

# AI\_term\_project

【人工智慧】期末作業

可以在 Github 上面[查看本專案](#)

## 開發環境

- 1. 初次使用請先同步環境 ( 電腦裡沒 uv 的[去安裝一下](#) )

```
uv sync
```

- 2. 後續開發：

```
uv run python main.py # 執行 main.py 檔案
uv run jupyterlab # 開啟 Jupyter Notebook
```

## 下載資料集

前往 [Google Drive](#) 下載資料集，解壓縮後放在 **data** 資料夾中。

data 資料夾應包含以下資料：

- **task1\_dataset\_kotae.csv** # 可由 **CityA Ground Truth Data.csv.gz** 解壓縮得到

## Jupyter Notebooks 說明

本專案的 Jupyter Notebook 分為以下幾類：

### 資料探索與基準模型

檔案	說明
<a href="#">model1.ipynb</a>	資料探索分析，計算人流統計與視覺化
<a href="#">model2.ipynb</a>	進階資料分析與人流預測模型開發
<a href="#">baseline_moving_average.ipynb</a>	<b>Baseline 模型</b> ：使用移動平均法 (Moving Average) 進行人流預測

### Seq2Seq 時間序列預測模型

檔案	說明
<a href="#">train_multivariate_seq2seq.ipynb</a>	<b>多變數 Seq2Seq 模型</b> ：結合多地點資料與週末標籤進行預測

檔案	說明
<a href="#">train_multivariate_timeperiod_seq2seq.ipynb</a>	多變數時段 <b>Seq2Seq</b> 模型：加入 One-Hot Encoding 的時段特徵 (早/午/晚/深夜)
<a href="#">hyperparameter_experiment.ipynb</a>	超參數實驗：Grid Search 測試不同 hidden_size、num_layers、learning_rate 組合

分類模型 (週間/週末判別)

檔案	說明
<a href="#">auto_label_weekend_and_train_dnn.ipynb</a>	<b>K-Means</b> 自動標籤 + <b>DNN</b> 分類器：使用非監督學習自動標記週末，並訓練 DNN 分類模型
<a href="#">model2_dnn_cnn_comparison.ipynb</a>	<b>DNN vs CNN</b> 比較：比較深度神經網路與卷積神經網路在時段分類任務的表現

其他

檔案	說明
<a href="#">test_script.ipynb</a>	測試用腳本，可查看當前環境是否能調用到 CUDA

## Models 檔案說明

訓練完成的模型與相關輸出檔案：

預測模型 (Seq2Seq)

檔案	說明
<a href="#">seq2seq_model.pth</a>	單變數 Seq2Seq 模型權重
<a href="#">seq2seq_multivariate.pth</a>	多變數 Seq2Seq 模型權重
<a href="#">seq2seq_multivariate_timeperiod.pth</a>	多變數 + 時段特徵 Seq2Seq 模型權重

分類模型

檔案	說明
<a href="#">dnn_time_classifier.pth</a>	DNN 時段分類器權重
<a href="#">cnn_time_classifier.pth</a>	CNN 時段分類器權重

資料預處理

檔案	說明
<a href="#">scaler.pkl</a>	單變數模型的 MinMaxScaler

檔案	說明
<a href="#">scaler_multivariate.pkl</a>	多變數模型的 MinMaxScaler
<a href="#">scaler_multivariate_timeperiod.pkl</a>	多變數時段模型的 MinMaxScaler

## 評估結果

檔案	說明
<a href="#">eval_log_baseline_ma.txt</a>	Baseline (Moving Average) 評估指標
<a href="#">eval_log_multivariate.txt</a>	多變數 Seq2Seq 評估指標
<a href="#">eval_log_multivariate_timeperiod.txt</a>	多變數時段 Seq2Seq 評估指標

## 視覺化結果

檔案	說明
<a href="#">prediction_result_baseline_ma.png</a>	Baseline 模型預測結果圖
<a href="#">prediction_result_multivariate.png</a>	多變數模型預測結果圖
<a href="#">prediction_result_multivariate_timeperiod.png</a>	多變數時段模型預測結果圖
<a href="#">training_loss_univariate.png</a>	單變數模型訓練 Loss 曲線
<a href="#">confusion_matrix_dnn.png</a>	DNN 分類器混淆矩陣
<a href="#">confusion_matrix_dnn_cnn.png</a>	DNN vs CNN 混淆矩陣比較
<a href="#">dnn_architecture_comparison.png</a>	DNN 架構比較圖
<a href="#">dnn_cnn_training_comparison.png</a>	DNN vs CNN 訓練過程比較
<a href="#">dnn_cnn_validation_comparison.png</a>	DNN vs CNN 驗證結果比較

## 超參數實驗結果

檔案	說明
<a href="#">hyperparameter_results/</a>	超參數搜尋結果資料夾
<a href="#">hyperparameter_results/hyperparameter_results.csv</a>	各超參數組合的實驗結果
<a href="#">hyperparameter_results/best_hyperparameters.txt</a>	最佳超參數配置
<a href="#">hyperparameter_results/hyperparameter_comparison.png</a>	超參數比較視覺化圖表