

SHIVAM JOSHI
Jersey City, NJ 07306
joshishivam047@gmail.com | (551) 260-8555 | [LinkedIn](#) | [GitHub](#)

Education

• Pace University, Seidenberg School of Computer Science and Information , New York, NY	May 2025
MS in Data Science	GPA: 3.82/4.00
• Mumbai University, BSE Institute Ltd , Mumbai, IN	June 2023
Post Graduate Diploma in Predictive Analytics (PGDPA)	GPA: 4.00/4.00
• Mumbai University, VES College Of Arts, Science & Commerce , Mumbai, IN	June 2023
Bachelors in Commerce (Banking & Insurance)	GPA: 3.12/4.00

Technical Skills

- **Programming Languages:** Python (NumPy, Pandas, SciPy), R, SQL
- **Data Science / Machine Learning / NLP:** Scikit-learn, TensorFlow, Keras, XGBoost, NLTK, SpaCy, OpenCV, OpenAI (API), Transformers
- **Data Visualization:** Power BI, Tableau, Excel (Advanced), Matplotlib, Seaborn, Plotly
- **Data Mining / Big Data / Cloud:** Apache Spark (PySpark), Hadoop, Google Cloud Platform, AWS (S3, EC2, Lambda)
- **Databases:** MySQL, PostgreSQL, Oracle, MongoDB, Firebase
- **Tools / Platforms:** Git, Docker, Jupyter, VS Code, GitHub Actions, REST APIs
- **Certifications:** Financial Markets (Yale), AI Agents in LangGraph (DeepLearning.AI), Prompt Design in Vertex AI (Google), Python Basic (HackerRank), DataBricks Fundamentals

Professional Experience

• 99 YARDS , New York, NY	Dec 2024 – Present
<i>Data Science Intern</i>	
– Developed and deployed a data-driven recommendation engine using TF-IDF, Firestore queries, and geolocation filtering; ensured data quality and accuracy for critical business insights across multiple data sources.	
– Enhanced geospatial intelligence by integrating geopy to prioritize recommendations based on ZIP code proximity, reducing irrelevant matches by 35%.	
– Designed a multi-tiered offer categorization logic using price banding, improving usability and clarity for B2B decision making.	
– Built an agentic AI-powered chatbot for the app using OpenAI APIs, enabling intelligent conversation flows and automated customer support.	
– Designed and fine-tuned a DistilBERT-based message blocker that detects and prevents contact-sharing attempts (phone, email, social handles) in chat, ensuring transactions remain on-platform; trained on 16K+ labeled messages for high accuracy.	
• Scienox Technologies , Mumbai, IN	Jan 2023 – Feb 2023
<i>Data Science Intern</i>	
– Conducted time series analysis on 500,000+ stock market records using Python (Pandas, Statsmodels), uncovering key trends and improving forecasting accuracy by 20%.	
– Engineered features and applied regression and classification models to enhance model stability, increasing reliability by 18% in investment risk analysis.	
– Experimented with Retrieval-Augmented Generation (RAG) techniques to enrich financial forecasting models with external economic indicators, improving context-aware insights for investment decisions.	
• Pipara & Co. LLP , Mumbai, IN	Jan 2022 – Mar 2022
<i>Data Analyst Intern</i>	
– Optimized data management processes for annual revenues exceeding INR 5000 Cr by implementing an analytics framework under IBC 2016, reducing decision-making time by 15%.	
– Identified compliance gaps and improved reporting workflows by 30%.	
• M/s Utangale & Co. Advocates & Solicitor , Mumbai, IN	Oct 2021 – Dec 2021
<i>Data Analyst Intern</i>	
– Analyzed and maintained financial records for 50+ clients, leveraging data interpretation skills to provide actionable insights and improve client satisfaction.	

Projects

• VinoSage: AI-Driven Wine Recommendation Chatbot	
– Built a chatbot using XGBoost, Elasticsearch, and FAISS with RAG and NLU workflows; streamlined personalized wine suggestions.	
– Fine-tuned DistilBERT for intent detection and BERT for NER to deliver accurate, context-aware recommendations.	
• AI Trip-Planner Agent	
– Developed an autonomous agentic travel planner using OpenAI + LangChain to suggest itineraries and book accommodations based on user preferences and constraints.	
– Implemented dynamic RAG pipelines to fetch real-time flight and hotel data, generating optimized trip proposals with user-specific recommendations, built for scalable deployment.	