

Vaidik Yadav | Data Scientist

+91 7974476349 | vaidiky90@gmail.com

 [Vaidik Yadav](#) |  [Vaidik26](#) | [Portfolio](#)

Jabalpur, Madhya Pradesh - 482001, India

OBJECTIVE

Data Analyst with 1.5 years of experience driving insights through data cleaning, visualization, and predictive modeling. Skilled in ML, NLP, and deployment, with hands-on projects in AI and end-to-end solutions.

SKILLS

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




EXPERIENCE

- **Platinus Technology Pvt Limited** September 2022 – August 2024
Data Analyst Hyderabad, India
 - Designed and deployed interactive Power BI and Excel dashboards that increased client placement visibility by 40%, providing real-time insights into recruitment metrics.
 - Improved placement rates by 20% by analyzing visa and market data for trend-based targeting.
 - Reduced job data retrieval time by 30% with database indexing and SQL query optimization.
 - Automated monthly reporting templates, saving 15+ hours of manual work.
 - Collaborated with cross-functional teams to implement data strategies, improving performance by 15%.

EDUCATION

Gyan Ganga College Of Technology
Bachelor of Technology (B.Tech) – Computer Science
Jabalpur, Madhya Pradesh
2021 – 2025

PROJECTS

- **Document QnA (LLM-powered Q&A System)**  [GitHub Repository](#)
Tools: LangChain, LlamaIndex, Hugging Face, FAISS, Streamlit
 - Built an intelligent document QnA system enabling context-aware answers from PDFs.
 - Integrated RAG with FAISS embeddings for semantic retrieval, improving accuracy by 25%.
 - Deployed interactive Streamlit UI for real-time document search and QnA.
- **Legal Advisory Chatbot – AI-powered Legal Assistance**  [GitHub Repository](#)
Tools: LlamaIndex, RAG, Hugging Face, Flask
 - Developed GenAI chatbot for legal document assistance with 85%+ retrieval precision.
 - Implemented contextual retrieval using LlamaIndex vector store for legal text embeddings.
 - Served 100+ concurrent queries with <2s latency via Flask API.
- **Milk Quality Classification**  [GitHub Repository](#)
Tools: Python, Scikit-learn, Docker, DVC, MongoDB
 - Achieved 89% accuracy on multiclass classification using Gradient Boosting RandomizedSearchCV.
 - Built reproducible ML pipeline with DVC and Docker for versioned deployment.
 - Automated preprocessing with label encoding and missing value handling for scalability.

CERTIFICATIONS & COURSES

- [PW Skills Data Analytics](#)
- [5 Minutes Engineering Data Science Noob To ProMax Batch 5](#)