

# JOHN CHEN

📧 johntiger1.github.io  
in johnn-chen/

## Skills

### DATA SCIENCE

PyTorch  
scikit-learn  
PySpark  
Tensorflow  
xgboost  
SparkSQL  
NumPy  
Pandas  
nltk  
SpaCy  
matplotlib

### FULL STACK PROGRAMMING

Javascript  
Node.js  
AngularJS  
React  
Python  
C  
Java  
Bootstrap  
PL/SQL  
Postgres  
MongoDB  
C#

## Education

University of Toronto  
Bachelor's of Science in Computer Science 2019  
Focus in Artificial Intelligence, cGPA: 3.98/4.00

Sept. 2014 to May 2019

University of Toronto  
PhD in Computer Science 2023

Sept. 2019 to Current

Research areas: Deep Learning and Natural Language Processing. Advised by Professor Frank Rudzicz

## Employment

Air Miles (a LoyaltyOne Company)

Toronto, Ontario

Data Science Intern

May 2019 to Aug. 2019

- Developed several machine learning models end-to-end, including flight redemption propensity model. FRP model achieved **0.97 AUC ROC**, and has **9x lift** compared to baseline. Improved model via **seasonality adjustments**, **model-free feature selection** algorithm and **model-based genetic feature selection** algorithm, as well as hyperparameter tuning via **Bayesian Optimization** using scikit-optimize. Model has been used in business and has delivered **40% e-mail cost savings** in an acquisition campaign targeting over **500 000 customers**.
- Research and development in recommendation engine prototype for leading Canadian grocery chain. Achieved **111% basket coverage improvement** over baseline by using market basket analysis and frequency-based model.

Vector Institute

Toronto, Ontario

Research Intern

June 2018 to Current

Developed and implemented several end-to-end deep learning solutions including:

- character-level recurrent neural network language model as part of industry partnership with DataX
- medical image captioning. **Doubled F1 performance** by reformulating problem as multilabel classification problem.
- paragraph-level style transfer in text (research ongoing)
- massively multilabel, highly imbalanced classification. **Improved F1 performance by 156% over baseline pretrained** network via image augmentation and finetuning.

University of Toronto

Teaching Assistant

Sept. 2016 to Current

- TAed for **5 different courses**: CSC209 (Software Tools and Systems Programming), CSC263 (Data Structures and Analysis), CSC165 (Mathematical Expression and Reasoning for Computer Science), CSC324 (Functional Programming), and CSC420 (Computer Vision and Image Understanding)

Ontario Teachers' Pension Plan

Software Developer Intern, DevOps and Risk Analytics

Sept. 2017 to May 2018

- Improved deployment set-up time by 30%** by developing full-stack RESTful C# Web API + AngularJS web app for managing UrbanCode agent and server registry information.
- Contributed to performance gains of 50x in financial software through developing hybrid Powershell/C# solution to automatically generate, format and send e-mail reports to senior management daily.

Centre for Computational Medicine at SickKids Hospital

Research Student

May 2017 to Aug. 2017

- Awarded \$6000 NSERC research grant to conduct summer research into Bayesian networks to improve automatic disease diagnosis algorithm used in Phenotips web platform. **Sped up naive algorithm by 500x** via dynamic programming and multithreading.

## Awards

Vector Research Grant · Vector Institute

June 2018

Awarded to conduct deep learning research under supervision of Dr. Frank Rudzicz

1st Place, Cisco DevNet Hackathon · Cisco

June 2016

Created an IoT app for real-time monitoring of plants, using Cisco's Tropo VoIP service and RelayR API.

Rotary Club of Toronto Community Champion Scholarship ·  
Rotary Club of Toronto and Toronto Argonauts

June  
2014

Full tuition scholarship to University of Toronto