NISHANT HEGDE

Rochester, New York | +1 (585)-559-7746 | nhegde2@ur.rochester.edu | LinkedIn - hegde-nishant | GitHub - hegde-nishant

EDUCATION

MS, Data Science **Aug 2024 to Dec 2025**

University of Rochester *GPA: 3.77/4.00*

Coursework : Data Science at Scale, End-to-End Deep Learning, Computational Statistics, Databases, Data Mining

B.E., Electronics and Communication **Aug 2016 to Jun 2020**

KLE Technological University *GPA: 3.67/4.00*

Coursework : Machine Learning, Digital Circuits, Analog Circuits, Digital Signal Processing, Control Systems

WORK EXPERIENCE

Machine Learning Intern | Laboratory for Laser Energetics **Aug 2025 to Present**

Rochester, New York

- Tackled a largely unexplored ML topic in modelling laser amplification process through OMEGA EP laser system, requiring rapid prototyping and development of novel approaches.
- Designed a U-Shaped Fourier Neural Operator and Attention Recurrent Residual U-Net to jointly predict output UV spatial laser beam profile and temporal pulse with physics-informed loss functions.
- Engineered HPC training infrastructure across A100 and H100 GPU clusters using model parallel, mixed-precision training, gradient checkpointing, and CUDA optimizations, cutting training time by 40% and VRAM usage by 30%.

Data Science Intern | EagleHawk One, Inc **Jun 2025 to Present**

Remote

- Designed a modular computer vision pipeline for real-time streetlight tracking in low-light nighttime drone imagery using an ensemble of YOLOv8 detectors, processing 5000+ images with 91% detection accuracy.
- Built a RAG-based emergency response training assistant using Docling for multimodal parsing and Chonkie for semantic chunking, cutting new employee training time by 70% and automating competency assessments.
- Deployed a hybrid retrieval architecture integrating BM25 with dense embeddings, enhanced by Zerank re-ranking and local Ollama LLM generation, enabling secure offline querying with high contextual precision.

Software Engineer | Toshiba (KIOXIA America, Inc) **Aug 2020 to May 2024**

Remote

- Developed data-driven C++ modules to validate the behavior of the NVMe management endpoint and NVMe controller management over interfaces such as SMBus and I2C for NVMe data-centric SSDs, reducing SSD operational latency by 31%.
- Designed validation, regression, stress testing, and failure analysis test suites in Java that improved firmware stability by 15%.
- Orchestrated the deployment of high-quality firmware validation modules by using Scrum and Agile methodologies.
- Automated validation pipelines with Jenkins, enabling seamless nightly builds causing a 20% reduction in manual QA effort.
- Streamlined collaboration through Confluence, improving documentation transparency and workflow efficiency by 33%.
- Built internal debugging and telemetry analysis tools in Python to streamline firmware failure root-cause analysis.
- Improved reliability of long-running endurance tests by adding fault-tolerant logging and auto-recovery in Java.

SELECT PROJECTS

NVMe RAG System [view project] - Built a RAG system for querying NVMe technical specifications using ChromaDB, sentence-transformers, and Ollama LLM, achieving 40% reduction in manual search time. Implemented semantic chunking and cross-encoder reranking via LangChain and FastAPI for scalable querying of technical standards.

Wafer Map Failure Pattern Detection [view project] - Deployed a Vision Transformer–based wafer map defect classification pipeline achieving 96.2% accuracy on 170K samples across 9 defect modes. Built a scalable inference workflow for fab integration using transfer learning, mixed-precision training, and high-throughput batch processing. Developed a comprehensive analytics platform with 40+ actionable visualizations, including PCA/t-SNE clustering and radial defect profiles for optimizing manufacturing decisions.

Social Media Intelligence Engine [view project] - Built a scalable streaming analytics platform using Apache Spark Streaming, Delta Lake, and MLflow to process 40K+ JSON files via medallion architecture on Databricks. Implemented end-to-end ETL with Spark UDFs and MLflow model registry for automated sentiment scoring and real-time monitoring.

TECHNICAL SKILLS

Machine Learning:	Neural Operator, U-Net, RNNs, Embeddings, Classical ML
Computer Vision:	RCNNs, GaNs, CNNs, Image Enhancement, Anomaly Detection, Semantic Segmentation
Languages and Databases:	Python, R, C++, MATLAB, Java, SQL, MongoDB, Postgres, Bash
Tools and Libraries:	PyTorch, TensorFlow, Apache Spark, Databricks, Tableau, Scikit-Learn, NLTK, CUDA
Cloud and DevOps:	Heroku, Docker, Kubernetes, MLflow, Delta Lake
Other Tools and Services :	JIRA, Confluence, Workday, LaTeX, MS Office, Slack, Git