CAPWIC 2025 Technical Program

CAPWIC is an ACM Capital Region Celebration of Women in Computing and provides a low-cost, regionally tailored, small conference for women and minorities in computing. The participants include women (students, faculty, and professionals), as well as all supporters of women in computing.

The conference will be hosted by The George Washington University, on March 29, 2025. All levels of experience are welcome.

Sponsors

CAPWIC 2025 would not be possible without the work of the Organizers and Sponsors. We are so grateful for our many sponsors (see below).

Platinum

THE GEORGE WASHINGTON UNIVERSITY
WASHINGTON, DC

Gold

GAMA-1 TECHNOLOGIES

JMU Department of Computer Science

Academic

Keynotes

Oh, The Places You'll Go!

Authors

Dr. Quincy K. Brown, Ph.D.

Abstract

Dr. Quincy K. Brown is an accomplished computer scientist, educator, former Policy Director in the White House, and expert on STEM education, policy development, cross-agency collaboration, and workforce initiatives in space, computing, technology, and STEM. Dr. Brown will discuss leveraging her background as a computer scientist across research, education, and policy to drive innovation and shape the future. Her expertise offers valuable perspectives and practical insights on innovation, collaboration, and career navigation.

Closing Remarks

Authors

Becky Robertson

Abstract

Becky Robertson is a Vice President at Booz Allen Hamilton, where she leads amazingly innovative teams building advanced technology solutions and products to accelerate outcomes for the Nation. In her more than 30-year career, Becky has blended her technical expertise, problem solving skills, and passion for collaboration in roles including web-based training development, software engineering project management, systems engineering and integration, commercial satellite vulnerability analysis, large contract management, competitive contract capture and negotiation, as well as leading corporate investments in next wave technologies. In her current role, Becky leads Booz Allen’s National Missions team, tackling priority national cyber issues with collective ingenuity and unmatched mission understanding. She lives in the Annapolis area with her family and spends almost all of her spare time reading mysteries, cooking for the people she loves, or on soccer and lacrosse fields.

Panels

Choosing a grad school...or not: Questions, reflections, and discussion

Authors

S. McCrickard, N. Basit, N. Andrus, F. Nikseresht, A. Shahid, A. Prakash

Exploring Various Career Paths in Computer Science and the Tech Industry

Authors

N. Perodin, N. Schneider, I. Gupta

Research Shorts

Plot'n Polish: Zero-shot Story Visualization and Disentangled Editing with Text-to-Image Diffusion Models

Authors

Kiymet Akdemir, Pinar Yanardag (Virginia Tech)

Abstract

Text-to-image diffusion models have demonstrated significant capabilities in generating detailed and diverse visuals across various domains, with story visualization emerging as a particularly promising application. However, as their use in real-world creative processes increases, the need for generating coherent story sequences, providing enhanced control, refinement, and the ability to modify images post-generation to meet specific creative demands has become an important challenge. Existing methods often lack the flexibility to apply fine or coarse edits while maintaining visual and narrative consistency across multiple frames, preventing the creators from seamlessly crafting, refining, and enhancing their visual stories. To address these challenges, we introduce Plot'n Polish, a zero-shot framework that enables end-to-end interactive story generation and provides fine-grained control over story visualizations at various levels of detail.

Context Canvas: Enhancing Text-to-Image Diffusion Models with Knowledge Graph-Based RAG

Authors

Kavana Venkatesh, Yusuf Dalva (Virginia Tech); Ismini Lourentzou (University of Illinois Urbana Champaign); Pinar Yanardag (Virginia Tech)

Abstract

We introduce a novel approach to enhance the capabilities of text-to-image models by incorporating Retrieval-Augmented Generation with a knowledge graph. Our system dynamically retrieves detailed character information and relational data from the knowledge graph, enabling the generation of visually accurate and contextually rich images. Furthermore, we propose a self-correcting mechanism within Stable Diffusion models to ensure consistency and fidelity in visual outputs, leveraging the rich context from the graph to guide corrections. To our knowledge, Context Canvas represents the first application of graph-based RAG in enhancing T2I models, representing a significant advancement for producing high-fidelity, context-aware multi-faceted images.

Use of Control Flow Graphs with Edges Consideration for Fault Localization

Authors

Zhuo (Cecilia) Chen (Bryn Mawr College); Christian Murphy (Swarthmore College)

Abstract

In the final stage of software development, debugging, especially fault localization, is one of the most time-consuming processes for both novice and senior software engineers. One way to locate the fault is by testing the code with different inputs and looking for statements that are predominantly covered by failing test cases. This process can be used to then calculate the suspiciousness of individual lines of code. Previous work introduced a Spectrum-Based Fault Localization (SBFL) technique with visualization to help developers identify suspicious lines of code. Here, we build on that work by displaying a Control Flow Graph (CFG) as a visualization technique for fault localization assistance, with the most suspicious line(s) of code highlighted to draw the programmer's attention, and other statements highlighted to indicate suspicious paths through the graph. Further, we include the consideration of edges in the graph while computing the suspiciousness of code lines, which helps localize faults that are in decision points in the CFG. In this paper, we provide an overview of our CFG-based approach with some examples of visualizations that are easy for developers to understand, as well as an explanation of how we

use edges in the graph to assess the suspiciousness of method decision points. We modified two existing approaches Tarantula and Ochiai to Tarantual+Edges and Ochiai+Edges. We also describe the results of experiments using both 71 various valid real software faults from three different libraries provided by \textit{Defects4J} benchmark and various Java methods from the Apache Commons Math library with mutation analysis, which demonstrates that our technique can be more effective at helping developers localize faults than previous approaches.

Emerging Technologies and Autonomous Interaction

Authors

Ileana Perez Ruiz (University of Mary Washington)

Abstract

This research explores how AI, virtual worlds, and decentralized autonomous organizations (DAOs) are creating novel paradigms of autonomous economic and social interaction. By integrating AI-driven avatars, blockchain-based land ownership, and smart contract governance, these technologies are reshaping digital engagement models. The study examines how these interconnected systems enable unprecedented levels of user autonomy, economic agency, and social coordination beyond traditional digital platforms. Preliminary analysis reveals potential transformative implications for business, social networking, and property rights in increasingly decentralized digital ecosystems.

Combining Open-Source Intelligence (OSINT) with AI for Threat Detection

Authors

Jackline Fahmy (Marymount University)

Abstract

Abstract The increasing sophistication of cyber threats, coupled with the expansion of social media platforms and the Dark Web, has significantly impacted global cybersecurity and intelligence efforts. Despite the growing use of Cyber Threat Intelligence (CTI), Open-Source Intelligence (OSINT), and Social Media Intelligence (SOCMINT), significant gaps remain in the automation, accuracy, and ethical governance of these intelligence-gathering methods. This study addresses the urgent need to enhance cyber threat detection, zero-day attack prevention, and cybercrime mitigation through artificial intelligence (AI), machine learning (ML), and Natural Language Processing (NLP) techniques. The primary research problem centers on the effectiveness, challenges, and ethical implications of AI-driven cybersecurity solutions in

monitoring and analyzing cyber threats across OSINT, social media platforms, and the Dark Web. The study is grounded in cybersecurity threat modeling and intelligence analysis frameworks, integrating machine learning-based classification, natural language processing techniques, and social network analysis (SNA) to evaluate threat actors, cybercriminal behaviors, and cyber-attack patterns. The key research questions explore how AI and NLP enhance zero-day attack detection, cybercrime classification, and intelligence gathering from social media and Dark Web platforms while addressing regulatory, ethical, and legal constraints. A systematic review and comparative analysis were conducted to assess existing AI-driven cybersecurity methodologies. The research utilizes text mining, clustering algorithms, sentiment analysis, and predictive modeling techniques to analyze threat intelligence extracted from real-time data sources such as Twitter, Facebook, Dark Web forums, and hacker marketplaces. Data analysis focuses on machine learning success rates, false positives, and precision-recall metrics in cyber threat detection. Findings indicate that AI-driven cybersecurity models can detect cyber threats with up to 80% accuracy, particularly in zero-day attack identification and cybercrime profiling. However, challenges remain in automating intelligence workflows, ensuring data reliability, and addressing privacy concerns. The study concludes that while AI significantly enhances cyber threat intelligence, more robust frameworks are needed to balance security imperatives with ethical considerations. Recommendations include developing standardized AI-based cybersecurity models, enhancing international regulations for SOCMINT and OSINT, and integrating human-AI collaboration for more effective cybercrime mitigation.

Machine Learning Insights into Academic Success in CS3: The Role of Mathematics and CS Coursework

Authors

Nawar Wali, Sara Hooshangi (Virginia Tech)

Abstract

This research paper aims to understand how mathematics and computer science education influence the academic trajectory and performance of students in Computer Science degree programs, with a particular focus on CS3. By analyzing data from 3905 students over a decade at an R1 institution, this study examines course terms, grades, course completion, and program participation for all Computer Science and Math courses. Through extensive statistical and machine learning analysis on this comprehensive dataset, we investigate the correlation between students' performance in mathematics and previous CS courses and their grades in the upper-level CS course, CS3. This study provides a nuanced understanding of the factors contributing to student success in advanced Computer Science coursework.

Enhancing Neuronal Connectivity Inference with Cross-Correlograms and Interpretations from Learned Model Representations

Authors

Xiaoqian Sun, Rahul Simha, Chen Zheng, Hui Lu (The George Washington University)

Abstract

A key challenge in computational neuroscience is to understand the architecture of neuronal circuits, which in turn determines their ability to process information, respond to external stimuli and generate behaviors. Understanding how neurons wire and interact is essential for uncovering the principles of neural computation. Recent imaging techniques enable in vivo observation of live animal brain and approximate recovery of neural signals, raising the question: Can these signals be used to infer neuronal connectivity? Since 1970s, cross-correlations (CCGs) have been used as a heuristic to infer synaptic connectivity by identifying a narrow peak at a lag of few milliseconds (Perkel et al., 1967). Leveraging this feature, Endo et al. (2021) proposed CoNNECT, applying a convolutional neural network (CNN) to CCG to infer connectivity, achieving superior performance in simulated data. While effective, plain CCG is limited by various inherent constraints in the signal including bursty activity, periodic firing, and slow transient background contributions. To address these limitations, we propose an enhanced preprocessing pipeline that refines the CCG by removing side lobes caused by intrinsic spiking patterns and disentangling slow background fluctuations, which could further improve upon CoNNECT‚Äôs performance. This approach could potentially provide insights into understanding the underlying mechanisms driving CNN performance. A detailed analysis of the feature maps and learned representations within the CNN layers could reveal how the model prioritizes key features from engineered CCG. Our investigation is based on a detailed neuronal simulation that accurately models neuronal networks, providing signals to assess an inference approach.

