Abhinav Jain

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

Languages: Python, C++, C, Rust, HTML, CSS, Typescript, Javascript, Java, Kotlin, Swift, Bash, SQL, R.

Technologies: Android, iOS, Linux, Docker, Jenkins, Git, GDB, NumPy, Pandas, TensorFlow, PyTorch, OpenCV, Node.js, CUDA.

Domains: Software Engineering, AI/ML, Research, Data Engineering, Embedded Software, Test Automation.

Experience

Amazon (Toronto, ON)

May 2024 - Present

Software Development Engineering Intern

Software Engineering- Kotlin, React.js, Typescript, AWS

Leveraging AI and ML to generate insights for sales teams, improving product recommendations for millions of users.

University of Waterloo (Waterloo, ON)

January 2024 - April 2024

Undergraduate Research Assistant

AI/ML Research- Swift, CoreML

• Implemented LLM-based sentence encoding, enabling qualitative analysis of health science documents with 5+ LLMs.

Qualcomm (Toronto, ON)

August 2023 - December 2023

Al Performance Analysis Intern

Data 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 Al 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 embedded debugging infrastructure, doubling debug speed for 40+ engineers.
- Created and deployed an Extract-Load-Transform (ELT) pipeline to analyze performance of CNN-based workloads.
- Built a performance analysis suite to manage 100+ KPIs and 5+ output formats, enabling robust data examination.

Cisco Systems (Ottawa, ON)

May 2022 - August 2022

DevOps and Analytics Intern

Software Engineering-Python, Docker, ElasticSearch

- Implemented, tested, and deployed an ML model to recommend code reviewers for pull requests, serving 3600+ users.
- 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

UW Reality Labs (Waterloo, ON)

May 2024 - Present

Designing a deep learning model to recognize custom hand gestures captured using a Meta Quest VR headset.

WAT.ai (Waterloo, ON)

October 2023 - April 2024

• Designed 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: hosted workshops on ML and AI basics, led a team of 5 analysts producing educational content.
- Reading group lead: presented papers covering neural networks, image processing, and NLP to 100+ club members.

Education

University of Waterloo (Waterloo, ON)

September 2020- May 2025