

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**) 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 connecting MySQL by Python (**pymysql**) and R (**RMySQL**).
- Experience with regression by R (**glm, glmnet, nls, sur, mle**).
- Experience with web crawling by Python (**request, BeautifulSoup, selenium**).
- Execting deployment MySQL on ubuntu. Changing IP address to entity address by No-IP and installing SSL certificates by Let's Encrypt.

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 machine learning 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.**
- Time series problem, eighty millions data size. Building models to predict inventory demand after 2 weeks.
- Rossmann Store Sales (Kaggle)** **Post-competition analysis, top 10% rank.**
- Time series problem. Building models to 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** **80 stars on github.**
- Automatic crawling PTT data daily, and providing open data, more than six millions article, in MySQL.
- FB-ChatBot**
- Automatic ordering Taiwan train tickets, and recognizing Taiwan train verification codes by CNN models.

Education

National Dong Hwa University
Master of Science, Sep. 2017.
Major: Mathematics and Statistics.

Tamkang University
Bachelor of Science, Sep. 2015.
Major : Mathematics.

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.