

Yilun Zhang - Data Science

Lab of Forward Thinking, Manulife

✉ zhangyilun.0719@gmail.com
<http://zhangyilun.github.io/>

SUMMARY OF QUALIFICATIONS

- 2 years experience in data analysis, machine learning and data visualization
- Superior capacity to quickly understand and apply new concepts developed through academic and work experience in different industries
- Extensive knowledge and experience in Excel, Matlab, Python (numpy, scipy, sklearn, pyspark, tensorflow, keras, seaborn, matplotlib), R (caret, ggplot2, dplyr, glmnet) and SAS
- Proficient in SQL languages and \LaTeX ; basic knowledge in Java and C++

PROFESSIONAL EXPERIENCE

- May 2015 - Aug 2015 **Data Scientist (Intern)**, *Athos*, Redwood City, CA, USA.
- Assessed existing algorithm, then individually invented, implemented and tested a new algorithm to calculate the Athos Score (currently in production app)
 - Implemented the first machine learning algorithm to predict lower body heart rate confidence
 - Wrote Python and MySQL scripts to automate data cleaning, transfer and storage processes
 - Created sensor signal and other data visualizations for problem detection and algorithm improvement
- Sep 2014 - Dec 2014 **Insight Analyst (Intern)**, *MD Financial Management*, Ottawa, Canada.
- Conducted statistical analysis on a product pre-launch survey; designed and implemented predictive models to identify future product purchasers and a visualization application for the survey data using R with Shiny
 - Mined customer and survey databases to explore gender differences in investment behaviour
- Jan 2014 - May 2014 **Bioinformatics Data Analyst (Intern)**, *Princess Margaret Hospital*, Toronto, Canada.
- Implemented an experiment database using Python with SQLite3 which greatly improved efficiency for data standardization and sharing
 - Automated experimental data storage to portable SQLite3 database using Python and Excel
 - Applied mixed effect models and forecasting on a set of therapeutic experiment data using R; wrote a detailed statistical report in \LaTeX
- Feb 2013 - Aug 2013 **Business Research Assistant (Intern)**, *Institute of System Science, NUS*, Singapore.
- Analyzed customer survey data using SPSS, R and Excel and produced visualization plots for presentation and final report to clients
 - Mined Singapore's transportation data for commuters' travel pattern analysis

EDUCATION

- Sep 2016 - Aug 2017 **Master of Mathematics, Statistics**, *University of Waterloo*, Waterloo, Canada.
- Research: detail review of sentiment analysis literature; implemented semi-supervised machine learning method to classify the sentiment of survey final comments
- Sep 2011 - Apr 2016 **Bachelor of Mathematics, Statistics and Computational Mathematics Co-op**, *University of Waterloo*, Waterloo, Canada.
- Research: usage of interaction terms in Poisson regression models with particular attention to potential misconceptions within the sociology literature; conducted simulation study to illustrate proper usage of interaction terms in Poisson models in R
- Sep 2014 - Aug 2015 **Coursera & edX, Online Courses**, Remote.
- Data Science Specialization | Machine Learning | Model Thinking | Data Visualization (Coursera)
 - Introduction to Big Data with Apache Spark | Scalable Machine Learning (edX)

OTHER EXPERIENCE

- Apr 2017 - May 2017 **Top 10%, Team Leader, Int'l Data Mining Cup 2017 Representing UW**, *Remote*.
- May 2016 **Top Writer 2016 on Quora (I write about statistics, data science, etc.)**, *Quora*.
- Aug 2015 **1st Place, 6sense Data Hack**, *6sense*, San Francisco, CA, USA.
- Feb 2015 **1st Place, Capital One Data Mining Cup**, *University of Waterloo*, Waterloo, Canada.