

Musfira Mubeen

Lahore, Pakistan — +92 322 4616539 — musfiramubeen553@gmail.com
[LinkedIn](#) — [GitHub](#)

Professional Summary

AI & Data Science enthusiast with hands-on experience in multi-modal machine learning, deep learning, and full-stack ML systems. Skilled in Python, SQL, Flask, Streamlit, ANN, CNN, LSTM, and Web Scraping. Passionate about building intelligent systems that deliver actionable insights in healthcare and business.

Projects

AI Personal Health Tracker – Multi-Modal Heart Disease Risk Prediction System (Deep Learning & Flask)

- Developed a healthcare AI system combining ANN (clinical data), CNN (X-ray images), and LSTM (ECG signals) for heart disease prediction.
- Built an embedding fusion model achieving ≈90% accuracy across multi-modal inputs.
- Created a Flask-based interface for real-time predictions with patient-level risk scoring and recommendations.

AI Customer Retention & Revenue Intelligence System (Full-Stack ML & Flask)

- Developed a full-stack ML system to predict customer churn, estimate revenue at risk, and recommend retention strategies using Python, Flask, scikit-learn, XGBoost, and SHAP.
- Built Gradient Boosting churn classifier (81% accuracy, 71% recall), revenue regression ($R^2=0.999$), and K-Means segmentation with PCA visualization.
- Integrated interactive Flask dashboard with batch/manual predictions, KPIs, SHAP explainability, and downloadable results.

Experience

BCG X – Data Science Virtual Internship

Feb 2025 – Oct 2025

- Conducted exploratory data analysis, feature engineering, and predictive modeling.
- Delivered a churn prediction model and provided actionable insights to enhance customer retention strategies.

Education

University of Education, Lahore

Dec 2020 – Dec 2024

BS Information Technology, specialized in Data Science

Skills

- **Programming & Tools:** Python, SQL, Flask, Streamlit, Joblib, Web Scraping
- **Libraries & Frameworks:** NumPy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras, OpenCV
- **Machine Learning & AI:** Regression, Classification, Clustering, Feature Engineering, Pipelines, Model Tuning, Deep Learning (ANN, CNN, LSTM), Multi-Modal Fusion
- **Concepts:** EDA, Model Deployment, Data Visualization, API Integration
- **Soft Skills:** Problem Solving, Analytical Thinking, Communication

Certifications

- Supervised Machine Learning: Regression and Classification – Stanford University (Coursera)
- Advanced Learning Algorithms – Stanford University (Coursera)
- Crash Course on Python – Google (Coursera)
- What is Data Science? – IBM (Coursera)
- Tools for Data Science – IBM (Coursera)