

Praneeth Reddy Mallupalli

+1 (224) 536-1466 | praneeth.nu.mscs@gmail.com | linkedin.com/in/praneethreddym | github.com/prm036

EDUCATION

Northwestern University, McCormick School of Engineering

Master of Science in Computer Science (Specialization in AI) | GPA: 3.8/4.0

- **Coursework:** Machine Learning, Deep Learning, GenAI, Reinforcement Learning, Agent AI, NLP, Parallel Computing.

Evanston, IL

(Expected) June 2027

Osmania University

Bachelor of Engineering in Computer Science and Engineering | GPA: 3.88/4

Hyderabad, India

June 2017 - Jun 2021

TECHNICAL SKILLS

Languages & Frameworks: Python(Keras, Scikit-learn, NumPy, Pandas, matplotlib), C/C++, Java, SQL, PyTorch.

Machine Learning: Transformers, Generative Models, LLMs (Fine-Tuning, RAG), Quantization, Model Evaluation, Agentic AI.

Data Engineering & Cloud: AWS, PostgreSQL, ETL, FastAPI, Distributed Search Engines, Data visualization.

Infrastructure & Systems: Docker, Kubernetes, Git, Linux, Jupyter, Distributed Systems, Hugging Face.

ACADEMIC PROJECTS

Fine-Tuning LLMs for Math Reasoning While Preserving Safety Alignment | [GitHub](#)

- Fine-tuned **Qwen2.5** model(1.5B & 7B) on **GSM8K** dataset using **LoRA**, improving math accuracy to **81%** while maintaining **88%** safety alignment score on **AILuminate** Safety Dataset, surpassing baselines.
- Performed ablation studies across **10** different hyperparameter configurations (learning rate, LoRA rank, dropout) to analyze performance trade-offs and mitigate catastrophic forgetting in fine-tuned models.

Post-Training Quantization (PTQ) for Diffusion Transformers | [GitHub](#)

- Conducted a head-to-head empirical study, comparing low-bit floating-point (FP) and integer (INT) quantization formats on a **PixArt- α** (0.6B) diffusion model under similar precision budgets.
- Observed that FP quantization yields better visual fidelity at low precision, improving **FID** from **42.4** \rightarrow **38.8** under **W4A8** compared to INT PTQ with comparable **CLIP** and **ImageReward** scores.

Deformable Object Manipulation with Vision-Language-Action Policies | [GitHub](#)

- Trained **SmolVLA** on teleoperated demonstrations using imitation learning for long-horizon control tasks, achieving **69%** success rate on garment folding, outperforming Diffusion Policy (**41%**) and ACT (**61%**) baselines.
- Enhanced model generalization by augmenting training with NVIDIA **Cosmos-Transfer** synthetic data, exposing the policy to diverse physics-grounded visual scenarios.

Rotation Invariant Multi-Object Detector | [GitHub](#)

- Addressed the limitations of traditional object detectors struggling to identify improperly oriented images by integrating eigenvector analysis and a custom decision criterion with a pre-trained YOLOv3 model.
- Achieved a **43%** accuracy improvement over the ResNet50 model on the Pascal VOC 2012 dataset for images deviated between **90 $^{\circ}$ -270 $^{\circ}$** , without any rotation-augmented training.

PROFESSIONAL EXPERIENCE

Arcesium (D.E. Shaw Group)

Senior Software Engineer

Hyderabad, India

July 2023 - Aug 2025

- Optimized the executor component of a **distributed job execution engine** (1M+ jobs/day) by introducing a host-level cache for eligible jobs metadata, resulting in a 70% reduction in DynamoDB reads, and enabling dynamic autoscaling.
- Developed a domain-aware alert aggregation feature for a large-scale alerting system, improving efficiency by **30%**.
- Secured APIs with **HMAC verification**, blocking 100% of unauthorized traffic while maintaining less than 0.5ms latency.

Software Engineer

June 2021 - June 2023

- Designed a backend service orchestrating a high-throughput log ingestion pipeline (40TB/day) operating with zero data loss.
- Developed a configuration-driven routing module that dynamically maps logs to Kafka topics and Elasticsearch indices using schema-aware rules and declarative configs.

Data Science Intern

Feb 2021 - May 2021

- Designed a statistical model to compute health scores for core infrastructure services by normalizing multi-source service metrics (1M+ data points/day) of non-uniform ranges.
- Leveraged ETS time-series models to capture trends and anomalies, improving health prediction accuracy by **30%** over heuristic methods.

LEADERSHIP & HONORS

- **Teaching Assistant**, Northwestern University. — CS496: Machine Learning and Sensing.
- **1st Position**, Arcesium Hackathon themed on “Data Science & Analytics at Scale” (2024) — Won among 130+ teams .
- **Winner**, Smart India Hackathon (SIH) 2020 software edition — Awarded by the Ministry of Education, Government of India.