Mohammad Sofyan Abdullah

0336-5052500 sofyanrajpoot567@gmail.com sofyan-abdullah

Summary

I am a Data Science undergraduate student with experience in data analysis, machine learning, and web development. I have successfully completed projects leveraging Python, neural networks, and AWS deployments. I am passionate about using my technical skills to solve real-world problems and continuously improve my knowledge in the field.

Technical Skills

- **Programming Languages:** Python, Java, C++, JavaScript, SQL
- **Web Development:** HTML, CSS, Flask, FastAPI, Node.js
- **Databases:** Firebase, MongoDB
- **Data Science & Analytics:** Pandas, NumPy, Matplotlib, Seaborn, Excel, Power BI
- **Machine Learning & AI:** Scikit-learn, TensorFlow, PyTorch, Keras
- **NLP & LLM Tools:** Hugging Face, Transformers, LangChain, LlamaIndex
- **Cloud Platforms:** AWS, Azure

Experience/Projects

Data Scientist – Data Forge **May 2025 –Present**

Data Science

- *Developed a DSaaS platform enabling AI engineers to automate data life cycle.*
- *Created agentic workflows for web scrapping formatting data in diverse fields.*
- *Built RAG based models for insight generation, visualization, and model finetuning.*
- *Enabled one click AWS deployment for end-to-end ML workflows.*

Virtual Internship British Airways **Apr 2024 – May 2024**

Data Analyst

- Scraped and analyzed customer reviews using Python (BeautifulSoup, NLP).
- Built a predictive model for customer booking behavior using scikit-learn.

Backend Web Dev Intern (Startup) **June 2023 – Aug 2023**

Web Developer

- *Worked with technologies like node.js and database MongoDB to make RESTFUL API for workflow.*

Projects

- **Scholarship Finder**

Generative AI, FastAPI, Grok, Tevily, Langchain, Firebase

- Developed a web-based system that uses the web to find scholarship opportunities according to the user's needs
- Integrated web search, document QA, memory and custom tools to create an end-to-end scholarship application pipeline.
- Designed and deployed a REST API backend using FASTAPI and lang serve with support for SoP generation, eligibility filtering and real time email.

- **RAG Based – BookSage**
  Generative AI, RAG, FastAPI, Gemini
  - Developed an intelligent book retrieval system that fetches books from the internet using ISBN input and process them using RAG.
  - Integrated Google Gemini model for contextual understanding and querying.
  - Enabled web search and books retrieval.
- **IMDb Sentiment Analysis:**
  Hugging face, Python, Transformers, PyTorch
  - Fine-tuned **bert-base-uncased** on IMDb dataset using Hugging Face **Transformers**
  - Performed **data preprocessing, tokenization**, and model training with **PyTorch**
  - Achieved **80%** accuracy on sentiment classification task
- **Alzheimer Disease Detection:**
  Kaggle, Scikit Learn, Keras, Flask, AWS
  - Using brain **MRI images**, build a neural network model to classify the type of tumor.
  - Created **Convolutional Neural Networks** and **RESNET** to classify the images.
  - Got **80+** accuracy on using CNN and after hyper tuning it using keras tuners increased the accuracy to **85+**.
  - Created a user-friendly frontend and for prediction used backend technology **Flask**.
  - Deployed the project on **AWS**.
- **Diabetic Prediction using Retinopathy images:**
  Kaggle, Scikit Learn, Keras, Flask, AWS, JavaScript
  - Using retinopathy, build neural network models to classify the intensity of diabetes.
  - Created **Convolutional Neural Networks** and **RESNET** to classify the images.
  - Got almost **80%** accuracy for the models.
  - Created a user-friendly frontend and for prediction used backend technology **Flask**.
  - Deployed the project on **AWS**.
- **Bangalore House Price Prediction:**
  Kaggle, Scikit Learn, Flask, AWS, JavaScript
  - Preprocessed the dataset for Bangalore house prices.
  - Using Machine Learning models such as **linear regression** and **decision tree** to predict the price.
  - Created a user-friendly frontend and used **Flask** for backend.
  - Deployed the project on **AWS**.

Courses and Certifications

- **CISCO Netacad | Introduction to Networking**
- **Coursera & Stanford University | Machine Learning**
- **AWS | Cloud Essentials**
- **Coursera | GenAi For Everyone**

Education

COMSATS University, Islamabad **Sep 2022 – Mar 2024**
Bachelor of Science in Data Science *CGPA 3.69*
90%