KAIVALYA KISHOR DIXIT

kaivalyawork@gmail.com | (862) 215-1490 | kaivalyadixit.vercel.app | linkedin.com/in/kaivalya-dixit | github.com/KaivDev4434

Education

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

Skills

Programming Languages: Python | Java | JavaScript | C/C++ | HTML/CSS | Shell Scripting | Assembly | LaTeX | TypeScript
Machine Learning & AI: PyTorch | TensorFlow | Scikit-learn | Keras | XGBoost | OpenMP | MPI | CUDA | LangChain | LlamaIndex | OpenAI Gym | Natural Language Processing (NLP) | Deep Learning | Neural Networks | Transformers | CNNs | RNNs (LSTM/GRU) | Transfer Learning | Ensembling | Multimodal Learning | AutoML | Time Series Analysis | Reinforcement Learning
Data Engineering & Analytics: SQL | Apache Spark | PySpark | Hadoop | Pandas | NumPy | MapReduce | Hive | MongoDB | ETL/ELT Pipelines | Data Warehousing | Data Lake | Feature Engineering | Statistical Analysis | Big Data | Tableau | Matplotlib | Seaborn | Plotly
Cloud & DevOps: AWS | Docker | Apptainer | Git/GitHub | Databricks | MLflow | Weights & Biases | CI/CD | Containerization
Frameworks & Tools: React | Next.js | Node.js | Flask | Streamlit | Spring Boot | REST APIs. Material-UI | Tailwind CSS | Bootstrap | SLURM | HPC Systems
Databases: PostgreSQL | MySQL | MongoDB | Redis | Data Modeling | Database Design | Query Optimization

Experience

HPC User Support Specialist Intern/ ***New Jersey Institute Of Technology*** *SEP 2024 – MAY 2025*

- Catalyzed research productivity for 400+ researchers via expert management of PyTorch/Conda environments, Docker containerization, and strategic GPU/CPU optimization across a hybrid cluster infrastructure.
- Orchestrated a SLURM-based benchmark suite to oversee node health using Python and Bash, diminishing manual checks by 80% and escalating cluster uptime by 200 hours quarterly.
- Built a Raspberry Pi-based HPC testbed for pre-deployment validation, lowering production cluster deployment risks by 30% through iterative testing.

DATA ANALYST INTERN | ***DASSAULT SYSTEMS*** *JAN 2023 – JUL 2023*

- Developed scalable Java ETL pipelines for 15+ SaaS product streams, processing enterprise-scale customer lifecycle data and supporting GDPR compliance across EU/NA teams, ensuring international regulatory standards.
- Designed and launched interactive dashboards for license conversion tracking with J2EE/Spring backends, improving decision-making efficiency by 25%.
- Automated data quality checks with SQL window functions and constraints, increasing pipeline reliability and reducing data inconsistencies by 40%.

Projects

DISTRIBUTED MACHINE LEARNING FRAMEWORK | C++, OPENMP, CUDA, MPI, PYTHON, PYBIND | *MAR 2025 – MAY 2025*

- Developed and optimized a parallel gradient synchronization solution and a parallel data ingestion pipeline leveraging OpenMP, MPI, and CUDA, resulting in a 15 ms reduction in synchronization time per iteration and a 26.6x acceleration in data loading on the MNIST benchmark, which collectively led to a 5.2x improvement in distributed deep learning training speed.
- Built a comprehensive benchmarking suite that demonstrated the custom implementation achieved 90% performance parity with PyTorch while using 20% less memory, enabling more efficient resource allocation.

ZIGGY - COMPANION CHATBOT | TYPESCRIPT, REACT, NEXT.JS, NODE.JS, EXPRESS, MONGODB, REDIS | *JUL 2025 – PRESENT*

- Revamped chatbot performance by integrating the Perplexity Pro API with a local LLM through Ollama on a custom server, enabling advanced query handling and reducing reliance on external APIs by 70%. Simultaneously, launched a mobile-first progressive web app with Docker containerization for seamless deployment, boosting mobile user engagement by 30%.
- Implemented a comprehensive testing framework for Ziggy Bot using Jest and React Testing Library, achieving 100%-unit test coverage and proactively discovering edge cases through property-based testing.

LMAN - AI-POWERED UNIX MANUAL ASSISTANT | PYTHON, NLP, RAG, CLI DEVELOPMENT | *MAY 2025 – PRESENT*

- Developed a CLI tool that leverages Retrieval-Augmented Generation (RAG) and a semantic search pipeline using sentence transformers and FAISS vector indexing to enable natural language querying and accurate command Q&A for Unix man pages.
- Designed a modular architecture with Poetry for dependency management and an intuitive CLI interface utilizing Click and Rich libraries, ensuring user-friendly operation and streamlined development.

QUANTITATIVE PORTFOLIO SIMULATOR | PYTHON, PANDAS, NUMPY, MATPLOTLIB, SEABORN | *FEB 2024 – APR 2024*

- Engineered a data pipeline consolidating 10 stock datasets and FX rates into an aligned time-series structure, including OHLC and Adjusted Close handling, and integrated a daily USD/JPY FX conversion system for accurate JPY-denominated performance tracking.
- Visualized strategy results against a synthetic tech index using matplotlib dual-axis plots with percentage change normalization, enabling comparative performance analysis.

Certifications

- NVIDIA Deep Learning Institute – Building RAG Agents with LLMs | 2025
- Kaggle Introduction to Machine Learning | 2022