Yousuf Khalid

US citizen | yousufkhalid.com | yousufkhalid653@gmail.com | linkedin.com/in/yousuf-khalid | github.com/SaluteTheNoot

Education

University of Texas at Dallas **Richardson, TX**

*Bachelor of Science in **Computer Science*** Expected Graduation: **May 2026**

Relevant Coursework: Machine Learning, Operating Systems, Databases, Computer Architecture, Data Structures and Algorithms

Experience

Software Engineer Intern @ Capital One June 2025 – Aug 2025

- Built a Python ETL pipeline that generates a device lock list from vulnerability scans across 70k devices belonging to 50k employees
- Wrote 200+ lines of code with a PyTest framework for unit, component, and live dependency tests, achieving 90%+ test coverage
- Aggregated and transformed 30 GB/week from Snowflake using YAML-defined connectors, Pandas transforms, and DQ checks
- Deployed to production on AWS Batch (Fargate) for weekly runs; sub 1 min end-to-end runtime, CloudWatch Logs + Slack alerts
- Freed 15 hours/week for associates, saved $50k+ annually, reduced manual errors completely and lock-list turnaround time by 98%

Machine Learning Engineer Research Intern @ UTD Aug 2023 – Dec 2024

- Analyzed political messaging by scraping and processing 50 GB of news data using Python, Selenium, and BeautifulSoup
- Contributed 400+ lines of code building data pipelines with Pandas to clean and merge news dataset
- Implemented custom tokenization tailored to English morphology with WordPiece, reducing out-of-vocabulary tokens by 30%
- Fine-tuned BERT LLM on HPC clusters with PyTorch, applied RAG with embedding-based retrieval; Achieved a 95% F1 scores

Software Engineer Intern @ Critical Start June 2024 – Aug 2024

- Developed an automated phishing investigation agent using LLMs and CNNs with 2000+ lines of code
- Trained a CNN network on 50,000+ emails in Azure Machine Learning
- Utilized RAG with Pinecone vector databases and OpenAI text embeddings to retrieve context for the LLM, like safe email domains
- Integrated CNN-based classifications into a LangChain-driven prompt engineering workflow across 3 LLMs, achieving 98% accuracy
- Automated the generation of comprehensive incident reports, saving SOC Analysts over 100 hours and $10,000 per month

Software Engineer Intern @ NASA January 2024 – April 2024

- Built an Apache Spark ETL pipeline to process large volumes (70 GB) of biometric data for health anomaly detection
- Built FastAPI services backed by Apache Cassandra; containerized with Docker and Kubernetes, reducing API latency by 25%
- Trained Transformer models for health anomaly detection with PyTorch, C++, and CUDA, optimizing for 20% faster GPU training
- Improved on baseline and achieved 90%+ anomaly detection accuracy, storing model outputs and analytics in PostgreSQL

Projects

UTD Trends **React, Next.js, TypeScript, Node.js, Tailwind, MySQL, HTML/CSS**

- Developed a 7,000+ user Next.js-based platform offering searchable course and professor data for 1,000+ classes at UTD
- Implemented class comparisons, RMP-based filtering, and grade distribution visualizations to help inform academic decisions

Polymarket Feed Engine **C++, WebSocket, SIMD, Linux, CMake**

- Built a low-latency C++ WebSocket feed handler maintaining live bid/ask order books across 100+ Polymarket markets in real-time
- Lock-free SPSC ring buffers decoupling ingestion and processing threads, achieving sub-50µs latency via rdtsc benchmarking
- VWAP and order imbalance signals using ARM NEON SIMD intrinsics, achieving 3x throughput over scalar implementation

Skills

Full-Stack: Python, C++, Go, React , JavaScript, TypeScript, RESTful APIs, SQL

Data Science: PyTorch, Scikit Learn, NumPy, Hugging Face, OpenCV

DevOps/Cloud: AWS Certified Solutions Architect Associate, AWS, Docker, Kubernetes, Terraform, Jenkins, Ansible, Linux

Other: Git/Github, Figma, Jupyter Notebook, Jira, Confluence, Agile, Kanban, Scrum, Codex, Claude Code