VENKATA SAAI PRANEETH THOTA

Los Angeles, California, US
https://github.com/ProPranu6 | (323) 640-9590 | thotav@usc.edu | https://www.linkedin.com/in/pranuthota/

EDUCATION

University of Southern California, *United States***.** *Aug 2023-May 2025*

Masters of Science in Computer Science - *GPA : 3.95/4.0*

- Coursework : Foundations of Artificial Intelligence, Machine Learning, Applied Natural Language Processing

SRM University, *India.* *May 2019-May 2023*

Bachelor of Technology in Computer Science and Engineering - *GPA : 9.72/10.0 - Department Gold Medalist*

- Specialisation : Artificial Intelligence & Machine Learning

CORE COMPETENCIES

Tools & Frameworks : Tensorflow, Keras, PyTorch, Scikit-Learn, Python, Pandas, Numpy, Django, Flask, MySQL, MongoDB, EEGLAB, MNE-Python, HTML, CSS, JavaScript, Bootstrap, OpenCV, Git, AWS, AutoML, MLFlow, Langchain, HuggingFace, CrewAI, PyAutoGen

EXPERIENCE

ML Engineer, *Modlee Eval.* *Jan 2025-Present*

Modlee Inc, New Haven, CT, US

- Contribute to backend development, machine learning pipelines, and dataset R&D, ensuring efficient data handling and model evaluation workflows.
- Develop and optimize large language models (LLMs) and deep learning (DL) architectures for domain-specific applications.

AI Engineer, *Share Labs.* *Jul 2024-Aug 2024*

Share Ventures, Los Angeles, CA, US

- Developed AI agents providing investment insights in the human performance industry, analyzing 100+ articles across 5 platforms using Hybrid-RAG with GPT-4o and Llama 3.1 LLMs.
- Created web interface for personalized investment recommendations with 10 KPIs and 4 filters, enabling quick market analysis through hybrid information retrieval systems

Research Assistant, *Big Data in Biotransport Center* *Oct 2023-Present*

University of Southern California, Los Angeles, CA, US

- Research techniques for identifying and verifying CVD biomarkers in retinal vasculature from retinal fundus and conjunctival images, focusing on extraction methods and association verification
- Designed TransU-Net for artery/vein segmentation and GCNNs for conjunctival vasculature segmentation, achieving ~75% validation precision and ~72% recall using multiple training configurations on multimodal inputs

Graduate Student Researcher, *Center for Artificial Intelligence in Society* *Nov 2023-Jun 2024*

University of Southern California, Los Angeles, CA, US

- Conducted NIH-funded research for predicting PTSD events using time series analysis and machine learning techniques
- Constructed preprocessing pipelines for aggregating digital phenotypes from 75 users' wearable devices over 84 days and developed algorithms for resolving duplicate intervals and merging data from 100+ streams.

PROJECTS

MusicGen, Machine Learning Course Project, USC *Jan 2024-May 2024*

- Developed LSTM, BiLSTM, and transformer based sequence-sequence architectures trained on 2.64M polyphonic human compositions encompassing 5 varied types of instruments, genres and artists for autoregressive generation
- Boasts 60% overlap with original composition in pitch type and pitch range histograms, 70% overlap in octave histogram and pitch transition matrix

Canny.ai - Autonomous Subjective Answer Grading System *Jan 2024-May 2024*

- Engineered Retrieval Augmented Generation (RAG) pipeline,using Microsoft's Phi-2 LLM to generate coherent answers
- Programmed unsupervised evaluations of subjective answers contextually embedded using DistilBERT, Paraphrase-MiniLM.
- Proposed Model has a MAE of 0.94 and Correlation of 0.69 with ground truth scores of 2700+ long form answers

PUBLICATIONS AND PRESENTATIONS

- Sharma, N., Thota, V. P., Yuvaraj, T., Tripathi, S., & Pandey, O. J. (2024). OptRISQL: Towards Performance Improvement of Time-Varying IoT Networks Using Q-Learning. IEEE Transactions on Network and Service Management (IEEE TNSM)
- Jury award from Chief Coordinating Officer of All India Council for Technical Education, Government of India, for research presentation on Vehicle Speed Estimation Using ConvLSTM CNNs at SRM University Research Day 2022