MUHAMMAD KASHIF SULTAN

+92 304 7992833 ⋄ Sahiwal, Pakistan

mk695870@gmail.com ⋄ [LinkedIn](#)

PROFESSIONAL SUMMARY

Computer Science undergraduate with hands-on experience in machine learning, data preprocessing, predictive modeling, and statistical analysis using Python. Built real-world regression and analytics projects involving feature engineering, model development, and data visualization. Skilled in scikit-learn with foundational experience in TensorFlow and PyTorch. Strong analytical mindset with research experience in small-dataset learning and AI applications.

SKILLS

Languages:	Python, SQL, Java, C++
Libraries:	Scikit-learn, TensorFlow (beginner), PyTorch (beginner), NumPy, Pandas, Matplotlib, Seaborn
Concepts:	EDA, Predictive Modeling, Statistics, Algorithm Design, Reporting and Documentation
Tools:	Jupyter Notebook, Google Colab, Git, MySQL

PROJECTS

Cyberbullying Detection System (ML + DistilBERT)   *SEP 2025*

- Built a cyberbullying classifier using traditional ML models and a fine-tuned DistilBERT transformer.
- Performed text preprocessing including tokenization, cleaning, and TF-IDF feature extraction.
- Achieved higher accuracy with DistilBERT and deployed a simple pipeline for batch or real-time predictions.
- *GitHub:* [Cyberbullying Detection System](#)

X Sentiment and Emotion Analysis (Batch + Real-Time ML Pipeline)   *DEC 2025*

- Developed a dual-mode sentiment and emotion analysis system capable of batch processing and real-time prediction.
- Designed preprocessing pipelines using tokenization, embeddings, and noise reduction for text streams.
- Trained ML and deep learning models to classify sentiment and emotion categories.
- Integrated real-time inference using Python API endpoints for fast prediction
- *GitHub:* [X Sentiment and Emotion Analysis](#)

Car Price Prediction System using Linear Regression   *Jul 2025*

- Designed a car price prediction model using multivariable Linear Regression.
- Performed feature engineering, encoding, outlier handling, and model evaluation using R^2 and MSE metrics.
- Visualized actual vs predicted prices, residuals, and feature impact using Matplotlib and Seaborn.
- *GitHub:* [Car Price Prediction System](#)

EDUCATION

BS in Computer Science, COMSATS University Islamabad, Sahiwal Campus   SEP 2025 – Present
Relevant coursework: Data Structures, Machine Learning, Probability & Statistics

RESEARCH & INVOLVEMENT

- **Volunteer Research Assistant – CRADLE Lab**: Contributing to research on edge AI and small dataset modeling for rural sectors. Collaborating with peers to draft technical research for data-driven solutions.
- **Community Mentor – Python & Data Science**: Guided junior students in the fundamentals of data analysis and machine learning practices through virtual mentoring sessions.

CERTIFICATIONS

- **Professional Certificate in Data Science Fundamentals with Python and SQL** — IBM, Coursera (2024)
- **Professional Certificate in Data Science** — IBM, Coursera (2024)