Alok Kumar

GitHub | LinkedIn | LeetCode | Kaggle | ay747283@gmail.com | +91 80901 75358

Summary

Data science and machine learning enthusiast with a strong foundation in statistical analysis and algorithm development. Passionate about leveraging data-driven insights to solve complex problems and enhance decision-making. Committed to continuous learning and eager to contribute to innovative projects that drive meaningful impact.

Education

Bachelor's Degree at ABV-IIITM — Gwalior
CGPA 7.5 — *2023*

Class 12th CBSE — 82%
2022

Class 10th CBSE — 91.2%
2020

Skills

Data Analysis: Python, Pandas, NumPy
Machine Learning: Scikit-learn, TensorFlow, Keras, Pytorch
Data Visualization: Matplotlib, Seaborn
Computer Vision: OpenCV, CVZone, YOLO
Natural Language Processing: Transformers, spaCy, NLTK, Gensim
Generative AI: LangChain, Hugging Face, LangGraph, Crewai
Databases: SQL, FAISS, Pinecone, Chroma, Qdrant
Web Frameworks: FastAPI, Flask
Tools: Git, GitHub, Jupyter Notebook, Docker, MLflow, DVC

Projects

LinkedIn Profile Parser — Github Link
Agentic AI system for automated candidate discovery and evaluation based on job descriptions.

- Built intelligent agent using LangChain and LangGraph to automate candidate sourcing and evaluation.
- Integrated Tavily API for LinkedIn searches and RapidAPI for detailed profile analysis.
- Developed autonomous decision-making to identify and rank best-fit candidates for specific roles.

Medical-Chatbot — Github Link
Retrieval-augmented generation chatbot using Python, LangChain, and Gemini AI with Pinecone vector database.

- Integrated LangChain for efficient document retrieval and response generation.
- Leveraged Gemini API for accurate predictions and enhanced context handling.
- Designed to address medical-related queries with precise and reliable answers.

Youtube Chrome-Plugin — Github Link
YouTube Comment Insights & Sentiment Analyzer Using Machine Learning

- Designed module to fetch YouTube comments and analyze sentiment with interactive charts.
- Employed ML algorithms with DagsHub and DVC for efficient data pipeline management.
- Integrated MLflow for monitoring and model registry, enhancing model management capabilities.

Achievements

1st Position, ISEC Data Challenge, Kaggle — Kaggle Leaderboard
Achieved 1st place among 28 teams.

3rd Position, Exploring Predictive Health Factors, Kaggle — Kaggle Leaderboard
Secured 3rd place among 139 teams.