Nikhil Chowdary Paleti

510-935-8895 | nikhilpaleti23@gmail.com | nikhil-paleti.github.io | linkedin.com/in/nikhil-paleti | github.com/Nikhil-Paleti

OBJECTIVE

Ambitious Data Science master's candidate with **6 published papers** in AI/ML. Specializing in **NLP**, **CV**, and **Generative AI**, I possess hands-on experience in building and deploying **scalable AI systems** on the cloud. I thrive in collaborative environments and seek to leverage my skills in a Summer 2025 internship to advance cutting-edge projects.

EDUCATION

University of California San Diego Sep 2024 – Jun 2026 (Expected)
Master of Science in Data Science

Amrita Vishwa Vidyapeetham University Oct 2020 – Jun 2024
Bachelor of Technology in Computer Science and Engineering (Artificial Intelligence) ***GPA: 9.15/10***

TECHNICAL SKILLS

Languages: Python (Advanced), SQL (Advanced), C++ (Intermediate).
Libraries: NumPy, Pandas, Matplotlib, Scikit-Learn, PyTorch, TensorFlow, Transformers.
Machine Learning: Supervised/Unsupervised Learning, Reinforcement Learning, Deep Learning.
Specializations: Natural Language Processing, Computer Vision, Generative AI (Large Language Models).
Tools & Cloud: Docker, Git, AWS, Google Cloud Platform (Vertex AI, Compute Engine, Vision API).

EXPERIENCE

Machine Learning Engineer Intern (Team Lead) Jan 2024 – Jun 2024
TechProfuse *Hyderabad, India*
Tech Stack: Python, Google Vision API, OpenCV (CV2), Large Language Models (LLM), Git, Docker

- **Led a team of 3** interns to convert English text to Dutch in scientific images using Google Vision API, CV2, and an LLM, ensuring scientific accuracy and increasing translation efficiency by **30%** to the client.
- Developed an **unstructured data extraction system** using Gemini 1.5 Pro LLM, processing **50k** bill of lading documents in **15 hours**, which reduced manual data entry by **96%** and streamlined client operations.
- Automated **email classification and summarization** using the Gemini 1.5 Flash model, enhancing support team efficiency by **40%**, reducing manual sorting time, and streamlining response workflows.

Research Assistant Jan 2023 – Sep 2023
Dr. Sowmya's Lab *Coimbatore, India*
Tech Stack: Python, Tensorflow, PyTorch, CNNs, Data Processing, Git

- Co-authored a research paper published in **Digital Signal Processing** (April 2024), presenting findings that enhanced PD classification accuracy & explainability using MRI data & Grad-CAM visualization. [Publication]
- Improved **Parkinson's Disease (PD) classification** by fine-tuning **12 pre-trained CNNs** on T1-weighted MRI, identifying **VGG19** as the best-performing model.
- Optimized MRI slice selection by reducing the number of slices from **182 to 87**, improving processing efficiency while preserving critical structures for diagnosis. Conducted an in-depth study on data leakage in existing research.

PROJECTS & PUBLICATIONS

Knee Osteoarthritis Severity Classification using Diffusion Augmented Images Aug 2023 – Dec 2023
Role: Team Lead | Team Size: 6 | Advisor: Aswathy P

- Presented at the **ICACECS 2023** conference before a 3,000 audience. [Springer Nature Publication]
- Attained a classification accuracy of **84%** for knee osteoarthritis by utilizing **EfficientNetB3**, alongside **CLAHE preprocessing** and data augmentation through **DDIM-based diffusion models** to mitigate class imbalance.

Improving RL Agent Training using Text-Based Guidance May 2023 – Sep 2023
Role: Team Lead | Team Size: 6 | Advisor: Dr. Sachin S

- Presented in the **DravidianLangTech 2023** Workshop held at Bulgaria. [ACL Anthology Publication]
- Enhanced multilingual understanding & generalization of RL agents by developing a methodology to train agents using **text-based instructions** in **4 languages** with a **SAC agent** for unseen path generalization.

Few-Shot Approach to Dysarthric Speech Classification Using Transformers Mar 2023 – Jul 2023
Role: Team Lead | Team Size: 5 | Advisor: Dr. Jyothish Lal G

- Presented findings in the **ICCCNT 2023 Conference** held at IIT Delhi. [IEEE Xplore Publication]
- Achieved **85% dysarthria detection accuracy** using **Whisper-large-v2**, a **10% improvement** over prior results and reduced computational costs by training a **1.5B parameter model** with **PEFT and LoRA**.