Ruize (Richard) Luo

3112 - 33 Sheppard Avenue East, North York, Ontario, Canada

☐ +1 (647) 339-5825 | ✉ ruize.luo@mail.utoronto.ca | ⌨ r-luo | 🔗 luoruize

Education

University of Toronto   *Toronto, Ontario, Canada*
Non-Degree Graduate Program, Computer Science   *Sep. 2017 - Dec 2017*
GPA: 4.0/4.0

- CSC2420: Algorithm Design, Analysis and Theory (A+)
- CSC2221: Introduction to the Theory of Distributed Computing (A+)

University of Toronto   *Toronto, Ontario, Canada*
Honours Bachelor of Science, Statistical Sciences   *Sep. 2011 - Jun. 2015*
CGPA: 3.94/4.00
Major Course Average: 95.4/100.0
Core Courses:

- STA414: Statistical Methods for Data Mining and Machine Learning (100/100)
- STA410: Statistical Computation (98/100)
- STA447: Stochastic Processes (95/100)
- STA457: Time Series Analysis (92/100)
- STA355: Theory of Statistical Practice (94/100)
- MAT357: Foundations of Real Analysis (99/100)

Work Experiences

Capital One Bank   *Toronto, Ontario, Canada*
Senior Data Scientist   *Sep. 2015 - Present*

- Fully owned the development of the next generation of our application fraud model
  - Migrated current model scoring pipeline from legacy platform to Python on AWS.
  - Leading the development on real time model with our streaming platform.
  - Exploring features generated from graph database (Neo4j).
  - Experimented with multiple hyperparameter tuning and feature selection techniques.
- Built a core business model and its end-to-end scoring pipeline for predicting customer default.
  - The whole process involved sample selection, data pull, data cleaning and validation, feature engineering and selection, model build and validation, model deployment, and documentation.
  - Worked with datasets on the magnitude of 100GB.
  - Used SQL, SAS and Tableau for data cleaning and validation; used GBM and Bayesian model selection techniques in R for feature selection; explored genetic algorithm along the way.
  - With careful data engineering (various variance stabilizing transformations, and splines) and model selection, our logistic regression model performed similarly to tree-based GBMs while being much more interpretable and stable over distributional shifts
  - Deployed the model on our internal scoring platform as a Python package
  - Guided by Agile principles: fast iterations of minimum viable product, quick adaptation to changes, and integration with business team for smooth communication
  - Demonstrated the project to the broader business team and presented at company's internal conference.
- Wrote Python package to perform large-scale feature engineering in parallel on AWS. The package is now used across the company to standardize feature generation and documentation.
- Built Monte Carlo simulation tool in Python to simulate customer call queue and waiting time. It has greatly improved call centre staffing to shorten customer waiting time while reducing cost.
- Continuous monitoring of the input and performance of all internal statistical models with Tableau.

Projects

Bone X-Ray Deep Learning Competition from Stanford ML Group   *Toronto, Ontario, Canada*
AI SQUARED FORUM   *Jun 2016 - Present*
- The challenge is to determine whether the bones presented in the x-ray pictures are normal. There are $\sim$38K images from $\sim$15K different studies.
- Led a team to learn computer vision techniques by participating in the challenge.
- Set up data pre-processing pipelines, training pipelines and designed evaluation metrics with Keras. Achieved good result with an ensemble of DenseNet trained with Adam using cyclical learning rate.
- **Github**: DeepMachineLearning/mura-team2

Wasserstien GAN   *Toronto, Ontario, Canada*
EXPLORATORY PROJECT   *Jan 2016*
- Learned TensorFlow with exploratory project.
- Explored autoencoder and Wasserstein GAN (WGAN) on MNIST data. Implemented Conditional WGAN.
- **Github**: r-luo/tensorflow-try

STA490: Statistical Consultation, University of Toronto   *Toronto, Ontario, Canada*
COURSE PROJECT   *Oct. 2014 - May. 2015*
- Collaborated as a group with a student on her research project in Ecology and Evolutionary Biology. The project inspects the behavioural patterns of different groups of golden headed lion tamarins (GHLT) in the presence of various predictors.
  - Applied various clustering methods on 10 years of field observational data ($\sim$140K observations) to group predictors based on the categories of their preys.
  - Used Markov Chain to model GHLT behaviour changes in time. Wrote R program to estimate the transition matrix, stationary distribution and confidence intervals.

Hart House Investment Club, University of Toronto   *Toronto, Ontario, Canada*
SENIOR ANALYST | QUANTITATIVE ANALYSIS TEAM   *Sep. 2014 - May. 2015*
- Used other teams' analyses as prior along with historical data to produce posterior distribution of stock prices, then hedge the portfolio based on the results with sector ETFs.

Honors & Awards

2012 - 2015	**Samuel Eckler Scholarship in Actuarial Science**, University of Toronto
2012 - 2015	**Victoria College Dean's List**, University of Toronto
2013	**5th place**, University of Toronto Undergraduate Mathematics Competition
2013	**The Edward Gladstone Schafer Memorial Scholarship**, University of Toronto
2011	**University of Toronto Scholars**, University of Toronto