Umar Saleem

+92 318 8784493 | umarsaleem0816@gmail.com GitHub | LinkedIn | Kaggle

About Me

Data Science undergraduate student with a strong foundation in programming, data analysis, machine learning, and business intelligence. Proficient in Python, SQL, data-driven solutions and data visualization, eager to apply analytical skills to solve real-world problems and drive data-driven decision-making.

Projects

Dream Vividness Prediction

- Built a Gradient Boosting model to predict vivid dream levels from sleep data.
- **Tools & Technologies:** Python, Pandas, Scikit-learn, Scipy, Seaborn, Matplotlib

QA and Summarizer App

- Developed a FullStack QA and Summarizer web app enabling passage-based question answering and text summarization.
- **Tools & Technologies:** React, Bootstrap, Python, FastApi, SQlite, transformers, torch, ML models (distilbert-base-uncased-distilled-squad, facebook/bart-large-cnn)

Medical Lab Dashboard

- Created an interactive Power BI dashboard to visualize medical lab test trends.
- **Tools & Technologies:** Power BI, Excel, DAX

Smart ROCF Scoring System

- Developed a Smart ROCF Scoring System using computer vision and machine learning to automatically extract shapes and structures from original and patient-drawn images for comparative scoring.
- **Tools & Technologies:** Python, OpenCV, Scikit-learn, NumPy, Matplotlib, React (for frontend)

E-commerce Sales Prediction Web App

- Built an interactive web app using Streamlit to visualize and analyze historical sales data to predict e-commerce sales.
- **Tools & Technologies:** Python (Streamlit, Pandas, Matplotlib, Seaborn, NumPy, Scikit-learn), CSS.

Education

Punjab University College of Information Technology Lahore, Pakistan

Bachelor of Data Science – Grade 3.52/4.00 *Dec. 2022 – Present*

Certificates and Courses

- Pandas Certification (*Kaggle*)
- Introduction to Machine Learning (*Kaggle*)
- Competitive Programming Certificates (Softec'24, ElectroCon'24, PUCON'23, CodeRush'25, PUCON'25).

Research

- **Forecasting Vivid Dreams –** Currently working on a machine learning–based research project focused on predicting vivid dream patterns using psychological and behavioral indicators. *(Manuscript in preparation for future conference submission).*

Skills

Programming and Scripting: Python, SQL, Java (Basic), HTML, CSS, JavaScript
Data Analysis & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Power BI, Excel
Libraries and Tools: Scikit-learn, K-means, Regression Models, Jupyter Notebook, Visual Studio, Streamlit, Git, Github
Databases: Oracle, DB Browser, SQL Server Management Studio (SSMS)
Relevant Coursework:

- Programming (Python), OOP, Data Structures & Algorithms
- Discrete Structures, Statistics & Probability, Advanced Statistics, DLD, Computer Organization and Assembly Language (COAL)
- Artificial Intelligence (AI), Database Systems, Operating Systems (Linux), Data Warehouse & Business Intelligence, Data Visualization, Introduction to Data Science
- Data Mining & Machine Learning, Computer Networks, Information Security, Computer Vision, Web Engineering (Java), Analysis of Algorithms