Muhammad Umer Mansoor

+92 301 5691704 | techsavvyumer@gmail.com | LinkedIn: techsavvyumer | GitHub: techsavvyumer

WORK EXPERIENCE

Machine Learning 1 | Data Scientist | Mar 2024 – Present

- Designed and implemented machine learning models for tasks including facial recognition, classification, and price prediction, leveraging libraries like Python and frameworks like TensorFlow.
- Conducted extensive data mining and scraping techniques using Selenium, Beautiful Soup, and hidden APIs to extract valuable data from diverse sources for training the model.
- Generated synthetic datasets to enhance training and testing of deepfakes detection models, addressing data scarcity, and improving model robustness.

Knowledge Streams | DS & ML Trainee | Oct 2023 – Jan 2024

- Learned exploratory data analysis, statistical analysis, data visualization, and machine learning algorithms, developing strong analytical thinking and research skills.
- Managed time effectively to complete 3 major projects while meeting tight deadlines, demonstrating strong project management abilities.
- Collaborated cross-functionally to implement machine learning solutions, exercising teamwork and communication skills.

PROJECTS

Automatic License Plate Detection and Recognition | Knowledge Streams | (YOLOv8, OpenCV, CNN)

- Implemented Yolo v8 models to detect vehicles and license plates, ensuring only relevant regions are processed further.
- Engineered an OpenCV pipeline to segment license plate characters for optimal recognition.
- Built a CNN recognition model that identified license plate characters, outperforming traditional approaches.

Duplicate Bug Report Detection and Classification | FYP @ VU | (NLTK, Spacy, Scikit-learn)

- Preprocessed real-world bug report data by tokenizing, removing stopwords and lemmatizing text for cleaner input and demonstrating NLP skills.
- Built a classifier using Naive Bayes, SVM and Random Forest that identified duplicate bug reports with 82% precision, allowing developers to prioritize unique issues.
- Recommended model improvements by analyzing misclassified examples, demonstrating analytical abilities.

Realtime Tweets Sentiment Analysis | Winner Project @ Datathon Competition | (Tweepy, NLTK, Streamlit)

- Created a live Twitter sentiment analyzer app with Streamlit that determined tweet sentiment with 75% accuracy, exercising programming abilities.
- Cleaned and preprocessed tweets using regex, NLTK etc. to prepare data for model ingestion.
- Streamlined data workflow from tweet acquisition to preprocessing to sentiment prediction, exhibiting excellent project management skills.

CORE SKILLS

Programming Languages: Python, C/C++
Developer Tools: VS Code Git, Conda, Linux
Frameworks: Scikit-learn, Keras, TensorFlow, PyTorch
Databases: MySQL, Microsoft SQL Server, Oracle SQL
Cloud Computing: Azure, AWS, GCP

CERTIFICATIONS

Google Advanced Data Analytics |Coursera | 2023

EDUCATION

Virtual University of Pakistan, Lahore
BS Computer Science, Artificial Intelligence | 3.40/4.0 | 2019 – 2023