Himanshu Parihar

LINKEDIN — GITHUB
Email: himanshuparihar853@gmail.com — Mobile: +91 9319264679

SKILLS SUMMARY

- **Languages:** Python , SQL , Pandas , numpy ,Seaborn
- **Frameworks:** Langchain, RAG, Scikit-Learn, TensorFlow
- **Tools:** MySQL, vector database (Pinecone, Chroma), AWS, LLM(openai ,mistral), ,GitHub
- **Platforms:** Jupyter Notebook, Visual Studio Code, PyCharm, GCP (Google Cloud Platform), AWS
- **Machine Learning:** Supervised Learning, Unsupervised Learning, Model Deployment, Model Optimization , Natural Language Processing(nlp)
- **DevOps:** Docker,Git
- **Soft Skills:** Strong Analytical Skills, Problem Solving, Collaboration

WORK EXPERIENCE

- **MACHINE LEARNING ENGINEER — Hivoco LINK** **August 2023 - Present**
  - Independently built a backend using Flask and NoSQL, integrating both open-source and proprietary large language models. Implemented a multi-modal Retrieval-Augmented Generation (RAG) system with a vector database (FAISS), capable of handling 82,000 users per day. Leveraged open-source solutions to reduce costs by 50%.
  - Developed and implemented machine learning models and algorithms (supervised) to improve product offerings.
  - Designed and conducted experiments to evaluate model performance, tuning and optimizing models for scalability and efficiency in production.
  - Collaborated with cross-functional teams to integrate ML solutions into products, enhancing user engagement and system performance.
  - Managed model deployment and monitoring, ensuring robust performance in production environments.
  - Maintained up-to-date knowledge of machine learning advancements to implement innovative solutions.

EDUCATION

- **University of Delhi**, Delhi, India **June 2020 - May 2023**
  Bachelor of Science (HONORS) - Mathematics: CGPA: 8.18

PROJECTS

- **Interactive QA marketing campaign** **March 2024 - May 2024**
  - Developed an interactive voice-based QA system using multi-modal Retrieval-Augmented Generation (RAG) with a vector database (FAISS) for efficient information retrieval.
  - Integrated machine learning models for speech-to-text and text-to-speech functionalities to enable seamless voice interactions, improving user engagement.
  - Utilized deep learning techniques and LLMs for accurate response generation, optimizing models for performance and scalability to support high traffic.
  - Deployed the system on AWS, ensuring reliability and scalability to handle up to 82,000 users per day.
- **Customer Service Chatbot for Saffron Queries** **November 2023 - April 2024**
  - Developed a customer service chatbot using Retrieval-Augmented Generation (RAG) to handle saffron-related queries with high accuracy, incorporating supervised learning models.
  - Integrated vector databases and the LangChain framework to optimize data retrieval and response accuracy.
  - Tuned machine learning models to improve the performance and response time of the chatbot, enhancing overall customer satisfaction.
  - Deployed and monitored the chatbot in a production environment, ensuring consistent performance.
- **NLP with Disaster Tweets** **October 2023 - November 2024** Project link
  - Developed an NLP model to classify disaster-related tweets, aiming to improve disaster response.
  - **Dataset:** Labeled tweets with text, location, and keywords.
  - **Method:** Data cleaning, EDA, text preprocessing, model building (Logistic Regression, Naive Bayes), and evaluation.
  - Achieved strong classification accuracy, aiding in faster detection of disaster-related content on social media.

CERTIFICATES

- **SQL (Advanced) Certificate (HackerRank) CERTIFICATE** **August 2023**
- **Data Science certification CERTIFICATE** **May 2023**