Rana Muhammad Talha

+92-333-7709594 Gmail Lahore, Pakistan GitHub LinkedIn Portfolio

Professional Summary

Final-year Computer Science student at UET Lahore with a strong focus on Machine Learning and Data Science. Skilled in Python, SQL, and data visualization tools, with solid hands-on experience in building and analyzing real-world datasets. Currently expanding my expertise in Deep Learning and excited about using data-driven insights to solve practical problems and build AI-powered solutions.

Education

University of Engineering and Technology, Lahore
BSc in Computer Science *Dec 2022 – Present*

Technical Skills

- **Programming Languages:** Python, C/C++, SQL(MySQL, PostgreSQL), HTML, CSS
- **ML Framework:** Scikit-learn, Streamlit, Joblib, FastAPI, Pickle, NLP
- **Data Processing:** Pandas, NumPy, Matplotlib, Seaborn, Tableau, Excel, Web Scraping
- **Tools:** Git, GitHub, Docker ,VS Code, Jupyter Notebook, Google Colab, Kaggle Notebook, SQL Server
- **Soft Skills:** Communication, Teamwork, Problem Solving

Projects

- **Clinical Risk Classification System** **Machine Learning**
  Developed a supervised learning pipeline to classify cardiovascular risk using structured clinical data. Implemented data preprocessing, feature scaling, and model evaluation using Logistic Regression, Random Forest, and SVM, achieving up to 87% accuracy through cross-validation and hyperparameter tuning.
  Tools: Python, Scikit-learn, Pandas, NumPy, Seaborn.
- **Academic Performance Prediction** **Machine Learning**
  Built an end-to-end machine learning system to predict students' academic performance using demographic and academic features. Implemented data preprocessing and feature encoding, trained regression models, and deployed the model using a FastAPI backend integrated with a Streamlit frontend for real-time predictions.
  Tools: Python, Scikit-learn, FastAPI, Streamlit, Pandas, NumPy
- **Insurance Premium Category Predictor** **Machine Learning**
  Designed an end-to-end ML inference pipeline to predict insurance premium categories. Deployed the trained Random Forest model using FastAPI, containerized the service with Docker, and published the image to Docker Hub.
  Tools: Python, Scikit-learn, FastAPI, Docker, Pydantic.
- **Sentiment Analysis on Pakistani Drama Subtitles** **NLP**
  Built multilingual sentiment analysis models on 160K+ Pakistani drama subtitles. Labeled 73K+ English and Urdu sentences using DistilBERT and M2M100, achieving 81% accuracy for English and 80% for Urdu.
  Tools/Skills: Python, NLP, Transformers, DistilBERT, M2M100, Pandas, NumPy.

Certifications

- Machine Learning Specialization – Coursera *October 2025*
- Python for Data Science, AI and Development – IBM *July 2025*
- AI for everyone – IBM *June 2025*
- Intermediate SQL – Data Camp *March 2025*

References

Available upon request.