Disha Thotappala Jayaprakash

thotappa@usc.edu | (408) 614-1508 | Los Angeles, CA| www.linkedin.com/in/dishajayaprakash/

EDUCATION

University of Southern California, Los Angeles August 2023 – May 2025
Master of Science (MS) in Electrical and Computer Engineering – Machine Learning and Data Science GPA: 3.9/4.0
Coursework: Natural Language Processing, Data Management, Digital Image Processing, Computational Introduction to Deep Learning, Machine Learning I: Supervised Methods, Linear Algebra, Probability

Dayananda Sagar College of Engineering, Bengaluru July 2019 - June 2023
Bachelor of Engineering (BE) in Electronics and Telecommunication GPA: 9.42/10
Coursework: Artificial Intelligence & Machine Learning, Data Science, Python/Java Programming, Complex variables and distribution, IoT and cloud computing, Information theory and coding, Digital Signal Processing

SKILLS

Languages: Python, Java, R, C/C++, MATLAB, VHDL, MySQL, Bash/Shell Scripting, JavaScript, PostgreSQL
Machine Learning & AI: TensorFlow, PyTorch, FastAI, TVM, LangChain, OpenAI, spaCy, NLTK, AllenNLP, Librosa, torchaudio, SpeechBrain, AWS SageMaker, MLflow, Kubeflow, Apache Spark, GCP, Snowflake, scikit-learn, SeamlessM4T, DeepSpeech
Software Development & Deployment: Flask, FastAPI, GitHub Actions, Docker, Kubernetes, CUDA, MongoDB

RESEARCH EXPERIENCE

NLP Graduate Research Assistant September 2024 - Present
Under Dr Yao Du, USC Keck School of Medicine

- Developing an AI-driven MERLS assessment system with speech emotion recognition for parental behavior classification
- Researching multilingual models and speech processing to enhance behavioral coding in Mandarin-English interactions
- Using speech-to-text models and LLMs for automated learning disorder detection in bilingual children, while reviewing literature and testing speech LLMs like Wav2Letter, DeepSpeech, Wav2Vec, and SeamlessM4T

WORK EXPERIENCE

Machine Learning Engineer Intern June 2024 – July 2024
WorkUp

- Developed AI agents that autonomously search and apply for jobs by parsing job descriptions using NLP techniques, matching them with user profiles for better alignment
- Analyzed speech prosody (pitch, tempo, intonation) using Praat and Librosa to extract linguistic features, improving the performance of an AI interview assistant through NLP-driven speech recognition
- Developed a recommendation system on AWS SageMaker using Nvidia Merlin, to deliver customized job matches and career advice based on user profiles on a TikTok style-based job application platform

Machine Learning Intern September 2022 - November 2022
DigiAdd Technologies

- Engineered a RNN using TensorFlow to analyze temporal flight data with an accuracy of 95% in predicting airport delays and reduced dependency on radar systems by 80%
- Conducted analysis of over 1 million flight records spanning five years through Power BI, extracting key patterns and insights and refined airport delay management algorithms
- Employed content-based filtering on a preprocessed dataset of 50k song lyrics using Pandas and NumPy to improve music system recommendations based on individual user preferences

ACADEMIC PROJECTS

Does Self-Feedback Improve Fairness in LLM Abstractive Summarization? August 2024 – Jan 2025

- Investigated if language models can improve fairness in summaries using self-feedback or external feedback
- Developing a new fairness metric to help models evaluate their summaries
- Compared self-feedback with external feedback to see which method better improves fairness in summarization

DefectDetectPro: Leveraging AI in Composite Analysis February 2024 – June 2024

- Achieved 98% accuracy using ResNet-50 and Inception-v3 (CNN) for microfracture detection in composites from a dataset of over 10,000 images
- Produced synthetic images using GANs to increase dataset diversity, leading to a 10% improvement in training model robustness

Stutter Classification November 2023 - December 2023

- Created an AI system to detect stutter in speech using SEP-28k and Fluency Bank datasets
- Utilized wav2vec embeddings and MFCC for feature extraction and transformed speech into a spatiotemporal format
- Adapted ResNet model and implemented ConvLSTM using librosa and Praat and achieved an F1 score of 0.71