

SAJJAD ALI SHAH

+92 316 9671878 | sajjadkhanyousafzai47@gmail.com | [Portfolio](#) | [Linkedin](#) | <https://github.com/SajjadKhanYousafzai>

PROFESSIONAL SUMMARY

Highly motivated Machine Learning Engineer and Data Scientist enthusiast specializing in Computer Vision, NLP, and predictive analytics. Successfully developed 8+ end-to-end ML solutions achieving 90-95% accuracy, including real-time ASL detection systems and intelligent recommendation engines. International collaborator on Omdena's AI Innovation Challenge (40-member global team across 20+ countries). Deep expertise in model development, fine-tuning, and deployment using TensorFlow, PyTorch, and modern ML frameworks. Passionate about leveraging AI and data-driven insights to solve real-world problems.

EDUCATION

Comsats University Islamabad (Wah Campus) **Sep 2022-Present**

Bachelor of Science, Software Engineering (GPA: 3.2/4.0)

• **Coursework:** AI, Linear Algebra, DSA, Statistics & Probability, Calculus, Database & Data Science

Army Public School & Degree College Malir Cantt Karachi **Mar 2018 - Aug 2020**

FSc, Pre Engineering (Grade: Mathematics (A+), Chemistry(A-), Physics(A).)

WORK EXPERIENCE

Omdena — Collaborator (AI Innovation Challenge: CropLogic) **Sep 2025 – Nov 2025**

Machine Learning Engineer

Omdena (Remote)

- Selected to join a global team of 40 collaborators from 20+ countries.
- Contributing to the development of **AI-powered farming intelligence solutions** for smallholder and medium farms.
- Responsibilities include **data preprocessing, machine learning model development, and documentation.**
- Collaborating in an agile, international environment using GitHub and Slack.

Digital Empowerment Network **Sep 2024 - Oct 2024**

Machine Learning Intern

Islamabad Pakistan

- Constructed and fine-tuned machine learning models for classification and regression tasks using TensorFlow and Scikit-learn, achieving up to 93% accuracy on structured datasets while leveraging Python.
- Conducted comprehensive data preprocessing, feature engineering, and model evaluation using Pandas to optimize predictive performance and derive actionable insights.
- Collaborated with cross-functional teams, including data engineers and analysts, employing SQL and Excel to validate and enhance data analysis.

PROJECTS

Sign Language Detection System

- Developed an end-to-end real-time American Sign Language (ASL) recognition system using TensorFlow, MediaPipe, and OpenCV. Preprocessed ASL datasets, built a CNN (90% accuracy), and integrated MediaPipe for hand landmark detection. Deployed a Tkinter GUI for live webcam detection with text-to-speech output via pyttsx3 and ported the model to TensorFlow.js for web deployment.

Movie Recommendation System

- Built a recommendation engine using collaborative and content-based filtering with Scikit-learn and Pandas, deployed via Streamlit. Delivered personalized movie suggestions, enhancing user engagement by 15% in testing.

Prediction Models

- Heart Disease Prediction: Developed a Scikit-learn classification model to predict heart disease, achieving 92% accuracy via feature engineering.
- Diabetes Prediction: Built logistic regression and decision tree models to assess diabetes risk, validated with 88% accuracy.
- Loan Approval Prediction: Designed a classification model with ensemble techniques, improving loan approval accuracy by 10%.
- Email Spam Detection: Implemented an NLP-based classifier with Scikit-learn, achieving 95%+ accuracy for spam detection.
- House Price Prediction: Created a regression model to estimate housing prices, reducing prediction error by 12% through cross-validation.

Customer Personality Analysis

- Performed customer segmentation using K-Means clustering with Scikit-learn, enabling targeted marketing strategies with a 20% improvement in campaign focus during testing.
-

SKILLS SUMMARY

- **Languages:** Python (Proficient), Flask, FastAPI, Dart, C++, Streamlit
- **AI/ML libraries:** TensorFlow, Keras, PyTorch, Hugging Face Transformers, OpenAI API, LangChain, CNNs, RNNs, LSTMs, Transformers, Fine-tuning Large Language Models (LLMs), Prompt Engineering, Retrieval-Augmented Generation (RAG), Pandas, Numpy, Scikit-learn, Matplotlib, Seaborn, Statsmodels
- **NLP & Computer Vision:** NLTK, SpaCy, OpenCV, MediaPipe
- **Databases:** Power BI, Excel, PowerPoint, Tableau, MySQL, SQLite, Git, Docker, Pinecone, Chroma DB, FAISS
- **Platforms:** PyCharm, Jupyter Notebook, Visual Studio Code, IntelliJ IDEA
- **Soft Skills:** Rapport Building, Strong Stakeholder Management, People Management, Excellent Communication

CERTIFICATES

- **AI Engineering:** [Link](#)
- **IBM Data Science:** [Link](#)
- **IBM Machine Learning:** [Link](#)
- **Deep Learning Specialization:** [Link](#)
- **Generative AI for Data Scientists:** [Link](#)
- **Machine Learning Specialization:** [Link](#)