

Sayed Hanan

📞 (+92) 343 4630086

✉ hananhaider.pk@gmail.com

🌐 linkedin.com/in/sayed-hanan

🐙 github.com/sayed-hannan

Education

Virtual University

Bachelor of Science in Software engineering

Expected September 2025

Faislabad, Pakistan

- **Relevant Coursework:** Data Structures and Algorithms (C++), C++ Language, Probability & Statistics, Dataware housing
- **Enrolled Coursework:** Cloud Computing, Software Project Mangement

Projects

Some ongoing projects |

- Machine Translation Urdu to English: built everything from scratch (model implementation, tokenizer, dataset scraping with 30,000 articles scraped from Wikipedia and translated; tokenizer based on Unigram).
- Retrieval Augmented Generation: building a research assistant using the arXiv dataset from Hugging Face(completed first iteration of 1000 samples).
- Fine-tuning a MiT-a vision transformer for segmenting the human body into different parts (semantic segmentation) complete project.

Churn Prediction with the Interactive Dashboard (Personal Project) | *Python (NumPy, Pandas), Scikit-learn, Streamlit*

- Developed a machine learning model to predict customer churn using transaction data from Kaggle.
- Achieved an 86% accuracy in predicting customer churn.
- Implemented various techniques including Logistic Regression, Random Forest, and ensemble methods using Python.
- Created an interactive dashboard with Streamlit to present results and provide actionable insights for enhancing customer retention strategies.
- [GitHub Link](#) [Dashboard Link](#) [Docker Image](#)

Email Classification (Spam and Ham) | *Scikit-learn, MLflow, NLTK, Docker*

- Developed a high-accuracy (96.77%) email classification model using Multinomial Naive Bayes to categorize emails as spam or legitimate (ham). Explored effective methods for email filtering and enhanced my understanding of machine learning and text processing (using Python, NLTK, and a TF-IDF vectorizer).
- Implemented the project as a Python script and containerized it using Docker for streamlined deployment. Additionally, leveraged DVC for version control and MLflow for model tracking.
- [GitHub Link](#) [Dashboard Link](#)

Tumor Classification Project | *TensorFlow, Keras, Matplotlib, TensorBoard, Google Colab*

- Leveraged transfer learning (VGG-16) to achieve 89% accuracy on a validation dataset for brain tumor classification using MRI scans (Kaggle dataset).
- Improved classification performance from 50% to 90% by addressing overfitting in the baseline model.
- Utilized TensorBoard to visualize model training and identify potential issues (e.g., overfitting).
- [GitHub Link](#)

GPT-2 Fine Tuning | *Hugging Face, GPT-2 (pre-trained), PyTorch, Wandb*

- Leveraged WARP steps for fine-tuning a pre-trained GPT-2 model using the Hugging Face Transformers library (ALPCA dataset).
- Utilized Weights and Biases (Wandb) to visualize training progress (loss curve).
- Conducted human evaluation (documented in the project README) to assess the model's ability to generate contextually relevant and engaging responses.
- [GitHub Link](#)

Technical Skills

Languages: Python

Technologies: TensorFlow, PyTorch, Keras, Hugging Face, LangChain, Scikit learn, Wandb, Mlflow, Docker, Django, DVC, Kubernetes

Concepts: Artificial Intelligence, Deep learning, Machine learning, AI system design(RAG, Agentic), GenAI(stable diffusion, LLMs) MLOps, DevOps, Agile Methodology, Software Development Life Cycle, Cloud Computing

Research Interests: Reasoning in LLMs, Contrastive learning, Chain of Thought, Multi-Model, CLIP(Multi-Modality)

Certifications

- Advanced Learning Algorithms by DeepLearning.ai
- Supervised Machine Learning: Regression and Classification