YUAN ZHANG

✉ yuanz2299@gmail.com ➤ yzhang0.github.io ⚙ yzhang0

SKILLS

Languages	Python (pandas, scikit-learn), R (dplyr, xgboost), SQL, C, C++, D3.js, HTML/CSS
Tools/OS	Hadoop (Spark, MapReduce, Hive, Sqoop), git, bash, LaTeX
Interests	Applied Machine Learning, Data Visualization, Deep Learning, Digital Marketing
Others	Native proficiency in English and French

EXPERIENCE

Scotiabank   Sep 2017 - Dec 2017
Data Science Intern

· Led a **cold start model** to predict digital banking likelihood of new customers with an accuracy of **73%**. Implemented the project using **supervised learning** in **Hive, Python** and **R**.
· Cleaned and analyzed millions of customer browsing data points using the **Hadoop Framework (HDFS, MapReduce, Spark)** to investigate the effeciency of third party data platform for marketing campaigns.
· Prepared data visualizations and presentations on how to leverage machine learning on the new digital platform for leadership of different technical familiarity.

IBM   Jan 2017 - Apr 2017
Data Analytics Developer

· Wrote **Python** scripts to detect anomaly and user trends using hourly warehouse data from **CouchDB**.
· Built an **automated** system to create scheduled data visualization using D3.js and Jinja2.
· Made recommendations on data visualizing libraries based on detailed experimentation with Elasticsearch, Kibana, Plot.ly, and NVD3.
· Built and tested a real-time data processing program using **reactive programming** to generate log files.

Agriculture and Agri-Food Canada   May 2016 - Aug 2016
Web Statistics Specialist

· Assisted users in interpreting **web statistics** reports and investigated possible discrepancies.
· Trained a team of two interns and five full time employees and wrote a 15-page reference manual on web analytics.

PROJECTS

Predict and Visualize Software Internships Trends at Waterloo   Jan 2018

· Built a reactive directed graph of 86 most recurrent companies and their number of interns using **D3.js**.
· Collected company ratings, revenue, etc. from Glassdoor using **Selenium** and **Requests** in Python.
· Implemented an **ensemble model** predicting with a MAE of **0.09** the score of companies to apply for given the student's work experience and 10 company features.

DSTL Satellite Image Feature Detection   In Progress

· Enhance the resolution of aerial layouts using panchromatic sharpening and georeference pixels.
· Predict the object type of satellite images using **Keras** and **TensorFlow** for **semantic segmentation**.

Best Selling Mangas Bubble Chart   Mar 2017

· Built a **bubble chart** with data from Wiki and external images from Mangapark using **D3.js** and Python.

ACTIVITIES AND VOLUNTEER EXPERIENCE

Future Blue Lead, *IBM*   Jan 2017 - Apr 2017
· Co-Organized bi-monthly career development workshops and social events for 30 interns.

Dept. of Statistics Representative, *Women in Math Committee*   Oct 2015 - Aug 2017
· Represented female students in the Statistics program during committee meetings.

AWARDS

GHC Scholarship 2016	Selected to attend Grace Hopper Celebration of Women in Computing.
Student Life Award 2015	Organized a bazaar and fundraised 1000$ for Nepal earthquake victims.

EDUCATION

University of Waterloo, Canada   2015 - 2019 (Expected)
Honours B. Math., Joint Major in Computer Science and Statistics
Coursework: Object-Oriented Programming, Data Structures and Algorithms, Forecasting, Applied Linear Models, General Linear Models, Intro to Combinatorics.