

JILL WANG

Data Scientist

📞 (647) 936-5468 @ jill.ycwang98@gmail.com 🔗 LinkedIn: linkedin.com/in/jill-ycwang98 | Portfolio: jillwang87.github.io

EXPERIENCE

Data Analytics Developer (via Adastra)

TD  05/2021 - Ongoing

Developed projects in the Data Governance and Analysis department of Business Banking

- Developing a Python based data lifecycle management (DLM) tool: a scanner that leverages either local files or Hadoop tables' metadata to predict security classifications and retention periods using ML, Computational Linguistics, and NLP
- **Technology: Impala, Python, pandas, NumPy, scikit-learn, NLP, tkinter, spaCy**
- ETL from Hadoop and developing tableau dashboards
- **Technology: Impala, Python, pandas, NumPy, Tableau, Oozie**

Data Science Consultant

Adastra North America  04/2021 - Ongoing

Data and AI consultancy

Data Scientist & Software Developer

Graphen (based in New York, USA)  06/2019 - 08/2019

Organization building machine learning platforms leveraging their graph database to mimic human brain functionality

- Developed an unsupervised Holt Winter model on time series data and a supervised XGBoost model on multivariate data, both for the purpose of performing finance behaviour anomaly detection
- Applied statistical methods to design a systematic way to calculate a risk (anomaly) score
- Advised company leadership on strategies to combine the company's core product, a fast and efficient graph database, with machine learning models that helped land a proof-of-concept project with one of the largest banks in Taiwan
- Developed the frontend of a React-based anti-money-laundering platform
- **Technologies: Python, NumPy, pandas, scikit-learn, XGBoost, Git, JavaScript, React, GraphQL, Elasticsearch, SASS, CSS3**

Software Developer

Entrepreneurship Society at University of Waterloo  01/2019 - 04/2020

University of Waterloo club connecting students to the Kitchener-Waterloo startup culture

- Transitioned the React-based static desktop website to a fully responsive, mobile friendly experience
- **Technologies: JavaScript, React, SASS, CSS3, Git**

PROGRAMMING LANGUAGES & TOOLS

Python	SQL
Tableau	Impala, Hive, Oozie (Hadoop)
R (Statistical Computing)	JavaScript, React
MATLAB	C, C++
Bash	Git

EDUCATION

Honours Bachelor of Mathematics in Computational Mathematics

University of Waterloo, 2020

Honours Bachelor of Mathematics in Statistics, Computing Technology Option

University of Waterloo, 2020

AWARD & PROJECTS

Spotify Analysis

- Use Spotify Web API to query for my streaming history and analyze my listening habits

• **Technologies: Python, Jupyter notebook**

 <https://github.com/jillwang87/spotify-analysis>

Jewellery Shop Prediction Top 4%

- Predicting whether an online jewellery shop visitor will purchase an item from the shop using XGBoost with 93.33% accuracy

• **Technologies: Python, scikit-learn, XGBoost**

 <https://www.kaggle.com/c/stat441datachallenge1/leaderboard>

Humpback Whale Identification

- Trained using a dataset with over 25,000 images and about 5000 classes
- Use a VGG19 inspired CNN and achieved a ranking of 1722/2121 on Kaggle

NOVA - MLH StarterHacks 2018 Best Pitch (Finalist) Award

An Android app with an AI chat-bot designed to connect university students to on-campus mental health resources

• **Technologies: Python, Java, Android, Facebook Messenger API, NLP**

 <https://devpost.com/software/nova-29i4zx>

VOLUNTEERING

Events Coordinator

StarterHacks

 08/2018 - 02/2019

Beginner-focused hackathon for over 1200 hackers