

# Salman Khan

📍 Islamabad Pakistan | **Email:** salmanrahim520@gmail.com | **Phone:** +923119365608 | **LinkedIn:** in/salman-khan5  
**GitHub:** https://github.com/KhanzadaSalman | **Portfolio:** https://khanzadasalman.github.io

Data Science Master's graduate with specialized research in Natural Language Processing (NLP) and Deep Learning. Experienced in building predictive models and analyzing text data using Python, with a focus on Roman Urdu hate speech detection. Dedicated to bridging the gap between academic research and practical AI applications through interactive project deployments.

## EXPERIENCE

### AI & Machine Learning Intern | DevelopersHub Corporation (Remote) Dec 2025 – Jan 2026

- Fine-tuned BERT for text classification and applied zero-shot learning for support ticket tagging.
- Developed a multimodal model combining image features and tabular data for housing price prediction.
- Built machine learning pipelines for customer churn and heart disease prediction using Scikit-learn.

### Graduate Researcher | Shifa Tameer-e-Millat University, Islamabad Nov 2024 – Aug 2025

- Conducted research on hate speech detection for Roman Urdu text.
- Designed and trained a lite Depth-wise Separable Convolutional Neural Network (lite DSCNN) for text classification.
- Experimented with multiple text representations, including FastText and transformer-based embeddings.
- Developed an end-to-end NLP pipeline covering preprocessing, training, and evaluation.
- Evaluated models using accuracy and F1-score metrics.

### Data Science Job Simulation Participant | British Airways (via Forage) Aug 2025

- Completed a data science simulation based on real-world airline business problems.
- Cleaned and prepared structured customer data provided as part of the simulation.
- Analyzed customer booking behavior to identify likely ticket purchasers.
- Segmented flight routes to support decisions on VIP lounge access.
- Prepared a presentation summarizing insights and recommendations for stakeholders.

## PROJECTS

### SpamGuard AI - Live Spam Detection System (Full-Stack ML Deployment) Feb 2026

**Technologies:** Python, TensorFlow, FastAPI, Hugging Face, HTML/CSS, NLP, Git.

- Built and deployed a deep learning web application for real-time spam message classification.
- Developed a FastAPI backend to serve model predictions.
- Implemented a responsive frontend using HTML, CSS, and JavaScript.
- Deployed the application with automated updates via GitHub integration.

### Live Spam Email Detection System (Full-Stack ML Deployment) Jan 2026

**Technologies:** Python, Flask, Render, HTML/CSS, uv, Scikit-learn, NLP, Git.

- Built a web application to classify emails as spam or non-spam.
- Implemented the backend with Flask and a simple HTML/CSS frontend.
- Deployed the application on Render with Git-based version control.

### Stock Market Trends Forecaster (Deep Learning) Dec 2025

**Technologies:** Python, Keras (LSTM), Pandas, Matplotlib, Scikit-learn

- Developed a time-series prediction model using historical stock price data.
- Implemented an LSTM neural network to capture temporal patterns.
- Engineered a data pipeline to clean, normalize, and visualize financial trends for analysis.
- Achieved low Mean Squared Error, demonstrating strong predictive capabilities for volatile markets.

### Lung Cancer Prediction Mar 2024

**Technologies:** Python, Scikit-learn, Random Forest, Decision Trees, Data Preprocessing

- Built a healthcare diagnostic tool to identify potential lung cancer risks using patient medical history.
- Applied feature engineering and hyperparameter tuning to optimize model performance.
- Achieved 90% accuracy, supporting early detection in healthcare applications.

**Personal Portfolio Website**  
**Technologies:** HTML, CSS, Git, GitHub Pages

- Designed and developed a personal portfolio website to showcase projects and skills.
- Managed source code versioning and automated deployment using Git and GitHub Pages.

Oct 2025 – Present

## EDUCATION

**MS Data Science | Shifa Tameer-e-Millat University, Islamabad**  
Islamabad, Pakistan  
CGPA: 3.5 / 4.0

Dec 2022 - Aug 2025

**BS Mathematics | Shaheed Benazir Bhutto University Sheringal**  
Dir Upper, KPK, Pakistan  
CGPA: 3.4 / 4.0

Oct 2018 - Dec 2022

## SKILLS

- Programming & Web:** Python, R, SQL, C++, HTML, CSS
- Frameworks & Libraries:** Pandas, NumPy, Scikit-learn, TensorFlow/Keras, NLTK, Matplotlib, Flask
- Machine Learning & NLP:** Sentiment Analysis, Text Classification, Word Embeddings, Transformers (BERT).
- Deep Learning:** CNN, RNN, LSTM, Transformers.
- Data Analytics & Visualization:** Tableau, Power BI, Excel (Pivot Tables, VLOOKUP), Seaborn.
- Databases:** PostgreSQL, MySQL.
- Tools & Platforms:** Git, GitHub, Render, uv, Jupyter Notebook, Visual Studio.

## CERTIFICATIONS

- AI & Machine Learning Intern** — DevelopersHub Corporation (Remote)
- Data Analysis Using Python** — University of Pennsylvania (Coursera)
- Analytics with SQL and Python** — Great Learning Academy
- Data Analytics Using Excel** — Great Learning Academy
- Basics of Python Programming** — OpenWeaver
- ChatGPT for Excel** — Great Learning
- Learning Data Analytics Foundations** — LinkedIn Learning

Jan 2026

Jun 2023

Jun 2024

Jun 2024

Apr 2023

Jun 2023

Jul 2024

## COURSEWORK

- Machine Learning, Deep Learning, NLP, Math & Statistics for Data Science.
- Information Retrieval, Programming Fundamentals, Research Methodology.
- Data Structures & Algorithms, Distributed Data Processing.

## REFERENCES

**Dr. Raja Habib,**  
Assistant Professor,  
Shifa Tameer-e-Millat University,  
Islamabad, Pakistan.  
Email: rajahabib.ssc@stmu.edu.pk  
Contact # +92331-5383504

**Dr. Muhammad Fahad Khan,**  
Associate Professor,  
Shifa Tameer-e-Millat University,  
Islamabad, Pakistan.  
Email: mfkhan@cs.qau.edu.pk  
Contact # +92 3009879721