DAWOOD SARFRAZ

Machine Learning Engineer

📞 +923061757838 ✉ dawoodsarfraz.cs@gmail.com in Dawood Sarfraz 🌐 DawoodSarfraz 📍 Lahore, Pakistan

EDUCATION

FAST National University of Computer and Emerging Sciences

Bachelor's in Computer Science *Sep 2020 - Sep 2024*

SKILLS

Languages: Python, C++, SQL
Tools: Git, Postman, Docker, Kubernetes, Jenkins, AWS
Libraries: FastAPI, Flask, Streamlit
ML Libraries & Frameworks: PyTorch, scikit-learn, TensorFlow, Keras, NumPy, Matplotlib, SciPy, Pandas, Seaborn, NLTK, spaCy, OpenCV
GenAI Libraries & Frameworks: LangChain, Weaviate DB, RAG, Crew AI, Llama3, LlamaIndex

EXPERIENCE

FAST NUCES   Sep 2023 – Sep 2024

Machine Learning Engineer (Research Intern) *On-site*

- Worked on a research project focused on classifying skin cancer using CNN, ShuffleNet, and NasNet models, gaining experience in medical data processing and deep learning architectures.

Anonymous Tree   July 2023 – Aug 2023

Machine Learning Engineer Intern *Remote*

- Developed a recommendation system using user data, item metadata, and techniques like content-based, collaborative filtering for personalized product suggestions.

PROJECTS

Skin Cancer Classification using NasNet and ShuffleNet

- Developed and optimized deep learning models (Custom CNN, NasNet, ShuffleNet) for multi-class skin cancer classification using the HAM10000 dataset. Improved model accuracy and stability through batch normalization, dropout layers, and Adamax optimizer. Achieved highest classification accuracy with NasNet, while ShuffleNet enhanced computational efficiency using grouped convolutions and channel shuffling.

LlamaAssist [Project Link]

- LLama Assist is an AI-powered application built with LLama 3.2 3B and LLama 3.2 Vision. It allows users to generate assignments with customizable output length and difficulty, save content into text files, and interact with images via AI-driven chat. Perfect for creating and managing assignments with ease and precision.

PaperScope [Project Link]

- PaperScope is an AI chatbot utilizing RAG, Mistral 7B, and Langchain to analyze and summarize academic papers. It leverages WeaviateDB for efficient document retrieval and context-aware responses.

RoboText Classifier [Project Link]

- Built a text classification model using RoBERTa and NLTK on a 500k-row dataset. Enhanced performance with dynamic masking, sentence packing, byte-level BPE vocabulary, and larger batch sizes for efficient, accurate classification of diverse text.

Enhancing Medical Education through Immersive VR [Project Link]

- Developed a VR-based medical training system with haptic feedback for realistic simulations, enhancing surgical training for medical students. Aimed to reduce costs and ethical concerns associated with traditional methods.

Duplicate Questions Pair [Project Link]

- Built a model to identify and detect duplicate question pairs using Random Forest, XGBoost, and Decision Tree classifiers. Achieved 90% accuracy with XGBoost classifier.

Text Generation [Project Link]

- Processed a 500MB unlabeled dataset with almost 3000 unique words and 1000 sentences. Applied data preprocessing techniques and used LSTM with Adam optimizer, achieving 93% accuracy.