CARL SHEN

github.com/**csheng8**
linkedin.com/in/**csheng8**
csheng8.github.io
+1 (226) 700-3906
carl.shen@uwaterloo.ca

SKILLS

Languages	–	Java, C++, C#, JavaScript, Python, C, SQL, Matlab
Concepts & Technologies	–	***Distributed/parallel computing**:* Spark, CUDA; ***NoSQL**:* Redis, Hadoop, Druid; ***Cloud**:* Azure; ***Data pipelines**:* Storm, Kafka, Gobblin; ***Machine Learning**:* Spark MLlib, Weka

EXPERIENCE

LinkedIn – Gobblin Team – Systems and Infrastructure Engineering Intern — Sunnyvale, *CA* | Sep – Dec *2018*

- Incoming intern on the Gobblin team, an open source distributed data integration framework for both streaming and batch data ecosystems.

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

- Owned and built a live projections feature to enhance the Daily Fantasy platform's live contest user experience by using an ***Storm topology*** and leveraging a ***Redis data store*** to implement a heuristic algorithm.
- Implemented an automated contest creation flow that replaced an hour-long, daily process by adding a set of APIs within a ***Jersey/Jackson & Datanucleus JDO/MySQL backend framework***.

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

- Developed a new advanced data preview interface within the Database Explorer web app by implementing APIs within a ***Node.js backend*** and leveraging ***SAP HANA***'s high speed in-memory operations.

Communications Research Centre Canada – Software Engineering Intern — Ottawa, *ON* | Sep - Dec *2016*

- Created an automated data engineering pipeline within a spectrum monitoring system that integrated ***Spark data analytics scripts*** using an ***Azure Data Factory***, reducing execution time by about 10x.
- Led a migration of my team's data analytics framework from Matlab to Spark, which involved porting existing data analytics scripts, setting up and deploying resources in Azure, as well as running weekly Spark tutorial sessions.

EDUCATION

University of Waterloo – BSE Candidate – 3A Software Engineering — Waterloo, *ON* | Expected Apr *2021*

- Overall GPA – ***3.981/4*** | Cumulative Average – ***93.42%***
- Awards – First in Class Scholarship (2018), Microsoft Tuition Scholarship (2017), President's Research Award (2017)
- Undergraduate Research Assistantships – Distributed Systems Lab, Stochastic Decoding Group

PROJECTS

Spark Hockey Analytics *2018*

- Using ***Spark*** and ***Kafka*** to process data from NHL APIs for hockey analytics projects, including a streaming script for providing recommendations for games to watch and a machine learning based projected points model.

Interactive Spatial Data Visualization Tool *2017*

- Created an interactive data visualization tool for dynamic filtering and aggregation of spatial data on a map view by leveraging a ***Geohash geocoding algorithm*** as a spatial index within an ***Druid NoSQL data store***.

User Behavior Prediction Model (Cluster Computing 2017 Vol 20 Iss 2 | DOI 10.1007/s10586-017-0749-z) *2016*

- Implemented components of a novel smart home user behavior prediction model using a ***MapReduce/Hadoop framework*** to parallelize the algorithms, improving execution speed by up to 5x.

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.