

HW3 共享腳踏車需求預測

Deadline:2025/05/28 23:55

隨著共享交通逐漸普及，如何預測未來的租借需求成為智慧城市、營運管理的重要議題。透過精準的需求預測，營運方可以更有效地調度腳踏車、優化資源配置，減少閒置與缺車的問題，提升使用者體驗與營運效率。本作業以華盛頓特區的共享腳踏車數據 (Bike Sharing Dataset) 為基礎，讓同學們練習利用時間序列預測模型 (LSTM)，進行短期需求預測。

請同學根據附件提供的**共享租借腳踏車資料集**，利用時間序列預測方法，預測 **2012/11/1~2012/11/20 二十天的腳踏車租借量**。本作業將著重於**實作 LSTM 模型進行需求預測**，並藉由參數調整、特徵選取來優化模型表現。

一、資料說明：

train.csv：2011/01/01 ~ 2012/10/31 每小時腳踏車租借數據

test.csv：2012/11/1 ~ 2012/11/20 每小時資料（不含租借量）

sample_submission.csv：作業繳交格式範例（預設為 0，請更新預測值後上傳 kaggle）

Readme.txt：相關參數說明供同學參考

注意：

雖然 test 資料是以每小時 (hourly) 為單位進行預測，但在提交作業時，請將「同一天」的 24 小時預測結果加總，轉換為「每日 (daily) 預測值」後，再填寫到 sample_submission.csv 中。

二、作業要求

I. 參與 Kaggle 競賽

- <https://www.kaggle.com/t/1dde80ad06d24d9a969e4e15a92c3ddc>
- 請根據 **train.csv** 訓練模型，並預測 **2012/11/1~2012/11/20** 每天的總租借量。請將預測結果上傳至 **Kaggle 作業競賽頁面** 進行評分（使用 MSE 評分）。
- 繳交格式請參考 sample_submission.csv

II. 報告與實作說明 (PPT/pdf)

請繳交一份簡報（不含封面封底，至多 6 頁），內容需包含以下部分：

- **實作流程與思路說明**（至多 3 頁）：
 - 模型架構
 - sliding window 設計
 - 輸入特徵選擇（是否有刪減不重要的欄位？為什麼？）
 - 預測流程說明

- **參數比較與討論（至多 3 頁）：**
 - 請針對至少 2 個參數（如：window size、LSTM 層數、重要特徵）進行調整實驗，並分析參數變動如何影響模型表現，說明你的觀察與結論。
- **其他有助於說明的分析（加分項）：**
 - 使用 SHAP 進行特徵重要性分析
 - 視覺化驗證集的預測結果與真實值對比

III. 程式碼 (ipynb)

繳交完整的 **ipynb** 檔案，內容包含：

- 訓練過程以及 loss 圖
- 其他輔助理解的圖表、分析皆可加分

三、評分方式

I. (25%)Kaggle 排名

II. (55%)PPT 實作說明

- (30%)實作流程與思路說明
- (20%)參數比較與討論
- (5%)其他加分項 (創意分析/視覺化等)

III. (20%)ipynb 實作內容 (完整度/清晰度)

四、繳交格式要求(請仔細詳閱，不符命名規定將酌予扣分)

I. Kaggle 請將名稱改為學號，無法分辨者不予計分

II. HW3_學號_姓名_實作報告.pdf

III. HW3_學號_姓名.ipynb

請將實作報告與程式碼壓縮為 HW3_學號_姓名.zip 後上傳至 E3。