

Kushal Varma Gangaraju

(585) 202-9561 | kgangara@ur.rochester.edu | [linkedin.com/in/kushalgangaraju](https://www.linkedin.com/in/kushalgangaraju) | github.com/kushalX13

EDUCATION

University of Rochester

Rochester, NY

Master of Science (MS) - Computer Science

Expected May 2026

Courses: Deep Learning in Imaging, Digital Image Processing, Data Science at Scale, Machine Vision, Data Structures & Algorithms, Intro to AI

Honors: Awarded 30% Merit Scholarship for Academic Excellence

BMS Institute of Technology and Management

Bengaluru, India

Bachelor of Engineering (BE) - Electronics and Communication Engineering

December 2020 - May 2024

Courses: Computer Organization and Architecture, Operating Systems, Cryptography, Signal Processing, Digital Signal Processing, Intro to C++

EXPERIENCE

ROC HCI Lab

Rochester, NY

Graduate Research Assistant

January 2026 - Present

- Implemented an AWS-based CI/CD pipeline for LLM-powered Python/Flask backend services supporting virtual systems architecture
- Enabled sub-10-minute CI/CD cycles and with quantitative analysis of parallel user study data for UX & system optimization evaluation

Earth Imaging Lab

Rochester, NY

Graduate Research Assistant

September 2025 - December 2025

- Designed and deployed a scalable cloud-native architecture using AWS Lambda, Batch, and EC2, containerized with Docker, enabling automated seismic signal preprocessing across trace batches, producing 2,000+ processed waveform windows per pipeline run
- Integrated Amazon S3 and DocumentDB to store processed outputs and index 10⁴+ metadata entries for efficient querying and retrieval

Global Health & Medical Device Laboratory

Rochester, NY

Graduate Machine Learning Research Assistant

March 2025 - July 2025

- Built an end-to-end fetal ultrasound pipeline combining video segmentation (MedSAM2) and image enhancement (CycleGAN): prompt-based (point/box) frame propagation over 400+ frames per video, with per-frame masks and overlay export via a Gradio interface
- Connected MATLAB ROI labeling workflow with a segmentation pipeline, enabling automation & reducing manual annotation by ~70%
- Evaluated fetal head segmentation using Dice and IoU, achieving Dice up to ~0.85 with temporal consistency across frames and sweeps

Indian Space Research Organization (ISRO)

Sriharikota, India

Machine Learning Research Intern

February 2024 - March 2024

- Developed weather forecasting models using SARIMAX, Random Forest, & Gradient Boosted Trees on real-time meteorological signals
- Enhanced forecast accuracy by 25% and reduced launch delays by approximately 15% using feature engineering and optimization

Camplain Innovations

Bengaluru, India

Quantitative Data Science Intern

April 2023 - June 2023

- Built stock prediction Dash app with real-time dashboards, integrated PyTorch + REST APIs, reducing manual analysis by ~30%
- Implemented LightGBM with lag/rolling features and TA-Lib, improving accuracy by 20% through feature engineering and tuning

PROJECTS

Medical Image Classification with KAN-Integrated Transformers | PyTorch, Transformers, CUDA

March 2025 - April 2025

- Implemented & scaled MedViTV2, a hybrid KAN-Vision Transformer architecture, developing 4 model variants with hierarchical Local and Global Feature Perception & evaluating cross-dataset performance on BloodMNIST & Czech WBC under hybrid training ratios
- Built a pipeline with preprocessing, ablation studies (4+ configs), Grad-CAM, delivering a 5M-parameter variant with ~97% accuracy

Real-Time Tweet Sentiment Pipeline | Spark, Databricks, Hugging Face

April 2025 - May 2025

- Designed and deployed a tweet sentiment pipeline on Spark Structured Streaming with Delta Lake (Bronze → Silver → Gold) using Databricks AutoLoader, processing 100,000 JSON tweets with triggers (30s/20s/60s), checkpointing, and medallion ETL architecture
- Integrated MLflow BERTweet model, achieving 95% precision and 0.615 F1-score, optimizing Spark execution (16 shuffle partitions)

3D OCT Latent Diffusion Model | PyTorch, CUDA, MONAI, 3D Generative Modeling

February 2026 - March 2026

- Architected and trained a ~530M-parameter 3D latent diffusion system for volumetric OCT, integrating a 66M VAE with a 464M MONAI MAISI/NV-Generate inspired UNet using Rectified Flow on 32×128×128 patches, generating 24.3M latent samples
- Engineered mixed-precision GPU training, memory-efficient patch scheduling, and reproducible volumetric benchmarking pipelines

CarDiag-RAG | PyTorch, FAISS, SentenceTransformers, BM25

March 2026 - March 2026

- Built a hybrid recall retrieval pipeline for NHTSA records (SentenceTransformer + BM25 + FAISS), tuned candidate pools at top-10 and top-100, and benchmarked 81 filtered diagnostic queries to improve robust fault-match coverage for paraphrased, real-world user inputs
- Developed an evaluation pipeline with 81 paraphrased diagnostic queries to benchmark retrieval quality, tune candidate-pool size, and compare ranking strategies; delivered Recall@1 of 0.94, Recall@10 of 1.00, and MRR of 0.967 on held-out test sets consistently

SKILLS & CERTIFICATIONS

- Programming Languages:** Python, C++, SQL, TypeScript, MATLAB, Bash
- Frameworks/Libraries:** PyTorch, TensorFlow, Hugging Face, Scikit-Learn, Pandas, NumPy, Flask, Apache Spark, MLflow, LangChain
- Systems & Platforms:** AWS (Lambda, Batch, S3, ECR, EC2, IAM), Docker, Kubernetes, CI/CD, GitHub Actions, Jenkins, Spark, MLflow, Delta Lake, Databricks, MongoDB, PostgreSQL, ETL pipelines, Server-less Architecture, Distributed Systems
- Certifications:** SQL for Data Science by UC Davis, Machine Learning by Stanford, Advanced Learning Algorithms by Stanford