github.com/**csheng98**
linkedin.com/in/**csheng98**

CARL SHEN

csheng98.github.io

(+1) 226-700-3906
c42shen@uwaterloo.ca

SKILLS

Languages – Java, JavaScript, C#, C++, C, Python, R, Matlab, SQL
Technologies – Apache Spark (Core, Streaming, MLlib), Apache Storm, Apache Kafka, Microsoft Azure, Hadoop, Redis, Node.js, Android, Weka Machine Learning

EXPERIENCE

Yahoo Sports — Sunnyvale, *CA*
Daily Fantasy Backend Team – Software Engineering Intern — Jan - Apr *2018*

- Backend platform and API development in Java, integrating with MySQL, Apache Storm, Redis, and a central messaging queue service to build out features in the Yahoo Sports Daily Fantasy product (>1 million users).
- Developed a suite of APIs to implement an automated flow to replace a manual set of admin operations related to creating contests for specific types of slates, which previously took over an hour each day.
- Built and integrated an Apache Storm processor within an existing topology that ingested streams of live stats and calculated and aggregated live projected fantasy points for a lineup, increasing user engagement during contests.

SAP — Waterloo, *ON*
Big Data Tooling Team – Software Developer Intern — May - Aug *2017*

- Full stack JavaScript/Node.js development of features for a customer facing Database Explorer web app.
- Implemented dynamic filtering capabilities and an advanced data preview interface by leveraging SAP HANA's high speed in memory querying capabilities to allow users to better explore and visualize data.

Communications Research Centre — Ottawa, *ON*
Network Applications Group – Software Engineering Intern — Sep - Dec *2016*

- Prototyping and development in Java and C# and using Apache Spark and Microsoft Azure for data engineering applications for a city-wide radio spectrum usage testbed system (>10 GB of data generated per day).
- Created an automated batch data analysis pipeline integrating Apache Spark analytics scripts (written in Java) and an Azure Data Factory (using a C# custom activity), reducing execution time from about 5 hours to 30 minutes.

EDUCATION

University of Waterloo — Waterloo, *ON*
BASc Candidate – 3A Computer Engineering — Expected Apr *2020*

- Cumulative Average: 93.42% | Overall GPA: 3.98/4 | 2B Class Rank: 1/164
- Undergraduate Research Assistantships – contributed to groups focused on distributed systems & LDPC decoding.

PROJECTS

Spark Hockey Analytics *2018*

- Using Apache Spark (Core, Streaming, Machine Learning lib) and Apache Kafka to process data from NHL APIs for hockey analytics projects, including a "Redzone" recommendations script and a projected points model.

Spatial Data Visualization Tool *2017*

- Created an interactive data visualization tool for dynamically filtering and aggregating spatial data on a map view based on zoom level and bounds by leveraging a Geohash spatial index within an Apache Druid data store.

Parallelized User Behavior Prediction Model *2016*

- Implemented components of a novel smart home user behavior prediction model using a MapReduce and Hadoop framework to parallelize the algorithms, significantly improving speed of execution.
- Published as secondary author of paper in Cluster Computing 2017 Vol 20 Iss 2 (DOI: 10.1007/s10586-017-0749-z)

Smart Bed Monitoring System *2015*

- Built a system to recognize and monitor bed and sleep related activity by implementing a decision tree generated using the Weka machine learning library on an experimentally gathered supervised training data set.