Paul Lin

248-607-4035 | linp40182@gmail.com | LinkedIn | GitHub | linpaul.com

Technical Skills

Programming Languages: Python, C++, C, Java, JavaScript, PHP, HTML, CSS, SQL
Frameworks/Libraries: PyTorch, JAX, CUDA, scikit-learn, LangChain, PySpark, NumPy, pandas, React, Node.js
Cloud and DevOps: Git, Github Actions, Docker, AWS, Google Cloud Platform

Education

Michigan State University **May 2026**

B.S. Computer Science, Honors College. GPA: 3.76

Work Experience

ECODE Lab | ***Undergraduate Research Assistant*** **Jun 2025 - Present**

- Developed black-box adversarial attacks in PyTorch with 99% success on state of the art vision models, exposing fundamental differences between human and machine visual perception.
- Streamlined large-scale ML training on Unix HPC clusters with Slurm jobs and maximized GPU utilization.
- Engineered an evolutionary neural network with Pytorch, achieving 30× faster training vs CPU through GPU vectorization.

BigCommerce | ***Software Developer Intern*** **May 2024 - Aug 2024**

- Built scalable ETL pipelines processing 1M+ records using Python and AWS CDK, leveraging Glue, AppFlow, S3, and Athena to automate ingestion, transformation, and querying.
- Worked in a 15-member crossfunctional Agile team, accelerating AWS pipeline development by ~40%.

SEIT Lab | ***Undergraduate Research Assistant*** **May 2024 - Aug 2024**

- Led a 4-person team to build a Python autonomous driving algorithm in SUMO, containerized with Docker.
- Achieved 99% safety compliance in simulations (2nd out of 10 labs) before real-world deployment.

Leadership Experience

CSE 331: Data Structures and Algorithms | ***Undergraduate TA*** **Aug 2024 - May 2025**

- Taught 50+ students weekly through helprooms and exam reviews, boosting average the exam score by 7%.

AI Club at MSU | ***Project Manager and Secretary*** **May 2024 - May 2025**

- Led 100+ participants across 10 projects spanning RAG-based LLM systems to an AI-powered fitness coach.
- Designed a dashboard to organize 500+ members data using Python, boosting E-Board efficiency by 50%.

Personal Projects

GPU-Accelerated Sequence Processing | Github Repository | **(C++17, CUDA)**

- Accelerated sequence-search algorithms using CUDA, achieving 30× speedup over CPU on 100GB datasets.
- Designed an asynchronous CPU-GPU streaming pipeline with batching and optimized memory allocation.
- Achieved 90%+ unit test coverage using GoogleTest and automated CI/CD with GitHub Actions.

Stock Prediction Pipeline | Github Repository | **(AWS, TypeScript, PySpark)**

- Developed a cloud ML pipeline on AWS to automate XGBoost model deployment for stock prediction.
- Trained and performed inference on 1M rows of financial data, ranking top 10% performance on NumerAI.
- Cut model training and deployment workload by 60% by automating large-scale batch inference with Glue and SageMaker and streamlining data ingestion GraphQL and Lambda with dataset storage in S3.