UMAIMA AMAN

Rawalpindi, Pakistan



+923315191782

umaimaaman6@gmail.com

https://github.com/umaima-aman

www.linkedin.com/in/umaimaaman

Education

FAST-NUCES Islamabad:

MS Data Science - CGPA: 3.7.0/4.00	2023(Research Stage)
Quaid-i-Azam university Islamabad:	
BS Mathematics- CGPA: 3.20/4.00	2023
Army Public School and College, Westridge 3, Rawalpindi	2019
Intermediate-874/1100, 79.45%	
Siddeeq Public School, Satellite Town, Rawalpindi	2017
Matriculation-1030/1100, 93.63%	

Skills

Technical Skills (Software and instruments):

- **Programming:** Python, R, MATLAB
- Data Science Tools and Machine Learning: Pandas, NumPy, Scikit-learn, TensorFlow, Keras
- Deep Learning Tools: Pytorch ,Lang Chain , Hugging Face Transformers
- Natural Language Processing: Text Preprocessing, Sentiment Analysis, Topic Modeling
- Data Visualization: Matplotlib, Seaborn, Excel
- Big Data: Hadoop, W&B (Weights & Biases) for Experiment Tracking

Soft Skills

Leadership, Teamwork, Time Management, Adaptability, verbal and written Communication, punctuality

Activities, Hobbies and Interests

Gardening, reading, event management, research work

Projects

Cardiovascular Risk Predictor

Designed and implemented a machine learning model to predict cardiovascular risk using medical data, determining if a patient is at risk of a heart attack (Yes/No). Applied Artificial Neural Networks to check performance.

Technologies Used: Python, scikit-learn, Pandas, Matplotlib

Diabetes Prediction Model

Developed a model to predict diabetes (Yes/No) based on patient health records. Experimented with multiple algorithms, including Support Vector Machine, Naive Bayes, Logistic Regression, and Artificial Neural Networks. Addressed data imbalance using SMOTE and validated results with k-fold cross-validation to enhance model reliability.

Technologies Used: Python, scikit-learn, Pandas, SMOTE, Matplotlib

Lung Cancer Detection Model

Created a robust classification system for detecting lung cancer using various supervised learning algorithms, such as SVM, Neural Networks, and Logistic Regression. Leveraged advanced data science techniques, including SMOTE for data balancing and k-fold cross-validation, to ensure consistent results across diverse datasets.

Technologies Used: Python, scikit-learn, SMOTE, Neural Networks

Sentiment Analysis Using Transformer Models

Fine-tuned transformer-based models like BERT to classify sentiments in tweets with high accuracy. Focused on advanced text preprocessing, tokenization, and hyperparameter tuning to improve the model's performance. Compared results with traditional approaches to highlight the advantage of pre-trained models in NLP tasks.

Technologies Used: Python, Hugging Face Transformers, TensorFlow, Keras

Sentiment Analysis Using Boosting Algorithms

Performed a comparative analysis of Gradient Boosting (GBM) and Adaptive Boosting (AdaBoost) algorithms for sentiment classification. Customized data preprocessing pipelines to optimize model performance, achieving robust and accurate predictions.

Technologies Used: Python, scikit-learn, Pandas

Disease Trajectory Modeling and Prediction

Conducted a systematic literature review of AI and data science techniques in disease trajectory prediction. Identified key methodologies and trends, summarizing findings into actionable insights for predictive healthcare applications.

Technologies Used: Research, Data Mining

Titanic Survival Analysis

Analyzed survival trends in the Titanic dataset, focusing on correlations between socioeconomic status, boarding location, and gender. Developed an R script to identify key patterns and create visualizations that enhanced data interpretation.

Technologies Used: R Programming

Question-Answering System

Built a user-friendly question-answering system combining RAG (Retrieval-Augmented Generation) with LangChain for efficient information retrieval. Improved accuracy and scalability by incorporating cutting-edge techniques in the chainlit framework.

Technologies Used: Python, LangChain, RAG

Online Academic Platform and College Management System

Designed and implemented a database-driven online academic management system. Created ER and EER diagrams, normalized database schemas, and executed complex SQL queries for data insertion, retrieval, and constraint creation.

Technologies Used: SQL, Database Schema Design

TA & Demonstrator Management System

Developed a database management system for organizing teaching assistant assignments. Implemented ER/EER diagrams and SQL queries to create, normalize, and manage relational schemas.

Technologies Used: SQL, Database Design

Languages

English, Urdu, Punjabi

Certifications

Writing in the Sciences

Stanford University (online course) 2024