SAJJAD ALI SHAH

+92 316 9671878 | sajjadkhanyousafzai47@gmail.com | LinkedIn | https://github.com/SajjadKhanYousafzai

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 & Discrete Mathematics

Army Public School & Degree College Malir Cantt Karachi **Mar 2018 - Aug 2020**
FSc, Pre Engineering (GPA: Grade: Mathematics (A+), Chemistry(A-), Physics(A).)

WORK EXPERIENCE

Digital Empowerment Network **Sep 2024 - Oct 2024**
Machine Learning Intern *Islmabad Pakistan*
- Developed machine learning models for classification and regression tasks using Python libraries including TensorFlow and Scikit-learn, achieving up to 93% accuracy on structured datasets while employing Pandas for data manipulation.
- Executed data preprocessing, feature engineering, and model evaluation to optimize predictive performance and synthesize actionable insights.
- Collaborated with cross-functional teams to iteratively refine analytics models, leveraging SQL and Excel for detailed data analysis and validation.

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, Java, C++, SQL, HTML, CSS,
- **Deep Learning:** TensorFlow, Keras, PyTorch, Hugging Face Transformers, OpenAI API, LangChain, LlamaIndex, CNNs, RNNs, LSTMs, Attention Mechanisms, Transformers, Fine-tuning Large Language Models (LLMs), Prompt Engineering, Retrieval-Augmented Generation (RAG)
- **NLP & Computer Vision:** NLTK, SpaCy, OpenCV, MediaPipe
- **Tools:** Power BI, Excel, PowerPoint, Tableau, MySQL, SQLite, Git, Docker, AWS
- **Frameworks:** FastAPI, Flask, Pandas, Numpy, Scikit-learn, Matplotlib, Seaborn, Statsmodels
- **Platforms:** PyCharm, Jupyter Notebook, Visual Studio Code, IntelliJ IDEA
- **Soft Skills:** Rapport Building, Strong Stakeholder Management, People Management, Excellent Communication

CERTIFICATES

- **AI Engineering:**: IBM
- **IBM Data Science**: IBM
- **IBM Machine Learning**: IBM
- **Deep Learning Specialization**: Deep Learning.AI
- **Generative AI for Data Scientists**: IBM
- **Machine Learning Specialization**: Deep Learning.AI & Stanford University