BUILDING TRUST AND TRANSPARENCY FOR GOVERNMENT CLOUD ADOPTION: A DATA SECURITY POSTURE MANAGEMENT (DSPM) PERSPECTIVE

Authors

Daniel Twum Gyamrah (Marymount University)

Abstract

Government agencies are increasingly migrating to the cloud for scalability, cost-efficiency, and innovation. However, data security and transparency concerns, especially within evolving frameworks like Trusted Internet Connections (TIC) 3.0, hinder broader adoption. This dissertation examines the role of Data Security Posture Management (DSPM) in fostering trust and transparency for government cloud adoption under TIC 3.0. The research explores current cloud security challenges faced by government agencies and analyzes DSPM capabilities in addressing them. A framework for implementing DSPM, aligned with TIC 3.0 principles, is developed to enhance trust and transparency. A mixed-methods approach combines a systematic literature review with empirical data gathered through surveys and semi-structured interviews with government agencies implementing cloud solutions. Findings contribute to understanding critical success factors for secure government cloud adoption within TIC 3.0 and offer practical guidance for policymakers and practitioners. The study reveals that strategically implemented DSPM, integrated with existing security frameworks and considering TIC 3.0 guidelines, significantly improves data visibility, control, and remediation, fostering stakeholder trust and transparency. The research emphasizes the importance of addressing organizational and cultural factors for effective DSPM implementation within TIC 3.0. Key areas where DSPM directly supports TIC 3.0 objectives, such as zero trust implementation and continuous monitoring, are also identified.

Mitigating In-Transit Vision Noise for Enhanced Vehicle Safety

Authors

Yichen Luo, Junzhou Chen, Xinyu Chen, Sidi Lu (William & Mary)

Abstract

Software-defined vehicles (SDVs) rely heavily on cameras for intelligent and safety-critical applications but face challenges from dynamic environmental noise, including weather and occlusions. Unlike static sensors, SDV cameras encounter noise patterns influenced by driving speed, a factor often overlooked in prior research. To address this, we analyze in-transit noise impacts using data from public datasets, the CARLA simulator, a robotic vehicle, and a real vehicle. Our findings suggest that maintaining a speed below 40km/h may serve as a threshold for ensuring reliable camera-based applications under noisy urban conditions. In addition, we propose TransitNet, a novel model to mitigate in-transit camera noise, particularly at high driving speeds. It outperforms baselines, improving the F-measure by 5.1%, mAP@50 by 3.6%, and increases FPS by 56.7% across all datasets. We also provide detailed observations and insights from extensive testing.

Web-based PDC Educational Video Games

Authors

Melissa Cameron, Siddhi Kasera (Virginia Tech)

Abstract

Parallel and distributed computing (PDC) use has been rising through-out most areas of industry. The demand for graduates that are skilled in PDC is naturally increasing. However, changes to the curricula for university computer science (CS) degrees have not kept pace with this new demand. Not only does the curriculum need to change to include more instruction in PDC, but it needs to include PDC concepts earlier in the curricula. Thus far the inclusion of PDC in introductory CS courses has been adopted by less than 15% of universities according to an informal review of course catalogs for US university CS degrees. The incorporation has been slow because there are real and perceived difficulties in teaching PDC concepts, particularly early in CS courses. To promote PDC inclusion in introductory courses, there is a need to produce educational materials that address these difficulties. The previously reported Parallel Islands, is a web-based educational video game tool designed for that purpose. Repetition of concepts, particularly using different methods, aids in understanding and retention of those concepts. Therefore Parallel Islands was designed to allow the use of additional mini-games within the overarching gameplay of Parallel Islands itself. One such game was developed to be able to be used in conjunction with Parallel Islands or as a stand alone web-based game, Knight and Dragon.

Examining Visual Attention in Gaze-Driven VR Learning: An Eye-Tracking Study

Authors

Yasasi Abeysinghe (Old Dominion University); Kevin Cauchi (St. John‚Äôs University); Vikas Ashok, Sampath Jayarathna (Old Dominion University)

Abstract

Virtual Reality/Augmented Reality (VR/AR) is a valuable tool for learning and training environments, offering unique characteristics such as immersion, a sense of presence, and the ability to simulate environments that are otherwise inaccessible or dangerous in real life. One promising feature of VR is gaze-driven context rendering, which dynamically presents virtual objects or information within the user's field of view (FOV) based on their gaze direction. This capability has the potential to enhance learning experiences by increasing interaction and adapting the learning environment to the user's attention. Understanding users' visual scanning behavior and focus in a gaze-driven VR environment can help improve their engagement. Eye tracking measures provide insightful cues to human visual search behavior, allowing for an understanding of how users interact with the VR space. In this study, we present an

eye-tracking user study for analyzing visual attention in a gaze-driven VR learning environment using a consumer-grade Meta Quest Pro VR headset. Eye tracking data were captured through the headset's built-in eye tracker. We then generated basic and advanced eye-tracking measures,Äîsuch as fixation duration, saccade amplitude, and the ambient/focal attention coefficient $\mathcal{K}$,Äîas indicators of visual attention within the VR setting. The generated gaze data are visualized in an advanced gaze analytics dashboard, enabling us to assess users' gaze behaviors and attention during interactive VR learning tasks. This study contributes by proposing a novel approach for integrating advanced eye-tracking technology into VR learning environments, specifically utilizing consumer-grade head-mounted displays.

A Hierarchical Deep Reinforcement Learning Chatbot for Cybergrooming Prevention

Authors

Heajun An, Qi Zhang, Arav Singh, Jin-Hee Cho (Virginia Tech)

Abstract

Cybergrooming presents a growing risk to young internet users, necessitating proactive, AI-driven intervention strategies. This work introduces an adaptive chatbot designed to enhance teenagers' resilience against cyber-grooming through dynamic, human-like interactions tailored to varying vulnerability levels. The chatbot integrates a large language model with hierarchical deep reinforcement learning to model flexible transitions across six cyber-grooming stages, incorporating both reward- and context-based transitions. The large language model, fine-tuned on the stage-tagged PJ dataset, ensures realistic conversational output aligned with actual chat dynamics. The decision-making process follows a hierarchical structure: a master policy determines the current stage based on contextual cues, while six sub-policies optimize conversational strategies tailored to the goals of each stage and the detected sentiment of the user. This approach allows the chatbot to dynamically navigate between stages, creating natural and contextually appropriate interactions. To ensure fluency, safety, and relevance before deployment in educational settings, the chatbot,Äôs performance will be evaluated through simulated interactions with another large language model. This evaluation will help identify potential risks, refine responses, and address ethical concerns before real-user implementation. Ultimately, this research provides an AI-driven, interactive learning tool for enhancing online safety education and preventing cybergrooming.

Increasing Digital Literacy Amongst Older Adults

Authors

Isabella Divietro (The Governor's School for Science and Technology)

Abstract

Many older adults struggle with low levels of digital literacy and few resources to consult. Due to this, it is common for them to lack the ability to fluently communicate online, take part in growing digital markets, and utilize appropriate safety measures while online. This research project outlines a survey that is designed to determine the biggest obstacles that they face while using technology, and the best strategies used to improve understanding. The surveys were administered to a sample population at a local senior living facility; participants read a brief description of 10 basic functions to complete on a cell phone. Referencing their personal experiences, they rated each task based on the level of difficulty, wrote a short description of why the task was rated that way, and described how it can be made easier. To analyze the trending problems/techniques through survey responses, the data from the experiment was interpreted using a chi-squared test and integrated into an Excel spreadsheet. After highlighting prominent concerns and challenges, the design goal is to create an app with an easy-to-navigate user interface that incorporates the most beneficial learning techniques and provides comprehensible descriptions of the most challenging tasks conducted on a cellular phone. Following the proposed success of the design goal, the app will be useful in continuing to advance the knowledge and understanding of older adults' struggles whilst using technology and will be an effective way to decrease the divide amongst age groups in digital literacy.

Assessing Public Perception of AI-Generated Social Media Content of the 2024 U.S. Presidential Debate

Authors

Rebecca Ansell, Prof. Leticia Bode, Dr. Sejin Paik, Prof. Lisa Singh, Autumn Toney-Wails, Dr. Tiago Ventura (Georgetown University)

Abstract

With the growing accessibility of large language models, AI-generated content is expected to become increasingly common in online political discourse, raising an important question: can social media users reliably differentiate between human- and AI-authored posts? In this study, we explore perceptions of AI-generated content discussing the 2024 U.S. presidential debate. We generate synthetic posts using OpenAI,Äôs GPT-4o with structured prompt engineering, incorporating five politically diverse personas to capture a range of ideological perspectives. Selecting real posts via hashtags (for X posts) and comments on related videos (for YouTube comments), we generate a set of 7,500 pairwise comparisons of human-authored and AI-generated posts that were labeled by 504 paid annotators on Connect. We then apply the Bradley-Terry statistical model, to rank and assign a perceived humanness score to the posts,

and linguistic and sentiment-based models, such as VADER, TweetEval, TweetBERT, and PoliBERT, to uncover textual features that influence perceptions of authorship. Our findings reveal that while users could often identify AI-generated content, their accuracy varied by platform (posts on X were more frequently correctly identified than those on YouTube), and sentiment was a discriminating feature (content with positive sentiment/optimism was more often labeled as AI-generated and content with negative sentiment/offensive language were more often labeled as human-authored). By identifying key factors shaping public perception of AI-generated content (e.g., emotional tone, linguistic markers, and variance in online platforms), this study offers valuable insights into the evolving challenges of generated content in political communication.

Strategic Questioning and Enhanced Engagement: Developing a Counselor Chatbot Using LLMs and Deep Reinforcement Learning for Deeper Self-Exploration and Therapeutic Dialogue

Authors

Qi Zhang, Heajun An, Arav Singh, Jin-Hee Cho (Virginia Tech)

Abstract

In the evolving landscape of mental health services, accessible and personalized psychological support is paramount. This paper introduces a specialized chatbot designed to function as a virtual counselor using a large language model (LLM) to bridge the gap between the demand for mental health services and the limitations of the current infrastructure. Key to our approach is the customization of the chatbot to individual users' personalities, domestic situations, and backgrounds, ensuring a tailored therapeutic interaction. The chatbot facilitates strategic questioning that enhances self-exploration and therapeutic dialogue by employing deep reinforcement learning to prompt the LLM and principles from counseling psychology. This personalization allows the chatbot to create a supportive and interactive environment that encourages clients to engage more deeply, providing a private, cost-effective, and accessible preliminary support tool that reduces the strain on human counselors and enriches the mental health infrastructure.

