

Joshua Sun

2B Software Engineering Student

☎ (+1) 415-881-0388 | ✉ joshuasun1998@icloud.com | 🌐 jsun98 | 🏠 joshsun.ca

QUALIFICATIONS

Languages:	Java, Python, JavaScript, C/C++
Machine Learning:	Vowpal Wabbit, TensorFlow, Scikit-learn, Gradient Boosting Decision Trees (GBDT)
Big Data Analysis:	Hadoop, MapReduce, HDFS, Pig, Oozie, Storm, Avro
Tools:	Git, Maven, Docker, Travis CI, Screwdriver, Jenkins, Heroku, AWS
Databases:	MySQL, MongoDB, NoSQL

EXPERIENCE

Yahoo! Backend Software Engineer | Mail Anti-Spam **Sunnyvale, California**

The Anti-Spam team analyzes tens of billions of SMTP requests daily for Yahoo!, AOL and Oath.

Jan 2018 - Apr 2018

- Built a terabyte-scale distributed model training pipeline using Java MapReduce API, HDFS, Pig and Oozie which is capable of processing over 10 billion training examples (70-100 dimensions) in under 30 minutes.
- Trained and tested various Vowpal Wabbit and Gradient Boosting Decision Trees models using Hadoop-Streaming, MPI AllReduce and the Minimum Spanning Tree, reducing training latency from days to minutes.
- Enhanced the existing Naive Bayes spam classifier by constructing a Random Forest classifier, resulting in an additional 250 millions spam messages captured daily and decreased the false-positive rate by 11%.
- Designed an Apache Storm duplicate detection bolt for an online spam detection algorithm using Guava Cache, which decreased traffic in the cloud message queue by 35%.

MentorMap Co-Founder

Toronto, Ontario

MentorMap is a startup that connects prospective students with current university students.

May 2017 - Nov 2017

- Architected a video scheduling web app for mentors and mentees using Node.js, MongoDB, HTML and CSS.
- Designed a scalable MongoDB schema that avoided transactions by leveraging embedded documents.
- Led a team of 5 in software development and marketing and recruited 100+ members.

OMERS Full Stack Developer | Test Automation Team

Toronto, Ontario

OMERS is Canada's largest pension fund with \$72 billion in net assets.

May 2017 - Aug 2017

- Developed RESTful React/Node.js web applications to simplify the mock data creation workflow for testers.
- Optimized the Selenium Grid infrastructure to efficiently run tests in parallel, reducing runtime by 45%.

PROJECTS

Stock Predictor + Trade-bot *Python, TensorFlow, Scikit-learn, MySQL*

- Applied supervised learning methods such as Neural Networks, Random Forest and Gradient Boosting regressors to forecast stock prices based on past returns and financial indicators.
- Developed a simulation trade-bot constructs a portfolio of stocks to diversify risks.

Lacs/LISP Compilers *Scala, C++*

- Implemented the CYK parsing algorithm, a garbage collector and functional programming concepts.

EDUCATION

University of Waterloo

Sep 2016 - Apr 2021

Bachelor of Software Engineering, Joint Combinatorics and Optimization

CGPA: 3.70

HONORS

Top 10 Finalist @ TreeHacks *Competed against 115 teams at the Stanford University hackathon* **Feb 2018**

Governor General's Academic Medal *Highest GPA in high school of 439 students*

Jun 2016