

NIKHIL CHOWDARY PALETI

510-935-8895 ◊ npaleti@ucsd.edu ◊ [nikhil-paleti.github.io](https://github.com/nikhil-paleti) ◊ [linkedin.com/in/nikhil-paleti](https://www.linkedin.com/in/nikhil-paleti) ◊ github.com/Nikhil-Paleti

OBJECTIVE

Highly motivated AI researcher and Data Science master's student with **6+ published research papers**, specializing in **Generative AI**. Experienced in designing, training, and deploying **scalable AI systems**, with hands-on expertise in **large language models (LLMs)** and **visual-language models (VLMs)**. Passionate about solving complex, real-world challenges. Seeking opportunities to contribute to cutting-edge projects at the intersection of AI research and impactful applications.

EDUCATION

University of California San Diego **Sep 2024 - Jun 2026 (Expected)**
Master of Science in Data Science **GPA: 4.0/4.0**
Relevant Coursework: Data Systems for ML, Advanced Data Mining, Machine Learning, Advanced Data-Driven Text Mining, Probability and Statistics for Data Science, Data Ethics

Amrita Vishwa Vidyapeetham University, India **Oct 2020 - Jun 2024**
Bachelor of Technology in Computer Science and Engineering (Artificial Intelligence) **GPA: 9.15/10**
Relevant Coursework: AI in Natural Language Processing, AI in Speech Processing, Deep Learning for Signal & Image Processing, Deep Reinforcement Learning, Python for Machine Learning, Big Data Analytics

SKILLS

Frameworks/Libraries	Transformers, PyTorch, Jax, TensorFlow, LangChain, crewAI
Specializations	Generative AI (Large Language Models), Natural Language Processing, Computer Vision
Programming Languages	Python (Advanced), SQL (Advanced), C++ (Intermediate)
Tools & Cloud	Docker, Git, Google Cloud Platform (Vertex AI, Compute Engine, Vision API)

EXPERIENCE

Waymo (formerly Google's Self-Driving Car Project), Alphabet Inc. **Jun 2025 - Sep 2025 (upcoming)**
Software Engineer Intern (ML Infra Team) *Mountain View, CA*

- Incoming intern to work on **foundational model training** on the **ML infrastructure team**.

Tech Profuse Pvt Ltd **Jan 2024 - Jun 2024**
Machine Learning Engineer Intern | *Python, Gemini 1.5 Pro LLM, OpenCV (CV2), REST APIs* *Hyderabad, India*

- Developed an **unstructured data extraction API** with **Gemini 1.5 Pro LLM**, processing **50k** bill of lading documents in **15 hours**, achieving a **90%** reduction in manual data entry requirements.
- Engineered a **RAG pipeline** for **email classification and summarization** using the **Gemini 1.5 Flash model**, increasing the support team's capacity to handle **70+** issues daily from **30**.

PUBLICATIONS

- **An Analysis of Data Leakage and Generalizability in MRI-Based Classification of Parkinson's Disease Using Explainable 2D Convolutional Neural Networks**, *Digital Signal Processing*, Elsevier, doi: [10.1016/j.dsp.2024.104407](https://doi.org/10.1016/j.dsp.2024.104407) Apr 2024
- **Enhancing Knee Osteoarthritis Severity Level Classification Using Diffusion Augmented Images**, *ICACECS 2023*, Springer, doi: [10.2991/978-94-6463-314-6.27](https://doi.org/10.2991/978-94-6463-314-6.27) Dec 2023
- **A Few-Shot Approach to Dysarthric Speech Intelligibility Level Classification Using Transformers**, *14th ICCNT*, IEEE, doi: [10.1109/ICCCNT56998.2023.10308067](https://doi.org/10.1109/ICCCNT56998.2023.10308067) Nov 2023
- **Improving Reinforcement Learning Agent Training Using Text-Based Guidance: A Study Using Commands in Dravidian Languages**, *3rd Workshop on Speech and Language Technologies for Dravidian Languages*, ACL Anthology, <https://aclanthology.org/2023.dravidianlangtech-1.5> Sep 2023
- **Face Mask Detection Using Transfer Learning and TensorRT Optimization**, *ICICC 2023*, Springer, doi: [10.1007/978-981-99-3315-0_63](https://doi.org/10.1007/978-981-99-3315-0_63) Jul 2023

PROJECTS

Indic Verse: Advanced Indic Language LLM System **Jan 2024 - Apr 2024**
Tech Stack: Python, PyTorch, Hugging Face Transformers, LORA [[Hugging Face](#)]

- Architected English-to-Indic translation and transliteration modules, processing over **1M sentences** from **6+ datasets** to develop a training and eval corpus.
- **The evaluation dataset was used by Hugging Face engineers** to assess Telugu capabilities in Fine-Web-2.
- Fine-tuned **3 LLMs** (Gemma 2B, Gemma 7B, LLaMA 3 8B) with PEFT, enhancing performance on Indic language tasks.