Ruize (Richard) Luo

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

+1 (647) 339-5825 | ruize.luo@mail.utoronto.ca | luoruize

Skills

Languages: Python (tensorflow, numpy, dask, scikit), R (data.table, ggplot2, Rshiny), SQL, SAS, VBA, Matlab
Models: Neural Networks (Autoencoder, MLP, CNN, GAN), Tree-based Models, Gradient Boosting Models, Markov Regime-Switching Models, Anomaly Detection, Model Interpretation
Techniques: Automatic Feature Creation, Large-scale Feature Selection, Data Cleaning, Data Pipeline, Scoring Pipeline
Expertise: Excellent Communication, Agile, Prioritization, Fast Learner, Problem Solving, High Performance under Stress

Work Experiences

Capital One Bank
Toronto, Ontario, Canada

Senior Data Scientist
Sep. 2015 - Present

- Built a core business model and its end-to-end scoring pipeline for predicting customer charge-off.
  - The whole process involved sample selection, data pull, data cleaning and validation, feature creation and selection, model build and validation, model deployment, and documentation.
  - Worked with datasets of 40K columns by millions of rows.
  - Used SQL, SAS and Tableau for data cleaning and validation, and R for feature selection and model build
  - Deployed the model on our internal scoring platform as a Python package
- Wrote Python package to perform of large-scale feature engineering in parallel on AWS. The package was 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 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.
  - Monitored distribution shift in model inputs and the performance of model outputs; detect and resolve model failures.
  - Automated the process to generate monitoring reports.

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)

Projects

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 (∼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.