Khalid Mehtab Khan

404-263-7813 | khalidmehtabk@gmail.com | linkedin.com/in/khalidm-khan | github.com/Kahl-d | khalidmkhan.vercel.app

EDUCATION

Master of Science in Data Science and Artificial Intelligence (GPA: 4.0)   Aug 2023 – May 2026
San Francisco State University   *San Francisco, CA*

SKILLS

Languages & Frameworks: Python, SQL, C++, PyTorch, Hugging Face, Git

ML & Modeling: Supervised Learning, Unsupervised Learning, Classification, Reinforcement Learning, Transformers (BERT, GPT, Mistral), Embeddings, Fine-tuning (PEFT, LoRA), Model Optimization

LLMs & AI Systems: Generative AI, Large Language Models, Retrieval-Augmented Generation (RAG), Semantic Search, AI Agents, Voice Agents, Real-time Inference, Vector Databases, WebRTC

Evaluation & Deployment: Model Evaluation, Error Analysis, Metrics, Azure (Azure ML, Azure OpenAI), High-Performance Computing (HPC), On-device Inference, vLLM

PROFESSIONAL EXPERIENCE

Founding AI Engineer   Aug 2025 – Dec 2025
Xuman.AI   *San Francisco, CA*

- Led 0-to-1 development as a founding AI engineer, working closely with the CEO to identify use cases, shape product direction, and ship agent-based systems.
- Built real-time voice agents using WebRTC for low-latency streaming, enabling interruptible, human-like conversations.
- Developed AI agents combining retrieval-based search, structured prompting, and tool use to generate grounded responses. Added tools to perform actions like appointment booking and app navigation.
- Designed evaluation pipelines assessing factual correctness, response quality, cross-turn consistency, and behavioral alignment using LLM-as-a-judge and manual validation.

Instructor – Data Science & Machine Learning   Aug 2025 – Present
College of Professional and Global Education, SFSU & Genentech   *San Francisco, CA*

- Instructed a structured 5-course Data Science & Machine Learning certificate program for Genentech professionals, delivering hands-on Python, machine learning, and neural network training tailored to biomedical use cases.

NLP Researcher   Aug 2024 – Aug 2025
ALMA Lab (SF BUILD)   *San Francisco, CA*

- Designed **AWARE**, a narrative approach for multi-label classification of cultural capital themes using essay-level context.
- Modeled narrative context with BiLSTM and attention pooling, aligning representations through domain-adaptive fine-tuning of DeBERTa using LoRA (PEFT).
- Improved performance by +2.1 Macro-F1 and built human-in-the-loop annotation pipelines, increasing labeling throughput by over 75%, with significant gains on rare classes.

Research Assistant   Jan 2024 – Aug 2024
Data Visualization & HCI Lab (SFSU)   *San Francisco, CA*

- Built a hierarchical image classification pipeline over 1,000 ImageNet-derived species using biological taxonomies. Traced inference paths through the hierarchy to identify where predictions diverged from ground truth and analyze misclassification trends.
- Used feature importance methods (LIME, SHAP) to highlight image regions driving predictions, revealing reliance on environmental correlations rather than object-specific features in both misclassified and correctly classified samples.

Data Analyst   Jan 2021 – Sep 2022
Innovaccer   *Noida, India*

- Led a team of three to integrate, process, and analyze EHR, claims, and operational datasets. Generated actionable insights to inform executive strategy and improve data reporting reliability for value-based care.
- Built automated analytical workflows and dashboards, reducing manual reporting effort by 30%. Accelerated data delivery for partners by translating complex metrics into actionable operational insights.

PROJECTS

SenseSecure | *SF Hacks 2025 Winner (Emerging AI Innovation)*   Mar 2025
- Built a privacy-preserving browser extension for real-time PII detection and redaction using on-device LLM inference.

Context-Aware Data Augmentation   Dec 2024
- Developed a context-aware synthetic data augmentation pipeline for low-resource biomedical text, improving downstream Random Forest classification performance by 3–5%.

PUBLICATIONS

- Khan, K. M., & Kulkarni, A. (2025). **AWARE**: Beyond Sentence Boundaries – A Contextual Transformer Framework for Identifying Cultural Capital in STEM Narratives. arXiv:2510.04983