Colin (Kejin) **Wan**

New York University| Master of Science in Data Science 2020-2022
University of Toronto| Mathematical Application and Statistics 2015-2020

LinkedIn GitHub
+1 647-927-4080 colin.wan@nyu.edu
2102 - 22 Wellesley St E. Toronto, Ontario, Canada Canadian Citizen

Passionate graduate student specializing in data science, theoretical mathematics, applied statistics, and machine learning. Seeking to enhance my knowledge and practice skills of analysis through internship programs.

Relevant Courses

Mathematics	Measure Theory, Advanced Real Analysis, Topology, Group THeory, Nonlinear Optimization, Complex Analysis, Partial Differential Equation, Linear Algebra, Advanced Calculus, Mathematical Theory of Finance
Statistics	Statistical Methods for Machine Learning, Stochastic Calculus, Probability Theory, Stochastic Processes, Time Series Analysis, Methods of Data Analysis
Computer Science	Object Oriented Computer Programming, Machine Learning

Professional Experience

Present – September 2018 — Teaching Assistant | University of Toronto, Depart. of Mathematics and Depart. of Statistics, Toronto
- Taught MAT137 (Calculus), STA257 (Probability Theory), STA261 (Statistics).
- Conducted tutorials to help students review current topic.
- Hosted office-hours to clarify questions from students.
- Graded exams, tests, quizzes, and assignments.

Teaching | Mathematics | Statistics

August 2019 – September 2018 — Data Scientist | PricewaterhouseCoopers, Advisory/Assurance, Toronto
- Designed and implemented segmentation models and service recommender systems.
- Developed statistical models for synthetic population.
- Implemented clustering algorithms for client analysis.
- Cleaned and pre-processed data for analysis.

Data Analysis | Machine Learning | Python | Presentation | Client facing

December 2018 – May 2018 — Quantitative Research Analyst | Universal Portfolio, Research Team, Toronto
- Researched the field of digital currency and sought potential profiting opportunities.
- Modeled and predicted the future trend by quantitatively analysing available data.
- Back tested the current investment strategy and optimized current portfolio.
- Conduct market research using variety of statistical and machine learning methods

Research | Data Analysis | Python | Algorithm Implementation | Cryptocurrency

Research Projects

Present – September 2019 — High Frequency and Algorithmic Trading, University of Toronto, Department of Statistics, Toronto
- Analyzed and visualized high frequency trading algorithms.
- Modeled strategy outcomes and contrasted results of different strategies.
- Implemented stochastic models and simulated trading process for each algorithm.

Algorithmic Trading | MatLab | Python | Stochastic Differential Equations

Present – September 2018 — Synthetic Population via Copulas Based Dependency Model, PricewaterhouseCoopers, Toronto
- Proposed a copulas based algorithm to capture conditional dependencies among features.
- Compiled a sample population for downtown Toronto to verify the model.
- Submitted an accompanying paper for AAAI 2020.

Synthetic Population | Copula | AAAI 2020 | Python

December 2018 – May 2018 | **Local Feature interpretability of Black-Box Model, PricewaterhouseCoopers, Toronto**

- Proposed an VAE based algorithm to understand the black-box model of any classifier.
- Designed a VAE based architecture to capture latent feature of a given classifier.
- Testing the results of proposed architecture on existing complex models.

Model Interpretability | Variational AutoEncoder | Python

Publications

2020 SynC : A Unified Framework for Generating Synthetic Population with Gaussian Copula. - Kejin (Colin) Wan, Zheng Li, Yue Zhao, PPAI 2020

Skills

Programming : Python, R,SQL SAS, MatLab, JavaScript

Languages

English ●●●●●
Mandarin ●●●●●
French ●○○○○

Interests

- Stochastic Modeling
- Model Interpretability
- Synthetic Population

Honors and Awards

2015 President's Entrance Scholarships, University of Toronto
2016 Dean's List Scholar, University of Toronto
2018 Dean's List Scholar, University of Toronto

Competition

Datafest Competition - University of Toronto May 2020

DataFest 2020
Interpreted public sentiment change during the quarantine period towards current administration, social welfare, international and interracial conflict by analyzing millions of Twitter data using various Natural Language Processing techniques.
Presented final finding through Tableau Dashboard

Data Analysis | Machine Learning | Framework Design | Presentation

Datafest Competition - Second place, University of Toronto May 2019

DataFest 2019
Modeled and interpreted athlete's performance fluctuation through analyzing nutritional information, sleeping pattern, mental state, etc.
Designed a framework for the committee to provide custom approaches to improve athlete's physical and mental condition depending on the daily condition.

Data Analysis | Machine Learning | Framework Design | Presentation

Financial Data Case Competition, University of Toronto March 2018

Financial Data Case Competion
Analyze credit and mortgage history of real clients to construct classification model for credit limit and mortgage interest rate.

Data Analysis | Financial data

Outreach and Volunteering

January 2018 – May. 2018 Project Assistant, *RiskLab, University of Toronto*
March 2017 – August. 2017 Statistic Lecturer, *Talent Education, Toronto*
September 2015 to August 2018 VP of HR, *UTFUN, University of Toronto*