Abhinav Jain

a252jain@uwaterloo.ca • abhij2706.github.io • linkedin.com/in/abhij2706/ • +1 343 558 2078

Languages: Python, C++, C, Rust, HTML, CSS, Javascript, Java, Kotlin, Swift, Bash, SQL, R.
Technologies: Android, iOS, Linux, Docker, Jenkins, Git, GDB, NumPy, Pandas, TensorFlow, PyTorch, OpenCV, Node.js, CUDA.
Domains: Embedded Software, AI/ML Model Development, Software Engineering, Full-Stack Development, Test Automation.

Experience

University of Waterloo (Waterloo, ON) **January 2024 - Present**
Undergraduate Research Assistant *Research- Swift, CoreML, iOS*
- Implementing LLM-based sentence encoding to enable qualitative analysis of health science documents.
- Benchmarking LLM speed on Apple hardware. Researching methods to accelerate LLMs running on the CPU and GPU.

Qualcomm (Toronto, ON) **August 2023 - December 2023**
AI Performance Analysis Intern *Software Engineering- Python, NumPy, Pandas*
- Led the development of a software profiling pipeline for analyzing the performance of CNNs running on Snapdragon.
- Increased speed of analysis tools by 500+%. Added reduction methods to accelerate preprocessing of large datasets.
- Designed a querying system and cache for large performance datasets (50k+ points), removing dataset size limitation.
- Designed and implemented a time-based search engine to enable nanosecond-level analysis of CNN execution.

Qualcomm (Toronto, ON) **January 2023 - August 2023**
Snapdragon AI Processor Intern *Embedded Software- C++, Python, Android, Jenkins*
- Enhanced features for an embedded framework that accelerates CNN-based image processing on Snapdragon.
  - Integrated data propagation features into production code. Wrote tests to ensure thread-safety and scalability.
  - Performed Inter-Process Communication (IPC) optimizations, achieving a 50% performance improvement.
  - Added runtime configuration for debugging infrastructure, doubling debugging speed for 40+ engineers.
- Created an Extract-Load-Transform (ELT) pipeline to analyze performance of CNN-based workloads on Snapdragon.
  - Built a scalable analysis suite to manage 100+ KPIs and 5+ output formats, enabling robust data examination.
  - Deployed the ELT pipeline to Jenkins, revealing bottlenecks that led to a 2x performance improvement when fixed.

Cisco Systems (Ottawa, ON) **May 2022 - August 2022**
DevOps and Analytics Intern *Software Engineering- Python, Docker, ElasticSearch*
- Designed, implemented and fine-tuned an ML model to recommend code reviewers for pull requests.
  - Employed Collaborative Filtering to make recommendations. Optimized for large datasets with 500,000+ entries.
  - Deployed a Jenkins job to automate recommendations. Integrated results into GitHub, serving 3600+ engineers.
- Optimized the Cisco Networking Bot's Natural Language Processing (NLP) pipeline, increasing its F1 score to 0.99.
- Reduced the Cisco Networking Bot's space usage by 90% through Docker image optimizations, saving $35,000+/year.

Cisco Systems (Ottawa, ON) **May 2021 - August 2021**
Segment Routing IPv6 (SRv6) Testing Intern *Test Automation- Python, TensorFlow, InfluxDB, Grafana*
- Developed a Recurrent Neural Network (RNN)-based model to detect anomalies in router telemetry data.
- Wrote automated unit tests for SRv6 features, to ensure correctness and scalability across various network topologies.
- Constructed RESTful APIs to collect and preprocess 60+ SRv6 telemetry metrics using YANG.

Extracurriculars

WAT.ai (Waterloo, ON) **October 2023 - Present**
Stock Forecasting Team *AI/ML Model Development- Python, PyTorch*
- Designing a deep reinforcement learning model to trade stocks using sentiment analysis and price predictions.

Waterloo Data Science Club (Waterloo, ON) **September 2022 - September 2023**
VP of Data Analysis/Reading Group Lead *AI/ML- Python, TensorFlow, MatPlotLib, Numpy, Pandas*
- Presented papers covering neural network architecture, image processing, and NLP to 100+ club members.
- Prepared resources and hosted workshops on Pandas, NumPy, MatPlotLib, and Tensorflow for 60+ club execs.

Education

University of Waterloo (Waterloo, ON) **September 2020- May 2025**
Bachelor of Computer Science (Co-op) with AI Specialization (4th year) *GPA: 89%*