Vaidik Yadav

Data Scientist | ML & GenAI | MLOps | Python | LLMs
Phone : +91 7974476349 | **Email** : vaidiky90@gmail.com
LinkedIn : Vaidik Yadav | **GitHub** : Vaidik26 | **Portfolio :** Link
Jabalpur, Madhya Pradesh

SUMMARY
Data Scientist with 1.5 years of analytics experience delivering ML, GenAI, and NLP solutions that enhance decision-making and operational efficiency. Skilled in building and deploying production-ready models with Dockerized pipelines, REST APIs, and vector search. Experienced in LLMs, MLOps, and cloud deployment. Proven record in model optimization for speed & accuracy and enabling real-time inference at scale.

SKILLS

Programming: Python, SQL
ML/DL/NLP: Scikit-learn, CatBoost, XGBoost, LightGBM, Random Forest, Gradient Boosting, TensorFlow, LSTM, Transformers, Hugging Face, LangChain, OpenAI APIs
Data Processing: Pandas, NumPy, Feature Engineering, EDA, Hypothesis Testing, Time Series Analysis, SMOTE, Time Series Forecasting, A/B Testing, PCA
MLOps & Deployment: Flask, FastAPI, Docker, DVC, GitHub Actions, MLflow, CI/CD
Cloud & Databases: AWS (EC2, S3, SageMaker), MySQL, MongoDB, Pinecone, FAISS
Visualization & BI: Matplotlib, Seaborn, Plotly, Power BI, Excel

---

Professional EXPERIENCE

Platinus Technology Pvt Limited — Data Analyst
Hyderabad, India | Sep 2022 – Aug 2024

- Increased client placement visibility by 40% using Excel & Power BI dashboards.
- Improved placement rates by 20% through visa data analysis & trend-based targeting.
- Reduced job data retrieval time by 30% via database optimization.
- Developed automated reporting templates, saving 15+ hours/month for the team.
- Collaborated with business teams to identify data-driven strategies improving team performance by 15%.

EDUCATION

Gyan Ganga College Of Technology
B.Tech – Computer Science | Jan 2021 – Jan 2025
Jabalpur, Madhya Pradesh

PROJECTS

Medical Chatbot | GitHub Repo : Medical Chatbot

Tools: LangChain, LLMs, RAG, Pinecone, Flask, Hugging Face

- Achieved 90% accuracy on 200+ medical FAQs using a RAG-based architecture.
- Served 100+ concurrent requests with <1.5s latency via Flask REST APIs and prompt routing.
- Embedded 10,000+ medical text chunks for semantic search, enabling precise context retrieval.

Stock Price Prediction using LSTM | GitHub Repo : Stock Price Prediction

Tools: TensorFlow, Pandas, Matplotlib

- Improved prediction stability by 18% via time-series preprocessing & sliding windows.
- Used walk-forward validation to prevent temporal leakage and improve robustness.
- Visualized predicted vs. actual prices for the last 100 trading days.

Milk Quality Classification | GitHub Repo : Milk Quality

Tools: Python, Scikit-learn, Docker, DVC, MongoDB

- Delivered 89% accurate multiclass model using Gradient Boosting & RandomizedSearchCV.
- Built fully reproducible ML pipeline with DVC and Docker for versioned deployment.
- Automated preprocessing with Label Encoding & missing value handling.

CERTIFICATIONS & COURSES

- PW Skills Data Analytics
- 5 Minutes Engineering Data Science Noob To ProMax Batch 5