

Abdullah Haroon

✉ Email ☎ +92 306 0590604  GitHub  Portfolio  LinkedIn  Google Scholar

Machine Learning · NLP · Backend & APIs · AI-Powered Systems

Education

Bachelor of Science in Computer Science
FAST-NUCES, Lahore, Pakistan

2023 - Present (Expected 2027)

Research

[\(view all\)](#)

Cross-Dataset Generalization in Urdu Fake News Detection: An Empirical Study with XLM-RoBERTa and a Length Confound Analysis

Sole Author · [View Publication](#)

- Fine-tuned a multilingual transformer (**XLM-RoBERTa**) across four cross-domain conditions on two Urdu corpora (**23,471 articles**); diagnosed a systematic length confound causing catastrophic zero-shot transfer failure (**macro F1: 0.005**), with a reusable diagnostic methodology for low-resource NLP.

Projects

FinSight: Financial Behavior & Risk Analysis Platform

[GitHub](#)

- End-to-end financial analytics platform (**FastAPI**, **scikit-learn**, **React.js**); led a team of 3 as Scrum Master with supervised risk scoring (**~92% accuracy**), SHAP-based explainability, and unsupervised behavioral clustering.

Urdu Fake News Detection: Transformer Cross-Domain Study

[GitHub](#)

- Fine-tuned multilingual transformer **XLM-RoBERTa** on two Urdu datasets; benchmarked against text vectorization + **LR/SVM** baselines and diagnosed shortcut learning failure in zero-shot cross-domain transfer.

NLP-Based Pakistani Fake News Classifier

[GitHub](#)

- Classified **4,813** political news articles; Logistic Regression achieved **79.2% accuracy** and **AUC-ROC 0.885** across 5-fold CV.

Pakistan Rainfall Morphology & Event Typology: Explainable ML

[GitHub](#)

- Extracted **4,563** rainfall events from **44 years** of data; GMM clustering and PELT changepoint detection identified a structural monsoon regime shift around 1996 (**Random Forest: 96.6%**).

Feedforward Neural Network Engine from Scratch

[GitHub](#)

- Built a complete neural network engine in pure **NumPy** (**Adam optimizer**, **backpropagation**); achieved **97.3%** benchmark accuracy, within **0.7%** of Keras.

Technical Skills

Languages: Python, C/C++, C#, SQL, HTML/CSS

ML/DL: Scikit-Learn, Keras, TensorFlow, Supervised & Unsupervised Learning, ANNs, CNNs, End-to-End Pipelines

NLP & LLMs: HuggingFace Transformers, Text Preprocessing, Text Vectorization, Text Classification

Backend & Infrastructure: FastAPI, SQLAlchemy, SQLite, MS SQL Server, REST API Design

Tools: Git, Jupyter, VS Code, Google Colab

Interpersonal: Problem Solving, Critical Thinking, Public Speaking, Collaboration, Team Leader, Team Player, Time Management

Leadership

- Scrum Master**, Semester Projects - led sprint planning, task delegation, and on-time delivery for **multi-member teams**.
- Deputy Head**, FAST Business Club - managed a team of **7**, coordinating events and workshops.

Certifications

- [Machine Learning with Python](#) - IBM (Coursera)
- [Data Analysis and Visualization with Python](#) - Microsoft (Coursera)
- [Python for Data Science, AI & Development](#) - IBM (Coursera)