MOHAMMED ANSARI

San Jose, California

📞 (669)-288-3655 ✉ mohammed.ansari@sjsu.edu 🔗 linkedin.com/in/imohammedansari 💻 github.com/bytemaverick
🌐 mohammedansari.dev

Education

San Jose State University	**Aug. 2023 – May 2027**
Bachelor of Science in Data Science — GPA: 3.94(Spring 25)	*San Jose, California*

- **Coursework:** Data Structures and Algorithms, Machine Learning, Advance Python, Database Management, Graph Theory, Object Oriented Design

Technical Skills

Programming Languages: Python, Java, R, Bash
Machine Learning & LLMs: scikit-learn, TensorFlow, Hugging Face Transformers, LangChain, LangGraph
Explainability & Data Analysis: SHAP, LIME, Pandas, NumPy, Matplotlib, Seaborn
Databases & Storage: MySQL, PostgreSQL, MongoDB, SQLite
Cloud & MLOps: AWS, Google Cloud, Docker, Apache Airflow, MLflow, DVC, GitHub Actions, Linux
Product & Collaboration Tools: Figma, Notion, Github Jira, Google Analytics
Concepts & Frameworks: RAG, Vector Search, RESTful APIs, A/B Testing, CI/CD pipelines, Flask

Technical Experience

Clausebit − AI-Powered Privacy Analyzer	**May 2025 – Present**
Creator	*San Jose, California*

- Designed and built an AI-powered browser extension and web app that summarizes Terms of Service, Privacy Policies, and Cookie Banners using user-defined privacy preferences; led product prototyping in Figma.
- Implemented an Adaptive RAG pipeline using LLM APIs + Vector Search to flag risky legal clauses (e.g., third-party data sharing, location tracking), allowing users to query and receive simplified explanations of policy clauses.
- Engineered a dynamic knowledge base that uses agentic AI agents to retrieve and store missing company data on demand, cutting reliance on expensive manual update workflows by over 60%.
- Achieved sub-3s latency through optimized chunking, document processing, and lightweight GCP deployment.
- Implemented pre-prompt hallucination control by validating clause relevance and filtering low-confidence vectors, reducing factual errors by 35%, and deployed the app on GCP using Cloud Run, Functions, Firestore, and MongoDB.

TripSafe: Traffic Accident Risk Prediction Tool	**April 2025 – May 2025**
Creator	*San Jose, California*

- Built a web application to predict traffic accident risk based on user inputs using real-world California traffic data.
- Generated interpretable risk labels and improved inference time using K-Means clustering and PCA.
- Trained and evaluated multiple ML models; selected best-performing model with 98% accuracy.
- Used LIME and SHAP to explain model predictions, enabling transparent risk scores and improving user trust.
- Deployed the Flask-based application on AWS with real-time inference and auto-scaling capabilities for reliable access.

PennyPilot: Intelligent Budget Management tool	**May 2025 – May 2025**
Co-Creator	*San Jose, California*

- Tackled the hassle of manual budgeting by creating a tool that automates expense tracking from PDF bank statements.
- Parsed financial data using pdfplumber, transforming unstructured PDFs into clean, tabular formats for analysis.
- Leveraged a BERT-based NLP model to categorize transactions, outperforming traditional classification models.
- Built a responsive PyQt5 GUI for users to view, filter, and manage their budget, with secure local SQLite storage.
- Collaborated on product design and planning using Figma for UI prototyping and Notion for project coordination.

Leadership/Extracurricular

Machine Learning Club @ SJSU	**Feb 2025 – Present**
Outreach and Recruitment Officer	*San Jose, California*

- Organized onboarding sessions and engagement activities to retain members and increase participation.
- Tracked weekly engagement metrics to guide outreach and improve retention.

Certifications and Awards

• **AWS Certified AI Practitioner**	**Jun 2025**
• **Machine Learning A-Z (Udemy)**	**Jul 2025**
• **Dean's List, San Jose State University**	**Fall 2023 – Spring 2025**