Kaivalya Kishor Dixit

github.com/KaivDev4434 +1(862)-215-1490 linkedin.com/in/kaivalya-dixit-2a25851b9 kaivalyawork@gmail.com

Education

New Jersey Institute of Technology May 2025
M.S. Data Science *GPA: 3.95/4.0*

Mahindra Ecole Centrale Jun 2023
B.Tech Electrical and Electronics Engineering *GPA: 3.54/4.0*

Skills

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

Projects

Hyperion, HPC testbed with rPIs | *RockyLinux, SLURM, Warewulf* Nov 2024 – Jan 2025
- Built a scalable Raspberry Pi-based HPC testbed to prototype cluster deployments and configurations
- Integrated SLURM to enable distributed task execution across 4 ARM64 nodes
- Achieved 100% provisioning reliability by debugging netboot failures across 50+ test iterations
- Reduced node deployment time using Warewulf and NFS automation

Reinforcement Learning-Based Autonomous Vacuum Cleaner | *Python, NumPy, matplotlib, Gymnasium* Mar 2025 – May 2025
- Developed a robot agent trained using SAC, PPO, and DDPG in a progressively complex simulation environment
- Built custom environments with modular reward structures to support scalable policy learning
- Designed a simulation pipeline from simple layouts to vision-based dynamic obstacle scenarios

High-Performance Distributed Machine Learning Framework | *OpenMP, CUDA, C++, pybind* Mar 2025 – May 2025
- Built a distributed deep learning framework optimized for HPC using CPU-GPU parallelism
- Implemented multi-process training pipeline with gradient synchronization
- Benchmarked on MNIST to evaluate speedup and communication overhead
- Achieved 26.6x data loading speedup and 5.2x training speedup

Experience

High Performance Computing | *HPC User Support Specialist, Student Intern* Sep 2024 – May 2025
- Support 400+ researchers with GPU/CPU performance tuning, containerization, and environment troubleshooting
- Developed an automated benchmark suite for node health monitoring using Slurm, Bash, and Python
- Provisioned 6 NVIDIA Grace Hopper nodes with InfiniBand, managing hardware setup and driver configurations
- Developed a Raspberry Pi–based testbed to validate updates and conduct pre-deployment testing before rolling changes out to the production HPC cluster

Dassault Systems | *Data Analyst Intern* Jan. 2023 – Jul 2023
- Engineered scalable Java-based ETL data pipelines for the Conversion Admin Service, enabling robust data ingestion and integration of enterprise-scale customer lifecycle data
- Automated data quality checks using SQL window functions and constraint validation, improving pipeline reliability and monitoring
- Designed and deployed interactive dashboards for license conversion tracking using internal visualization tools, supporting end-to-end data engineering workflows
- Contributed to infrastructure management by optimizing data flow and ensuring high availability of internal analytics systems

Certifications

- AWS Certified Machine Learning Engineer – Associate (MLA-C01)
- NVIDIA DLI – Building RAG Agents with LLMs
- Bloomberg Market Concepts – Bloomberg for Education