Who Gets In? The Role of AI in Shaping the Next Generation of Computer Scientists

Authors

Ananya Prakash, Mohammed Seyam (Virginia Tech)

Abstract

The number of graduate students has been increasing rapidly to meet industry demands, with over 200% increase in competitive fields like computer science (CS) in the past decade. Universities have adopted AI in their admissions processes for various tasks like evaluating transcripts, extracting critical information from essays, and scoring applications. With the minimal change in the composition of CS graduates over the past decade and the recent ban on affirmative action, AI-based admissions may further exacerbate bias.¬†In this work, we present the tradeoffs of leveraging AI to increase admissions efficiency with its potential risks to diversity. We study three research questions to assess the effect of the affirmative action ban on admissions, identify the independent features that contribute to bias and examine the potential intersectional attributes that could induce bias in a model.¬†We propose a framework to systematically detect bias that may be inferred by a machine learning model using methods like exploratory data analysis, clustering, subgroup discovery, and feature importance. By implementing our proposed framework on a dataset of applications to two graduate CS programs from 2014 to 2024, we investigate the research questions and demonstrate how universities may tackle the challenges of using AI for admissions.¬†Our work introduces a framework to discover inferred biases before applying machine learning models for admissions, provides evidence that demographic features can induce bias in AI-based admissions automation and highlights the importance of bias detection to foster diversity in graduate CS education.¬†

Data Collection Pipeline to Diversify AI Training Data

Authors

Victoria Wiegand, Sarah Cooney (Villanova University)

Abstract

Artificial intelligence (AI) vision-language models have improved rapidly in recent years. However, these improvements have not mitigated all biases. AI vision-language models are commonly trained on data from the internet. In turn, these models develop a perspective from predominantly Western countries such as the United States, Germany, Japan, etc. As a result, AI vision-language models perform poorly in image classification and image generation for non-Western perspectives. These models demonstrate low classification accuracy and generate images that reflect underdevelopment and impoverishment unaligned with the attitudes of the community being portrayed. As AI vision-language models will continue to expand and be used throughout the world, it is imperative to narrow the already present digital divide between Western and non-Western countries. Collecting and annotating data globally is an expensive and time-consuming process. In this project, we propose an alternative solution and examine

the feasibility of a data collection pipeline to assist with the diversification of AI vision-language models' training data. In collaboration with a university service trip, we trained participants on ethical data collection. We then asked participants to submit images of spaces or objects of the visiting community through a webserver form linked to a WhatsApp community. The images will be gathered and assembled into a public online dataset. This dataset will be made available for developers to supplement their own datasets to make their models more adaptable to underrepresented cultures. In the future, we will employ WhatsApp to allow community members to have data autonomy and submit their own images.

Driver Facial Expression Classification: A Comparative Study of Computer Vision Techniques

Authors

Yeana Lee Bond, Syed Talal Hasan, Muhammad Hamza, Mughees Ur Rehman (Virginia Polytechnic Institute and State University)

Abstract

Driving is an essential part of daily life, and technologies such as Advanced Driver Assistance Systems (ADAS) may incorporate driver state detection in the future, as road safety is influenced by various driver states. More than 80\% of the drivers in the US reported that they feel safer with ADAS. Future intelligent driver assistance systems will likely integrate ADAS with Natural User Interfaces and Affective Computing using computer vision techniques. In this project, we investigate how many images state-of-the-art machine learning models require to effectively classify three key emotions that human drivers can express while they are manually driving: happy, angry, and neutral facial expressions. Our evaluation compares EfficientNet, Vision Transformer, and CNN-based models on the Multiple Light Intensities Driver Emotion Recognition dataset. Results show that EfficientNet, with only 4M parameters, achieves 96\% accuracy when trained on more than 1,000 images, making it the most efficient model due to its faster training time. In contrast, we find that Vision Transformer, despite having 86M parameters, reaches 87\% accuracy with just 600 images after fine-tuning. Additionally, a comparison of our custom CNN-based models (9M parameters) against a baseline (3M parameters) reveals that architectural design primarily influences performance rather than the number of parameters. These findings underscore two key machine learning principles: the critical role of dataset distribution and the importance of data quality.

Advanced Gaze Measures for Analyzing Joint Visual Attention

Authors

Kumushini Thennakoon, Yasasi Abeysinghe, Bhanuka Mahanama, Rochana R. Obadage, Vikas Ashok, Sampath Jayarathna (Old Dominion University)

Abstract

Joint visual attention (JVA) is a crucial element of effective collaboration that indicates how participants synchronize their focus, cognitive engagement, and physical actions during shared tasks. This study explored JVA between dyads leveraging egocentric data combined with eye-tracking technology. While conventional measures of JVA rely on overlapping gaze points and gaze-following behavior, advanced gaze measures remain unexplored with the potential for a deeper understanding of the phenomenon. We conducted experiments using a wearable eye-tracking device with multimodal data collection, capturing gaze data and conversational exchanges. This setup enabled us to analyze how participants coordinated their visual attention while engaging in a screen-based collaborative visual search activity. We analyzed the dynamics of ambient/focal attention with coefficient K as a metric to understand the attention alignment beyond simple gaze overlap. By integrating this measure, we could discern detailed patterns in attention behavior, revealing how individuals dynamically shift their attention in response to their partner's actions and verbal cues. Our findings suggest that users who maintained similar attention behaviors (ambient/focal) over time exhibited more frequent and sustained moments of joint attention compared to those with differing attention behaviors. The study highlights the potential of advanced gaze measures for evaluating attention processes in dyadic interactions and it offers valuable insights for applications in human-computer interaction, education, and teamwork analysis. Future work will further refine the methodology and explore the integration of machine learning techniques to automatically identify and classify different patterns of ambient and focal attention during collaborative tasks.

Evaluating Children's Ability to Distinguish Between Traditional and AI-Generated Media

Authors

Sanjna Kumari, Faraz Ulhaq Shah, Sehrish Basir Nizamani (Virginia Tech)

Abstract

Evaluating Children's Ability to Distinguish Between Traditional and AI-Generated Media Background: The rapid integration of artificial intelligence (AI) in media production has led to an influx of AI-generated content in children's daily consumption. This shift raises critical questions about children's cognitive ability to discern between traditional (human-generated) and AI-generated media. Limited research exists on how these media forms impact children's perception, comprehension, and decision-making processes. Objective: This study aims to evaluate children's capacity to differentiate between traditional and AI-generated content across

various media types, including stories, images, and songs. The key objectives include assessing recognition accuracy, confidence levels, and reaction times, alongside understanding the developmental implications of media differentiation. Methodology: We have collected a curated set of real and AI-generated stories, images, and songs. This dataset will serve as the foundation for future analyses to measure children's ability to identify the content's origin. The next phase will involve designing questionnaire-based surveys to assess recognition skills, confidence, and reaction times. The collected data will enable us to develop tailored survey tools and experimental protocols. Preliminary Insights: The data collection process has highlighted variations in media characteristics that may influence children's recognition abilities. These insights will guide the development of future assessment tools and experimental designs to evaluate children's capacity to differentiate between AI-generated and traditional media. Significance: This research contributes to understanding how AI-generated content affects children's media literacy. The outcomes have potential implications for educational curricula, content creation guidelines, and regulatory policies aimed at enhancing children's critical thinking and safeguarding against media manipulation in the digital age. Keywords: AI-generated media, traditional media, media literacy, cognitive development, children, media differentiation

SCVI: Bridging Social and Cyber Dimensions for Comprehensive Vulnerability Assessment

Authors

Shutonu Mitra (Virginia Tech); Thomas Nguyen (Department of Computer Science, Virginia Tech); Qi Zhang (Virginia Tech); Hyungmin Kim (Department of Computer Science, Virginia Tech); Hossein Salemi (Department of Information Sciences and Technology, George Mason University); Chen-Wei Chang (None); Fengxiu Zhang (School of Policy and Government, George Mason University); Michin Hong (School of Social Work, Indiana University); Chang-Tien Lu (Department of Computer Science, Virginia Tech); Hemant Purohit (George Mason University); Jin-Hee Cho (Virginia Tech)

Abstract

The rise of cyber threats on social media platforms neces- sitates advanced metrics to assess and mitigate social cyber vulnerabilities. This paper presents the Social Cyber Vulnera- bility Index (SCVI), a novel framework integrating individual- level factors (e.g., awareness, behavioral traits, psychological attributes) and attack-level characteristics (e.g., frequency, consequence, sophistication) for comprehensive socio-cyber vulnerability assessment. SCVI is validated using survey data (iPoll) and textual data (Reddit scam reports), demonstrating adaptability across modalities while revealing demographic disparities and regional

vulnerabilities. Comparative analyses with the Common Vulnerability Scoring System (CVSS) and the Social Vulnerability Index (SVI) show SCVI‚Äôs superior ability to capture nuanced socio-technical risks. Monte Carlo- based weight variability analysis confirms SCVI‚Äôs robustness and highlights its utility in identifying high-risk groups. By addressing gaps in traditional metrics, SCVI offers action- able insights for policymakers and practitioners, advancing inclusive strategies to mitigate emerging threats such as AI- powered phishing and deepfake scams.

DBWorkout: A Collaborative and Gamified Web-Based Platform for SQL Learning

Authors

Snehitha Ravella, Ashutosh reddy Pochamreddy, Janice Wang, Qiuya Chen, Sehrish Nizamani, Sally Hamouda, Saad Nizamani (Virginia Tech)

Abstract

SQL education in traditional classroom settings often lacks engagement, immediate feedback, and interactive collaboration, making it challenging for students to develop a deep understanding of database concepts. To address these limitations, DBWorkout, a web-based interactive SQL learning platform, integrates real-time collaboration, automated feedback, and gamified engagement to enhance the learning experience. This research presents the development and evaluation of DBWorkout, designed to improve SQL mastery through hands-on practice and competitive challenges. DBWorkout features an SQL editor that allows students to write and execute SQL queries while receiving instant, actionable feedback. It incorporates WebSocket-based collaboration, enabling students to work together dynamically. Gamification elements, including leaderboards and achievement badges, enhance engagement and motivation. Students earn points based on query accuracy and participation in collaborative activities, fostering a competitive yet educational environment. The platform‚Äôs cloud-based infrastructure ensures scalability and real-time collaboration. For security, JWT authentication is implemented to ensure secure access and compliance with institutional policies. For instructors, DBWorkout provides an analytics dashboard offering insights into student performance and learning trends. Instructors can create and manage SQL assignments, streamlining classroom workflows and reducing the logistical burden of setting up individual database environments. By integrating real-time collaboration, gamification, and automated feedback, DBWorkout modernizes SQL education, enhancing student comprehension and retention. This research contributes to the development of technology-enhanced learning environments that equip students with industry-relevant SQL skills in an engaging and interactive manner.

