

Shivansh Mahajan

✉ shivansh.m2003@gmail.com +91 9798940184 www.linkedin.com/in/shivansh1976 www.github.com/shivansh-2003

SUMMARY

Aspiring Data Scientist and AI Engineer with strong foundation in Python, SQL, and data visualization. Hands-on experience building machine learning models, particularly in Natural Language Processing using LangChain and Transformers. Interest lies in Generative AI, with practical skills in RAG techniques and LLM fine-tuning, demonstrated through diverse project experience in medical prediction, Data Analyst Assistant, and Ultimate summarization.

EDUCATION

Jaypee University of Engineering and Technology

B Tech Computer Science - 7.2/10 GPA

09/2022 – 05/2026

Kautilya senior secondary school

12th – 78%

04/2019 – 05/2021

PROJECT

Ultimate Summarization

- Implemented specialized LangChain summarization algorithms to optimize content processing while maintaining semantic integrity and integrated multiple API services with secure authentication.
- Engineered a multi-modal content summarization platform integrating OpenAI, Google Gemini, and Assembly AI APIs with Hugging Face models to process websites, documents, videos, and speech.
- Built a responsive Streamlit interface supporting multiple file formats (PDF, TXT, MP3, MP4, WAV) with customizable summarization parameters.

Data Analyst Assistant

- Developed an AI-powered data analysis platform with multi-format support, using Google Gemini Flash for OCR extraction and OpenAI GPT-4-turbo for automated insights generation. Implemented custom data processing pipelines to reduce decision-making time from days to minutes for business stakeholders.
- Built specialized LangChain/LangGraph agentic frameworks that process natural language inputs to deliver statistical analysis and visualizations, using Streamlit for an interactive user interface.

Medical App

- Developed several medical prediction tools, including Lung Disease, Heart Failure, Kidney Disease, Obesity, and Diabetes Prediction models that enables healthcare providers to identify high-risk patients .
- Leveraged Python, Scikit-Learn , pandas , matplotlib and seaborn with ML model with above 97% accuracy .

SKILLS

- **Programming Languages** : Python, C++, C , HTML .
- **Python Libraries & Frameworks**: Pandas, Matplotlib, Seaborn, Scikit-Learn, TensorFlow, Py Torch , Fast API, NLTK, Spacy, Beautiful Soup, Selenium and Neo4j.
- **AI/ML Technologies**: Machine Learning, Natural Language Processing, LLM Fine-Tuning , RAG Techniques , Data Science , Deep Learning ,AI Engineering and Data Analytics .
- **Cloud & Databases**: AWS (IAM, EC2, S3, Lambda, Bedrock, Sage Maker), MongoDB Atlas, MySQL, Supabase(database) , Pinecone , Chroma DB .
- **Generative AI Stack**: Lang Chain, Lang Graph, Crew AI, Agno AI and Hugging Face.
- **Tools & Software**: Tableau, Excel, GitHub, Docker, ML Flow and GitHub Actions.

ACHIEVEMENTS

- Awarded 2nd Runner-up at Ride Hacks Hackathon and Top 10 at Hack Out Hackathon, demonstrating competitive technical skills and led a team of 4 in time-constrained innovation challenges.
- Coordinator for Bitwise Club, a competitive programming club to enhance algorithmic skills.