topic: AI Applications

Landsifier: a Python library to estimate likely triggers of mapped landslides using machine and deep learning

Kamal Rana and Nishant Malik

Emails: *kamalrana520@gmail.com; nxmsma@rit.edu*

Landslide hazard models aim at mitigating landslide impact by providing probabilistic forecasting, and the accuracy of these models hinges on landslide databases for model training and testing. Landslide databases at times lack information on the underlying triggering mechanism, making these inventories almost unusable in hazard models. We developed a Python-based library, landsifier, that contains three different Machine-Learning frameworks for assessing the likely triggering mechanisms of individual landslides or entire inventories based on landslide geometry. Two of these methods only use the 2D landslide planforms, and the third utilizes the 3D shape of landslides relying on an underlying Digital Elevation Model (DEM). The base method extracts geometric properties of landslide polygons as a feature space for the shallow learner—Random Forest (RF). An alternative method extracts topological properties of 3D landslides through Topological Data Analysis (TDA) and then feeds these properties as a feature space to the Random Forest classifier. The last framework relies on landslide-planform images as an input for the deep learning algorithm—Convolutional Neural Network (CNN). We tested all three inter-changeable methods on several inventories with known triggers spread over the Japanese archipelago. To demonstrate the effectiveness of developed methods, we used two testing configurations. The first configuration merges all the available data for the k-fold cross-validation, whereas the second configuration excludes one inventory during the training phase to use as the sole testing inventory. Classification accuracies for different testing schemes vary between 70 % and 95 %. Finally, we implemented the three methods on an inventory without any triggering information to showcase a real-world application.

Topic: AI Applications

Using Deep Learning to Measure Galaxy Redshifts

Rohan Pattnaik

Email: *rp2503@rit.edu*

Redshift is one of the most fundamental and crucial measurements in observational extra-galactic astronomy. It is used to measure the distance to galaxies and allows us to convert spectra from these objects into their rest-frame. Converting spectra to rest-frame is crucial as it allows us to identify intrinsic properties of these objects. These properties help us study how galaxies in the Universe evolved over cosmic time. Current methods of spectroscopic redshift measurement involves visually inspecting each individual spectrum to identify spectral features or cross-correlating observed spectra with galaxy templates at varying redshifts. Both these processes can be either time and labor-intensive or prone to errors. We explore the use of Deep Learning techniques such as a Convolutional Neural Network (CNN) to measure redshifts directly from an input spectrum as a fast and reliable alternative to the current methods.

Topic: AI Applications

Unsupervised Machine Learning of Galaxy Morphologies using Vector-Quantized Variable Auto Encoder

James Liu

Email: *jjl4166@rit.edu*

Since 1926, when the Hubble Sequence was first invented, astronomers have visually classified galaxies based on morphology. With the exponential growth of astronomical data from new state-of the art telescopes, such as the recently launched James Webb Space Telescope (JWST), astronomers have turned to artificial intelligence to automate image classification tasks. We utilize a Vector Quantized Variable Auto Encoder (VQ-VAE), a machine learning architecture consisting of an encoder and decoder, to compresses image data into a smaller discrete representation (latent space). A clustering algorithm is used on the compressed data to group them into a number of clusters, where each cluster represents a type of galaxy morphology. We present preliminary classification results from the algorithm when applied to galaxies in simulated JWST images.

Topic: AI Applications

Computational Implementation and Demonstrations of LUSI

Zi-Jia Gong and Ernest Fokoue

Emails: *zg3988@rit.edu; epfeqa@rit.edu*

LUSI (Learning Using Statistical Invariants) is a new machine-learning paradigm proposed by Vapnik and Izmailov. In LUSI, a classification function is searched in the reproducing kernel Hilbert space (RKHS) by minimizing the loss function, while a set of 'predicates', functionals on the training data that incorporate the specific knowledge of the machine-learning problem, are preserved invariant. In this project, we implemented a few experiments of LUSI in Python in order to evaluate its classification performance. We use a 2D banana-shape dataset and the MNIST dataset to fit various LUSI models, while the performances of different LUSI variants and different predicates are compared. Our LUSI code is designed to be compatible with scikit-learn, and is open-source on GitHub.

Topic: AI Applications

Artificial Intelligence in Space Exploration

Michael Barbosu

Email: *mxbsma@rit.edu*

Artificial Intelligence (AI) was first used in space exploration over 20 years ago. Since then, the number of applications in this field has continued to grow exponentially.

AI has become an essential tool in many applications, such as autonomous spacecraft navigation, Moon and Mars rovers' operations, satellite data and image processing, tracking artificial satellites and space debris, training and supporting astronauts is space missions, and discovering new planets.

Here we will review some of these applications and will present current projects of RIT's Space Exploration Group (SPEX).

Topic: AI Applications

Improving the robustness of AI-enabled software systems

Mohamad Fazelnia

Email: *mf8754@rit.edu*

Artificial intelligence (AI) plays an essential role in various software systems. However, this popularity of AI techniques in various software applications attracts malicious actors and adversaries. To design secure and robust AI-enabled systems, developers need to be aware of AI algorithms' weaknesses and vulnerabilities, as well as the potential defensive strategies that can be employed to mitigate the vulnerabilities and prevent cyber-attacks. To this end, first, we present a framework to characterize AI-specific vulnerabilities, the attacks associated with AI-enabled systems, and the corresponding defensive strategies. This framework aims to support AI developers to take proactive measures in developing AI-enabled software systems, understand such systems' attack surfaces, and develop robust systems against various emerging attacks associated with AI techniques.

In the next step, built upon the proposed framework, we aim to represent a technique to enable AI developers to take proactive security-related measures during software development by analyzing the source code. This method utilizes static program analysis and natural language processing techniques to support AI developers to automatically find AI-specific weaknesses, reason about the possible attacks against the system, and integrate appropriate mitigation strategies to improve the system's robustness. Due to the automatic and static nature of this approach, it requires fewer computational resources than the existing methods and can support secure AI-enabled software development from the early stages of development.

Topic: AI Enablers (SW/Infrastructure/data)