Detecting Stealthy Manipulations of Control logic in Programmable Logic Controllers

Authors

Adeen Ayub (James Madison University); Irfan Ahmed (Virginia Commonwealth University)

Abstract

In industrial control systems (ICS), programmable logic controllers (PLCs) are directly connected to physical processes such as nuclear plants, power grids, and gas and oil pipelines. These controllers run control logic programs that define how physical processes are controlled and monitored. Attackers often target PLCs to disrupt these processes, making it essential to protect PLCs against cyberattacks. Existing detection systems have some limitations. Network-based detection systems primarily focus on identifying control-logic attacks by analyzing payload and header information to detect the transfer of control logic over the network. However, these systems are ineffective against attacks that do not involve sending malicious control logic. Instead, attackers may manipulate firmware objects to covertly alter control logic. Memory-based attacks can be detected by defensive and forensic approaches that monitor changes in memory. However, these approaches often fail to detect attacks that do not modify invariants but instead reuse existing code within the memory and update dynamic structures such as the stack. To address these limitations, we propose a system called Control fLow verification and Attacks Detection (CLAD). Our system performs a static analysis of the benign control logic to construct a call graph, followed by a dynamic analysis in which it logs data structures such as the stack and registers to create a dynamic call graph. Notably, CLAD is the first detection scheme capable of identifying manipulations in dynamic and static data structures. Our evaluation results show that CLAD successfully detects different types of stealthy control logic attacks, whereas existing approaches fail to do so.

Infrastructure for Tracking Information Flow from Social Media to U.S. TV News

Authors

Himarsha R. Jayanetti (Old Dominion University); Alexander C. Nwala (William & Mary); Michael L. Nelson, Michele C. Weigle (Old Dominion University)

Abstract

We examine the intersection between social media and mainstream television (TV) news to understand how TV broadcasts amplify the reach of social media content. The widely accepted way of quantifying exposure focuses solely on interactions within the social media platform,

such as engagement metrics like likes and shares. However, this approach underestimates the total impact of information or disinformation since it excludes TV viewers who consume social media content reported by local, broadcast, and cable TV news. Although we anecdotally know that information flows from social media to TV news, e.g., to provide evidence, to amplify popular memes, or to share people's experiences, the scale and nature of this information flow remain largely unexplored. These gaps motivate our research questions: (1) To what extent and in what ways is social media content featured in local and national TV news? and (2) How can we quantify the amplification effect of TV news coverage on social media messages? To answer these questions, we analyze TV news data from the Internet Archive, examining closed captions, on-screen text, and social media logo detection to identify when and how social media is referenced. We are developing machine learning algorithms to detect these instances, categorize their topics and virality, and quantify the amplification effect, providing the first large-scale analysis of how TV news amplifies social media messages.

Code-Decomp: Automated Prompt Decomposition for Code Generation

Authors

Ebtesam Al Haque, Anuridhi Gupta, Brittany Johnson (George Mason University)

Abstract

As AI-enabled programming assistants become increasingly integrated into software development workflows, developers face new bottlenecks in effectively utilizing these tools. Our prior work shows that while these assistants help developers solve problems faster and work with unfamiliar technologies, they lead to overreliance on AI-generated code without a deep understanding of its underlying logic. This tendency to treat AI outputs as a black box can potentially hinder code maintainability. To address this challenge, we introduce Code-Decomp, a novel prompt decomposition technique that automatically breaks down complex programming prompts into structured sub-problems, guiding AI assistants to generate more interpretable and incremental solutions. By structuring prompts in a way that mirrors human problem-solving approaches, Code-Decomp encourages a more systematic approach to engagement with AI-generated code, which could lead to better comprehension and reduce the risks associated with overreliance. Our empirical evaluation using the EvoEval benchmark shows that Code-Decomp achieves improvement in code generation accuracy compared to traditional zero-shot and chain-of-thought prompting approaches across multiple state-of-the-art language models. This improvement in automated prompt decomposition could transform how developers learn and integrate new technologies, enabling them to use AI programming assistants more effectively while maintaining deeper understanding of the generated solutions.

Accelerated Community Detection using Directed Graphs on GPUs

Authors

Beenish Gul (University of Virginia)

Abstract

Community detection methods are widely used across various fields, including gene expression analysis and social network studies, to efficiently identify natural divisions or functional modules within biological and social networks. These methods play a crucial role in identifying functional modules in gene expression networks, social connections in networks, and neural circuits in the brain. However, their computational demands have grown significantly due to the increasing dataset sizes in these fields. The Leiden algorithm, known for its inherent parallelism, remains underutilized on modern multi-core CPUs, which typically have only 32-64 cores. GPUs, with their high memory bandwidth and extensive parallel processing units, better map to Leiden‚Äôs parallelism and offer significant performance gains. To the best of our knowledge, CuGraph is the only existing GPU implementation of Leiden, using a blend of Python and C, but it only supports undirected graphs, discarding valuable edge directionality information. Moreover, Python-based GPU implementations are generally slower than those in C/C++, which better optimize performance. This work presents a lightweight CUDA C++-based GPU implementation of the Leiden algorithm, the first to support directed graphs, which typically require nearly double the computational time and memory compared to undirected graphs. We also demonstrate the scalability and effectiveness of our approach on large biological and social network datasets.

Using virtual reality to simulate wilderness search and rescue 'clue-finding' tasks

Authors

Paean Luby (University of Richmond); Ashley Mathis, Bryan White, Rebecca Penn (New Mexico State University); Arryn Robbins (University of Richmond); Michael Hout (New Mexico State University)

Abstract

To complement two-dimensional laboratory search tasks, some researchers have begun to explore three-dimensional environments that simulate realistic scenarios. Our prior work simulated an open-terrain outdoor search and rescue task. Unfortunately, this type of task requires a large investment of equipment, personnel, space, training, and coordination. Such tasks are also limited by the terrain available to researchers. The current investigation sought to

establish virtual reality as a feasible alternative, allowing for tighter experimental control/manipulation, three-dimensional interactive environments, and varied real-world target categories that better capture wilderness search and rescue scenarios. We found that targets with higher visibility (reflected in a higher mean squared error) captured from the AerialFar vantage point were less likely to result in target misses ($p < .001$), reinforcing its validity as a visibility metric in virtual reality environments. Higher target distance from the player also consistently increased both timeout likelihood ($p < .001$) and response times ($p = .021$). These results mirror those from previous visual search literature, where increased target distance and lower visibility are associated with longer response times and higher miss rates, supporting the ecological validity of the virtual reality task.

An Investigation into Maintenance Support for Neural Networks

Authors

Fatema Tuz Zohra, Brittany Johnson (George Mason University)

Abstract

As the potential for neural networks to augment our daily lives grows, ensuring their quality through effective testing, debugging, and maintenance is essential. This is especially the case as we recognize the prospects for negative impacts from these technologies. Traditional software engineering techniques like testing and debugging have been effective for maintaining software quality; however, they reveal significant gaps in research and practice for maintaining neural networks. Specifically, there is limited understanding of how practitioners currently address challenges related to understanding and mitigating undesirable behaviors in neural networks. In our research, we explore the current state of research and practice in maintaining neural networks, focusing on insights from practitioners through a mixed-method approach involving surveys and interviews. Our initial findings suggest that existing tools mostly focus on building and training models. While these tools can be helpful, they often fall short in supporting practitioners to understand and address the underlying causes of unexpected model behavior. By evaluating current procedures and identifying limitations of conventional methodologies, our study provides a developer-centric perspective on where current practices fall short and highlights opportunities for improving maintenance support in neural networks. The goal of our work is to bridge the gap between available maintenance tools and techniques and their use in practice.

Human AI Collaboration for Scalable Rationale Generation to Support Explainable Stance Detection Systems

Authors

Anuridhi Gupta, Hemant Purohit (George Mason University)

Abstract

Recent years have demonstrated the transformative capability of Natural Language Processing (NLP) systems to process social media information and gather public consensus across debatable and significant topics. The application of this process has value in many areas, ranging from policy-making and educational campaigns of government agencies to marketing strategies for businesses based on stance from public opinions. Over the past few years, numerous machine learning (ML) models have been curated to achieve this; however, with limited exploration on designing explainable ML systems. Given the sensitive nature of topics (gun safety), it is critical to ensure that the model employs sound reasoning when interpreting these social media messages. To enhance explainability while maintaining the model's performance, explainable ML models typically require substantial human-annotated rationales. Our work proposes a novel human-AI collaboration approach that generates rationale annotations at scale and improves performance on the downstream task of explainable stance detection. We employ a combination of human and algorithmically generated rationale annotations to fine-tune a BERT-based stance detection model using attention weights, attaining inherent explanations. Specifically, we utilize multiple large language models (LLMs) to generate algorithmic, human-like rationales for training instances. Further, we propose a novel framework that integrates algorithmic and human rationales to choose their combination for training cases. Our experimental findings demonstrate the feasibility of leveraging algorithmic rationales to offset the requirement for costly, large-scale human rationales. The application of this research presents a promising direction to support the design of explainable ML systems for tasks beyond stance detection.

Mitigating Cyber Threats in V2V and V2I Networks: A Security-Centric Approach

Authors

Susan Zehra, Syed R Rizvi (Old Dominion University)

Abstract

A novel framework is introduced to address security and privacy concerns in Vehicle-to-Vehicle (V2V) and Vehicle-to-Infrastructure (V2I) communication. Operating under a semi-honest threat model, the framework utilizes public-key cryptography for secure communication and dynamic key management, effectively countering threats from compromised Roadside Units (RSUs) and malicious vehicles while ensuring user privacy. Synchronized communication among RSUs ensures consistent security credentials across the network, enforcing uniform security

standards. A partial trust model, supported by a blockchain-based fallback mechanism, strengthens the reliability and integrity of RSUs. An adaptive deregistration strategy efficiently revokes vehicle credentials, minimizing overhead in high-traffic environments, while dynamic key management helps maintain vehicle anonymity. To protect privacy, the framework uses pseudonymous public keys and encrypts both message content and metadata. The system‚Äôs performance is evaluated using metrics such as Message Delivery Ratio (MDR), Privacy Preservation Rate (PPR), and Security Breach Rate (SBR). Results show a consistently high MDR, indicating reliable message delivery even in congested scenarios, along with a high PPR, ensuring robust privacy protection, and a low SBR, demonstrating strong resistance to security breaches. This comprehensive framework offers a secure and privacy-preserving solution for V2V and V2I networks, mitigating cyber threats effectively.

