Afnan Mir

Mobile: 469-703-9636
afnanmir@utexas.edu

Github: github.com/afnanmmir
Portfolio: afnanmmir.github.io
LinkedIn: linkedin.com/in/afnan-mir/

Education

- **University of Texas at Austin** — Austin, TX, United States
  Bachelor of Science - Electrical and Computer Engineering; GPA: 4.00 — *August 2020 - May 2024*
  Courses: *Algorithms, Data Science Laboratory/Principles, Computer Vision, Natural Language Processing, Operating Systems, Data Structures, Probability, Linear Algebra, Edge Machine Learning and Artificial Intelligence*

Experience

- **Amazon** — Seattle, WA, United States
  Software Development Engineering Intern — *May 2022 - Aug 2022*
  - Designed and built log analysis tool which ingests, parses, and processes 32K+ logs per second
  - Used Java, TypeScript, Logstash, Elasticsearch, and AWS to provide a dashboard that allows engineers to search/filter through logs based on fields and create visualizations based on log data.
  - Deployed tool internally to reduce shadow validation and debugging time for team from hours to minutes.
- **Applied Research Laboratories** — Austin, TX, United States
  Machine Learning Research Intern — *Jun 2021- Aug 2021*
  - Finetuned natural language processing models such as BERT for the task of named entity recognition using PyTorch and HuggingFace
  - Designed and executed experiments to evaluate the robustness and bias of premade language models in the task of named entity recognition.
  - Designed and executed experiments to evaluate the robustness and bias of premade language models in the task of named entity recognition which later contributed to the GEM Natural Language Augmenter paper

Projects

- **UT Course Search (Python, Vector Databases, LLMs, LlamaIndexing)**: Built a semantic search engine for UT courses using OpenAI embeddings API and Pinecone Vector Database. Additionally, used LlamaIndexing with ChatGPT API to provide users with a chat-like answer to their queries about courses.
- **LoFi Music Generator (Python, TensorFlow, Deep Learning, Generative Models)**: Used deep learning architectures such as Recurrent Neural Networks (RNNs) and Variational Autoencoders (VAEs) to generate music in the style of lofi hip hop. This model was trained to predict the next note in a sequence of notes given the previous notes in the sequence, and this data was found in a github repository of lofi hip hop samples.
- **Basketball Shot Predictor (Python, Tensorflow, Computer Vision)**: Trained a gradient boosted decision tree model to predict the outcome of a player's free throw. Feature vector inputs were generated using deep learning based pose estimation to extract the keypoints of the player's body during the shot. Model achieved an accuracy of 0.726 and ROC-AUC score of 0.794.
- **Research Paper Summarizer (Python, Natural Language Processing, LLMs)**: Finetuned a base T5 model that would take technical abstracts from research papers and generate a readable summary of the paper. The model was finetuned using the HuggingFace transformers library and the Trainer API, and it achieved a ROUGE-1 score of 0.403590 and a ROUGE-2 score of 0.124948. A demo of this was also created using Streamlit.

Organizations

- **Longhorn Racing Solar** — *Power Generation Lead* — *Oct 2020 - Present*
  - Researched, implemented, and modified various candidate maximum power point tracking (MPPT) algorithms in the array simulator and helped transcribe these algorithms into the firmware of a maximum power point tracker.
  - Helped design and maintain an array simulator made in python, which is used to test and optimize MPPT algorithms.
  - Oversee a team of 5+ members in the design and manufacturing process of solar cell modules and their accompanying electronics.
- **Machine Learning and Data Science Organization** — *Member* — *January 2021 - Present*
  - Participated in annual MLDS competition. Used CNN to classify images as characters of a text document
  - Participated in weekly meetings to discuss various ML topics as well as reading groups for new ML papers

Publications

- **NL-Augmenter: A Framework for Task-Sensitive Natural Language Augmentation (Research Paper)**: Submitted contributions to NL-Augmenter paper to provide language augmenting experiments to extensively test large language models in various tasks

Honors and Awards

- Engineering Honors Scholarship - Fall 2020
- Recipient of Fall 2020 Undergraduate Research Fellowship Grant for research in MPPT algorithms - December 2020
- Distinguished Honors Scholar 2x - 2022, 2023

Skills

- **Languages**: Python, Java, JavaScript/TypeScript, C/C++, HTML/CSS, LaTeX
- **Frameworks**: Scikit, SpaCy, TensorFlow, Flask, NodeJS, React, PyTorch, HuggingFace
- **Tools**: AWS, Git, BASH, NumPy, Pandas
- **Interests**: Machine Learning, NLP, Fitness, Basketball, Music, TV Shows