Rohit Yadav

(+91) 6394254022 | rohiityadav6394@gmail.com
[LinkedIn](#)

Education

- **Indian Institute of Technology, Roorkee** ⟶ 2021 – 2025
  B.Tech, Chemical Engineering

Skills

Machine Learning and AI: Machine Learning , Deep Learning , Natural Language Processing (NLP), Large Language Models (LLMs), Retrieval-Augmented Generation (RAG)
Programming and Data Handling: Python, MySQL, Pandas, NumPy, Data Cleaning
Model Deployment - Tools: TensorFlow, PyTorch, Scikit-learn, Hugging Face
Visualization Tools: Power BI, Microsoft Excel, Seaborn ,Matplotlib

Experience

- **Data Analyst Intern | PhysicsWallah** ⟶ November 2023 - February 2024
  - Developed a classification model using Random Forest achieving a 0.76 F1 score in forecasting customer conversion
  - Identified that preview interaction, time spent on site, and clicking on instructor profiles drive 65%+ conversions
  - Recommended enhancing SEO and content marketing efforts to address the issue of 90% of users missing ads

- **Data Science Intern | KhetAI** ⟶ May 2024 - August 2024
  - Developed LLM-based chatbot offering precise answers to farming questions for building Agritech startup KhetAI
  - Devised a RAG system to handle queries on over 500 topics, ranging from crop selection to disease management
  - Implemented LLM with Qdrant vector database to dynamically select relevant data based on farmer queries
  - Integrated and designed the chatbot feature in KhetAI App, receiving 90% positive feedback in the pilot run

Projects

– LangChain-Based PDF Chatbot ⟶ March 2025 - April 2025
  - Built a smart chatbot to query multiple PDFs using LangChain and open-source embedding techniques
  - Utilized Instructor and Sentence Transformers for chunk-level semantic retrieval from uploaded document content
  - Created a Streamlit web application enabling real-time PDF-based question answering with memory support

– AI/LLM Generated Text Detector ⟶ December 2023 - January 2024
  - Fine-tuned a DistilBERT model on 14,000 labeled text samples for robust and accurate AI text detection tasks
  - Deployed a Streamlit web application using Hugging Face Transformers for real-time content classification
  - Implemented sentence-level NLP analysis achieving high accuracy distinguishing AI from human written text

– Automated Image Captioner ⟶ May 2023 - June 2023
  - Constructed a CNN-LSTM model with Pytorch using 8k images for automated image captioning
  - Utilized VGG16 for feature extraction and generated comprehensive textual descriptions from the image dataset
  - Converted captions to audio and braille using text-to-speech, improving accessibility for visually impaired users

Leadership & Extracurriculars

– Senior Executive Member | Student Technical Council, IITR ⟶ August 2023 - May 2024
  - Organized TechShila event with 80+ teams; led hackathons and workshops promoting applied ML and AI adoption
  - Managed STC's social media sites, enhancing outreach and engagement; produced Inter-Bhawan Tech Meet video

Additional Coursework

Completed IBM Data Analyst | Machine Learning Specialization course on Coursera