Analysis of Subtelomere and Telomere Regions of Cancer Genomes on the Cloud

Authors

Eleni Adam, Desh Ranjan, Harold Riethman (Old Dominion University)

Abstract

Cancer continues to affect millions of people worldwide, nearly 40% at some point in their lives, with around one-third of cases having potentially a lethal outcome. Many cancers impact specific populations more severely than others and a systematic analysis of genetic features is needed to understand these cancer health disparities. Analysis of patterns of changed genes and rearranged chromosomes characteristic of cancer types will lead to more effective therapies. The telomeres, which form the ends of human chromosomes are a vitally important part of the cell, required for proper replication and stability. Their dysfunction is a major early event leading to cancer. For that reason, we focus our research in the telomeric and subtelomeric areas of the human genome. To study genome maintenance in cancer, we use the TCGA large genomic cancer datasets. The subtelomeric analysis of cancer genomes consists of three main parts. In the first part, we use computational methods to extract the telomeric and subtelomeric information out of the large genomic datasets. The computational pipeline is run on the AWS HealthOmics infrastructure. Thereafter, based on the extracted data, we define the telomere and subtelomere-associated features. Given these newly defined features, machine learning methods will be used to correlate them to clinical data. We have completed the development of our subtelomeric/telomeric computational pipeline and successfully applied it to the metastatic prostate cancer dataset of 101 normal/tumor paired individual genomes. Currently, we are in progress of optimizing it and refining it for the next step of feature identification.

Usability Heuristics and Large Language Models: Enhancing University Website Evaluations

Authors

Marissa Hirakawa, Malavika Saritha, Mayank Goel, Siddhesh Bapat, Neren Dawar, Rahul Bose Reddy, Sehrish Basir Nizamani (Virginia Tech)

Abstract

Usability evaluations typically rely on expert reviewers, who can be difficult to find and expensive to employ. In this study, we explore how large language models (LLMs) assisted in the usability evaluation of 10 university websites using Nielsen‚Äôs 10 usability heuristics: visibility of system status, match between system and real world, user control and freedom, consistency and standards, error prevention, recognition rather than recall, flexibility and efficiency of use, aesthetic and minimalist design, help users recognize and recover from errors, and help and documentation. By leveraging LLM-driven methodologies, this study evaluates how effectively LLMs can identify usability issues in university websites. Each website was analyzed using two different LLMs, and only the usability issues identified by both models were selected for further validation. Findings highlight that LLMs can assist human evaluators by identifying additional usability issues that might be overlooked in manual assessments. However, the results also revealed that LLMs occasionally hallucinate usability issues, making human verification essential. The validation process showed that while LLM-generated assessments often aligned with expert insights, human reviewers were necessary to confirm the accuracy and relevance of the identified issues. This underscores the role of LLMs as a supportive tool rather than a standalone replacement for expert evaluations. Future work includes comparing LLM-generated usability evaluations with expert human evaluations to assess their relative effectiveness and reliability. Additionally, we aim to refine LLM evaluation processes, integrate multimodal data,Äîsuch as user action logs and screenshots,Äîto enhance their usability assessment capabilities further, and develop systematic methods for filtering hallucinated issues.

Agentic AI for the Rescue: Factual Summarisation of Crisis-Related Documents

Authors

Hajra Klair (Virginia Tech)

Abstract

This research proposes a novel entity-centric abstractive summarization framework for crisis-related documents that integrates targeted question-answer formulation through LLMs to generate factually correct and comprehensive document summaries. The framework was evaluated using the dataset from CRISISFacts Track (TREC 2022-2023, McCreadie et al.), which comprises news articles and microblog content from 18 distinct real-world crisis events. System performance is evaluated using ROUGE and BERT metrics against three reference summaries: NIST incident documentation, Wikipedia entries, and NIST assessor-curated fact pools. The system employs a two-phase architecture that first performs entity-based document retrieval, followed by a constrained LLM summarization system designed to minimize hallucination while maximizing information coverage. The methodology specifically addresses the challenge of generating summaries that align with emergency response officials' information requirements through structured query-driven content formulation. In the first document retrieval phase, an entity-centric ranking algorithm prioritizes documents with higher entity mention frequencies, showing improved retrieval effectiveness. In the second abstractive summarization phase, the entities are used to generate targeted questions aligned with crisis-specific information requirements (encoded as queries in the dataset). The questions generated are then employed to guide another LLM agent in generating factually consistent responses, which are then restructured into cohesive summaries for each event. This dual-agent approach of question-answer generation maintains temporal and causal relationships while preserving source fidelity, achieving superior performance in both automated metrics and human evaluation of factual accuracy. The framework demonstrates domain adaptability beyond crisis informatics, with preliminary experiments indicating enhanced performance in general document summarization tasks compared to baseline approaches.

Leveraging Smartwatch Sensors For Detecting Off-Task Behaviors Of Neurodivergent Individuals

Authors

Anika Binte Islam, Vivian Genaro Motti (George Mason University)

Abstract

Neurodivergent individuals experience a high unemployment rate due to a noticeable gap between accommodation needs and available support. Despite their exceptional ability and potential, they often experience challenges with organizing, communicating, and sustaining attention, which often prevents their workplace success. Even most traditional accommodations are often costly, time-consuming, and inconsistent, hindering neurodivergent individuals from performing up to their potential. Moreover, neurodivergent employees often need positive reinforcement, reminders, and aid with concentration. In fact, data-driven assistive wearable

technologies have shown promise in supporting neurodivergent young adults in the workplace. Along this direction, our study explores how smartwatch sensor data can detect off-task behavior in neurodivergent individuals and provide them with personalized interventions to improve workplace success. The data were collected from 25 neurodivergent young adults during a 30-minute manual task in a controlled lab setting via smartwatch. The task designed for this study resembled activities involved in stockers and order fillers, making it applicable to similar job roles and responsibilities. Our next step is to apply machine learning models to extract insights that contribute towards the identification of off-task behavior and thereby design tailored real-time interventions to improve workplace productivity.

Posters

Home Tech Care: Symbiotic Tech Caregiving

Authors

Natalie Andrus (Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA); Onyinye Mbanefo (Department of Human Development, Virginia Tech, Blacksburg, VA 24061, USA); Vee Pettit, Wei Lu Wang, D. Scott McCrickard (Department of Computer Science, Virginia Tech, Blacksburg, VA 24061, USA)

Abstract

By 2050, 2 billion older adults will exist worldwide. In 25 years, we'll face the largest older adult demographic in history. We must prepare the youth of today to support future older adults. This won,Äôt be easy and will require cultural resilience to confront society,Äôs mutual intergenerational bias. Yes, mutual. Ageism affects everyone, with biases flowing both from old to young and young to old. These negative narratives create a hostile environment for aging and darken the days of current older adults. We have created a student-led Home Tech Care initiative at Warm Hearth Village [W.H.V] , Blacksburg, to understand how to reduce mutual intergenerational bias. We aim to answer two research questions: RQ1: How is intergenerational bias impacted through Home Tech Care sessions? RQ2: How can the Home Tech Care program be improved? We will share our preliminary findings in our CAPWIC research short.

The Web in Schools

Authors

Abby Kozlowski (None)

Abstract

Today, the web is used in classrooms all around the world for learning and creative purposes. Various students utilize the web to take advantage of its learning materials and apply them to their learning process, catering to their learning styles. However, the Internet poses many dangers and distractions for elementary and high school students that can impact their learning process and mental health. Often, students are uneducated on different Internet safety topics, which leads to young children being at risk when it comes to keeping personal information private on dangerous websites. Internet education and web dangers are being implemented into elementary school curriculums so that children learn this information at a young age so that they can travel with this information into their adulthood. Older students find themselves becoming distracted during long lectures or challenging assignments, where they rely on entertaining websites, social media, or webpages used for cheating to speed up assignments and the school day. While the Internet is being used in schools, all students can become exposed to distractions, cyberbullying, cheating websites, privacy risks, extended screentime, and many other harms, if the Internet is not being filtered or specific web expectations are not set in school systems. Ultimately, elementary and high school systems push for web use within the classroom to allow students to fully benefit from different educational resources, although it comes with many threats and dangers.

US Computer Science Faculty Pipeline Dashboard

Authors

Lesley Frew, Michele C. Weigle (Old Dominion University)

Abstract

The US Computer Science Faculty Pipeline Dashboard's purpose is to evaluate the operations (graduates, employees) of different CS graduate programs with respect to gender. The dashboard helps users compare the graduate student gender diversity between two different universities, examine historical trends in female PhD CS programs, and examine trends in regional CS hiring. The three datasets used to create this dashboard are IPEDS 1984-2011 for CS PhDs awarded per year per institution, by gender, Clauset et al.,Äôs ‚ÄúSystematic inequality and hierarchy in faculty hiring networks‚Äù for 2012 CS tenure-track faculty, by gender, job title, alma mater name and region, employing institution name and region, and a SPARQL query on Wikidata for US Universities, by name and IPEDS code, used to align the previous two data sets. The D3 dashboard is composed of a scatterplot, 3 bar charts, and a line graph, along with six sets of selection widgets. The scatterplot shows universities correlating percent of female CS tenure-track faculty alma mater versus percent of female CS tenure-track faculty employer. The first bar chart compares different selected schools‚Äô CS PhD enrollment,

alma mater, and employer institutions. The second bar chart compares regional hiring trends and can be further zoomed for a specific institution. The third bar chart shows the number of years with no female graduates, by either count or streak. The line graph shows the count of CS PhD graduates by gender between 1984 and 2011. These interactive charts help the user compare programs,Äô gender diversity and examine employment trends.

Enhancement of Deep Learning for Segmentation of Protein Secondary Structures from Cryo-EM

Authors

Thu Nguyen, Yongcheng Mu, Jiangwen Sun, Jing He (Old Dominion University)

Abstract

Advances in cryo-electron microscopy has made it become the key technique for protein structure determination. DeepSSETracer, a deep learning framework for the segmentation of protein secondary structure was integrated to the molecular visualization program - ChimeraX to provide a convenient and user-friendly interface for the cryo-EM segmentation tasks. However, due to the nature of deep learning memory requirement, the process can be both memory intensive and time consuming for the users, which also limits the size of input that can be processed on their machines. To address this issue, we performed volume processing by splitting the input cryo-EM maps and weighted merging the segmentation output. Our results showed that the partitioning of volumetric data accelerates the execution time and can reach comparable result to the original framework settings, which enables DeepSSETracer to handle large input data more efficiently.

