AASHISH KUMAR SAINI

Data Scientist | Data Analyst | ML Engineer | Generative AI Engineer

aashishkumarsaini2003@gmail.com | +918764954136 | Alwar, Rajasthan

Linkedin | GitHub | HackerRank

Data Scientist skilled in Machine Learning, Data Analysis, and Generative AI, with hands-on experience developing intelligent applications using LangChain, LangGraph, and RAG pipelines.

EDUCATION

Laxmi Devi Institute of Engineering and Technology   Alwar, Rajasthan, India
Bachelor of Technology (B.Tech) in Computer Science and Engineering   August 2022 - July 2026
CGPA: 9.54 / 10

EXPERIENCE

Suvidha Foundation | Artificial Intelligence Intern   Maharashtra | April 2025 - May 2025
   Fin Advisor: Personalized Financial Advice Generator (AI Intern Solo Project)
- Built an end-to-end NLP system to generate personalized investment advice from real-time financial news.
- Implemented data pipeline: news scraping (RSS + Feedparser), preprocessing (pandas), NER (spaCy), sector & ticker mapping(yfinance), sentiment analysis (FinBERT).
- Created embeddings (Sentence Transformers) and stored in ChromaDB for semantic search and user profile matching.
- Applied Retrieval-Augmented Generation (RAG) with LLMs to deliver tailored financial insights based on user profiles.

SKILLS

Programming Languages:	Python, C++
Libraries/Frameworks:	LangGraph, MCP, CrewAI, LangChain, LangSmith, TensorFlow, HuggingFace Hub, Scikit-learn, Pandas, Numpy, Seaborn, Matplotlib, BeautifulSoup, Docker, FastAPI
Tools / Platforms:	n8n, MLflow, Excel, Git, Power BI, Jupyter Notebook, Google Collab, Streamlit, Microsoft Word
Databases:	SQL, MongoDB, ChromaDB, FAISS

PROJECTS

Smart DB Agent | Link   Python, Langgraph, Langchain, Streamlit, MySQL

1. A Smart DB Agent that connects seamlessly with hosted databases, users can execute natural language queries, it will return the Structured Result and SQL query used to get results.

2. Built with LangGraph, LangChain and Streamlit, it simplifies database interaction with intelligent automation. (Note: Here, I have used Super Store Dataset as Database).

InsightGenie | Link   Python, Langchain, Streamlit, LLM, Vector Database

1. RAG-based Question Answering System that supports multi-document input (TXT, PDF, DOCX) for contextual Q&A.
2. Includes additional features: text summarization and language translation.
3. Built with Streamlit and provides a clean, user-friendly interface with easy navigation and file handling.

Crop Recommendation System | Link   Python (Numpy, Pandas, Scikit-learn), Flask, HTML, CSS

1. Developed a machine learning system to recommend optimal crops based on soil, weather, and environmental conditions using Python, Pandas, and Scikit-learn.
2. Ensured accurate, data-driven recommendations to assist farmers in decision-making.

AWARDS

- 5 stars in Python (HackerRank)
- 5 stars in Problem Solving (HackerRank)
- 5 stars in MySQL (HackerRank)