

Syed Muhammad Taha Imam

Islamabad | simam.bsccs22seecs@seecs.edu.pk | 03340518269 | [linkedin.com/in/syed-muhammad-taha-imam](https://www.linkedin.com/in/syed-muhammad-taha-imam)
github.com/Syed-MuhammadTaha

Skills

Transformer Models | NLP | Computer Vision | Deep Learning | Advance TensorFlow | CI/CD pipelines | Prompt Engineering | LLM Fine-Tuning | Multi-Agent workflows | RAGs | Optimization Techniques | Distributed Training strategies | Open-source LLMs Inference

Experience

Machine Learning Intern, RapidsAI – SINES, NUST Sept 2024 – Present

- Integrated RAG-based chatbot on Kalambot website with Streamlit UI hosted on Streamlit Cloud.
- Improved responses by adding prompt engineering techniques like COT and contextual prompting
- Cut OpenAI API costs by 50% with a RAG-based model for specialized responses and a lightweight model for generic queries.
- Designed FastAPI endpoint for chatbot responses, enabling user session management and contextual interactions.
- Executed inference on specialized open-source models for protein sequence analysis and interaction-based queries.

Projects

Attention Is All You Need paper implementation github.com/transformer

- Trained on the Shakespeare dataset containing 1 million characters.
- Utilized character tokenization to generate positional and token embeddings for sequence representation.
- Developed a baseline Bigram model for initial evaluation and benchmarking.
- Integrated multi-head attention with masking to generate weights while hiding future tokens.
- Improved the model further by incorporating layer and batch norm, residual connections and feed forward layers.
- **Tools Used:** Tensorflow, Numpy

CI/CD pipeline for HuggingFace spaces github.com/mlops

- Developed a basic text summarizer using transformers and streamlit
- Deployed the app on HuggingFace spaces
- Engineered a CI/CD pipeline using Github Actions to build and push file to HuggingFace spaces on updates to GitHub repository
- **Tools Used:** Github Actions, HuggingFace, Streamlit

Twitter Sentiment Analysis

- Utilized a pre-labeled dataset of 100,000 tweets with sentiments for sentiment analysis.
- Improved model performance through robust pre-processing techniques such as tokenization, stop-word removal, and lemmatization.
- Leveraged pre-trained word embeddings to project words into vector space for enhanced feature representation.
- Integrated attention blocks into an LSTM architecture with 150 units, achieving a 90% F1 score.
- **Tools Used:** Tensorflow, Pandas, Matplotlib

Education

National University of Science and Technology, BS in Computer Science Sept 2022 – Jun 2026

- GPA: 3.48/4.0
- **Coursework:** Linear Algebra, Probability and Statistics, Machine Learning