

# ALYSSA ZHANG

## aspiring data scientist

✉ alyssa.zhang@uwaterloo.ca | ☎ (416)-436-5885 | 🌐 [alyssaazhang](#) | [in/alyssaazhang/](#) | [alyssaazhang.github.io/](#)

## SKILLS

Languages & Data Formats: Python, R, SQL, C, C++, HTML/CSS, JSON, CSV

Tools & Software: Oracle SQL Developer, Jupyter Notebook, Google Colaboratory, IBM Watson, Git & GitHub, Apache Hadoop, Keras & TensorFlow, MS Word/Excel/Access/SharePoint/Power BI, Libraries (Pandas, NumPy, Matplotlib, Scikit-Learn, SciPy), Linux

## EXPERIENCE

### MACHINE LEARNING SYSTEMS ENGINEER

*Toyota Motor Manufacturing Canada (TMMC) – Innovation Lab* | Sep-Dec 2021

- Trained convolutional neural networks (CNN) with YOLOv3 for 2 live computer vision (CV) systems: used image classification and object detection to identify defects in vehicle parts on the assembly line
- Introduced new workflow for training deep learning model with Mask-RCNN and reduced model training time by **50%**
- Proposed and presented ideas to TMMC General Managers regarding the prevention of vehicle part inventory waste
- Assisted with media production and photography for **40+** engineers and the TMMC President during the annual Innovation Showcase

## PROJECTS

### TWITTER STATISTICS ([GitHub](#))

- Performed **data mining** on Twitter for trending topics via **Python Tweepy** and the **Twitter API**
- Conducted **data munging** and used **natural language processing** (NLP) techniques to create a word cloud
- Implemented **object-oriented programming** to collect live tweets related to the COVID-19 vaccine; displayed their locations on an interactive map via **MapQuest Geocoding API**
- Retrieved tweets as **JSON** objects and stored in a **MongoDB** database to display user information and activity on an interactive choropleth map

### FORECASTING WEATHER AND HOUSING PRICES ([GitHub](#))

- Estimated past and predicted future temperatures using **time series forecasting** on NOAA's weather data
- Implemented **simple linear regression** two ways: **Python SciPy** and a **Scikit-learn machine learning estimator**
- Generated sophisticated house price predictions with 8-variable **multiple linear regression** on Scikit-learn's California Housing data

### MNIST DIGITS ([Github](#))

- Trained and tested a model for predicting handwritten digits (from MNIST database) using **k-nearest neighbors algorithm**
- Measured and evaluated model accuracy with a **confusion matrix**, **classification report**, and **k-fold cross-validation**
- Performed **dimensionality reduction** and visualization with **Python Scikit-learn** and **Seaborn**

### COVID-19 DASHBOARD ([Github](#))

- Visualized COVID-19 statistics around the world on an interactive **Power BI** dashboard

## VOLUNTEER EXPERIENCE

### MAIN INFORMATION DESK

*Mississauga Hospital (Trillium Health Partners)* | 2019-2020

- Utilized patient database to locate patients, answered visitor inquiries (in-person and phone), assisted in wayfinding and deliveries; took extra shifts in patient discharge, gift shop management, and water deliveries to patients; accumulated **160+** hours
- Trained **3** new volunteers efficiently and comprehensively to increase productivity during peak visiting hours

### SWIM TEAM COACH

*St. Francis Xavier Secondary School* | 2018-2020

- Planned and designed training sets for **20+** swimmers, adjusting based on personal needs; corrected technique and introduced new skills; organized, led, and mentored swimmers at meets
- Number of medals at OFSAA (the highest level of high school swimming) increased by **50%**

## EDUCATION

### UNIVERSITY OF WATERLOO

*BCS, Honours Computer Science (Co-op)* | 2020-2025

- Varsity Swim Team – Athletic Financial Award recipient
- Relevant Courses: Computational Statistics & Data Analysis, Probability, Data Structures & Data Management