PRANAV KALE

Los Angeles, California | 213-272-3490 | pakale@usc.edu | https://www.linkedin.com/in/pranav-kale

SUMMARY

2+ years of professional and research experience across Natural Language Processing, Computer Vision, Machine Learning, Data Science, Generative AI and Software Development. **Seeking Fulltime opportunities starting May 2024 / Coop (Spring 2024).**

EDUCATION

University of Southern California (USC) **Los Angeles, California**
Master of Science (MS) in Computer Science August 2022-May 2024
Coursework: *Analysis of Algorithms, Deep Learning, Machine Learning, Database Systems, Applied NLP, Information Retrieval*

University of Pune **Pune, India**
Bachelor of Technology (B.Tech) in Computer Engineering August 2017-May 2021

TECHNICAL SKILLS

- Programming Language: Python, C, C++, Flask, Django, TensorFlow, PyTorch, Power BI, MySQL, NLTK, Pandas
- Environments: GitLab, Linux, AWS, Microsoft Azure, GCP, Docker, OpenCV, Scikit learn, LangChain, Streamlit

PROFESSIONAL EXPERIENCE

Ford Motor Company, GDIA, AI Advancement Center **Dearborn, Michigan**
AI-ML Intern May 2023-August 2023

- Conducted in-depth assessment and experiments with 5+ **evaluation** metrics for **open-source large language models (LLMs)** such as Vicuna, Llama, Falcon for **question answering task**.
- **Created** a clean vector database of **60,000 Q&A pairs** from vehicle owner manual using **Vicuna-13B** which proved to be pivotal resource enabling collaboration between cross-functional team to validate other LLMs.
- **Finetuned LLMs** on domain specific use-case using proprietary Q&A pairs, to provide accurate responses **without external references**. Showcased **7.8x accuracy** of finetuned model compared with base model.

bizAmica Software Pvt. Ltd. (TiE'50, Silicon Valley Award Winner Mid-Scale AI Startup) **Pune, India**
Machine Learning/ Data Scientist Engineer January 2021-June 2022

- Designed and deployed **model training APIs** hosted on **AWS EC2 platform** with 8+ configurable parameters having auto start/stop functionality. Resulted in 25% cost benefits with end-to-end customer solution.
- Built **on-premise** product for data privacy with integrating customized trained **BERT** transformer model for 99.3%+ performance on **unstructured datasets** for banking sectors to keep **data confidential** utilizing Microsoft Azure on-premise service.
- Led model development, testing, validation, and optimization of **3+** deep learning models **improving accuracy and efficiency** by 35% with 50% **reduced inference time** by implementing Spacy and Transformer models for **NLP tasks**.

RESEARCH EXPERIENCE

USC Information Science Institute – Visual Intelligence Multimedia Analytics Lab **Los Angeles, California**
Machine Learning Research Assistant September 2022-Present

- Architect **end-to-end pipeline** for **temporal video segmentation** to extract information from presentation video with summarization module for speech. **Project funded by NSF, DARPA**.
- Research and develop **state-of-art Video Vision Transformer model** to temporarily segment frames with 97.89% accuracy by using **spatio-temporal tokens with positional frame embeddings**. Research Paper publication in progress at **NeurIPS**.
- Successfully developed, trained, and tested a multi-modal transformer model on a 2300-slide presentation, **improving f1-score** by 20% leveraging **textual and image embeddings**.

ACADEMIC PROJECTS

Traffic Sign Classification with Speech Output using CNN June 2020-June 2021

- Classified 43 Sign boards (Speed Limit, Turn Signs, Stop) using convolutional neural networks; built a GUI with speech output.
- Achieved a 2x speedup in terms of execution time and 3x better memory utilization when benchmarked on GTSRB dataset**.**
- Granted an Innovation Patent and presented a research paper in an international journal.

PUBLICATIONS

- Pranav Kale," A STREAMLINE TRAFFIC SIGN CLASSIFICATION SYSTEM UTILIZING CONVOLUTIONAL NEURAL NETWORK MODEL", Innovation Patent, IP Australia, Patent number: 2021101273, April 21, 2021.
- Pranav Kale, "Traffic Sign Classification Using Convolutional Neural Network", International Journal of Scientific Research in Computer Science (IJSRCSEIT), ISSN: 2456-3307, Volume 7, Issue 6, pp.01-10, https://doi.org/10.32628/CSEIT217545