

Venkat Kumar Laxmi Kanth Nemala

New York, NY | +1 347-798-7171 | vn2263@nyu.edu | [LinkedIn/nvklaxmikanth](https://www.linkedin.com/in/nvklaxmikanth) | [GitHub/nvklaxmikanth](https://github.com/nvklaxmikanth)

EDUCATION

New York University

Master of Science in Computer Engineering | CGPA: 3.8/4.0

Vignan's IIT

Bachelor of Technology in Computer Science and Engineering | CGPA: 3.7/4

New York, NY

August 2024 – May 2026

Visakhapatnam, India

August 2019 – May 2023

EXPERIENCE

Research Assistant

AI4CE Lab, New York University

September 2024 – Present

New York, NY

- Developed and improved the **Multiview Scene Graph (MSG)** model using the large-scale **ScanNet++** image and video dataset for multiview visual perception and 3D scene understanding, improving **MSG edge IoU by 7%**.
- Implemented a **VQ-VAE**-based quantized variant of **DINO-WM**, leveraging pretrained **DINOv2** visual representations to enable zero-shot planning and manipulation, achieving a **4% performance improvement**.
- Explored diffusion-based generative models using **Diffusion Transformers (DiT)** for structured scene representation and visual synthesis in 3D environments.
- Teaching Assistant for **Prof. Yann LeCun's** DS-GA 1008 Deep Learning and Section Leader for DS-GA 1011 **Natural Language Processing**, mentoring 200+ students in deep learning architectures and multimodal ML systems.

Member of Technical Staff - I

Alphastream.ai

August 2023 – August 2024

Bengaluru, India

- Engineered an end-to-end **AutoML pipeline** for dataset preparation, model training, evaluation, and deployment, reducing retraining time from **9 days to 3 days** and improving extraction accuracy by **9%**; led **dataset curation and annotation workflows** using **Label Studio** with a team of 7 engineers.
- Fine-tuned and integrated the **Co-DETR** model to extract complex hierarchical patterns from tables, addressing **95%** of prior model limitations and analyzing model robustness across evaluation datasets.
- Applied **multimodal LayoutLMv3** for document visual content understanding, identifying table text and images in PDFs and achieving a **14% accuracy improvement**.

Python Developer Intern

Alphastream.ai

December 2022 – July 2023

Bengaluru, India

- Developed an NLP-based parser integrated with **Amazon Textract OCR** to mask sensitive information, achieving **99% data masking accuracy** and improving extraction performance by **13%**.
- Built a **RAG pipeline** for news article processing, benchmarking LLMs (**LLaMA, Qwen, Mistral**) and achieving **97% summarization accuracy**.
- Designed a personalized article recommendation and ad-targeting pipeline based on user reading behavior, improving overall user engagement by **30%**.

PROJECTS

Multimodal Action Localization | Python, PyTorch, Omnivore, LaViLa

[GitHub](#)

- Enhanced VSLNet with **LaViLa vision-language embeddings**, stacked encoders, and attention for Ego4D action localization, achieving **12.14 mIoU**.

Predictive Steering with I-JEPA | Python, PyTorch, CARLA, Waymo, ADL-JEPA

[Github](#)

- Pretrained the I-JEPA model on **Waymo** and **CARLA** datasets and fine-tuned with **ADL-JEPA** for **label-efficient** steering angle prediction, achieving **99.3% accuracy** and validating on a **CARLA simulator**.

SKILLS

Languages: Python, C/C++ , SQL, TypeScript, JavaScript

Frameworks: PyTorch, TensorFlow, JAX, torchvision, Core ML

GPU / Low-Level Programming: CUDA, OpenCL, TensorRT, TVM, XLA

Deployment / Cloud: AWS SageMaker, GCP Vertex AI, Kubernetes, Docker

Development Tools: Linux, Git, Terraform, Ansible, MLflow, Prometheus, Grafana, Wandb