

ARPITA THAKUR

Prayagraj , India

☎ 8445396329

✉ arpitathakurxib@gmail.com

in [Linkedin](#)

🐙 [Github](#)

ABOUT ME

Aspiring AI & Data Science Intern with strong foundations in Machine Learning, Deep Learning, and Generative AI. Skilled in Python, SQL, TensorFlow, PyTorch, and data visualization tools. Seeking an opportunity to apply technical and analytical skills to solve real-world problems, contribute to data-driven decision-making, and grow in a collaborative environment.

EDUCATION

Bachelor of Computer Applications

University of Allahabad | 2023-Present

PROJECTS

May | 2025

Stock Market Analysis and Forecasting | Python, RNN

- Conducted end-to-end predictive modeling on 5+ years of stock market OHLC data, focusing on time-series forecasting.
- Applied Exploratory Data Analysis (EDA) to identify patterns, seasonality, and correlations in stock data.
- Built and trained a Long Short-Term Memory (LSTM) network using PyTorch, achieving 89% prediction accuracy.
- Implemented hyperparameter tuning (learning rate, batch size, dropout rate) and 5-fold cross-validation to enhance generalization.
- Designed a data preprocessing pipeline (normalization, handling missing values, sliding window generation) that reduced preprocessing errors by 10%.
- Visualized predicted vs. actual prices using Matplotlib & Seaborn, providing clear insights into model performance.

June | 2025

SQLGenie: Natural Language to SQL Query Generator

- Designed and built an AI-powered interface that enables business users to query relational databases using plain English commands.
- Utilized Gemini 1.5 Flash and SQLCoder to translate natural language into optimized SQL queries.
- Reduced dependency on technical teams by 80%, improving turnaround time for data-driven business queries..
- Stored query logs in MySQL to analyze user behavior and improve system accuracy over time.

TECHNICAL SKILLS with Achievements & Extras

- Programming: Python, SQL
- ML/DL Frameworks: TensorFlow, PyTorch, Scikit-learn, OpenCV, YOLO
- Specialized Skills: Data Preprocessing, Feature Engineering, EDA, Model Training & Evaluation, Hyperparameter Tuning, NLP (LLMs), Model Deployment (Streamlit, Flask, FastAPI)
- Visualization Tools: Power BI, Tableau, Matplotlib, Seaborn
- Databases: MySQL, PostgreSQL
- Cloud Exposure: AWS (S3, SageMaker), GCP (Vertex AI)
- Completed multiple data visualization dashboards (Tableau/Power BI) for real-world datasets.
- Participated in Kaggle data science competitions and university hackathons.
- Strong foundation in Statistics & Probability for Data Science.