

Ajay Krishna Vajjala

akrish@gmu.edu | profile.ajaykv.com | linkedin.com/in/ajay-kv | github.com/ajaykv1 | (703) 303-4282

EDUCATION

- **PhD in Computer Science** Aug 2021 - Present
George Mason University - Advised by Dr. David S. Rosenblum Fairfax, VA
- **B.S. & M.S. in Computer Science** Aug 2017 - Dec 2021
George Mason University Fairfax, VA

RESEARCH INTERESTS

- My research interests lie in machine learning, with a focus on recommender systems and applied ML in software engineering. I am passionate about driving progress in machine learning, information retrieval, and software engineering research, particularly through real-world applications of LLMs to solve complex challenges, improve developer workflows, and enhance user experiences.

PUBLICATIONS

- **A. Krishna Vajjala**, D. Meher, S. Pothagoni, Z. Zhu, and D. Rosenblum. "Vietoris-Rips Complex: A New Direction for Cross-Domain Cold-Start Recommendation." In Proceedings of the SIAM International Conference on Data Mining (SDM 2024), Houston, TX, U.S. 18th-20th April, 2024. (29.2% acceptance rate).
- **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.
- Ar. Krishna Vajjala, **A. Krishna Vajjala**, Y. Yan, S. Pothagoni, D. Poshypanyk, and K. Moran. "FRAME: Enhancing Multimodal GUI Embeddings with Structural Information" (**Under Review**)

WORK EXPERIENCE

- **Microsoft Research (MSR)** Redmond, WA
Research Intern - Advised by Dr. Tom Zimmermann, Dr. Christian Bird, and Dr. Nicole Forsgren June 2024 - August 2024
 - **Large Language Models for Automating Release Engineering Workflows**
 - * Worked on the SAINTES team at MSR, to implement ML solutions to streamline release engineering workflows.
 - * Collaborated with Microsoft's IDNA product team to explore AI-driven productivity improvements for engineers.
 - * Interviewed on-call engineers to identify workflow inefficiencies and areas for AI enhancements.
 - * Developed a novel model using Large Language Models (LLMs) to automatically label differences between test and production environments as "noise" or "critical," improving productivity and efficiency for on-call engineers.
 - * Fine-tuned GPT-3.5 and GPT-4 using Azure OpenAI Studio, achieving 98% labeling accuracy in offline evaluations.
 - * Conducted online evaluation by testing fine-tuned model during on-call sessions, consistently achieving over 90% accuracy.
 - * Reduced labeling time from 10 hours to just 15 minutes, significantly improving developer productivity.
 - * Presented results to the Corporate Vice President of IDNA and senior leadership at Microsoft, receiving executive review.
 - * Conducted a successful tech transfer, integrating the model into the product team's production workflow.
- **George Mason University** Fairfax, VA
Graduate Research Assistant May 2022 - Present
 - (**Research Project**) **Large Language Models (LLM) for Transfer Learning in Recommender Systems**
 - * Introduced zero-shot and few-shot prompting methods tailored for LLMs to optimize cross-domain recommendation.
 - * Utilized GPT-3.5 and GPT-4, and fine-tuned open-source models such as LLaMA-2 across 7B, 13B, and 70B parameters to evaluate their cross-domain recommendation performance using parameter-efficient fine-tuning (PEFT) methods.
 - * Used QLoRA for fine-tuning LLMs on dual-domain data to enhance cross-domain recommendation capabilities.
 - (**Research Project**) **Vietoris-Rips Complex for Cross-Domain Cold-Start Recommendation**
 - * Leveraged the Rips Complex, from computational geometry, and deep learning to transfer user preferences across domains, resulting in new user profiles for personalized recommendations.
 - * Achieved a performance increase of over 20% in extreme cold-start scenarios, surpassing leading methods.
 - * Employed 5-fold cross-validation for hyper-parameter tuning to improve model performance on unseen data.
 - * Utilized Python and Tensorflow for algorithm development, executing computations on the Nvidia A100 GPU.
 - (**Research Project**) **Domain Similarity for Cross-Domain Recommender Systems**
 - * Used Python, GloVe pre-trained embeddings, and BERT from NLP to develop domain similarity metrics.
 - * Developed baseline cross-domain recommendation algorithms using the open-source Recbole-CDR library and conducted a comprehensive evaluation across 18 domain combinations with three leading cross-domain algorithms.
 - * Findings indicated recommendation performance wasn't significantly affected by domain combinations, using a paired t-test.
- **Center of Adaptive Systems of Brain and Body Interaction** Fairfax, VA
NSF National Research Trainee Fellow May 2021 - May 2022
 - (**Research Project**) **Reentry and Corrections**

- * Led an interdisciplinary team to develop a web application for incarcerated individuals to access reentry services information.
- * Managed web application development with React (front-end), Node/Express (back-end), and MySQL (database).
- * Initiated and led pivotal partnership discussions with both the Director of DC Jail and the Vice President of American Prison Data Systems (APDS). This collaboration aims to trial the app on APDS tablets.

