

Rohit Kumar

rohitmehra.ggh@gmail.com — +91 9893983586 — [LinkedIn](#) — [GitHub](#) — [Portfolio](#)
Jabalpur, Madhya Pradesh, India

Professional Summary

Analytical and result-driven Data Science professional with hands-on experience in machine learning, NLP, and predictive analytics. Skilled in developing and deploying scalable AI/ML models, data pipelines, and dashboards that enhance data-driven decision-making.

Technical Skills

Languages: Python, SQL, JavaScript, HTML, CSS

Libraries/Tools: Pandas, NumPy, Matplotlib, Seaborn, Scikit-learn, Streamlit, TensorFlow

ML/AI: Regression, Classification, Transformers (T5, BERT), Text Summarization, Chatbots

Other: Data Cleaning, APIs, Git, Cloud, Linux, Model Deployment

Education

B.Tech, Information Technology
Jabalpur Engineering College, India

Oct 2021 – Mar 2025
CGPA: 7.08 / 10

Professional Experience

Data Science Intern — Sarvarth

May 2025 – sept 2025

- Developed and deployed ML models, improving reporting accuracy by 15%.
- Built automated data pipelines reducing manual effort by 30%.
- Designed interactive Streamlit dashboards serving 100+ users for real-time insights.

AI Development Trainee — CEDMAP MSME

Jul 2025 – Sept 2025

- Implemented transformer-based NLP models using Hugging Face.
- Integrated vector databases and orchestration tools, cutting response latency by 20%.

Key Projects

Transformer Text Summarizer — [GitHub](#)

- Built an NLP model using **T5 Transformer** for automated text summarization tasks.
- Fine-tuned on custom datasets with **Hugging Face Transformers** and **PyTorch**.
- Implemented **beam search decoding** for high-quality output generation.
- Deployed a user-friendly **Streamlit** app enabling real-time document summarization.
- Reduced manual content review time by **40%** through efficient model inference.

RAG-based AI Chatbot — [GitHub](#)

- Developed a **Retrieval-Augmented Generation (RAG)** chatbot using **LangChain** and **FAISS**.
- Implemented **OpenAI embeddings** for semantic search and contextual query understanding.
- Enabled multi-turn conversational memory with precise PDF data retrieval.
- Deployed using **Streamlit**, providing interactive responses with minimal latency.
- Enhanced contextual accuracy by **30%** via optimized document chunking and vector search.

Student Data Analytics Dashboard — [GitHub](#)

- Designed a **data analytics dashboard** using **Pandas**, **Plotly**, and **Streamlit**.
- Processed large student datasets from **CSV/SQL** sources with automated cleaning pipelines.
- Visualized metrics such as attendance trends, grade distribution, and performance outliers.
- Enabled academic teams to make **data-driven decisions**, improving insight efficiency by **18%**.

Certifications

Cloud Computing Bootcamp (Yhills)
Django Web Development (Udemy)

Leadership & Activities

Branch Representative, TEDx JEC (2024): Led outreach initiatives impacting 300+ students.

Sports Coordinator, Aavahan Fest (2024): Managed 250+ event participants and event logistics.