Uyen Hoang

(416) 838-1310 u.hoangth@gmail.com https://uhoang.github.io/ uhoang

SUMMARY

A result-oriented data scientist with hands-on experience in building statistical models to address business problems via regression, segmentation, design of experiments and other forms of multivariate analysis. Major strengths include attention to detail, excellent problem solving skills, the capacity to work as a team player, and effective communication skills.

TECHNICAL SKILLS

Data mining techniques:

- Clustering methods, factor analysis, neural networks, ensemble learning, etc.
- Linear and logistic regression, multi-level/hierarchial models
- Text mining, non-parametric models

Programming languages:

- R (ggplot2, shiny, tidyr, devtools, data.table, h2o, caret, xgboost, etc.)
- Python (TensorFlow, Keras, SciKit-Learn, SciPy, NumPy, Pandas)
- SQL, and shell script

PROFESSIONAL EXPERIENCE

- **Mountain Equipment Co-op** **Toronto, ON**
  Freelance Data Scientist *Sept 2017 – Oct 2017*
  - Built a customer lifestyle segmentation model to construct scalable, customer-centric marketing solutions
  - Developed a R framework to automatically produce descriptive statistics for each cluster and carry out a pairwise comparison task on an extensive set of profiling questions
  - Applied a raking technique to re-balance a survey sample to improve its representation

- **Vox Pop Labs** **Toronto, ON**
  Data Scientist *Jul 2014 – Jul 2017*
  - Developed and implemented the individual ranking and clustering algorithms for the Toronto Star Sentimeter application, which attracted fifty thousand users in two weeks
  - Assisted in the development of internal tools that reduced the running time of the data analyses in half
  - Worked with political scientists to analyze Vote Compass survey data and provide insights into public opinion on a variety of key issues. Compiled reports that were covered weekly by national public broadcasting partners during election campaigns
  - Created a hybrid model for weighting non-representative samples to reduce the variability of sample estimators, and increase the model's prediction accuracy by 3%
  - Modified a crowdsourcing algorithm to improve an innovative method of predicting vote intention at the electoral level

- **University of Toronto** **Toronto, ON**
  Research Assistant *May 2010 – Nov 2013*
  - Carried out a power analysis to calculate the ideal sample size required to detect a given effect size in a large scale randomized control experiment to test interactive voice recording intervention in an industry project
  - Provided statistical advice, helped to compile data files, and successfully completed the analyses for the academic paper: "Automatic planning prompt reduces credit card data delinquency"
  - Performed graphical computation in R, coauthored the research paper "Vector Exponential Models and Second Order Inference", and the paper "A tutorial on tangent exponential models"

EDUCATION

Certificate: Sequence Models (100%) by deeplearning.ai on Coursera *Mar 2018*

PhD in Statistics (Incomplete). University of Toronto. Toronto, ON 2011 – 2013

Master of Science in Statistics. University of Toronto. Toronto, ON 2010 – 2011

Bachelor of Science (Honours) in Statistics. University of Toronto. Toronto, ON 2006 – 2010