林子軒 Lin Sam · ✉ samlin266118@gmail.com · 📞 0987198234 · **in** sam lin · ⦿ f496328mm ·

Summary

Experience with data mining, machine learning, and web crawling.
Hopes to focus more on data science and data engineer in future career.

Skills

- Experience with data mining and feature engineering by Python (**numpy, pandas, sklearn, multiprocessing, joblib**) and R (**parallel, dplyr, data.table, mice**).

- Experience with machine learning by Python (**xgboost-gpu**) and R (**xgboost, svm, random forest, knn**).

- Experience with deep learning by Python (**kears-CNN**).

- Experience with regression by R (**glm, glmnet, nls, sur, mle**).

- Experience with web crawling by Python (**request, BeautifulSoup, selenium**).

Projects

Bosch Production Line Performance (Kaggle)   **Post-competition analysis, top 6% rank.**

- Highly imbalance data, ratio is 1000 : 1, 10 GB dataset size. And the data is 50% missing value. More than 4000 variables, but I build models by only 50 features.

- The goal is to predict which parts will fail quality control.

Grupo Bimbo Inventory Demand (Kaggle)   **Post-competition analysis, top 8% rank.**
Rossmann Store Sales (Kaggle)   **Post-competition analysis, top 10% rank.**

- Time series problem, eighty millions data size. Building models predict inventory demand after 2 weeks and predict sales after 48 days.

Instacart Market Basket Analysis (Kaggle)   **Real competition, top 25% rank.**

- Predicting which products will an consumer purchase again.

Open Source of PTT Data   **99 stars on github.**

- Automatic crawling PTT data daily, and providing open data, more than six millions article, in MySQL.

FB-ChatBot (In development)

- Automatic ordering Taiwan train tickets, and recognizing Taiwan train verification codes by CNN models.

Financial Open Data (In development)

- Taiwan Stock Prices, Financial Statements, Stock Dividend, Institutional Investors buy and sell. Oil price. Exchange Rate of 53 countrys. Center bank Interest Rate of 12 countrys . Automatic update daily.

Education

National Dong Hwa University   **Tamkang University**
Master of Science, Sep. 2017.   Bachelor of Science, Sep. 2015.
Major : Mathematics and Statistics.   Major : Mathematics.
GPA : 3.87

Work Experience

NDHU - RA   **Mar. 2016 - Aug. 2017**

- Analysing G7 financial data. Model validation and parameter estimation by regression models (SUR, MLE, Bootstrapping). And comparing single equation estimators and confidence interval with system equation.

NDHU - TA   **Sep. 2015 - Jul. 2017**

- Calculus, Linear Algebra, Statistics.

Languages

R, Python.
Basic in English and proficient in Chinese.