Shakil Ansari

New Delhi, India | ansarishakeel006@gmail.com | +91-9650118976 | Portfolio | LinkedIn | GitHub | Kaggle

About

Data Scientist specializing in Generative AI (GenAI), LLMs, and end-to-end ML solutions with expertise in NLP, time series forecasting, and deep learning. Skilled in transformer fine-tuning, prompt engineering, and deploying AI solutions with a systems-thinking approach from a Mechatronics and Life Sciences background.

Experience

Data Science - Intern, First Quadrant Labs, [Remote]   **July 2025 – Sept 2025**
- Developed ML models for fuel efficiency prediction (target MAE < 1.5) and stellar object classification (target accuracy > 90%) using Python, scikit-learn, and TensorFlow. Created data visualizations and reports using Power BI and Excel to communicate insights and support model evaluation.

Educator, Vidhyanjali Rising Point, Delhi   **June 2019 – Present**
- Developed structured learning modules aligned with CBSE to enhance analytical thinking and performance.

Research Trainee, Abyom SpaceTech and defense Pvt. Ltd.   **June 2022 – Aug 2022**
- Performed collection and analysis of research papers and Planetary data on the topic of research

Education

Guru Gobind Singh Indraprastha University, B-Tech [Mechatronics Engineering]   **Oct 2020 – June 2024**
- GPA: 8.0/10.0

Indira Gandhi National Open University, B.Sc - Life Sciences [Botany]   **July 2018 – Sept 2021**
- GPA: 6.0/10.0

Certifications

Diploma in Data Science – Boston Institute of Analytics (BIA), 2025

IBM Data Science Professional Certificate (Coursera, 2025)

Projects

Abstractive Dialogue Summarization using Transformers   [Try Live Demo]
- Fine-tuned FLAN-T5 on the DialogSum dataset to generate abstractive summaries from multi-turn dialogues using prompt-based sequence modeling.

Domain-Specific NLP: Indian Recipe Generation   [Try Live Demo]
- Developed a prompt-based NLP model for Indian recipe generation, mastering dataset preparation and fine-tuning for domain adaptation and text generation.

Fake News Detection (GloVe + LSTM)   [Try Live Demo]
- Built a fake news classifier using GloVe embeddings and an LSTM network, achieving 95% accuracy.

Movie Recommendation (IMDB)   [Try Live Demo]
- Built a collaborative filtering system for personalized suggestions.

Stock Forecasting (ARIMA/SARIMA):   [Try Live Demo]
- Predicted stock trends using time series models with seasonal analysis.

Stellar Object Classifier   [Try Live Demo]
- Classified celestial objects using supervised ML models and neural networks on SDSS-17 dataset, accuracy-97%.

Technologies

Programming Languages: Python, R, SQL, HTML, CSS, JavaScript

Technologies: Jupyter Notebook, MySQL, Git, Power BI, Tableau, Excel, Kaggle, Hugging Face, GenAI, Cursor, Docker, Amazon Sagemaker

Data Science: Data Analysis, Machine Learning, Deep Learning[ANN, RNN, CNN], NLP, Time Series, Prompt Engineering, Statistical Analysis, Python(scikit-learn, pandas, seaborn, matplotlib, PyTorch, TensorFlow etc.), DSA