

Vivaan Jain

+91 9920520903 | vivaan.jain246@gmail.com | linkedin.com/vivaan | github.com/vivaan

EDUCATION

Dwarkadas Jivanlal Sanghvi College Of Engineering

Bachelor of Technology in Computer Engineering, Honor's in Data Science

CGPA : 9.4 / 10

Sep 2024 – May 2028

PROJECTS

MindVault | *LangChain, FAISS, NetworkX, FastAPI, Ollama, Next.js*

[Link](#)

- Engineered a production grade RAG system using LangChain LCEL with hybrid semantic and graph-based retrieval, reducing query latency by 60% by collapsing multi step LLM chains into single inference calls, featuring knowledge graph construction from LLM extracted typed entity relationships across 50+ concept nodes.
- Built a multi chain AI architecture with intelligent intent classification, context-aware query rewriting, cross session conversational memory, and mode-adaptive prompt engineering demonstrating end to end ownership of embedding pipelines, vector retrieval, knowledge representation, and local LLM optimization on constrained hardware.

AuditMind AI | *PyTorch, Whisper, SpeechBrain, SoundFile, NoiseReduce*

[Link](#)

- Engineered an end to end multilingual audio intelligence pipeline processing noisy field interviews across 13 Indian languages, leveraging Whisper large-v3 + SpeechBrain ECAPA-TDNN for ASR and speaker diarization delivering speaker labelled transcripts with zero paid API dependency.
- Architected a quantitative preprocessing evaluator testing 16 signal processing configurations (bandpass, noise reduction, RMS normalisation) scored via Whisper confidence metrics, identifying the optimal pipeline variant.

Multi Horizon Transformer for Systematic Equity Forecasting | *TensorFlow, Pandas, NumPy, Scikit learn*[Link](#)

- Achieved 0.62 AUC and 0.6+ non overlapping long short Sharpe by building a Transformer encoder from scratch with 60 day temporal self attention for multi horizon systematic equity direction forecasting, and integrating a fully vectorized long short backtesting engine to validate economic performance beyond drift.
- Engineered 16 quantitative technical factors and implemented leakage free scaling, walk forward validation, and decile based strategy evaluation to ensure statistical robustness and mitigate temporal bias using TensorFlow.

BLOGS AND ARTICLES

From Stills To Motion: AI Powered Video Creation Secrets | *150+ Impressions*

[Link](#)

- Authored a comprehensive technical blog that demystifies the process of AI powered video generation, explaining how generative models such as GANs and diffusion networks transform static images into dynamic video sequences and outlining prompt engineering strategies for high quality outputs.

Natural Language Processing (NLP): From Fundamentals to Modern Applications | *100+ Impressions*[Link](#)

- Authored a comprehensive technical blog explaining core NLP concepts and workflows, covering preprocessing, representation, training, evaluation, and deployment for real world applications.

EXTRA CURRICULAR

DJS ACM

Aug 2025 – Present

Research Mentee

- Authoring research on credit scoring under faculty mentorship, proposed solution during Ideathon(50+ audience) following upskilling sessions on GitHub, Hackathons and Paper Writing to strengthen overall analytical depth.
- Proposed and developing a novel, risk aware credit approval architecture by extending baseline credit scoring models with advanced feature engineering and systematic optimization, positioning the work as a research oriented contribution beyond traditional XgBoost approaches.

TECHNICAL SKILLS

Languages: Python, C/C++, MySQL, Java

Development Frameworks: Flask, Streamlit, Fast API

Libraries: LangGraph, LangChain, NetworkX, Ollama, Tensorflow, Keras, NLTK, SpaCy, Pandas, NumPy, Matplotlib, Scikit Learn

Tools: Docker, Git, Github, Jupyter, Linux, Excel, VS Code