Kolade Alabi

(956) 494-0935 | College Park, Maryland | kalabi1@umd.edu | kalamojo.github.io | Kalamojo | kolade-alabi

Summary

ML Engineer & Graduate Researcher at University of Maryland, College Park with 2+ years experience shipping production-grade ML features and resilient backend systems. Currently focused on developing theoretical frameworks for robust optimization in ML systems.

Education

University of Maryland, College Park — *College Park, MD*

Master of Science in Computer Science — *Expected May 2027*

- **Key Coursework**: Theory of Robust Machine Learning, Systems for Machine Learning, Uncertainty Communication for Decision-Making, Interactive Technologies in Human-Computer Interaction, Advanced Computer Graphics

Skills

Machine Learning & Data TensorFlow, Keras, PyTorch, ONNX, Scikit-learn, SciPy, Pandas, Polars, NumPy, Matplotlib, LangChain, LangGraph, SQLite, MySQL, PostgreSQL, PgVector, MongoDB, OpenSearch/ElasticSearch

Languages & Web: Python, Flask, Django, FastAPI, Streamlit, C/C++, C#, .NET, JavaScript/TypeScript, React, Webpack, R, Java, Spring Boot, Apache Maven, SQL, HTML/CSS

CI/CD: Git, Terraform, AWS, Docker, Apache Jmeter, GitHub Actions, Jenkins, SonarQube, Spinnaker

Experience

University of Maryland, College Park — *College Park, MD*

Graduate Research Assistant — *January 2026 - Present*

- Developing theoretical Machine Learning metric frameworks for outperforming Empirical Risk Minimization (ERM) in multi-class settings
- Investigating robust optimization techniques for minimizing sample complexity and training in practial settings

Teaching Assistant — *September 2025 - December 2025*

- Facilitated communication between professor and students, and assisted with management of Unity-based Game Programming course, leading grading organization and course announcement activities
- Tutored students individually to improve understanding of course topics, including Unity Development, C# programming, and game physics

JPMorgan Chase — *Houston, TX*

Software Engineer — *August 2023 - August 2025*

- Spearheaded the design and development of an Agentic LLM Assistant for presentation creation, fit with Q&A support capabilities as well as integrations with newly created VectorDB internal knowledge base for Retrieval Augmented Generation (RAG), content library APIs, and internal financial applications
- Developed the core platform for the AI Assistant, a .NET PowerPoint add-in that streamlined business presentation creation and unrestricted content retrieval for over 80,000 users
- Modernized core APIs to AWS hosting, including the refactoring of Apache Solr-powered search to AWS OpenSearch full-text querying — increasing scalability and allowing for retirement of global physical servers
- Execution Excellence Award Q1 2024: One of 5 recipients out of ∼200 JPMC SEP engineers for contributions to the modernization of key application modules and sharing of expertise on AWS ECS and OpenSearch

Carnegie Mellon University — *Remote*

Research Intern (Part Time) — *July 2023 - December 2024*

- Designed k-d tree false positive particle filtering system, reducing noise in particle data by more than 40% to improve clustering performance
- Created Singularity container to run the Deep Iterative Subtomogram Clustering Approach (DISCA) deep learning pipeline end-to-end and furthermore trained CMU collaborators from multiple universities on the usage of said container, reducing setup time by nearly 80%

Capital One — *McLean, VA*

Software Engineering Intern — *June 2023 - August 2023*

- Engineered and deployed Universal Deep Link servicing across web, email, and mobile application channels — powering seamless integration for partners like Walmart, Google Autofill, and Zelle — driving increased traffic to the EASE mobile application while reducing navigation time by up to 75% for ∼38 million existing EASE users
- Deployed a PostgreSQL database on Amazon Aurora, eliminating the use of a middleman service and reducing expenses by 19%

Projects

By Its Cover | PyTorch, ONNX, OpenAI CLIP, FastAPI, PgVector, GraphQL, .NET 9, AWS (Batch, Lambda, RDS)

- Architecting a multi-vector embedding-based search engine to support for vague natural-text searches for books by way of a Late-Interaction approach, enabling fine-grained spatial queries for both cover images and book metadata
- Developing a CLIP embedding-based hybrid Collaborative Filtering model for recommending book covers to users, taking advantage of features from both user relationships (by ratings) as well as aesthetic qualities from cover images
- Implemented an automated database ingestion pipeline, fit with book detail sourcing from both GoodReads and the HardCover API, to populate an AWS RDS PostgreSQL database of book cover CLIP embeddings and metadata, enabling semantic search and recommendation capabilities
- Deployed a serverless CLIP embedding endpoint to AWS Lambda, utilizing ONNX to quantize and optimize the CLIP model, reducing cold-start overhead and inference times by more than 80%