Arun Krishna Vajjala

akrishn@gmu.edu | arunkv.com | 703-303-6284 | linkedIn/arunkrishnavajjala

EDUCATION

Ph.D. Computer Science (GPA 3.9/4)   George Mason University | Aug 2021 - May 2025
M.S. Computer Science (GPA 3.7/4)   George Mason University | Jan 2021
B.S. Computer Science (GPA 3.5/4)   George Mason University | Jan 2020

PERSONAL STATEMENT

My goal as a researcher is to work at the intersection of software engineering and machine learning to create developer-facing tools and enhance the software development workflow, bridging the gap between innovation and practical implementation. I aim to collaborate with research labs to produce impactful research that reaches many people.

PUBLICATIONS

- **A. Krishna Vajjala**, Aj. Krishna Vajjala, Y. Yan, S. Pothagoni, D. Poshyvanyk, and K. Moran. "FRAME: Enhancing Multimodal GUI Embeddings with Structural Information". **(In Review ICSE)**.
- **A. Krishna Vajjala**, H. Mansur, J. Jose, and K. Moran. "MOTOREASE: Automated Detection of Motor Impairment Accessibility Issues in Mobile App UIs". In Proceedings of the 46th ACM *International Conference on Software Engineering* (ICSE 2024), Lisbon, Portugal April 14-20, 2024.
- **A. Krishna Vajjala** and K. Moran. "Engineering Accessible Software". In Proceedings of the 39th IEEE *International Conference on Software Maintenance and Evolution* (ICSME'23), Doctoral Symposium, Bogota, Columbia, 2023.
- A. Krishna Vajjala, **Ar. Krishna Vajjala**, Z. Zhu, and D. Rosenblum. "Analyzing the Impact of Domain Similarity: A New Perspective in Cross-Domain Recommendation." In Proceedings of the IEEE International Joint Conference on Neural Networks (IJCNN 2024), Yokohama, Japan, June 30th to July 5th, 2024.

WORK EXPERIENCE

MICROSOFT RESEARCH | RESEARCH INTERN   Redmond, WA | June 2024 - August 2024

- Mentored by Dr. Christian Bird, Dr. Nicole Forsgren, and Dr. Rob Deline, I identified bottlenecks in the software deployment and build process, leveraging Machine Learning and Artificial Intelligence to automate and streamline workflows for developers.
- Collaborated with developers to gather requirements, designing and building a user interface that integrates K-means clustering on build failures using Azure OpenAI embeddings. This groups failures for easy access and triage by on-call developers. Built with Python, Node.js, Flask, React, and supported by an Azure Kusto Database.
- Deployed custom Large Language Models (LLMs) within an information-secure Azure environment (OpenAi GPT-4o) to proactively tackle explainability, and traceability, significantly reducing manual inspection and fatigue.
- Extracted a set of generalizable design rules to gamify and redesign UI workflows, empowering teams to leverage AI for automating repetitive tasks, enhancing productivity, and minimizing manual efforts.
- Presented this project at an executive review with a Corporate Vice President (CVP), where the partner product team requested an immediate push to production and initiated a successful tech transfer due to its potential to improve developer efficiency. I aim to publish results demonstrating productivity increases for developers.

SAGE RESEARCH LAB | GRADUATE RESEARCH ASSISTANT   Fairfax, VA | May 2021-Present

- **(Research Project)** Created FRAME, which addressed an industry need for enhanced UI layout comprehension, by implementing a Neural Graph based approach to create a structurally motivated GUI embedding. Under review
- FRAME integrated state-of-the-art CLIP, BERT, and Graph embedding techniques, coupled with structure enhancing mathematical concepts in the Rips-Complex and Embedding Propagation to infuse existing embedding techniques with structural semantic context of a GUI
- FRAME embeddings outperformed CLIP embeddings by 8-12% in UI testing and retrieval tasks with statistically significant results, demonstrating its ability to abstract GUI layouts across various screens
- **(Research Project)** Designed and implemented GUIFix, a developer tool leveraging LLMs to detect and automatically repair accessibility issues in Android applications, streamlining the process for developers.
- Developed a novel workflow utilizing Python and CLIP embeddings to localize and repair accessibility issues within applications, enabling automated detection and code modifications with minimal developer intervention.

