

Yilun (Tom) Zhang

Waterloo, Ontario, Canada
☎ +1 (519) 998 3545
✉ zhangyilun.0719@gmail.com
<http://zhangyilun.github.io/>

SUMMARY OF QUALIFICATIONS

- Over 7 years experience in statistical analysis, machine learning, deep learning and data visualization
- Superior capacity to quickly understand and apply new concepts developed through academic and work experience in different industries
- Highly experienced in Python (Transformers, PyTorch, scikit-learn, XGBoost, Keras, Matplotlib, BentoML, etc.), and moderate experience in R and Matlab
- Experienced with pipeline automation using Jenkins, Docker and Kubernetes for ML service hosting, and various SQL languages

PROFESSIONAL EXPERIENCE

Sep '19 - Now **Engineering Manager/Machine Learning Engineer**, *SoundHound*, Toronto, Canada.

- Managed a team of machine learning scientists and software engineers to build innovative solutions in natural language processing and understanding using traditional approaches and LLMs
- Led development of end-to-end machine learning research projects in areas of natural language processing, text (pair)/token/audio classification, phrase segmentation, and natural language/expression generation to solve business problems including failed query handling, ASR error detection and automatic training data generation
- Built automatic and scalable data processing pipelines and production-level machine learning solutions that are integrated into company's product offering (cloud & embedded)
- Dropped production NLU failure rate by over 40% and greatly improved user experience for in-car and voice ordering environment
- Filed several patents for our novel solutions
- Received top impact award in company's quarterly meeting

Aug '17 - Sep '19 **Data Scientist**, *Manulife Lab of Forward Thinking (LOFT)*, Toronto, Canada.

- Led multiple data science and machine learning projects, built various statistical models and deep learning models (NLP & CV) to provide insights to business or to semi-automate current repetitive business processes
- Researched on computer vision adversarial attack and defenses, used Grad-CAM to visually understand the effect of attacks on confusing the attention of deep neural networks and effectiveness of available defenses, demonstrated how adversarial attack will potentially influence insurance industries by using transfer learning to generate industry-specific cases, and presented research at Vector Institute ESS #2
- Worked closely with back-end and front-end engineers to design databases, create ETL pipelines for data flow and deploy RESTful APIs for exposing developed machine learning models

- May '15 - Aug '15 **Data Scientist (Co-op), Athos, Redwood City, USA.**
- Assessed existing algorithm, then individually invented, implemented and tested a new mathematical algorithm to calculate the Athos Score (currently in production app) representing the overall work-out score for the session
 - Implemented the first machine learning algorithm to predict lower body heart rate confidence (currently in production app)
 - Wrote Python and MySQL scripts to automate data cleaning, transfer and storage processes
 - Developed various sensor signal and other data visualization tools for problem detection and algorithm improvement

EDUCATION

- Sep '16 - Aug '17 **Master of Mathematics, Statistics, University of Waterloo, Waterloo, Canada.**
- **Relevant Courses:** causal inference, spatial data analysis, longitudinal data analysis, stochastic processes, machine learning
- Sep '11 - Apr '16 **Bachelor of Mathematics, Statistics and Computational Mathematics Co-op, University of Waterloo, Waterloo, Canada.**

OTHER EXPERIENCE | PROJECTS

- Nov '17 **Presented Research at Vector Institute Endless Summer School #2 Representing Manulife, Toronto, Canada.**
- **Title:** Adversarial Attacks and Defenses on Computer Vision Systems and Their Impact to Regulated Industries
- Apr '17 - May '17 **Top 10% & No.1 in NA, Int'l Data Mining Cup 2017 Representing University of Waterloo, Host in Europe, Remote Participation.**
- **Problem:** predict the revenue generated from online shopping actions
 - **Challenge:** 3 types of actions (view, basket, order) but only order will generate revenue, however the action type is not given in test data; exact time order of actions are not clear within a day if performed by the same customer (only the date is available); lots of missing values
 - **Model:** a single XGBoost regressor to predict revenue directly
- Aug '15 **1st Place, 6sense Data Hack, 6sense, San Francisco, USA.**
- **Data:** 1.6M reviews, 500k tips by 366k users for 61k businesses from Yelp
 - **Solution:** *FoodFates*, an end-to-end web application with map for users to choose potential restaurant location, then it outputs the predicted star rating and existing positive/negative reviews for similar restaurants in nearby area by correlating various characteristics of restaurants
 - **Model:** a single Adaptive Boosting regressor for rating prediction; NLTK for sentiment analysis
- Feb '15 **1st Place, Capital One Data Mining Cup, University of Waterloo, Waterloo, Canada.**
- **Problem:** predict the break-even bid price for search engine advertisement
 - **Challenge:** highly imbalanced dataset; need to calculate/predict multiple intermediate values before calculating the break-even bid ($= \text{product revenue} \times \text{conversion rate} \times \text{approval rate}$)
 - **Model:** a classification and regression tree (CART) to predict conversion rate
- Oct '16 **Code and the City Idea Jam, Mississauga, Canada.**
- **Problem:** if Mississauga is planning to extend subway lines, where should the stations be built
 - **Solution:** used various open source geospatial data including residence, office buildings, eateries and parking ticket locations, applied clustering algorithm with fine-tuned weighted importance of different data dimensions to identify crowded and vital locations where a subway station will have the most benefits and impact