Usman Yaqoob

Pakistan | usmanyaqoob49@gmail.com | https://www.linkedin.com/in/-usman-yaqoob/ | (+92) 306-8147112
https://github.com/usmanyaqoob49 | https://usmanyaqoob49.github.io/portfolio/ | https://www.kaggle.com/usman49

Professional Experience

JMM TECHNOLOGIES, *Data Scientist* 2023/09 – present
Peshawar, Pakistan

A. ProofLingo: https://prooflingo.com/

- Developed a AI system that can detect contextual errors, missing translation, consistency errors, Incorrect translation from translation document (Spanish, French, German, Arabic) after comparing it with source document using openai GPT-4o LLM. Enhanced the accuracy of system from 50% to 90% using advanced prompt engineering techniques.
- Developed RAG system for same product to help the user to know more about the documents.
- Build the backend of the Website and restful APIs for AI system using python and Django.
- Deployed the Backend system to AWS using LightSail instance and container for the integration with the frontend.

B. Cryptocurrency Trading Bot Development:

- Trained models (XGBoost, Random Forest, Decision Tree) for forecasting Ethereum's closing value over the next 4 hours by using historical Ethereum data and market fear and greed data to enhance prediction accuracy. Selected the XGBoost model for the final implementation due to its lowest RMSE (Root Mean Squared Error) loss.
- Developed a Random Forest model to predict the best APR (Annual Percentage Rate) by analyzing other protocols' APRs, coin-specific APRs, and news sentiment extracted from cryptocurrency news data.
- Created a prediction pipeline that extracts real-time coin APR data, protocol APR data, news sentiment data, and market fear and greed data. The pipeline predicts APR based on the amount, duration, and token specified by the user.

C. Riyadh's Different Places Reviews Sentiment Analysis and Analytics:

- Analyzed 2.1 million reviews of different places in Riyadh to extract sentiments from the reviews and identify specific aspects being discussed. Determined whether the sentiment was positive or negative, helping to understand what people liked or disliked about various aspects. Achieved 50% cost savings by utilizing the OpenAI GPT-4 batch API for sentiment extraction.
- Conducted detailed analysis on the results to provide actionable insights. Identified top aspects that people liked about popular places and pinpointed negative aspects that needed improvement. This information was aimed at helping place owners understand and address specific issues.

D. HRM:

- Developed a Resume Scoring System for the HRM Product that automates the initial selection of the resume by scoring the resumes according to Job Description.
- Used DeepFace Library to build Employee Face Recoginition System for same product. Improves the storage efficiecy by 50% by utilizing OpenFace model instead of Vgg16.
- Developed Backend APIs for both systems using Flask.

Lab-D Tech Solutions, *Machine Learning Engineer* 2023/07 – 2023/08
Islamabad, Pakistan

- Developed and optimized a human face recognition system for a school utilizing MTCNN and ResNet. The system identifies authorized individuals and alerts the administrator upon detecting an unauthorized person, prompting them to authorize the individual by entering their name.

Education

Computer Systems Engineering, *University of Engineering and Technology, Peshawar* 2019 – 2023
CGPA: 3.41 (CGPA) Peshawar, Pakistan

- Extensively studied Data Analytics, Database Management Systems such as MySQL, Operating Systems, Data Structures and Algorithms, and Digital Signals Processing.

Skills

Machine Learning/Data Analysis:: Python, TensorFlow, Keras, OpenCV, Pytorch, HuggingFace, Scikit Learn, Numpy, Pandas, Matplotlib

Data Engineering and Cloud Computing: SQL, ETLs, Docker, AWS(EC2, Lightsail Lambda, S3, Cloud9, SageMaker), GCP(Artifacts Registry, Cloud Run, Google Compute Engine, VertexAI)

Generative AI: LLMs Finetuning, LangChain, LLamaIndex, OpenAI (Assistant API, Batch API), Prompt Chaining, Prompt Engineering

Backend Technologies: Flask, Dash, Django, FastAPI

Projects

Human Emotion Detection and Enhancement via Recommendation Systems (FYP), 2022/05 – 2023/06
Tech Stack: *Python, Numpy, Pandas, Matplotlib, NLP, DistilBert, Bert, Deep Neural Networks, Streamlit*

- Trained a model with an accuracy of 96% for four emotions detection using BERT and Deep Neural Networks.
- Made a Recommendation System for recommendation of the Exercises from our customized exercise dataset to do mood enhancement.

Pakistan Suicide Bomb Attacks Analysis 2023/09 – 2023/09

- After Cleaning the data, I performed detailed Analysis and Visualization to get some Insights from the Dataset. Like:
  - Total Number of People Killed, Cities Where Most Number of Attacks Happened, Targetted Sect etc.

Certificates

Machine Learning Specialization
Coursera | Stanford Online | DeepLearnig.ai

Mathematics for Machine Learning
Coursera | Imperial College London

Foundations of Data Science
Google

Neural Networks and Deep Learning
DeepLearning.ai