AI-Powered Enhancement of Student Learning in Large Programming Courses

Authors

Tianyu Zhan, Barton Hou, Sehrish Nizamani (Virginia Tech)

Abstract

Enhancing student learning experiences through AI-powered tools has become an essential area of research, particularly in addressing challenges in coding-related courses. A major issue in large programming classes is the overwhelming volume of student inquiries, making it difficult for instructors to provide timely and personalized responses. In our analysis of last year,Äôs Piazza data for level 3000 Data Structures and Algorithms Course, we observed that students posted over 300 questions per project, with four projects throughout the semester. Many of

these questions were repetitive or focused on similar topics, suggesting an opportunity to streamline responses and improve student support. To address this, we conducted a systematic categorization of past student inquiries, identifying common themes and frequently asked topics. Building on these insights, we developed an AI-driven extension for Piazza aimed at optimizing student support in two key directions. First, a retrieval-based system was designed to surface similar past questions and high-quality responses. By generating keywords from historical student queries and employing a key-value storage approach, the system efficiently retrieves relevant responses, reducing redundancy in common inquiries. Second, a guided AI response mechanism was implemented using OpenAI models to provide step-by-step explanations rather than direct answers, fostering deeper understanding and critical thinking. Although the system has been developed, it has not yet been deployed in a real classroom setting. Future work will focus on evaluating its effectiveness in a natural learning environment, assessing its impact on student engagement, learning outcomes, and instructor workload. Through this AI-powered enhancement, we aim to create a more efficient and supportive learning experience in large-scale programming courses.

Analyzing Ground-lightning Dataset Using the Density Based Spatial Clustering of Applications with Noise (DBSCAN)

Authors

James Agresto, Zhuojun Duan, Mace Bentley, Tobias Gerken, Dudley Bonsal (James Madison University)

Abstract

With over seven million warm-season, cloud-to-ground lightning flashes in Washington, DC to analyze, producing a clustering algorithm capable of resolving individual thunderstorm events completes a major step in the analysis and visualization of lightning data. Then, by examining the spatiality of urban thunderstorm events and its relationship with aerosols and thermodynamics, a multi-variable investigation of thunderstorm environments can be conducted in the geographic region. We utilized the Density Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm to cluster lightning flashes. This choice was made because DBSCAN does not require a priori specification of the number of clusters and can identify clusters of arbitrary size and shape, aligning well with the natural characteristics of thunderstorm events. We will conduct experiments using lightning flashes collected by the National Lightning Detection Network (NLDN) to demonstrate whether our clustering approach effectively identifies, and groups lightning events.

P2VFL: Performance and Privacy-enhanced Incentive Mechanism for Vertical Federated Learning

Authors

Sindhuja Madabushi (Virginia Polytechnic Institute and State University); Haider Ali, Ahmad Faraz Khan, Jin-Hee Cho (Virginia Tech)

Abstract

Federated Learning (FL) is a distributed learning framework that allows multiple participants to collaboratively train a machine learning model without sharing their private data. While significant research has focused on incentive mechanisms in horizontal FL (HFL), vertical FL (VFL) has received less attention. Clients may hesitate to participate in the federation due to concerns over privacy loss and the high costs of computing and resources. We introduce a novel Performance and Privacy-enhanced incentive mechanism for VFL, P2VFL. This mechanism is designed to address these concerns, motivating participants to join the federation while enhancing the global model's performance and providing robust privacy guarantees for clients. Clients are rewarded based on the quality of their data and features, ensuring those who contribute most to the overall model accuracy receive greater incentives. To preserve privacy, the server employs differential privacy to protect against reconstruction attacks from other clients. Through extensive experiments, we demonstrate that P2VFL effectively defends against data poisoning and inference attacks with minimal impact on prediction accuracy.

Roots for STEM Organization

Authors

Gaelan Venturi, Lana Quick, Sofia Leyva, Haley Mroczkowski (None)

Abstract

Roots for STEM (RFS) is a student-led, female oriented organization located at the Governor‚Äôs School for Science and Technology (GSST). RFS aims to empower young children from the Hampton Roads community (seven school divisions) and the Natasha House (a shelter for women and their children) to pursue their interests in Computational Science and other STEM fields. At least once a month, the RFS leadership organizes and runs free, educational STEM camps in which first to ninth grade students are invited to learn in a hands-on environment surrounded by like-minded peers on the GSST campus. RFS offers children the chance to explore different aspects of STEM by spreading awareness of STEM-related opportunities in the community, such as computer programming in Python, building and programming robots, and engaging in other engineering and science focused events. In order to

accommodate the hundreds of students that participate in these STEM camps, juniors and seniors from the Governor‚Äôs School volunteer after school. Whether answering a question, connecting the correct part of a robot, or even walking them through a difficult piece of code, these volunteers help the children through each new challenge. By presenting STEM educational opportunities in a fun, interactive format, RFS creates a more invested and interconnected community that can continue to inspire the new generation of computer scientists, engineers, and chemists.

Exploring QUIC Load Balancing Computation Offload to the Kerne

Authors

Donia Ghazy, Quanzhi Fu, Dan Williams (Virginia Tech)

Abstract

QUIC, designed as a successor to TCP + TLS for HTTPS, is a connection-oriented encrypted transport protocol that offers enhanced performance and privacy features. However, its ro- bust security design introduces challenges for load balancing QUIC traffic. While QUIC-LB protocol has been proposed to address these challenges, it remains at the Proof-Of-Concept stage as an Internet Draft, warranting further performance analysis. Modern high-performance networking commonly employs kernel bypass techniques or kernel extensions such as DPDK and XDP. In this work, we present SHRIMP, a QUIC load balancer that leverages XDP to offload Connec- tion ID (CID) extraction to the kernel. Our experimental eval- uation demonstrates significant performance improvements, with kernel offloading reducing CID extraction latency by 70%. These results highlight the effectiveness of kernel-based optimization for QUIC load balancing and provide valuable insights for future in-kernel QUIC-LB implementations and performance enhancement strategies.

Characterizing GPU Memory Errors: Insights from a Cross-supercomputer Study

Authors

Zhu Zhu (George Mason University)

Abstract

My name is Zhu Zhu. I am a second-year Ph.D. student at George Mason University. My major is computer science and focused on system reliability especially error characterization and CPU fault tolerance. In this poster, I present research on error characteristics. Understanding memory error patterns in modern graphic processing units (GPUs) is vital to research

productivity and supercomputer utilization, especially given the ongoing trend of large foundation model pretraining that can take thousands of GPUs for months. Here we investigate the reliability behavior of GPU memory errors by analyzing the error logs of 20 million Ampere GPU hours from the Delta, Polaris, and Perlmutter supercomputers. We quantify GPU memory error rate, Mean-Time-Between-Errors (MTBE) in clusters, and comparison of observations among different clusters. Our observations and insights can significantly benefit fault-tolerance software/hardware design and machine operation to accommodate the increasing need for HPC for modern supercomputing applications, such as large foundation models, more efficiently.

Investigating the Impact of AI-Assisted Tools on Practitioner Well-Being

Authors

Fairuz Nawer Meem, Brittany Johnson (George Mason University)

Abstract

The increasing adoption of AI-assisted tools, such as ChatGPT, in software development presents both opportunities and challenges. While these tools enhance productivity and streamline tasks, they also introduce new job demands, such as cognitive overload, and impact practitioners,Äô well-being. Well-being in this context includes mental, emotional, and physical health factors, such as job satisfaction, stress, burnout, and engagement. Our research proposes a study grounded in the Job Demands-Resources (JD-R) model to explore how AI tools influence practitioners,Äô well-being. A comprehensive survey is designed to investigate key factors like job demands, organizational and social resources, and their interplay. As a preparatory study, it anticipates significant contributions, offering theoretical insights into the dual-edged nature of AI tools and practical recommendations for ethical and effective integration. By emphasizing practitioners' well-being, this research aims to ensure sustainable and equitable adoption of AI-assisted tools in software development.

Investigating the Effectiveness of Using K-Nearest-Neighbors and Random Forest in Hybrid Recommendation Systems to Improve Personalized User Experience

Authors

Alice Zhang, Natale Gray, Nada Basit (University of Virginia)

Abstract

Machine learning algorithms are typically at the core of recommendation systems that are used to increase user experience and satisfaction. There are various algorithms that can be used for recommendation systems, but it is important to discover which algorithms perform the best at giving meaningful and effective recommendations. The three main types of recommendation systems used are: Collaborative Filtering, Content-Based Filtering, and Hybrid Systems. Our research mainly focuses on Hybrid Systems, which combines Collaborative Filtering with Content-Based Filtering to utilize similar users and similar items to produce customer recommendations. The two Hybrid System machine learning algorithms we used for our research are K-Nearest-Neighbors (K-NN) and Random Forest (RF). By comparing and analyzing the results of both of these algorithms we aim to make recommendation systems more personalized and effective for users. Finally, we contrast our findings to existing work in literature to assess the applicability of using K-NN and RF in this hybrid recommendation system.

Voices in Code

Authors

Madison Van Buren, Jessica Zeitz (University of Mary Washington)

Abstract

Despite increasing efforts to improve diversity in computer science, underrepresented groups continue to face barriers to inclusion, confidence, and career persistence. This study builds on previous research to explore how undergraduate computer science students experience belonging, confidence, and representation over the last 6 years at the University of Mary Washington. Through a new survey incorporating updated and more inclusive questions, we aim to understand students' perceptions of their classroom experiences, mentorship opportunities, group dynamics, and how those perceptions may have changed over the last 6 years. Key focus areas include the role of gender identity, racial and ethnic representation, and the intersectionality of student experiences in shaping their sense of belonging and intent to pursue a career in technology. By analyzing survey responses through statistical and qualitative methods, we seek to identify trends in students' comfort levels, feelings of being seen and valued, and confidence in their academic and professional aspirations at the University of Mary Washington. Our findings will inform recommendations for fostering a more inclusive and supportive learning environment for our computer science program and hopefully others.

Redirects Unraveled: From Lost Links to Rickrolls

Authors

Kritika Garg (Old Dominion University); Sawood Alam (Internet Archive); Michele C. Weigle, Michael L. Nelson (Old Dominion University)

Abstract

URI redirections are essential for managing website structure, improving search engine rankings, and enhancing security. However, they also introduce challenges that affect user experience, web performance, and long-term content accessibility. This study analyzes 11 million unique redirecting URIs, following each redirection path for up to 10 steps to uncover common patterns and their implications. We found that while half of all redirections successfully led to their intended destination, the other half resulted in errors, including a small fraction (0.06%) that exceeded the 10-hop limit, causing inefficiencies. Standard redirections, such as HTTP-to-HTTPS upgrades, generally followed best practices, but many involved domain or path changes, reflecting website migrations, rebranding efforts, and security risks. A surprising discovery was the presence of "sink" URIs,Äîendpoints where multiple redirects converge,Äîused for traffic consolidation by major websites or even playful misdirections like "Rickrolling." Additionally, we identified 62,000 custom 404 error pages, many of which were "soft 404s," where missing content was incorrectly treated as valid, leading to wasted resources. These findings highlight the significant role of URI redirections in shaping the web and reveal critical challenges such as outdated links, server instability, and improper error handling. By examining large-scale redirection data, this research provides valuable insights to web developers, digital archivists, and researchers, helping them improve website efficiency, optimize resources, and ensure the long-term accessibility of online content.

