MOHAMED FAHEEM M

✉ mdfaheempasha@gmail.com   in Mohamed Faheem M   ⊙ MohamedFaheemM

Education

Amrita Vishwa Vidyapeetham   **Oct 2021 – May 2025**

Bachelor of Technology, Computer Science and Engineering (Artificial Intelligence) - CGPA: 8.73   *Chennai, Tamil Nadu*

Experience

GSSoC | Link   **May 2023 – August 2023**

Open Source Contributor   *Remote*

- Contributed to the integration of Parkinson's Disease Metadata into the main project repository, enhancing the dataset's ability to provide insights for patients with Parkinson's Disease.

Projects

Deepfake Detection in Tamil Speech using Hybrid CNNs | *SkLearn, TF/Keras, Librosa* | **Link**   **Apr 2024**

- Assembled a diverse Tamil speech dataset, acquiring 1160 samples per class from public repositories and online platforms.
- Enriched the dataset with deepfake samples, crafting an additional set of 1160 synthetic audio recordings using Eleven Labs and Open-Audio TTS Powered by OpenAI.
- Extracted high-quality features from 2320 audio samples using advanced techniques such as MFCC, Mel spectrogram, and Chromagram; resulted in a 20% improvement in audio classification accuracy.
- Employed CNN-LSTM, CNN-GRU, and SVC models, yielding a CNN-GRU accuracy of 97% with MFCC features.

Defending LLMs using Multi-Agent Debate | *PyTorch, NLTK, Transformers* | **Link**   **Apr 2024**

- Researched and mitigated language model vulnerabilities to adversarial attacks during inference, reducing harmful content through implementation of a multi-agent debate methodology for enhanced content safety.
- Implemented extensive experimentation and analysis, and validated using Perspective API to evaluate toxicity scores.
- Demonstrated superior performance with a 20% increase in overall model accuracy, indicating the efficacy of multi-agent debate in enhancing language model robustness and suggesting further exploration for better safeguards.

Real-Time Helmet Detection & Plate Extraction | *Ultralytics, Detectron2, Roboflow* | **Link**   **Apr 2024**

- Developed and curated a balanced dataset of 5 classes through manual web scraping and segmentation.
- Analyzed the performance of YOLOv8 segmentation model, achieving a precision (P) of 0.975.
- Conducted performance analysis of the Mask R-CNN model, resulting in an accuracy of 0.968.
- Integrated EasyOCR to extract license plate data from non-helmeted riders, streamlining real-time processing and storage on Google Firebase.

Flora Snap: Plant Disease Detection App | *Flutter, Google Gemini API* | **Link**   **Aug 2024**

- Developed a Flutter application that allows users to diagnose plant diseases by analyzing images of their plants using the Google Gemini API.
- Implemented features such as image capture, gallery upload, and text-based queries to facilitate easy disease diagnosis and plant health management.
- Integrated automatic disease detection and personalized care recommendations to enhance user experience and plant management.

Publication

Comparative Evaluation of Large Language Models for Abstractive Summarization

Confluence-2024:14th International Conference on Cloud Computing, Data Science and Engineering   *Jan 2024*

Certifications

Ask Questions to Make Data-Driven Decisions   *Coursera - Google*

Skills: Decision-Making, Spreadsheet, Data Analysis, Problem-Solving, Questioning ● View Certificate

Foundations: Data, Data, Everywhere   *Coursera - Google*

Skills: Spreadsheet, Data Analysis, SQL, Data Visualization, Data Cleansing ● View Certificate

Achievement

- Ranked among the top 63 teams out of 2000 in the TiDB Future App Hackathon 2023, conducted globally by PingCap.
- Runner up in Cryptoshield 2024 National Level Hackathon sponsored by SkillsDA.

Technical Skills

Domain: Machine Learning, Deep Learning, Computer Vision, NLP, AI
Languages: Java, Python, SQL, C
Frameworks: PyTorch, TensorFlow, Keras, Scikit-learn, NLTK, Transformers
Developer Tools: Git, AWS, VS Code, Matlab