

Aakash Singh

+91 8433821672
aakashs11.github.io

aakashsingh1129@gmail.com
linkedin.com/in/aakash-singh

EDUCATION & COURSEWORK

- IIT Bombay** Mumbai, India
B.Tech. in Metallurgical Engineering & Materials Science March 2021
- IIT Hyderabad** Hyderabad, India
Certification in Artificial Intelligence and Machine Learning (Gold Medallist) June 2019

WORK EXPERIENCE

- Nytro.ai (acquired by MarketStar)** San Francisco (Remote)/ Bangalore, India
Founding Software Engineer (Applied AI/NLP) May 2022 – Dec 2024
 - Domain-Specific LLM Adaptation for Sales Enablement:** Fine-tuned GPT series with domain-specific corpora and programmable guardrails to align outputs with enterprise sales workflows. Achieved a 14% reduction in SDR onboarding time and an 11% uplift in close-win rates for 2,000+ agents across sales cohorts.
 - Retrieval-Augmented Dialogue Agent for Multi-turn Sales Simulations:** Architected RAG-based memory to ground LLM outputs in vectorized domain knowledge using LlamaIndex, Langchain, and PostgreSQL. Enhanced coherence in AI-generated responses and reduced hallucination rates by 40% across 50+ sales simulation paths.
 - Contrastive Feedback Engine for Semantic Pitch Re-ranking:** Built a semantic similarity engine combining BERT, OpenAI and Gemini embedding models with contrastive learning to match sales rep pitches with expert baselines. Deployment led to a 40% improvement in feedback precision and reclaimed 625 hours of manual effort.
 - Multi-Metric Evaluation Suite for Dialogue Systems:** Built a systematic evaluation framework for AI-generated conversations using ROUGE, BERTScore, and task-specific coherence checks (topic derailment, redundancy loops) to benchmark response quality and ensure safety across services. This evaluation framework drove a 20% gain in dialogue clarity and contextual relevance through iterative model improvements.
 - Scalable Topic Segmentation for Contact Center Transcripts :** Implemented topic segmentation for multi-modal long-form conversations using DistilBERT and a sliding window mechanism. This reduced manual evaluation latency by 62%, accelerating batch feedback analysis from 4 minutes to 1.5 minutes on 100K+ calls.
 - Prompt Optimization and Stability Testing Infrastructure:** Engineered an automated prompt testing suite incorporating zero-shot, few-shot, and policy-gradient strategies. GradSUM-based reliability checks ensured response stability, achieving 98% fidelity across edge and fallback cases. Integrated human-in-the-loop strategies.
- Unacademy** Bangalore, India
Educational Content Manager June 2021 – Feb 2022
 - Content Strategy for Learning Platform and YouTube:** Spearheaded curriculum development and optimization for K-12 STEM subjects. Used learner analytics to drive syllabus updates and content publishing, resulting in 20M+ cumulative watch hours and contributing to a 23% increase in platform revenue in 3 months.
 - User Behavior and Retention Analysis:** Conducted exploratory analysis on user behavior using NLTK. Built sentiment-driven models, leading to a 46% reduction in churn and a 60% increase in session-level engagement

PROJECTS

- Real-Time Sign Language Detection | Guide: Ph.D. Manisha Dubey, IIT Hyderabad | Project Link:** Constructed a real-time gesture recognizer with 97% accuracy using Deep CNNs (VGG-16, VGG-19) trained on a dataset of 50K+ hand gesture images. Deployed using TensorFlow with transfer learning for 7 custom hand gestures.
- Measurement of Weld Quality | Guide: Prof. Satish Vitta, IIT Bombay:** Designed a supervised learning system to correlate arc sound features with weld quality using decision tree models (80% accuracy). Compiled a 25-page research report and demonstrated ML-based non-invasive weld quality assessment.
- Credit Card Fraud Detection | Self Project:** Built an anomaly detection model with 99% precision over 250K transactions using resampling techniques. Visualized trends through correlation analysis and Matplotlib plots.
- RAG-based QA System | telegram-bot-akask:** Developed an reflexive agentic system using OpenAI and FAISS, deployed with FastAPI on GCP. Achieved an F1-score of 97% on student queries.