Khalid Mehtab Khan

San Francisco, CA • 404-263-7813 • khalidmehtabk@gmail.com
LinkedIn: linkedin.com/in/khalidm-khan • GitHub: github.com/Kahl-d • Website: khalidmk.vercel.app

EDUCATION

Master of Science in Data Science & Artificial Intelligence | GPA: 4.0
San Francisco State University, San Francisco, CA   Expected May 2025
Coursework: Machine Learning, Deep Learning, NLP, Data Mining, Software Development, Algorithm Analysis

SKILLS

Programming: Python, SQL, JavaScript, C++, PyTorch, scikit-learn, React, Flask, REST APIs
ML & AI: Supervised/Unsupervised Learning, Model Evaluation, Feature Engineering, LLM Fine-Tuning, Gen-AI
NLP: Named Entity Recognition, Semantic Search, Masked Language Modeling, (DeBERTa, Mistral, LLaMA)
Tools & Platforms: AWS, Azure, Docker, Ollama, MongoDB, Weights & Biases, Tableau, D3.js, t-SNE

RELEVANT EXPERIENCE

Research Assistant
Tacit Alma Lab, San Francisco, CA   Aug 2024 – Present

- **Multi-Label Text Classification**: Spearheaded development of an NLP pipeline to automate qualitative coding of STEM student reflections. Fine-tuned DeBERTa with LoRA on A100 GPUs and utilized Bayesian hyperparameter optimization (Optuna) to maximize F1-score on imbalanced data.

- **Retrieval-Augmented Generation (RAG):** Engineered context-aware RAG workflows to identify nuanced Cultural Capital themes in student text. Combined domain-adapted language models (DeBERTa and Mistral) with semantic retrieval and prompt engineering to improve detection of subtle themes.

- **Model Evaluation:** Improved labeling efficiency and interpretability of model outputs. Achieved a 12% recall gain and a 75% reduction in manual annotation effort by introducing t-SNE–guided diagnostics for error analysis.

Instructional Student Assistant
San Francisco State University, San Francisco, CA   Aug 2024 – Present

- CSC 620/820 NLP: Deliver lab sessions and develop course materials for topics including Regression, Classification, Neural Networks, and Model Evaluation, serving 70+ upper-division students.

Healthcare Associate
Innovaccer, Noida, India   Jan 2021 – Sep 2022

- **Data Pipeline Development:** Led a team of 3 in building robust ETL pipelines to unify healthcare payer-provider datasets. Accelerated business intelligence capabilities by 40% through improved data quality and automation.

- **Data Integration & Analysis:** Consolidated cross-platform healthcare data, reducing data latency by 30%. Uncovered key trends in patient-provider interactions, informing client outreach and product strategy.

- Developed data-driven KPI reports that enhanced client outreach strategies. Resulted in increased user engagement and system performance improvements, as evidenced by client adoption of recommended practices.

PROJECTS

Secure Sense *(Winner: Emerging AI Innovation Hack, SF Hacks 2025)*   Mar 2025 – Present

- **Real-Time Sensitive Data Masking:** Developed a Chrome browser extension using on-device compact LLMs to detect and mask sensitive user data (emails, personal info) during web browsing in real time, enhancing privacy.

- **Model Optimization:** Led model compression using distillation (LLaMA) reducing model size from approximately 2GB to less than 400 MB while preserving performance. Deployed compact models locally via Ollama for efficient inference.

- **System Integration & Deployment:** As part of a 48-hour hackathon, collaborated on building a hybrid inference pipeline using React (frontend), rule-based logic, and a Flask API backend to serve predictions with low latency.

Context-Aware Data Augmentation Tool for Low Data Scenarios   Aug 2024 – Dec 2024

- Developed a data augmentation pipeline to address class imbalance in health survey datasets (NHANES) by fine-tuning BioBERT with domain-specific knowledge graphs.

- Achieved a 5% accuracy improvement through controlled A/B experiments, validating relevance of augmented data.