Jonathan Jiang

jiangjonathan.com jonathanrsjiang@icloud.com +1-647-408-3030

Skills

Languages: C++, Python, Java, TypeScript, JavaScript, Golang, Lua, SQL, HTML/CSS, Rust

Technologies: Git, Hadoop, Hive, Spark, NumPy, scikit-learn, Jenkins, Vite, Linux, React, LLVM, PostgreSQL, Node.js

Experience

Huawei Technologies Toronto, ON

Software Engineer Intern (Big Data Platform) *May 2023 – May 2024 · 1 year*

Hybrid Execution SQL Query Engine (Java, C++):

* Translated core Java SQL engine components to C++ using JNI to improve hybrid execution performance.
* Built a unified development environment to accelerate onboarding and streamline collaborative workflows.
* Enhanced overall code quality by contributing thorough reviews, documentation, testing, and debugging efforts.

Benchmarking and Profiling Pipeline (Python):

* Developed a full benchmarking pipeline to automate profiling and end-to-end performance analysis.
* Aggregated Spark cluster outputs and extracted detailed metrics including hotspots and operator usage per node.
* Automated flamegraph generation and ensured compatibility with perf, Vtune, arthas, and async profilers.

OLAP SQL Operator Optimization (C++):

* Optimized SQL hash aggregation and join operators using efficient data structures and high-performance hashing.
* Benchmarked a wide range of operator variants across dataset scales to identify best-performing configurations.
* Applied SIMD, vectorization, caching, and JIT compilation techniques to improve overall operator speed.
* Improved operator performance by 19%, and 7% end-to-end, measured by the 10TB TPC-DS dataset benchmark.

Data Cleaning Pipeline for MoE Pre-training (Python):

* Designed scalable MoE data-cleaning pipelines to prepare expert-specific datasets for large-model training.
* Automated normalization, missing-value handling, and outlier filtering optimized for massive corpora.
* Dynamically tuned filtering parameters using judge LLMs, curriculum learning ideas, and ablations.
* Improved score of prototype model trained on filtered dataset by 4% on MMLU Benchmark STEM subset.

Projects

Portfolio Website (Typescript):

- Built an interactive 3D portfolio website using Three.js, vanilla TypeScript, and Vite, featuring custom models.
- Implemented a modular multi-scene architecture with object-based navigation and seamless 2D/3D UI integration.
- Designed interactive turntable components, YouTube player integration, and automated music metadata retrieval.

Counter Strike: Arbitrage (JavaScript, Python):

- Developed Python scripts to scrape third-party CS:GO item exchanges to identify arbitrage opportunities.
- Utilized heuristics such as trade volume, previous sales, transaction costs and others to evaluate relative risk of items.
- Traded $500 CAD of initial capital to $12,000 CAD using the script to select optimal transactions.

Rotator Cuff Rehabilitation Device (Arduino, Python):

- Prototyped at-home rehabilitation device used to remotely monitor post-surgery rehabilitation of the shoulder.
- Integrated muscle repetition detection and recovery projections into result graphs with NumPy and Matplotlib.
- Designed application UI, and implemented user account system to store and share patient records securely.

Education

McMaster University Hamilton, ON

Bachelor of Engineering, Software and Biomedical Engineering Co-op *Sept 2020 – Present*

Expected Graduation: May 2026