

Jiawei (Charlie) Yu

Data Scientist

519-729-9479

Jiawei.yu@mail.utoronto.ca

<https://jiawei-yu-97.github.io/covid19-dashboard/>

Pursuing M.S. in Applied Computing at University of Toronto (data science concentration).

1 year of industry experience in data science.

2 years of teaching assistantship experience.

Languages: Python, SQL, R, JavaScript, C++, VBA

Education

Master of Science in Applied Computing (Data Science)

University of Toronto

2020-present

Bachelor of Mathematics in Data Science (Statistics)

University of Waterloo

2015-2020

Work Experience

Teaching Assistant

University of Toronto

09/2020 - present

Data Scientist

Carta Inc., Waterloo

10/2019 – 12/2019

- Prediction of customer turnovers using feature engineering and GLM models.
- Answering business intelligence questions in complex SQL queries.

Data Scientist

Kik Interactive, Waterloo

09/2019 – 10/2019

- Machine detection of spammers (feature engineering and XGB modeling).
- Data exploration and design of manual rules to detect and categorize spammers.

Software Developer (Data Engineering)

Miovision Technologies, Waterloo

01/2019 – 04/2019

- Computer Vision: development and assessment of object detection models using a variety of datasets and techniques. Examples include models that improves night performance, class-balanced

models that improves under-represented classes, and models fitted to specific backgrounds. Gains ranged from 15% to 30%.

- Implemented tight rotation (of images and bounding boxes) as an addition to TensorFlow's data augmentation library for object detection in Python.
- Designed protocols to expand datasets with continuous data inflow, while maintaining train-valid integrity and fair grounds for model comparisons.
- Front-end development in Python and JavaScript for data and model visualizations.
- Relational database design, database transition from MongoDB to PostgreSQL and data backfill.

Teaching Assistant

University of Waterloo

05/2017 – 12/2018

- Courses: MATH135, MATH235, STAT231
- Grading assignments and tests. Held office hours to answer questions.

Junior Data Scientist

Intelligent Mechatronic Systems, Waterloo

01/2018 – 04/2018

- Analysis on telematics signals to gain data insights and explore new opportunities on driver ratings.
- Laid out, implemented, and tested internal python libraries for data quality control and sanity check.
- Designed and verified algorithms to detect and predict unlabelled events such as gear shifts.
- Applied K-Means to distinguish hybrid and traditional cars, achieved 97% accuracy.
- Data cleaning & transformation, creation and modification of static databases for quality review.

Rail Operations Co-op

Metrolinx, Toronto

01/2017 – 04/2017

- Created new reporting templates and mechanisms in Excel VBA.
- Improved reporting style by substituting macro commands for pivot tables and human effort.
- Cleaned old VBA codes to shorten runtime by 30-50%.
- Analyzed the on-time performance of GO trains, summarized major incidents, and provided statistics for decision-making in platform adjustments and schedule modifications.

Project Management Associate

PureLiving China (Shanghai) Ltd, Beijing

05/2016 - 08/2016

- Delivered audits, proposals and work reports for clients including Siemens, Google, P&G, McKinsey, etc.
- Installed, configured, debugged and repaired air monitoring devices at clients' site.