- **(Research Project)** Designed and implemented MotorEase, an automated tool to detect motor-impairment accessibility issues in mobile applications using Java and Python programming languages, Accepted ICSE 2024
- Integrated state-of-the-art techniques in PyTorch computer vision, pattern-matching, and static analysis to detect various accessibility violations through application screenshots and XML data. Achieved an 87% accuracy when detecting accessibility guidelines at runtime

ALCON | Research & Design Intern   Fort Worth, TX | May 2021-Aug 2021

- Collaborated with a multi-disciplinary team of researchers, engineers, and surgeons to prototype a surgical voice assistant, focusing on improving end-user interactions (surgeons) and enhancing intraoperative workflows through machine learning and software integration.
- Led the design and development of a wake-word detection model using TensorFlow, AWS SageMaker, and PyTorch. Developed a robust audio processing and feature extraction pipeline with Python, Librosa, and PyAudio, achieving 80% accuracy in detecting the wake-word "Hey, Alcon" in real-time input streams. This work pioneered voice assistant technology tailored to the high-stakes requirements of surgical environments.
- Successfully deployed the voice assistant across multiple operating room devices in the U.S., significantly improving user interaction and reducing manual input during surgeries, surpassing initial performance expectations.

INVITED TALKS
- DiffViewer: Infusion of AI/ML in developer workflows to improve productivity, Microsoft Research 2024
- MotorEase: automated detection of Motor Impairment Issues, ICSE 2024
- Engineering Accessible Software, ICSME 2023 Doctoral Symposium
- MIRACLE: Automated Testing in Android Apps, GMU 2022
- Emerging Research Poster Presentation: MIRACLE, GMU 2021

MEDIA
- "The Power of Positionality - Why Accessibility? An Interview With Kevin Moran and Arun Krishnavajjala" Read

CONFERENCE CO-REVIEWING
- ICSE 2025
- MSR 2024
- ICSE 2023
- SANER 2023
- ASE 2022
- MSR 2022
- SANER 2022

MENTORSHIP
- Samar Karanch | University of Central Florida
- Sophia Lin | Thomas Jefferson HS
- Justin Jose | South Lakes HS
- Emma Tan | Bishop Moore Catholic High School

PERSONAL PROJECTS

GITCHAT   ChatGPT-4, Python, Code Abstraction, LangChain, DeepLake

Surveyed developer needs and developed a command-line based approach to querying and interacting with any GitHub codebase to increase developer productivity. GitChat is a Retrieval-augmented generative (RAG) model that enables developers to link public GitHub projects and interactively query a GPT-based large language model (LLM) to comprehend source code. It has a built in caching system to revisit codebases to document, modify, or comprehend code. Built with Python, OpenAi API, DeepLake Vector Store, and LangChain to aid in GPT comprehension

DIABETIC RETINOPATHY IMAGE CLASSIFICATION ⬈   Python, PySpark, TensorFlow, Databricks

Performed a multi-level classification on images of retinas to determine diabetic retinopathy severity. Built using 16 giga-bytes of data, AWS EMR, and EC2, achieving 97% accuracy in detecting diabetic retinopathy severity, demonstrating the effective use of big data and advanced machine learning techniques to address critical healthcare challenges. Provided a robust solution for early detection and treatment planning of diabetic retinopathy

NYC TAXI TIME PREDICTION ⬈   Python, AWS EC2 & EMR, Hadoop, Big Data, PySpark

Enhanced a decision tree model with linear regressors in its leaves, achieving a 75% improvement in RMSE and more accurate trip duration predictions. Utilized PySpark, Hadoop, and 13 giga-bytes of data on AWS EMR and EC2. This approach allowed for efficient handling of big data, leveraging the scalability and computational power of AWS services

SKILLS

Technology: Java, Python, C, SQL, R, AWS, Android, Docker, LaTeX, MongoDB, Kubernetes, Jenkins, Hadoop, Adobe Suite, Unix, Git, Azure, Kusto Query Language (KQL), AzureDevOps, Azure OpenAi Studio