Jack Wang

437-983-4168 | jj22wang@edu.uwaterloo.ca | github/Jacqueuela

Skills

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

Tools Keras, Tensorflow, Scikit Learn, Pandas, Spark, Hadoop, Kafka, RabbitMQ, Elastic Search, AWS, GCP, Heroku, Docker, Git

Experience

Cardiogram

San Francisco, California

Software Engineer Intern

May 2017 - August 2017

- Improved neural network infrastructure to automate machine allocation, training and performance analysis of disease detection models
- Updated existing neural network architectures to be trainable on multiple GPUs
- Developed a neural network to predict sleep with a 97% area under receiver operating characteristic curve
- Designed and implemented a service to generate and serve sleep predictions to consumers using the Cardiogram mobile app
- Performed query and index optimizations to reduce the database load from Cardiogram mobile app workers

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 adverstising campaigns
- Deployed projects to local Hadoop Clusters and AWS EC2 clusters of over 300 machines

Education

University of Waterloo

Waterloo, Ontario

Candidate for Bachelor of Software Engineering

2014 - 2019 (Expected)