Computational Analysis of Urbanization's Impact on Water Quality Using Earth Observation and GIS

Authors

Zahra Rizvi (College of William & Mary); Sophia Rizvi (Grafton High School)

Abstract

Urbanization's impact on water quality in the Chesapeake Bay, particularly in the Hampton Roads region, is analyzed through a computational framework that integrates Earth Observation (EO) data and statistical modeling. Landsat 8 imagery and Copernicus Global Land Cover Layers quantify urban expansion, while a well-established algorithm retrieves Total Suspended Material (TSM) from EO data. Multiple linear regression modeling examines the relationship between TSM and key urbanization indicators,Äîpopulation density and percentage of built-up land. This computational approach enables efficient data integration, processing, and analysis, revealing a significant correlation between urbanization and TSM levels. While causation is not established, the study highlights the power of EO-driven computational methods in assessing

environmental impacts, offering a scalable and cost-effective methodology for future research.

Crowdsourced Litter Mapping: A Smart App for Community Cleanup

Authors

Zahra Rizvi (College of William & Mary); Sophia Rizvi (Grafton High School)

Abstract

Litter pollution poses a significant environmental challenge, affecting the aesthetics and health of communities. To address this issue, a mobile application has been developed that leverages crowdsourcing, machine learning, and geospatial technology to empower individuals in tracking and reducing litter. Users can photograph litter, automatically geo-tag images, and receive AI-generated recommendations for categorization. A Smart Litter Grabber, integrated with Bluetooth functionality, enhances the user experience by automating data collection during cleanup efforts. The app also provides real-time analytics to optimize waste management strategies. This technology-driven approach demonstrates how computation and community engagement can drive impactful environmental solutions.

A Solution for Seizures: Developing Machine-Learning Algorithms For Use In Real-Time Seizure Detection Applications

Authors

Riva Jain, Ameen Harandi, Malaysia Schaffer, Ishita Punna, Sehrish Basir Nizamani (Virginia Tech)

Abstract

Approximately 1 in 26 people develop epilepsy throughout their lifetime, accounting for 0.5% of worldwide disease, with an average cost-per-patient of $30,000 a year on epileptic healthcare. Beyond the financial cost, epilepsy takes a physical and emotional toll on those affected, creating the need for improved patient-centered care. Due to the lack of accessible, patient-operated real-time seizure prediction, this research focuses on creating a prediction platform to mitigate the impact of seizures and improve the quality of life for epileptic patients. The aim is to develop a mobile application connecting real-time data from a take-home electroencephalogram (EEG) device to a continuously running machine learning model. The model detects the preictal state that occurs before an unprovoked seizure, allowing users to be alerted of an upcoming seizure. To develop and optimize this model, we explore linear regression, long short-term memory networks, and other solutions. The CHB-MIT Scalp EEG

Database, alongside additional publicly available datasets regarding epileptic seizures, are utilized. Developing this application and machine-learning model could have a far-reaching impact on epileptic symptom management. Knowledge of when unprovoked seizures are about to occur would give patients time to take precautionary measures such as finding a safe position and environment. It would also permit those dealing with epilepsy to perform activities they would not have been able to otherwise, such as operating certain types of machinery, and accelerate an overall improvement in epileptic quality of life.

Praxly: An Online IDE for the Praxis CS Test Pseudocode

Authors

Ellona Macmillan, Chris Mayfield (James Madison University)

Abstract

CodeVA offers a 15-week professional development course for Virginia K-12 teachers to prepare for the Praxis CS test. Teachers who pass this standardized test fulfill licensure requirements to teach advanced CS courses in public schools. The test uses a pseudocode language that is quite different from traditional programming languages. Previously, teachers have not had an easy way to trace the execution of pseudocode programs when learning and practicing. To help, we have created an online development environment, named Praxly, that visualizes and runs pseudocode programs, providing a space in which teachers can practice. This talk highlights Praxly's features and implementation details, outlines early research results, and discusses next steps for the project. Praxly supports bidirectional synchronization between both text-based and block-based editors, a step debugger, and dozens of built-in example programs. We have recently embedded Praxly into Canvas as the primary means for teaching pseudocode in CodeVA‚Äôs course. Initial data collection from the current cohort of teachers indicates greater engagement and a deeper understanding of pseudocode compared to previous cohorts. Praxly is open source and freely available for anyone to use at https://praxly.cs.jmu.edu.

Campus Health Tracker: A Location-Based Disease Monitoring App for College Students and Highly Populous Areas

Authors

Samhita Gupta, Man Patel, Owen Stuckman (None)

Abstract

Disease spread is a constant battle, and is something that will always occur as humans interact. The spread is heightened by denser populations, and causes small outbreaks to greatly expand if precautions are not taken. In particular college students fit into this category, as they live in densely populated areas, and as part of being a student interacts with many, increasing disease spread. Around 34 percent of students get diagnosed with a cold/virus or other respiratory illness every year, which detracts from the time they can spend in class or activities. Including other diseases, the numbers greatly increase. Though for most students the few days of class they miss can be inconsequential, students or individuals with immunodeficiencies and other impairments can be greatly affected. The aim of this project is to develop an application which notifies students or those in densely populated areas of diseases that are beginning to spread, and to be an assistant to assist in assisting one,Äôs own health situation, and inform the users on possible diseases they may have contracted and precautionary steps to take depending on disease spread in the surrounding area. Increasing knowledge of the data, and providing information to better inform users on the specific diseases will help users understand their symptoms and aid as a preventative tool to prevent contracting in the first place. Additionally, the data collected in the app can provide a platform to initiate alerts and get a real time assessment of disease spread from in app reporting.

Dynamic Quantization: An Era of Reduction

Authors

Farhana Amin (Virginia Tech)

Abstract

With the advancement of time, the size of machine learning models also has enhanced a lot as the parameters used with it. As a result, the memory footprint and latency have also been increased for these models. Quantization can play a major role in minimizing the memory footprint to some extent. But the presence of outliers in different weights and activation layers makes the static quantization somewhat ineffective leading to quantization error. Hence adapting dynamic quantization scheme may lead to minimizing quantization error. In our work we propose a dynamic quantization scheme where instead of applying static quantization scheme, we propose a nonlinear quantization scheme which is proportional to memory footprint also. Depending on the availability of memory, we opt for multiple quantized versions of the models which can be made available based on the available memory of mobile devices.

Flash Talks

FoodieSafety

Authors

Temitope Emokpae, Pranav Ramkumar (Virginia Tech)

Abstract

Food recalls in the U.S. reached a 10-year high in 2022 and continued to rise in 2023, with an 8% increase in recalls and record-breaking cases from the USDA. High-profile incidents, such as lead-tainted applesauce and Salmonella-contaminated cantaloupes, highlight the urgent need for a more effective food recall notification system. Government agencies, grocery stores, and manufacturers struggle to communicate recalls efficiently, and consumers often fail to act on alerts. To address this, we propose a web application aggregating food safety and recall data, empowering grocery shoppers with real-time alerts and expiration tracking. The project will span three semesters, each with a minimal viable product (MVP). Semester 1: Pull and display food recall data from the OpenFDA API, allowing users to subscribe to weekly recall newsletters by zip code. Stretch goal: Integrate USDA and CDC data. Semester 2: Implement barcode scanning and image upload to track grocery purchases in a backend database. Stretch goal: Personalized recall alerts. Semester 3: Notify users of upcoming food expirations and provide AI-generated recipes,Äîstretch goal: Personalized price comparisons based on shopping habits. While existing solutions like "Stop Foodborne Illness" consolidate recall data, they lack engagement and personalization. Our app will offer real-time updates, barcode tracking, and AI-driven suggestions. Security and data privacy will be prioritized, ensuring user trust. Key stakeholders include consumers, government agencies, manufacturers, and retailers. By enhancing food safety awareness, our solution aims to reduce waste, improve consumer decision-making, and support public health.

Fairness Tools in Practice: The Researcher Perspective

Authors

Sadia Afrin Mim, Brittany Johnson (George Mason University)

Abstract

As our society grows increasingly reliant on AI, the need for fairness in our models is paramount. A biased model can be devastating to oppressed people. In response, researchers and practitioners have rallied to develop and use tools that support fairness. To understand their practical implication, we designed an interview to curate experiences with fairness tools from researchers. In this paper, we discuss insights from our first round of interviews with researchers from Academia. Although numerous fairness tools have been developed, only a few industry developed(e.g, AIF360 and Fairlearn) are practically usable and commonly

employed by researchers due to regular maintenance. The existing toolkit landscape is primarily equipped to solely handle textual data and lacks sufficient resources for language models. Our findings thus far provide insights into one perspective on fairness tool engagement; our future efforts will investigate experiences and perspectives on fairness tool support beyond research.

LLM-Driven Heuristic Evaluation of Code Snippets from Website Homepages

Authors

Nolan Platt, Ethan Luchs, Sehrish Basir Nizamani (Virginia Tech)

Abstract

Usability evaluations are critical in ensuring that applications meet user needs, yet traditional heuristic evaluations rely on human expertise and are often conducted late in the development cycle. With the increasing role of AI in software development, can we leverage large language models (LLMs) to detect usability flaws earlier? In this flash talk, we will share our experience using LLMs to evaluate code snippets for usability concerns at the programming stage of website homepages. We will discuss insights gained, the effectiveness of AI-driven analysis, and the challenges encountered in integrating LLMs into the usability evaluation process. This talk aims to spark a discussion on the role of AI in early-stage design assessments and its potential to reshape usability evaluation methodologies.

Exploring The Role of Electrodermal Activity (EDA) in Estimating Attention Via Wearable Devices

Authors

Farina Faiz, Vivian Genaro Motti (George Mason University)

Abstract

Attention refers to an individual's alertness or ability to engage with their surroundings. It is closely associated with physiological arousal and can be assessed using Electrodermal Activity (EDA). Daily activities that involve cognitive, visual, and auditory stimuli can influence EDA and indicate variability in individuals' attention levels. Since EDA patterns are unique to each individual, it is essential to explore its role in detecting personal attention levels. Existing studies on this topic are primarily conducted in controlled environments, which limit their application to real-world, non-invasive, and unobtrusive data collection. In addition, acquiring labels for different attention levels remains a significant challenge, as it often depends on annotators or domain experts to label sessions. With In-Situ and Ecological Momentary Assessment (EMA)

responses, understanding correlation between EDA and attention levels is necessary. This study aims to investigate the role of EDA in detecting attention levels during ongoing tasks. Specifically, we focus on determining how EDA data collected from non-invasive wearable devices is related to attention levels. To accomplish this, we initially use EDA recordings from 11 subjects collected from a publicly available Dataset. Furthermore, we explore how EDA patterns (Phasic and Tonic) vary within individuals during their daily activities. Future work involves the development of machine learning models to detect different attention levels using EDA. This study can contribute to enhancing user engagement and productivity in both educational and professional contexts.

