

MOHAMMED ANSARI

San Jose, California

📞 (669) 288-3655

✉️ mohammed.ansari@sjtu.edu

.linkedin.com/in/imohammedansari

github.com/ByteMaverick

mohammedansari.dev

Education

San Jose State University

Bachelor of Science in Data Science

Aug. 2023 – May 2027

San Jose, California

- **Coursework:** Data Structures and Algorithms, Advanced Machine Learning, Advanced Python, Processing Big Data, Object Oriented Design, Linear Algebra, Statistics, Calculus

Technical Skills

Programming & Core: Python, Java, SQL(PostgreSQL, MySQL),Bash Scripting

Machine Learning & GenAI: TensorFlow, Hugging Face, OpenAI/Vertex AI APIs, LangChain, LangGraph

Cloud & DevOps: AWS, Google Cloud, Docker, MLflow, Apache Airflow, GitHub, GitHub Actions (CI/CD)

Frameworks & Concepts: FastAPI, REST APIs, Unit Testing (PyTest/JUnit), Agile, System Design, A/B Testing, Jira

Experience

Spartan Knowledge-Driven Intelligence & Learning Lab @ SJSU

Sep 2025 – Present

Research Assistant

San Jose, California

- Designing and implementing data pipelines to **collect, clean, and integrate structured graph data** for large-language-model research aimed at improving reasoning and transparency.
- Experimenting with different approaches to **knowledge-graph schema design**, data preprocessing, and **early reinforcement-learning prototypes** to advance graph-guided LLM experiments in healthcare and financial domains.

Artificial Intelligence & Machine Learning Club @ SJSU

Feb 2025 – Present

Lead Outreach Officer

San Jose, California

- Lead a **five-person team** to organize events on generative models, responsible AI, and applied machine learning, helping students explore both theory and practice.
- Work with faculty and guest speakers on workshops and talks connecting academic learning to real AI/ML use cases.

Projects

DockFlow – Real-Time Bike Dock Availability Predictor

Aug 2025 – Present

Creator

San Jose, California

- Developed a real-time **ML platform** predicting Bay Wheels dock availability 15 min ahead using live data & weather.
- Built a scalable **Apache Airflow ETL pipeline** on **GCS/BigQuery** to automate ingestion, cleaning, and feature engineering, delivering reproducible, low-latency updates for the time-series model.
- Implemented continuous **model-drift monitoring** with **Evidently AI** and deployed a containerized **FastAPI** service to **GCP Cloud Run** with CI/CD via GitHub Actions for automated retraining and seamless releases.

ClauseBit – AI-Powered Privacy Analyzer

May 2025 – July 2025

Creator

San Jose, California

- Built an **AI-powered browser extension and web app** that scrapes and analyzes privacy policies to flag risky clauses based on user preferences, enabling real-time clause-level Q&A and concise legal summaries.
- Designed an **Adaptive RAG pipeline** with LLM APIs, vector search, and clause-relevance filtering to cut hallucinations by 40%, combining dynamic retrieval with vector confidence scoring and fallback rejection.
- Engineered a self-updating **knowledge base** with on-demand web scraping agents and Pydantic models, deploying the full-stack system on **GCP Cloud Run/Functions/Firebase** with async caching for sub-25s latency.

TripSafe – Traffic Accident Risk Prediction Tool

Apr 2025 – May 2025

Team Lead — Course Project – A Grade

San Jose, California

- Led end-to-end development of a **multi-class accident risk classifier** using 7M+ U.S. records, handling data cleaning with **Bash**, feature engineering, and model deployment to guide traffic-safety and route-planning systems.
- Created interpretable risk labels (Low, Medium, High) through **K-Means** clustering and applied **PCA** to reduce dimensionality for efficient, high-performing multi-class classification.
- Tuned **Logistic Regression, SVM, XGBoost** with GridSearchCV and achieved **95%** accuracy using cross-validation.

Certifications and Awards

- **AWS Certified AI Practitioner**

Jun 2025

- **Dean's List, San Jose State University**

Fall 2023 – Spring 2024