

# Jack Wang

647-687-4168 | jj22wang@uwaterloo.ca | jj22wang

## Skills

---

**Languages** Python, Scala, Java, C++, SQL

**Tools** Spark, Hadoop, Kafka, Scikit Learn, Pandas, Elastic Search, AWS, Docker, Git

## Experience

---

### Yelp

*San Francisco, California*

**SOFTWARE ENGINEER INTERN**

*September 2016 - December 2016*

- Improved business location quality by increasing the precision and recall of a bad location detector by 15% and 8%
- Increased machine learning model performance by adding features, tightening training data criteria and developing a gold data set for testing
- Created feedback loop using in app survey questions to validate model predictions
- Optimized when survey questions about businesses were asked in order to maximize meaningful response rate and reduce bias towards common attributes

### Capital One Canada

*Waterloo, Ontario*

**DATA SCIENTIST INTERN**

*January 2016 - April 2016*

- Developed data driven web applications to help customers manage their spending
- Implemented a parallelized data pipeline to ingest terabytes of credit card transactions to create meaningful insights for customers using pySpark, Hive and SKlearn
- Added Docker support to web server for easy deployment and scaling on AWS

### Loyalty One

*Toronto, Ontario*

**SOFTWARE DEVELOPER**

*May 2015 - August 2015*

- Developed a data aggregation tool using Spark to eliminate unnecessary boilerplate code for feature generation and data obfuscation
- Parallelized existing data aggregation processes to reduce run times by 100 times
- Executed large scale performance testing on over 80 terabytes of data to optimize resource allocation when performing data aggregation on AWS
- Prototyped an end to end data streaming pipeline using Spark, Kafka and Elastic Search to do real time analysis on advertising campaigns
- Deployed projects to local Hadoop Clusters and AWS EC2 clusters of over 300 machines

## Personal Projects

---

### Tagger PYTHON

An automatic mp3 attribute tagger

- Performs filename analysis to infer artist and song title, based on research done by Sony CSL
- Queries and matches against data from online music databases

### Straights C++ | GTK+

A card game that supports upto 4 human/AI players and features a complete GUI.

### MeetMe JAVASCRIPT | GOOGLE MAPS API

A Google Maps application that finds the closest meeting location, such as a coffee shop, between multiple points

## Education

---

### University of Waterloo

*Waterloo, Ontario*

CANDIDATE FOR BACHELOR OF SOFTWARE ENGINEERING

*2014 - 2019 (Expected)*