TIN NGUYEN

https://www.linkedin.com/in/tin-nguyen-879896125/ ⋄ https://github.com/tinnguyen96
732-858-2956 ⋄ tdn@mit.edu

EDUCATION

Massachusetts Institute of Technology — August 2018 - (Anticipated) December 2023
Doctor of Philosophy in Electrical Engineering and Computer Science — Cambridge, MA
· Thesis: Computational Methods and Sensitivity Detection for Bayesian Unsupervised Learning
· Completed en route Master's degree in June 2020
· Coursework: Bayesian Modeling and Inference, Automatic Speech Recognition

Princeton University — August 2014 - May 2018
Bachelor of Science in Engineering in Operations Research and Financial Engineering — Princeton, NJ
· Coursework: Machine Learning and Pattern Recognition, Neural Networks: Theory and Applications

RESEARCH EXPERIENCE

Massachusetts Institute of Technology — April 2021 - present
Research Assistant — Cambridge, MA
· Collaborated with an economist in a Mexican microcredit study to estimate the treatment effect and quantify robustness.
· Fitted Bayesian linear models using rstan and discovered that the treatment arm had higher profit.
· Visualized the relationship between profit and demographic information using Tableau.
· Programmed an R tool to identify outliers whose omission changes conclusions of Bayesian generalized linear models.
· Reported influential outliers in the microcredit study: omitting 1 datum out of 16,000 changed the treatment effect's sign.

Massachusetts Institute of Technology — October 2020 - March 2022
Research Assistant — Cambridge, MA
· Lead a 3-person research group from project inception to peer-reviewed publication.
· Designed an algorithm that reduces the time taken to fit probabilistic clustering models.
· Fitted Dirichlet process mixture models on agricultural and biological datasets in Python using numpy and scikit-learn.
· Discovered meaningful clusters in a gene expression dataset and visualized the clusters using principal component analysis.
· Won a Student Poster Award at New England Statistical Society 2022 meeting and an Oral Presentation at AISTATS 2022.

PERSONAL PROJECTS

Improving sports outcomes through supervised learning on simulated game footage — October 2021-present
Main Contributor
· Generated an image dataset with 10,000 observations by recording and processing FIFA game footage.
· Trained convolutional neural networks in PyTorch to predict the where the ball will travel to.
· Created an interactive app that supports the defensive team by highlighting pitch areas likely to receive the ball.

Predict ride time for Boston's public bike sharing system — May 2021-November 2021
Main Contributor
· Joined bike trips data table with weather data table using Spark SQL.
· Visualized ride time across 1,200,000 trips, grouping by gender, age and customer type.
· Fitted quantile regression in Databricks to predict ride time given time of day, weather, gender, age, and travel distance.

INDUSTRY EXPERIENCE

IBM Research — June 2019 - August 2019
Research Intern — Cambridge, MA
· Fitted time-varying Poisson process models on Los Angeles freeway traffic data.
· Designed a tool to approximate leave-one-out cross-validation in hidden Markov models and conditional random fields.
· Reduced the runtime by 2 orders of magnitude, while giving highly accurate estimate of cross-validation error metric.
· Conducted experiments in Python using numpy, scikit-learn, and JAX.

QUALIFICATIONS

Advanced beginner with SQL, Apache Spark, Tableau, and JavaScript.
Advanced with Python (numpy, scikit-learn, pandas, PyTorch, JAX, matplotlib) and R (lme4, rstan, tidyverse, ggplot2).