

EDUCATION

國立東華大學
應用數學系 - 碩士，2017 年 9 月。

私立淡江大學
數學系 - 學士，2015 年 9 月。

WORK EXPERIENCE

- 東華大學經濟系-研究助理 2016 年 3 月 - 2017 年 8 月
- 協助分析國際金融數據 (G7)，模型驗證，參數估計 (SUR, MLE, Bootstrapping)，Lag Length 估計。比較 System Equations 與 Single Equation 估計參數的準確性。
- 東華大學數學系-教學助理 2015 年 9 月 - 2017 年 7 月
- 微積分、線性代數、統計學。

PROJECTS

- Grupo Bimbo Inventory Demand (Kaggle) 賽後分析進入前 8% 排名
- 時間序列問題，預測未來兩個星期，使用 XGBoost 進行建模。
- Rossmann Store Sales (Kaggle) 賽後分析進入前 10% 排名
- 時間序列問題，預測未來 48 天，使用 XGBoost 和 glmnet 進行集合建模。
- Bosch Production Line Performance (Kaggle) 賽後分析進入前 6% 排名
- 高度不平衡資料，不良品約占 0.6%。
 - 由於是生產線，資料具有高度遺失值 (超過 50%)。
 - 接近 1000 道製程，最後藉由其中 50 道製程，並使用 XGBoost 進行建模。
- Instacart Market Basket Analysis (Kaggle)
- 預測哪些商品會再次被購買，使用 XGBoost 進行建模。
 - 正式比賽，在進行一個月後，曾進入前 10% 排名，最後由於方向錯誤，排名落在前 25%。
- Open Source of Ptt Data
- Python 進行自動化爬蟲，每日更新，並架設 MySQL 對外開放。
 - 超過 600 萬篇文章，在 Github 獲得 80 個 Stars。
- CIFAR - 10
- 使用 Python - Keras - VGG16 建立 CNN 模型 (GPU 版本)。
 - 87.4 % 準確率。
- 台鐵驗證碼辨識
- 將驗證碼進行影像處理 (灰度化、去雜點、切割、旋轉)。
 - 使用 Python - Keras - VGG16 建立 CNN 模型 (GPU 版本)。
- Others
- 在 Kaggle 比賽上，主要使用 R 進行分析，並使用 parallel 和 dplyr 進行 Feature Engineering。
 - 爬蟲主要使用 Python 的 selenium、requests、BeautifulSoup 套件。

LANGUAGES

R, Python