Rishabh Prajapati

+91-7860634479 | prajapatirishabh071@gmail.com
GitHub | Portfolio | LinkedIn

Education

Degree/Certificate	Institute/Board	CGPA/Percentage	Year
Bachelor's Degree	Indian Institute of Technology, Patna	CPI: 8.33	2023–2027

Experience

- **Data Science Intern — GUVI HCL** *Jun. 2025 – Sep. 2025*
  Data Science Intern Online
  - Built and analyzed end-to-end data science projects including **Titanic EDA**, **handwritten digit recognition**, and **movie review sentiment analysis**, applying data cleaning, feature engineering, and model evaluation techniques.
  - Implemented machine learning, deep learning, and NLP models using **Python**, **Pandas**, **Scikit-learn**, and **TensorFlow**, gaining hands-on experience in classification, model tuning, and data visualization.

Projects

- **Medical ChatBot (RAG-based Healthcare Assistant) (GitHub)**
  - Developed an AI-powered medical chatbot using RAG architecture, Gemini LLMs, and LangChain to deliver reliable healthcare insights. Integrated Pinecone for semantic search and Flask for a web-based interface.
  - **Technologies:** RAG, Gemini LLMs, LangChain, LangSmith, Pinecone, Flask, Python, HTML/CSS, JavaScript
- **Duplicate Question Detection using NLP & Machine Learning (GitHub)**
  - Developed a model to identify semantically similar or duplicate questions using NLP preprocessing, fuzzy matching, and machine learning. Built a pipeline for text cleaning, tokenization, and classification using RandomForestClassifier with strong accuracy and visualization through Matplotlib and Seaborn.
  - **Technologies:** NLP, BeautifulSoup, FuzzyWuzzy, CountVectorizer, RandomForestClassifier, Python, Seaborn, Matplotlib
- **Brain Tumor Classification using VGG16 & Transfer Learning (GitHub)**
  - Developed an end-to-end deep learning model for brain tumor detection using VGG16 and transfer learning. Implemented a complete pipeline including data preprocessing, model fine-tuning, training, and evaluation with MLflow tracking.
  - **Technologies:** TensorFlow, Keras, VGG16, Transfer Learning, MLflow, NumPy, Pandas, Matplotlib, Python
- **Car Price Prediction using Machine Learning (GitHub)**
  - Built a predictive model to estimate car prices based on key factors such as model, company, and fuel type using Linear Regression and Pipeline techniques. Conducted data preprocessing, feature analysis, and model tuning for improved accuracy. Deployed an interactive Streamlit web interface for real-time price prediction.
  - **Technologies:** Python, Pandas, NumPy, Scikit-learn, Matplotlib, Streamlit

Technical Skills

- **Programming:** Python, SQL
- **Machine Learning & AI:** Supervised & Unsupervised Learning, Deep Learning (ANN, CNN, RNN, LSTM), NLP, Retrieval-Augmented Generation (RAG), Generative AI, Google Gemini API
- **Libraries & Frameworks:** NumPy, Pandas, Scikit-learn, TensorFlow, PyTorch, Keras, LangChain, Pinecone, LangGraph, LangSmith
- **Data Analysis & Visualization:** EDA, Feature Engineering, Data Cleaning, Statistical Analysis, Hypothesis Testing, Matplotlib, Seaborn
- **Tools & Platforms:** Git, GitHub, MLflow, Jupyter Notebook, Google Colab, Hugging Face, Microsoft Excel
- **Deployment & Web:** Streamlit, Flask, HTML, CSS, JavaScript

Positions of Responsibility

- **Member,** Technology Club under the Student Technical Council, IIT Patna *Mar. 2025 – Dec. 2025*
  - **As part of the AIgnition Hackathon, contributed to the evaluation of participants' projects in the domains of Artificial Intelligence, Machine Learning, and Deep Learning, focusing on innovation, technical implementation, and real-world impact.**

Certifications

- Data Science – Guvi HCL (June 2025)
- Pytorch Ultimate 2024: From Basics to Cutting-Edge – Udemy (Dec. 2024)
- Data Visualization – Kaggle (Jan. 2025)
- Pandas – Kaggle (Dec. 2024)