

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

EXPERIENCE

Zeex AI

April 2026 – Present

Full Stack AI Intern

- Architected and built a production grade end to end audio intelligence pipeline on AWS integrating audio preprocessing, Automatic Speech Recognition (ASR), speaker diarization, and semantic question mapping to transform unstructured conversational audio into structured outputs at scale.
- Improved question mapping accuracy by 20% and enhanced transcription quality by 15% through preprocessing optimization, architecture level experimentation, speaker segmentation refinement, and integration of speech processing frameworks, while reducing overall pipeline latency by 30%.

PROJECTS

ApplyPilot AI | Langgraph, Kubernetes, PostgreSQL, FastAPI, Next.js

[Link](#)

- Engineered an autonomous AI Agent system that discovered relevant opportunities, generated personalized applications, and executed end to end application workflows, reducing manual effort by 90% and compressing application completion time from 1 hour to under 5 minutes per application.
- Spearheaded the design of a scalable multi-agent decision and execution pipeline that automated resume adaptation, application tracking, and workflow coordination, enabling users to process 100+ targeted applications with minimal intervention while improving consistency and throughput by 20x.

MindVault | LangChain, FAISS, RAGAS, NetworkX, Pydantic, Supabase, FastAPI, Next.js

[Link](#)

- Engineered a production grade hybrid retrieval pipeline semantic search, BM25 keyword matching, and FlashRank cross-encoder reranking fused via Reciprocal Rank Fusion eliminating single-method retrieval blind spots and slashing query latency 40% through parallel async retrieval and selective LLM call pruning, rigorously validated via automated RAGAS evaluation (faithfulness, context precision/recall) across 25+ test cases.
- Architected and shipped a multi-tenant RAG system (FastAPI, Next.js, Supabase/pgvector) serving real-time streaming responses, multimodal document/image ingestion, and shareable conversation links — diagnosing and resolving 6+ critical production incidents (memory leaks, race conditions, rate-limit cascades) to achieve zero-downtime reliability on constrained 512MB infrastructure.

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.

EXTRA CURRICULAR

DJS ACM

Aug 2025 – Present

Research Mentee

- Authoring research on credit scoring, 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, Next js

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

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