AnushKrishna VenkataKrishnan

av3278@rit.edu | github.com/anushkrishnav | linkedin.com/in/anushkrishnav | (585)-981-1189 | New York

Education

Rochester Institute of Technology | Masters, Data Science | GPA: 3.94/4.0 **May 2026**
Coursework: Statistical Machine Learning, High Performance Data Science(MPI & CUDA), Applied Statistics
Bharathiar University | Bachelors, Computer Science | GPA: 3.44/4.0 **August 2022**

Technical Skills

Python | C++ | CUDA | Apache Airflow | Azure | Celery | Docker | FastAPI | Flask | Git | Google Cloud | GraphQL | JAVA | Kubernetes | Lightning | Numpy | PostgreSQL | Pytorch | Redis | Spacy | SQL | Tensorflow

Experience

Machine Learning Engineer August 2024 – Dec 2025
Metabob San Francisco, CA

- Spearheaded the development of a custom graph extraction framework to support a **GNN-based bug detection model**, addressing the challenge of representing complex code structures accurately.
- Created a scalable pipeline to generate over **800K high-fidelity AST and CFG graphs weekly**, solving code parsing and extraction inconsistencies to ensure robust model input.
- Diagnosed graph sparsity and edge inconsistency issues across noisy real-world codebases, building **validation and fallback mechanisms** that increased graph quality and model convergence rates.
- Enhanced GNN training robustness by implementing graph cleaning, sampling strategies, and augmentation techniques, directly improving learning stability and bug detection accuracy.
- Deployed a **Prefect-based orchestration system on Azure Kubernetes Services**, accelerating large-scale graph processing workflows and reducing model experimentation time by over 80%.

Lead Backend Engineer September 2023 – Present
RIT Student Government Rochester, NY

- Revamped the **PawPrints platform** serving **3,000+ RIT students**, enabling efficient petition creation, signing, and distribution.
- Implemented OAuth authentication for PawPrints using Shibboleth, enabling secure single sign-on (SSO) for users while integrating with institutional identity providers.
- Optimized database design and restructured SQL queries, leveraging **Postgres full-text search** to improve search performance by **10x**.

Data Engineer Dec 2021 – Jul 2023
Metabob Mountain View, CA

- Built a scalable data collection pipeline using **Celery & Redis** for queuing, **Postgres** for storage, and **Kubernetes** for scaling, improving code and commit data extraction.
- Achieved a **52%** reduction in preprocessing and training time by deploying a **Dask-based NLP pipeline** on **GKE and AKS** with batch parallelization.
- Migrated Dask ML and data pipelines from **GCP to Azure**, ensuring seamless transition with zero downtime.

Publications

- **V. AnushKrishna**, B. Alkhazib, and M. W. Mkaouer. "LibShift Search: A Similarity-Driven Model for Mapping Deprecated Python API Methods to Current Alternatives." Under review, 2025.
- E. A. AlOmar, **V. AnushKrishna**, M. W. Mkaouer, C. Newman, and A. Ouni. "How to refactor this code? An exploratory study on developer-ChatGPT refactoring conversations." In Proceedings of International Conferences on Mining Software Repositories, 5 pages, 2024 [MSR].

Projects

LibShift: Improving LLM-Based Code Generation | *Python, Graph Embeddings, Research Collaboration*

- Initiated based on findings from prior research analyzing **17,000+ developer-LLM interactions**, identifying key pitfalls in LLM-driven software refactoring and third-party API migrations.
- Leading an ongoing research effort to enhance **code intelligence for deprecated API migrations**, collaborating with Prof. Mohamed Wiem Mkaouer (University of Michigan-Flint) and Prof. Bader Alkhazi (Kuwait University).
- Project is **currently supported by a research grant from Kuwait University** and focuses on developing graph-based techniques and embedding models for improving LLM code generation.
- Actively working towards publishing **two first-author papers** based on this research initiative.