

Kaivalya Kishor Dixit

 github.com/KaivDev4434  +1(862)-215-1490  [linkedin.com/in/kaivalya-dixit-2a25851b9](https://www.linkedin.com/in/kaivalya-dixit-2a25851b9)  kd454@njit.edu

EDUCATION

New Jersey Institute of Technology <i>M.S. Data Science</i>	May 2025 GPA: 4.0/4.0
Mahindra Ecole Centrale <i>B.Tech Electrical and Electronics Engineer</i>	Jun 2023 GPA: 3.54/4.0

RELEVANT COURSEWORK

Courses: Machine Learning, Introduction to Big Data, Applied Statistics, Deep Learning, Database Management Systems, Time Series Analysis and Forecasting, Reinforcement Learning, Applications of Parallel Computing

SKILLS

Languages: Python, Java, JavaScript, C, HTML/CSS, Shell Scripting, Assembly, \LaTeX
Frameworks: React, Node.js, Flask, Streamlit, Docker, Apptainer
Data Skills: SQL, Tableau, Apache Spark, Hadoop, Pandas, Map-Reduce, Hive, MongoDB
Tools: Git/GitHub, Unix Shell, VS Code, IntelliJ, Vim, bash, zsh, Linux, AWS, Tableau, Postman
Technical skills: Pytorch, Tensorflow, Pandas, NumPy, Matplotlib, ScikitLearn, LangChain, OpenGym, OpenMP, MPI, plotly, seaborn, XGBoost
Soft skills: Data Storytelling & Visualization, Problem-Solving, Critical Thinking, Technical Communication

PROJECTS

Hyperion, HPC testbed with rPIs | *RockyLinux, DHCP, TFTP, NFS, Warewulf* Nov 2024 – Jan 2025

- Designed Raspberry Pi HPC testbed to prototype cluster management fundamentals, reducing future deployment risks through iterative testing
- Integrated SLURM job scheduler, enabling parallel task execution across 4 ARM64 nodes
- Resolved iPXE netboot failures by reverse-engineering TFTP configurations through 50+ GitHub/forum tests, achieving 100% node provisioning reliability
- Automated compute node deployment via Warewulf/NFS (2hr/node → 15min/node)

Reinforcement Learning-Based Autonomous Vacuum Cleaner | *Python, NumPy, matplotlib, OpenAI Gymnasium* Mar 2025 – Current

- Developed an autonomous vacuum cleaning robot using Python and OpenAI Gym, implementing and comparing three reinforcement learning algorithms—Soft Actor-Critic (SAC), Proximal Policy Optimization (PPO), and Deep Deterministic Policy Gradient (DDPG)—within custom-built environments.
- Engineered a phased simulation framework progressing through increasing complexity: starting from empty rooms, introducing static obstacles, adding dirt, incorporating dynamic objects, and culminating in a vision-enabled agent.
- Designed modular training environments and reward structures to facilitate scalable policy learning and performance benchmarking across diverse scenarios.

High-Performance Distributed Machine Learning Framework | *OpenMP, MPI, CUDA, C++, pybind* Mar 2025 – Current

- Developed a high-performance distributed deep learning framework integrating OpenMP, MPI, and CUDA to optimize training across CPUs and GPUs on HPC systems.
- Implemented parallel data loading, training, gradient synchronization, and checkpointing, inspired by PyTorch Distributed Data Parallel (DDP).
- Benchmarked scalability and performance using datasets like MNIST and CIFAR-10, analyzing computational intensity and communication overhead.

EXPERIENCE

High Performance Computing | *HPC User Support Specialist, Student Intern* Sept 2024 – Current

- Provide expert troubleshooting for 400+ researchers on PyTorch/Conda environments, Docker containerization, and GPU/CPU performance optimization across the clusters
- Developed automated benchmarking suite for node health checks using Slurm hooks and custom Bash/Python scripts (CPU stress tests, GPU memory validation)
- Upgraded HPC infrastructure by provisioning 6 NVIDIA Grace Hopper nodes with InfiniBand NDR200, including rack/cable management, and driver/firmware updates.

Dassault Systems | *Data Analyst Intern* Jan. 2023 – Jul 2023

- Engineered Java-based ETL pipeline for Conversion Admin Service, processing enterprise-scale customer lifecycle data
- Designed interactive dashboard for license conversion tracking using internal visualization frameworks
- Automated data quality checks streams using SQL window functions and constraint validation