+1 (571) 547-9928
nikhilreddyk024@gmail.com

Nikhil Reddy Kodumuru

in/Nikhil-Reddy-Kodumuru

nikhilreddy024.github.io
github.com/Nikhilreddy024

EDUCATION

George Washington University **08/2023 – Present**
MS in Computer Science - GPA: 3.90/4.0

Vasavi College of Engineering **05/2019 – 04/2023**
Bachelor's in Information Technology

EXPERIENCE

Research Assistant **01/2024 – Present**
Integrative Nucleic Acid Frameworks - GWU (Advisor: **Dr. Xiangyun Qiu**)

- Developed RNA-focused **AlphaFold models** using **graph transformers** and **diffusion techniques**, improving structure prediction accuracy by 30% for biomedical applications.
- Engineered sequence and structure-aware **RNA generative models** with optimized **GPT-like architectures**, reducing processing time by 70% without sacrificing accuracy.
- Implemented molecular simulation pipelines using **PyTorch, JAX** on **High Performance Computing**, enabling scalable parallelization of complex workflows.

Machine Learning Intern **01/2023 – 06/2023**
Brane Enterprises

- Engineered a facial recognition system combining **MTCNN** and **ArcFace**, improving detection accuracy by 15% via model optimization.
- Developed computer vision tools with **PyTesseract** and **OpenCV** for warehouse navigation; deployed on **NVIDIA Jetson Nano** with full implementation in 10 days.
- Advanced self-driving capabilities using **reinforcement learning** and **SLAM** for dynamic environment navigation.

Machine Learning Intern **06/2022 – 10/2022**
Edmound Software Solutions

- Built an end-to-end **NLP pipeline** using **Hugging Face Transformers** and **MLflow**, automating sentiment analysis with 85% accuracy.
- Built scalable **data pipelines** with **Apache Airflow** and **pandas**, reducing feature engineering time by 50%.

PROJECTS

Project EGRET: Emotion-Aware Conversational AI

- Designed **emotion-aware chatbot** combining **LLaMA-3B** with custom **GCN architecture**, reducing perplexity by 57% over baseline.
- Developed **multi-task learning framework**, improving **BLEU-4** by 57.1% over Facebook AI's baseline.
- Implemented **GCN-based emotional context tracking** with **LoRA optimization**, achieving 0.4101 **ROUGE-L score**.

Project InterACT: Multimodal AI System

- Architected a system integrating **YOLO-World** (94% accuracy) with **LLM-powered** conversation for task automation.
- Fine-tuned **Llama-2-7B** for domain-specific use; led full **ML pipeline deployment**.
- Optimized using **QLoRA** and **PEFT**, reducing resource use by 15% while increasing accuracy by 12%.

SKILLS

Languages: Python, Java, C, SQL, Shell Scripting, MATLAB
ML & AI: Generative AI, NLP, Computer Vision, Reinforcement Learning, LLMs, GNNs
Frameworks: PyTorch, TensorFlow, Keras, JAX, Transformers, scikit-learn, LangChain, crewAI
Tools & Cloud: Docker, Kubernetes, Git, W&B, Grafana, AWS (SageMaker, EC2), GCP (Vertex AI, Vision API)

ACHIEVEMENTS

- Submitted **Project InterACT** research to **ICMLAS 2025**.
- Published work on **Mechanistic Interpretability in LLMs** at **ICSADL 2025**.
- Awarded **SEAS Merit Scholarship** at GWU (top 10%, 50% tuition coverage).
- Secured **2nd place** in university hackathon for building **LLM-powered Chain-of-Thought RAG** using **Claude-Sonnet** on graph-based physics data.

INTERESTS

Physics, History, Geopolitics