

Group 30 Project Report

選擇的 Model1: Using Regression

Model chosen: xgboost 中的 xgbregressor

資料處理：

- 使用輕量版的資料
- 用最後一個星期做為 Train data 的 Y
- 把剩下的當作 Train data 的 X
- 去掉 User_ID 的 Column

建構模型：

- 為每一個 timeslot 建構一個模型(共建 28 個模型)
- 使用 XGBRegressor 做 Training

預測：

- 用每一個 timeslot 的模型去預測對應的 test data

結果：最高可達到 0.74181

選擇的模型 2：使用 RNN 中的 LSTM

Model chosen: tensorflow 中 keras 所提供的 LSTM 模型

資料處理：

- 使用輕量版資料
- 把 User_ID 去掉
- 把 training data 和 test data reshape 成(37,28,1)的三維資料

建構模型：

- 我們用 LSTM 來建構模型
- 我們設定 LSTM 跟 Dense 的 activation 都使用'sigmoid'
- 設定 LSTM 的 Dropout 為 0.375
- 設定 Learning rate = 0.0001

預測：

- 將 Test data 丟進去預測，得到輸出的 csv

結果：最高可達到 0.75263