

# RAHUL RAI

Sahadatpura, Maunath Bhanjan, Uttar Pradesh - 275101 [rahulraimau5@gmail.com](mailto:rahulraimau5@gmail.com) | +91 95600 36286

GitHub: [github.com/rahulraimau](https://github.com/rahulraimau) | LinkedIn: [linkedin.com/in/rahul-rai-629554245](https://linkedin.com/in/rahul-rai-629554245)

## Professional Summary

Aspiring Data Analyst with hands-on experience in data wrangling, statistical analysis, machine learning, and dashboard development. Proficient in Python, SQL, and visualization tools. Experienced in predictive modeling, computer vision, and data storytelling. Eager to contribute to real-world, data-driven decision-making. **Technical Skills**

- **Programming & Tools:** Python, SQL, Git, Excel, Streamlit, OpenCV
- **Data Analysis & Visualization:** Pandas, NumPy, Seaborn, Matplotlib, Tableau, Plotly, Excel PivotTables
- **Machine Learning:** Regression, Classification, Clustering, XGBoost, SMOTE, KMeans, Random Forest
- **Time Series Forecasting:** ARIMA, Prophet, Moving Averages
- **Computer Vision:** CNN, YOLOv11, ResNet50, Cosine Similarity
- **NLP:** Text Cleaning, TF-IDF, Sentiment Analysis, Topic Modeling (LDA/NMF)
- **Deployment:** Streamlit, Jupyter, GitHub Pages
- **Other:** EDA, Feature Engineering, Hypothesis Testing, GitHub Collaboration

## Projects

- **Fraud Transaction Detection:** Developed a fraud detection pipeline using XGBoost and SMOTE on imbalanced data (1.7M records). Achieved 94% ROC-AUC with advanced feature engineering.
- **Fish Image Classification (Computer Vision):** Built a multi-class fish classifier using YOLOv11 and ResNet50 on a custom dataset. Achieved high accuracy with cosine similarity-based validation.
- **E-commerce Segmentation Dashboard:** Performed RFM segmentation and visualized clusters using Tableau and KMeans. Improved marketing targeting using CLV analysis.
- **Electricity Demand Forecasting:** Forecasted energy demand using ARIMA and Prophet. Compared residuals and seasonal trends over time.
- **Customer Complaint Dashboard (Excel):** Created an Excel dashboard using PivotTables and VLOOKUP to analyze resolution trends by region.

## Professional Experience

**HSIL Ltd., Bahadurgarh Maintenance Engineer**

Aug 2013 - Sep 2016

Plant Engineering Department

- Reduced downtime by 15% via root cause analysis.
- Interpreted maintenance data and led reliability improvements.

## Internships

**InnovateClouds Data Science Intern**

Jun 2025-Jul 2025

Worked on data preprocessing, model building, and deployment workflows in cloud environments.

**LabMentix AI/ML Intern**

Apr 2025-May 2025

Participated in AI model experimentation and research for healthcare-related data solutions.

**NullClasses Data Analyst Intern**

Feb 2025-Mar 2025

Conducted EDA, visualization, and predictive modeling for business datasets.

**UnifiedMentor Data Science Intern**

Aug 2025-Sep 2025

Supported development of dashboards and analytics tools for student performance tracking. Worked on projects including Animal Classifications, Forest Cover Prediction, Lung Cancer, and ASL Detection.

## Education

Data Science 360 Certification Program, AnalytixLabs, Noida (2022- 2023)

Full-time UPSC Preparation (2016- 2020)

B.Tech in Mechanical Engineering, GLA Institute of Technology, Mathura (2009 – 2013)

## Additional Strengths

Strong analytical thinking

Effective communicator

Team player

Fast learner

Self-motivated