+92 334 7777081
Lahore, Pakistan
faridkhan7777081@gmail.com

Farid Khan

Data Scientist/ ML Engineer

kaggle/farid75
github/faridkhan081
linkedin/faridkl

Passionate Machine Learning Engineer currently working at SDSol Technologies. Skilled in Deep Learning, Computer Vision, Natural Language Processing, and Data Science, Generative-AI and Large Language Models (LLMs) with a strong commitment to research and development.

Skills

Programming Languages	Python, R (Basic), SQL
Generative AI	LangChain, LlamaIndex, OpenAI , HuggingFace, Gemini
Deep Learning	Keras, TensorFlow, OpenCV, FastAI, PyTorch, Transformers, NLTK
Machine Learning	Numpy, Pandas, scikit-learn, Matplotlib, Seaborn
Tools	Linux, VS Code, Jupyter-Notebook, Google Colab, Kaggle
Technologies	Git, FastAPI, Gradio, Streamlit
Soft Skills	Time Management, Leadership, Communication, Team Work

Technical Experience

Machine Learning Engineer **June 2024 — Present**
SDSol Technologies *Lahore, Pakistan*

- Developed a deep learning pipeline for classifying aerial images using datasets such as AID, UC Merced, and RESISC45 (12 classes). Integrated models like ResNet, EfficientNet, DenseNet, SqueezeNet, and ConvNeXt, optimizing performance with minimal hyperparameter tuning.
- Created a Label Studio ML backend using YOLOv8 to automate the image labeling process, enhancing the efficiency and accuracy of annotation workflows.
- Implemented sentiment analysis on the Roman Urdu Sentiment Analysis Dataset (RUSAD) using NLTK for preprocessing, Sentence Transformers for embeddings, and SVM for sentiment classification.
- Designed and developed APIs with FastAPI for various machine learning models, including regression, image classification, and YOLOv8-based detection models. Ensured scalable and efficient deployment in production environments.
- Generated numerical datasets using the Gemini API for various machine learning applications, developing synthetic data solutions to enhance model training and evaluation.
- Developed a querying agent that retrieves data from a website and uses Retrieval-Augmented Generation (RAG) techniques combined with GPT-Neo models to generate detailed and contextually relevant responses.
- Developed a real-time route optimization API using OpenAI SDK for optimal delivery methods. Utilized Geopy for latitude/longitude calculations and deployed the solution using FastAPI.Also created an interactive map using Folium to visualize driving routes and stop locations. Integrated the OSRM API to calculate and display estimated travel times from a base location to each stop.
- Developed "Auto Essentials," an AI-powered multi-vendor e-commerce platform for the auto industry, featuring a tire health predictor, community support chatbot, and user-friendly dashboards for administrators and sellers. Implemented advanced AI solutions as part of this project. **(Final Year Project at CUI)**

Education

Bachelor of Computer Science, *COMSATS University* September 2020 — July 2024
Talent Hunt Merit Scholarship, COMSATS University *CGPA: 3.35*
Relevant Courses: Machine Learning, Artificial Intelligence, Data Science, Computer Vision, Bioinformatics

Certifications

- **Machine Learning Specialization**, Stanford Online — October 2023
- **AI for Everyone**, Stanford Online — August 2023
- **Neural Networks and Deep Learning**, Stanford Online — April 2024
- **Generative AI with LLMs**, Stanford Online — July 2024

References

- **Muaaz Hafiz**
  Project Manager, SDSol Technologies
  Email: muaaz@sdsol.com
- **Qaiser Abbas**
  Senior Machine Learning Engineer, SDSol Technologies
  Email: abbas.qaixer@gmail.com