INVITED TALKS

- **Automating Release Engineering Workflows using LLMs** — Microsoft Research 2024, Redmond, Washington
- **Domain Similarity in Cross-Domain Recommender Systems** — IJCNN 2024 Main Conference Session, Yokohama, Japan
- **Computational Geometry for Cross-Domain Recommendation** — SDM 2024 Main Conference Session, Houston, TX
- **Transfer Learning and its Impact in Recommender Systems** — GMU 2024 Software Engineering Seminar, Fairfax, VA
- **Improving User Experience Through Context-Awareness** — GMU 2023 Software Engineering Seminar, Fairfax, VA
- **Leveraging Software Engineering to Enhance Reentry Experiences for Incarceration** — CASBBI 2022, Fairfax, VA

MENTORSHIP

- **Dipak Meher** (PhD Student)— Summer 2023 to Present
 - Affiliation: George Mason University | Fairfax, VA
 - Project: Generating images for items based on user preferences, by leveraging LLaVa and DALL-E, to improve recommendations.
- **Roshan Ram** (High School Student)— Summer 2024
 - Affiliation: Westfield High School | Fairfax, VA
 - Project: Empirically analyzing the performance of context-aware recommendation models under differing circumstances.
- **Rishab Nanduri** (High School Student)— Summer 2024
 - Affiliation: Oakton High School | Vienna, VA
 - Project: Improving context-aware recommender systems by leveraging LLMs to enhance quality of side-information.
- **Soham Gangrade** (High School Student)— Summer 2024
 - Affiliation: Westfield High School | Fairfax, VA
 - Project: Novel dataset for context-aware recommendation, which leverages LLMs to generate contextual features for users.
- **Shaurya Mallampati** (High School Student)— Summer 2024
 - Affiliation: Mariotts Ridge High School | Howard, MD
 - Project: Empirical study to evaluate LLM enhanced context-aware recommendation models with deep learning models.

ACADEMIC SERVICE

- **Social Event Coordinator** — 2024 GMU Software Engineering Research Group, Fairfax, VA
- **Seminar Coordinator** — 2023 GMU Software Engineering Seminar Series, Fairfax, VA
- **Seminar Coordinator** — 2022 GMU Software Engineering Seminar Series, Fairfax, VA
- **Student Volunteer** — 2022 ACM Conference on Recommender Systems (RecSys), Seattle, WA

PERSONAL PROJECTS

- **GitChat**
 - Developed a GPT-4 powered RAG tool letting developers ask questions and query their GitHub repository for code insights.
 - Built using Python, and integrated the OpenAI API, DeepLake Vector Store, and Langchain for enhanced querying.
 - Offered developers a user-friendly command line tool to easily ask, understand, and clarify doubts directly within their codebase.
- **Amazon User Review Sentiment Analysis**
 - Leveraged BERT pre-trained embeddings, a transformer-based NLP technique, to fetch embeddings for product reviews.
 - Represented each review as an average of its word embeddings and used K-Nearest Neighbors (KNN) to predict sentiment.
 - Achieved 80% accuracy in predicting sentiment of product reviews using the BERT and KNN approach.
- **Collaborative Filtering Recommender System**
 - Developed a deep learning single-domain recommender system using an MLP network, implementing the tower method with progressively smaller dense neuron layers to output rating scores between 1-5.
 - Conducted 5-fold cross-validation to tune hyperparameters, optimizing the number of dense neurons, regularization terms, dropout rates, and learning rates for improved model performance.
 - Evaluated the model using Root Mean Square Error (RMSE), achieving performance on par with state-of-the-art models.
- **NBA Winner and Score Prediction**
 - Developed a neural network model to predict the score and winner of NBA games based on first-quarter statistics, achieving 85% accuracy in predicting winners and final scores within 5 points of the actual outcome
 - Created a user-friendly GUI application allowing users to input first-quarter stats such as steals, rebounds, and blocks for both teams, with the model predicting the game's outcome.
 - Trained the model using NBA game statistics from 1970 to 2017, leveraging historical data to enhance prediction accuracy.

TECHNICAL SKILLS

- **Languages:** Python, Java, PySpark, React, Node, Express, SQL, Kusto Query Language (KQL)
- **Tools:** Tensorflow, Scikit-Learn, RecBole, GIT, Azure OpenAI Studio, Azure DevOps (ADO)