Dawood Sarfraz

Machine Learning Engineer

+923061757838 dawoodsarfraz.cs@gmail.com 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, MATLAB
Tools: Git, Postman, Docker, Kubernetes, Jenkins
Libraries: FastAPI, Flask, Streamlit
ML Libraries and Frameworks: PyTorch, scikit-learn, TensorFlow, Keras, NumPy, Matplotlib, SciPy, Pandas, Seaborn, NLTK, spaCy, OpenCV, Langchain

Experience

Research Assistant **Sep 2023 – Sep 2024**
Machine Learning Engineer *Remote*

- 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*

- Worked as a Machine Learning Engineer, assisting beginners in learning core concepts of Machine Learning.

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.

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 80% accuracy with XGBoost classifier.

Text Generation Project Link

- Processed a 500MB unlabeled dataset with almost 3000 unique words and 800 data entries. Applied data preprocessing techniques and used LSTM with Adam optimizer, achieving 93% accuracy.

Cyber Attacks Classification using Machine Learning Project Link

- Developed a machine learning model to classify and identify different types of cyberattacks. Preprocessed and transformed data for model use. Achieved 93% accuracy using MLP.