DEEP RABADIYA

New Delhi | 9510108275 | deeprabadiya8@gmail.com | Linkedin | Github

About

I am a passionate Data Scientist with strong expertise in Python, Machine Learning, Deep Learning, and Data Analytics. I specialize in transforming complex data into actionable insights through advanced modeling, automation, and visualization. With hands-on experience in SQL, Excel, LangChain, I build scalable data-driven solutions and intelligent systems. I'm particularly interested in applying AI and data science techniques to solve real-world challenges across diverse domains. Eager to collaborate on innovative projects that bridge data, technology, and decision-making for meaningful impact.

Technical Skills

Programming Languages: Python, R, SQL
Frameworks and Libraries: Numpy, Pandas, TensorFlow, Keras, Scikit-Learn, pytorch, OpenCV, Langchain, Streamlit
Technologies:AutoML
Data Visualization: Excel, Power BI, Tableau, Matplotlib, Seaborn
Area of Interest: Statistics, Machine Learning, Deep Learning, Data Visualization, NLP, Web Scrapping, Predictive Analysis, Generative AI, Large Languuage Models, RAG
Miscellaneous: Git and GitHub, Fine-tuning, Prompt-tuning, Object-Oriented Programming (OOPS), JavaScript, Time Series Analysis
Soft Skills: Project Management, Cross-functional Collaboration, Problem-solving, Effective Communication
Languages: English, Hindi, Gujarati

Experience

Data Analyst(Ass. Manager) Jan 2025 - Present
SBI General Insurance *New Delhi*

- Built an ML pipeline to classify district-level rainfall anomalies (Deficit, Normal, Excess) based on historical IMD and ERA5 data.
- Developed AI-driven crop classification models using AutoML and Deep Learning with spectral signature analysis and Achieved **87% kappa coefficient.**
- Implemented semantic segmentation models (U-Net, YOLO) for field boundary detection, improving mapping accuracy by 50%.
- Built machine learning Stacking model for crop yield forecasting.
- Designed workflows for weather data extraction, transformation, and integration from APIs and gridded datasets to support crop modeling and risk assessment.

Data Analyst May 2024 – July 2024
Semantic Technologies *Pune, Maharashtra*

- **Hybrid AI Model for Oncology Drug Screening**
  - Developed Hybrid classification model for cancer compounds using fragmented data augmentation techniques.
  - Achieved **80% accuracy** on dataset which has Millions of chemical compounds, enabling more efficient drug discovery.
  - Collaborated with Charité University scientists (Berlin) to validate model performance for real-world cancer research applications.

Machine Learning Intern March 2024 – April 2024
Coding Samurai *Remote*

Machine Learning Intern March 2024 – March 2024
Mentorness *Remote*

Education

Dhirubhai Ambani University (DAIICT) Gandhinagar, GJ
Master in Data Analytics (8.2/10) *2023 – 2025*

Projects

Sentiment Analysis and Visualization of Tech Product Reviews

- Scraped Amazon reviews for smartphones, laptops, smartwatches, and earbuds. Applied Hugging Face sentiment analysis, generated model descriptions, and summarized sentiments using LLMs. Created an interactive Power BI dashboard to visualize insights, showcasing data analysis and communication skills.

Chatbot Platform for SQL Queries and PDF Q&A

- Developed a Streamlit application featuring two chatbots: one for SQL querying that connects to a MySQL database and another for Q&A from uploaded PDFs. Integrated multiple LLMs such as Llama, Gemini, and Mistral AI, allowing users to select their preferred model.

Fraud Detection using SQL and Python

- Designed and implemented an end-to-end fraud detection system, leveraging SQL for data extraction and transformation and Python for machine learning-based anomaly detection. Built predictive models to analyze transaction patterns, detect fraudulent activities, and enhance financial security. Demonstrated expertise in big data handling, time-series forecasting, and advanced analytics.

Price Forecasting for Tomato using Various Statistical Models

- Forecasted tomato prices using 14 years of historical data on prices and arrivals. Applied statistical models in Python to demonstrate the effectiveness of accurate market trend prediction.

Certifications

Advance Excel, Power BI and Tableau | *The Pioneer Tech*

Generative AI with Langchain and Huggingface | *Udemy*

Python for Computer Vision with OpenCV and Deep Learning | *Udemy*