

MOHIT SHARMA

✉ mohitsharma782000@gmail.com

☎ 9079032677 📍 15 Shiv Colony Tijara Fatak 📅 07/08/2000

🐙 github.com/Moh1235

🌐 [mohit-sharma-985a8b226/](https://www.linkedin.com/in/mohit-sharma-985a8b226/)

PROFILE

Detail-oriented and certified Data Science with expertise in Python, R, SQL, and end-to-end machine learning workflows. Proven projects knowledge in data pipelines, building predictive models, and deploying analytical dashboards. Skilled in both structured and unstructured data analysis with an aptitude for problem-solving and delivering actionable business insights. Passionate about leveraging data to optimize decision-making and drive innovation in a collaborative environment.

CORE SKILLS

Languages & Libraries: Python (Pandas, NumPy, Scikit-learn, TensorFlow, Keras), R, SQL

Data Analytics: EDA, Feature Engineering, Data Preprocessing, Statistical Analysis, Hypothesis Testing

Data Science: Machine Learning, Deep Learning (CNN, ANN, RNN), NLP, Transformers, Neural Networks, Fast-API, AWS, Django, LSTM

GEN-AI: LANG-CHAIN, OPEN-AI, GROQ-API, Embedding Systems, RAG, LLM, Fast Api,

Data Visualization: Power-BI, Tableau

CERTIFICATES

- **Complete Data Science, Machine Learning, Deep Learning & NLP Bootcamp** *Issued: Mar 2025* — Certificate [🔗](#)
- **Google Data Analytics Professional Certificate** — Coursera *Issued: Sep 2024* — Certificate [🔗](#)
- **Advanced Data Analytics Certification** — IRA Edu-Tech *Issued: Oct 2023* — Certificate [🔗](#)

PROJECTS

Theme-Identifying Document Analysis Chatbot (Wasser Stoff Generative AI Internship) [🔗](#)

- Developed an AI-powered chatbot using Fast API and Lang Chain to analyse PDFs/images for regulatory compliance themes.
- Integrated Groq API with LLaMA3-8B model for natural language processing and thematic summaries.
- Built vector search with ChromaDB and Hugging Face embeddings to enable semantic query responses.
- Implemented OCR pipeline with PyMuPDF and tesseract for text extraction from documents.

Kidney Disease Detection using CNN [🔗](#)

- Designed a convolutional neural network (CNN) with TensorFlow to identify kidney diseases from medical imaging data.
- Achieved high classification accuracy by fine-tuning CNN layers and applying data augmentation techniques.
- Visualized activation maps to interpret model predictions and detect key indicators.

Movie Recommendation System (Content-Based Filtering) [🔗](#)

- Conducted data preprocessing on The Movie Database (TMDb) dataset, cleaning and structuring movie attributes like genres, descriptions, and ratings for analysis and modelling.
- Utilized Python with Pandas and Scikit-learn to implement content-based filtering, leveraging TF-IDF vectorization and cosine similarity to recommend movies based on user preferences.

EDUCATION

BECOLOR OF TECHNOLOGY

Modern Institute of Technology & Research Center
Rajasthan Technical University, Kota

08/2017 – 09/2021

Alwar, India

DECLARATION

I hereby declare that the given above information are true to the best of my knowledge and belief and can be supported with reliable documents when needed.