

# Musfira Mubeen

Lahore, Pakistan — +92 322 4616539 — [musfiramubeen553@gmail.com](mailto:musfiramubeen553@gmail.com)

[LinkedIn](#) — [GitHub](#)

## Professional Summary

---

AI & Data Science enthusiast with experience in multi-modal ML, deep learning, and full-stack ML systems. Skilled in Python, SQL, Flask, Streamlit, ANN, CNN, LSTM, and Web Scraping. Builds intelligent systems delivering actionable insights.

## Projects

---

### AI Personal Health Tracker – Multi-Modal Heart Disease Risk Prediction System

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

### AI Customer Retention & Revenue Intelligence System

- Full-stack ML system predicting churn, revenue at risk, and recommending 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 EDA, feature engineering, and predictive modeling.
- Delivered churn prediction model and actionable insights for retention.

## 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)