AI-Based Usability Evaluations of Low-Fidelity vs. High-Fidelity Prototypes

Authors

Sheikh Moonwara Anjum Monisha (Virginia Tech); Khairatun Hissan (Islamic University of Technology); Sehrish Basir Nizamani (Virginia Tech)

Abstract

Low-fidelity (lo-fi) prototypes, such as wireframes and sketches, are essential in early-stage design for exploring user interactions before high-fidelity (hi-fi) development. Traditional usability evaluations rely on human experts assessing designs against usability heuristics, such as Nielsen‚Äôs 10 Usability Heuristics. With advancements in artificial intelligence (AI), particularly large language models (LLMs) and computer vision-based tools, automated usability evaluation is becoming a viable alternative. However, the effectiveness of AI in identifying usability issues across different levels of prototype fidelity remains an open question. This study investigates how well AI-driven models, specifically LLMs, detect usability issues in lo-fi prototypes compared to hi-fi prototypes. We analyze AI-generated evaluations using Nielsen‚Äôs heuristics, comparing them with expert human assessments. Key research questions include whether AI can effectively interpret low-fidelity designs, the accuracy of its heuristic-based feedback, and the potential gaps between AI and human evaluations. Findings from this research will provide insights into the strengths and limitations of AI for early-stage usability testing and inform how designers can integrate AI-based evaluation into iterative design workflows.

Beyond the Surface: Rethinking Pulse Oximetry for Equitable Healthcare

Authors

Arshnoor Bhutani, Mahi Sanghavi, Kristina Kramarczuk (University of Maryland, College Park)

Abstract

Pulse oximeters, which measure patient blood oxygen levels, are inherently biased. They demonstrate significant inaccuracies for patients with darker skin tones due to melanin interference, compromising a fundamental and preliminary step to all medical check-ups and processes. This heightens risks of occult hypoxemia among historically marginalized communities. Our previous research examined how these biases are perceived and addressed by medical and technical students and professionals at UMD, revealing an awareness gap that needs attention and calling for interdisciplinary collaboration. This work also underscored the importance of incorporating diverse datasets and other adjustments to mitigate racial disparities in SpO,ÇÇ measurement. Building on these findings, our current study investigates the intersection of engineering design choices and clinical usage of pulse oximeters. Through semi-structured interviews with medical professionals and engineers, we seek to identify communication gaps and points of divergence in how these groups approach pulse oximeter accuracy across different skin tones. By analyzing how engineers rationalize wavelength optimization and how doctors interpret potentially flawed readings, we aim to uncover underlying biases and systemic barriers. Our findings will provide data-driven recommendations for manufacturers, advocating for more inclusive design strategies and fostering a collaborative framework that bridges technical feasibility with clinical applicability.

Birds of Feather

Economic Impact of Decision-Making Software on Minorities

Authors

Erika Olimpiew (Virginia Tech)

Abstract

I am interested in understanding more about the economic impact on minorities and vulnerable populations of decision-making software that calculates car insurance rates and prices of goods and wages. Does this decision-making software unfairly target these populations? If so, are there ways to make these decisions more transparent and to mitigate this bias? Can software companies be held accountable for biased decision-making software?

CAPWIC Male Allies

Authors

J. Hott, S. McCrickard, S. Sheth, M. Stewart, M. Seyam

Women and non-binary undergraduate students

Authors

N. Andrus

Technical Talks

Explaining in Diffusion: Explaining a Classifier Through Hierarchical Semantics with Text-to-Image Diffusion Models

Authors

Tahira Kazimi (Virginia Tech); Ritika Allada (Vriginia Tech); Pinar Yanardag (Virginia Tech)

Abstract

Classifiers are important components in many computer vision tasks, serving as the foundational backbone of a wide variety of models employed across diverse applications. However, understanding the decision-making process of classifiers remains a significant challenge. We propose DiffEx, a novel method that leverages the capabilities of text-to-image diffusion models to explain classifier decisions. Unlike traditional GAN-based explainability models, which are limited to simple, single-concept analyses and typically require training a new model for each classifier, our approach can explain classifiers that focus on single concepts (such as faces or animals) as well as those that handle complex scenes involving multiple concepts. DiffEx employs vision-language models to create a hierarchical list of semantics, allowing users to identify not only the overarching semantic influences on classifiers (e.g., the 'beard' semantic in a facial classifier) but also their sub-types, such as 'goatee' or 'Balbo' beard. Our experiments demonstrate that DiffEx is able to cover a significantly broader spectrum of semantics compared to its GAN counterparts, providing a hierarchical tool that delivers a more detailed and fine-grained understanding of classifier decisions.

Ethical AI for Healthcare Systems: Uncertainty-Aware, Fair Federated Learning

Authors

dian chen, Qi Zhang (Virginia Tech); Lance Kaplan (US DEVCOM Army Research Laboratory); Audun Josang (Department of Informatics University of Oslo); Donghyun Jeong (Department of Computer Science and Information Technology, University of the District of Columbia); Feng Chen (The University of Texas at Dallas, Richardson); Jin-Hee Cho (Virginia Tech)

Abstract

This paper proposes {\tt U-FARE}, an uncertainty-aware fair federated learning (FL) framework aimed at improving disease prediction in healthcare, with a specific focus on Alzheimer's disease detection. {\tt U-FARE} incorporates evidential neural networks (ENN) to quantify uncertainty, enhancing both model fairness and accuracy. The framework ensures group-level fairness, providing consistent model performance across diverse healthcare environments despite data heterogeneity. We evaluate {\tt U-FARE} on three real-world healthcare datasets,ÄîNACC, OASIS, and ADNI,ÄîcomparingÄî its performance to several state-of-the-art fairness-aware FL methods. Experimental results demonstrate that {\tt U-FARE} outperforms baseline methods in both prediction accuracy and fairness, effectively balancing these two crucial aspects. The results also reveal the trade-off between fairness and accuracy, where higher fairness levels may compromise prediction accuracy. {\tt U-FARE} achieves the highest accuracy (0.928) on the NACC dataset, consistently outperforms the competitive baseline q-FedAvg by 46\%, particularly when higher fairness constraints are applied, and outperforms methods like Ditto and q-FFL with minimal accuracy variance and loss disparity. This is the first approach to simultaneously optimize fairness and accuracy in FL for Alzheimer,Äôs disease detection, providing a novel solution to the challenge of fair and effective AI in healthcare. The framework demonstrates the potential to address data heterogeneity while ensuring privacy and fairness in real-world applications.

Enhanced methods for incorporating resiliency in UAV swarms

Authors

Abhishek Phadke (Christopher Newport University)

Abstract

UAV (Unoccupied Aerial Vehicle) swarms are excellent examples of a dynamic system whose target operating environment is filled with disruptions, making operations challenging. The swarm has to be resilient to failures and adapt, extend, bounce back, or withstand disruptions to prevent cascaded system failure. While there are several novel techniques for increasing the resiliency of mobile multi-agent robotic systems, this research addresses the challenge from two crucial standpoints. The first is a swarm-specific SAR (Search and Rescue framework). This framework deploys rescue robots to rescue other swarm agents that have faced disruptions. With advanced features such as an agent heartbeat signal detector and insights on agent

failure, rescue agents can find and recover fallen agents, thereby preventing the loss of agents and increasing swarm rebound parameters. The second objective addresses the resiliency problem from the opposite standpoint by examining the disruptions that cause swarm failures. Multiple external disruptions that cause UAV failure are addressed, and techniques for simulating them are showcased. The advantages of disruption simulation and performance benchmarking of UAV swarms to disruptions are key takeaways of this section. These two techniques have demonstrated a significant methodology for researchers to understand the complex resiliency problem and create viable solutions to increase resiliency in Cyber-Physical Systems and networks.

HealthLens: AI-Powered Insights for Healthcare Data

Authors

Nikhil Gaddam, Sreenidhi Gurunathan,Ä¨

Abstract

The growing complexity of healthcare data demands efficient, accessible, and intelligent analytics. MediQuery is an AI-powered healthcare insights platform that enables medical professionals, administrators, and policymakers to seamlessly interact with vast healthcare datasets using natural language queries. By integrating Large Language Models (LLMs) and Retrieval-Augmented Generation (RAG), MediQuery transforms unstructured data into actionable insights,Äîwithout requiring technical expertise. Our system allows users to track patient vitals, analyze disease trends, manage prescriptions, and optimize ambulance response times, all within a secure, role-based environment. Powered by Flask, PostgreSQL, and AI-driven visualization tools, MediQuery dynamically generates SQL queries and Python-based visualizations, ensuring real-time, data-driven decision-making. Beyond functionality, our research explores the ethical and technical challenges of AI in healthcare, including bias in LLM-generated insights, data privacy concerns, and the need for robust access control mechanisms. At CAPWIC, we aim to showcase how AI can bridge the gap between medical expertise and data analytics, enabling smarter healthcare decisions, reducing inefficiencies, and **empowering

Technical Workshops

Breaking Machine Learning Models to understand Security Implications

Authors

Pavan Reddy (The George Washington University)

Abstract

Deep learning models are widely used in critical applications, from image recognition to autonomous systems. However, these models are highly vulnerable to adversarial attacks,Äîsmall, often imperceptible modifications to input data that can cause misclassification. This hands-on workshop is designed for beginners and introduces adversarial machine learning, focusing on attacking deep learning models to understand security implications. In the first part, attendees will learn the fundamentals of deep learning models, how they process inputs, and why adversarial attacks pose a security risk. Through real-world examples and a demonstration of pixel-based perturbations, we highlight the risks introduced by these attacks. Participants will set up Google Colab, install Adversarial Lab, and run their first Fast Gradient Sign Method (FGSM) attack on the Inception V3 model, observing how small changes deceive the model. The second section examines additive attacks using PGD and C&W methods, explaining key parameters such as epsilon and scale while visualizing noise evolution through plots and videos. We explore the mathematical foundations of adversarial noise generation and optimization, enabling participants to manipulate attack strength and analyze model vulnerabilities. Finally, we introduce black-box attacks, where an adversary has no knowledge of the model's parameters. By simulating a black-box scenario and launching a ZOO-based attack, attendees will understand how real-world attackers operate. We conclude with discussions on security implications and open challenges in adversarial machine learning. This workshop is intended towards beginners and no prior experience with adversarial attacks is required, however, basic math and python experience is recommended.

Cybersecurity: Strengthening Workforce Education Excellence in Programming Securely (SWEEPS)

Authors

D. Kariuki, I. Ngambeki