Weihao Li

Clayton, Victoria, Australia

+61 404 591 219 | wlii0039@student.monash.edu | TengMCing

Education

Monash University *Clayton, Australia*

BCom (Hons) in Econometrics *Mar. 2020 - Nov. 2020 (expected)*

- Studied units include collaborative and reproducible practices, Bayesian and frequentist econometrics and exploratory data analysis.

Monash University *Clayton, Australia*

BCom in Business Analytics *Mar. 2018 - Nov. 2019*

- GPA: 3.5/4.0
- Over 90/100 for 2 units: Algorithms and Programming in Python and Derivatives.
- Studied a variety of units covering different aspects of data analysis include data access, cleaning, visualization and modelling.
- Minored in Finance enhanced my investment analysis skills in currency, equity, debt and derivatives markets.

Guangdong University of Finance *Guangzhou, China*

BFin in Finance *Sep. 2015 - Jun. 2020*

- GPA: 3.4/4.0
- Best in class for 4 units: Microeconomics, Fundamentals of Finance, Human Resources Management and Managing Intercultural Communication.

Awards

2020 Econometrics Honours Memorial Scholarship, Monash University *$15,000*

Research interest

- Applications of machine learning in business, finance, economic and environmental science
- Computer vision and Natural Language Processing
- Data visualization
- Bayesian machine learning

Research experience

Using Remote Sensing Data to Understand Fire Ignition During the 2019-2020 Australian Bushfire Season

Monash University, 2020-present

- Honours degree research project, supervised by Prof Di Cook and Emily Dodwell

This project was using remote sensing data to understand the cause of 2019-2020 bushfires ignition and predict the fire risk in the neighbourhood. Developed a clustering algorithm to track bushfires from satellite hotspots data remotely. Proposed both statistical and computational models incorporating weather, vegetation and anthropogenic factors to predict the cause of bushfire ignition during 2019-2020 Australian bushfire season, which was unknown at the moment. Established a dynamic model predicting short-term fire risk in the neighbourhood primarily using satellite hotspots data. This allowed fast prediction in the condition of lack of field observations.

Data analysis projects

Predicting Wikipedia Daily Click Volume *2019*

Deep Learning Project

- Used deep learning to predict the next year daily click volume of a Wikipedia website given the data of the previous 500 days.
- Tested different deep learning architectures include pure Fully Connected neural network, LSTM, GRU and 1D Convolution neural network for this time-series task.
- Tuned Hyperparameters using cross-validated grid-search.
- Prediction performance was much better than ETS and ARIMA in most of the cases.

Predicting How Points End in Tennis *2019*

InClass Kaggle Competition

- Used machine learning algorithms to predict outcome categories of tennis points given 3D coordinates of the ball position and 2D coordinates of the player position throughout a match.
- Applied ensemble learning to aggregate the prediction of several neural networks, XGBoost, CatBoost and Random forests models which are trained independently.
- Top 10 of the private leaderboard.

Modelling And Predicting The Performance Of Portfolios Consisting of ASX200 Stocks *2019*

Financial Econometrics Project

- Used ARIMA-ARCH-type models to predict the log return and volatility of portfolios given the past performance.
- Successfully constructed a hypothetical portfolio with a high Sharpe ratio and passed a 3-months review.

Data Analysis Of Melbourne Airbnb Market *2018*

Data Modelling Project

- Used classical regression models to explore the factors affecting Melbourne Airbnb price and rating.

Volunteer activities

R workshops for Beginners *2019*

Helper

- A series of student ran workshops on data analysis in R.
- Workshop topics include introduction to R, Rstudio and Rmarkdown, data visualization in ggplot2, data types and import data in R, data wrangling with dplyr and handling missing data with naniar.

Publications

1. Li, W, E Dodwell, and D Cook (2020). A clustering algorithm to organize satellite hotspots data for the purpose of tracking bushfires remotely. *(in preparation)*.