

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

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.

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

[Link](#)

- Engineered a private document intelligence system using LangChain LCEL and on-device LLMs that enables 100% secure querying of sensitive financial documents with zero cloud exposure, reducing manual retrieval time by 70% through hybrid semantic and graph based search across specialized RAG pipelines.
- Built a multi-chain architecture delivering source-attributed answers traceable to exact document sections, 60% faster query resolution via a combined classifier chain, and compliance-aware response generation directly applicable to contract analysis, regulatory search, and internal knowledge retrieval at scale.

Sentiment Driven Quantitative Alpha Modeling | Python, Pandas, NumPy, XGBoost, Scikit learn

[Link](#)

- Architected and validated a leakage controlled financial ML pipeline that fused historical price, volume, and API sourced news sentiment data to rigorously test whether sentiment provides incremental predictive signal beyond pure price based models.
- Quantified a 5 day return correlation improvement from 0.43 to 0.46 using sentiment features, supported by MAE and RMSE diagnostics and shuffle testing, revealing that gains were modest, unstable, and highly sensitive to validation design, highlighting the fragile nature of financial signals.

BLOGS AND ARTICLES

From Motion To Prompts: AI Powered Video Creation Secrets | 120+ 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, 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

ACHIEVEMENTS

- First Prize, Urban Empire: Built winning data driven city budget allocation strategy**