Kaivalya Kishor Dixit

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

Education

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

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

Relevent 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, OpenMPI, 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 by 30% through iterative testing
- Integrated SLURM job scheduler with 95% resource utilization efficiency, 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 $\rightarrow$ 15min/node)

Quantitative Portfolio Simulator | *Python, pandas, NumPy, matplotlib, seaborn* Feb 2024

- Developed backtesting engine implementing 5-day rebalancing strategies (”buy low” vs ”buy high”) with dynamic position sizing, tracking MTM through 50+ rebalancing cycles
- Engineered data pipeline processing 10 stock datasets + FX rates into unified time-series structure using pandas, handling OHLC/Adjusted Close alignment and corporate actions
- Visualized strategy performance against synthetic tech index using matplotlib dual-axis plots with percentage change normalization for comparative analysis
- Integrated USD/JPY FX rate conversion system with daily close alignment, enabling JPY-denominated performance tracking

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 hybrid 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, driver/firmware updates, and NCCL validation for multi-GPU communication

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 (J2EE/Spring backend)
- Automated data quality checks across 15+ SaaS product streams using SQL window functions and constraint validation
- Developed cross-functional documentation for GDPR-compliant data handling procedures adopted by EU/NA teams