

# 利用深度學習進行的竊電行為檢測





**01**

**背景說明**

**02**

**資料說明與觀察**

**03**

**資料前處理方法**

**04**

**流程介紹**

**05**

**分析方法**

**06**

**實驗方法**

**07**

**建議之決策方式**

**08**

**驗證結果**

**09**

**輔助決策工具DEMO**

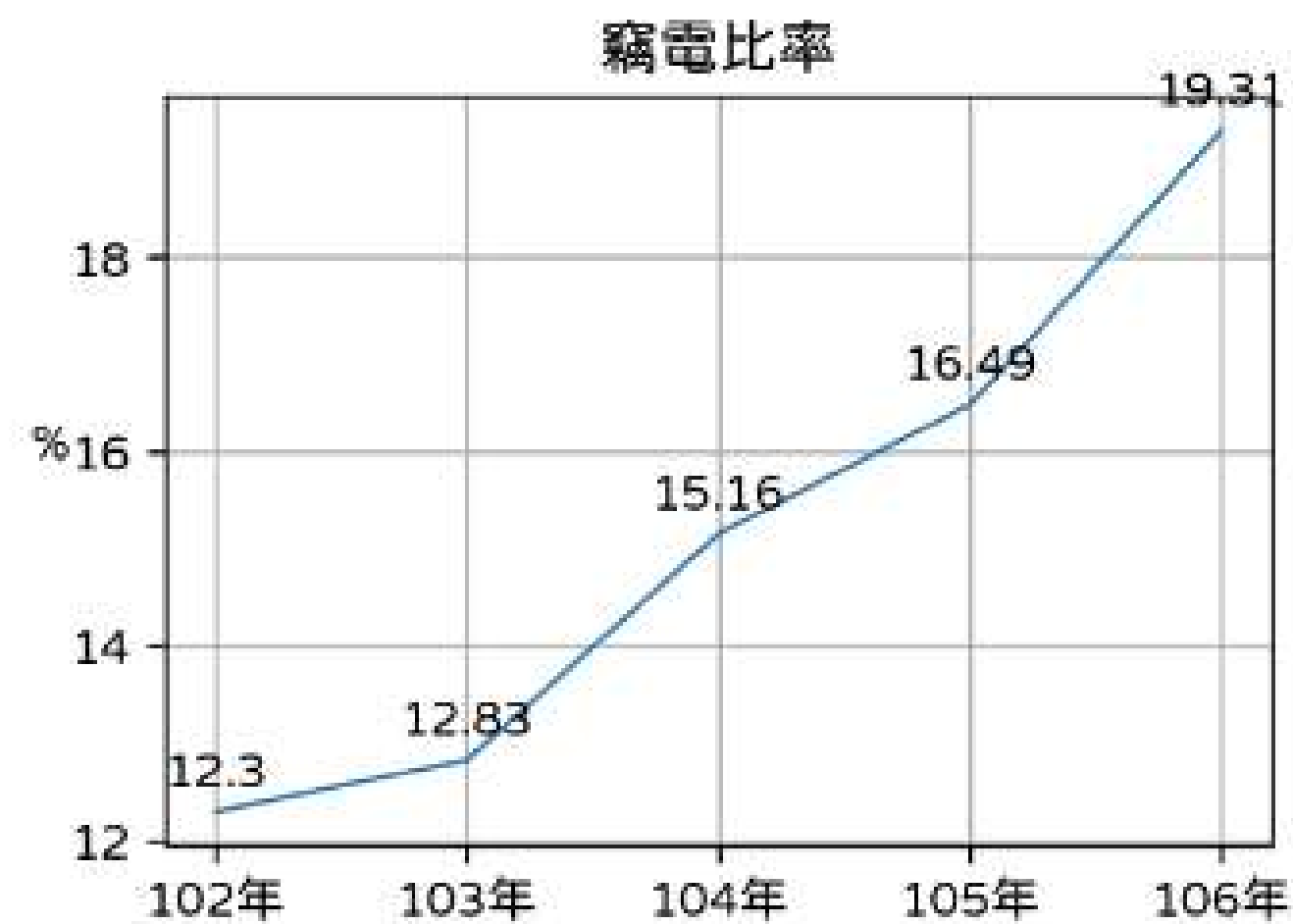
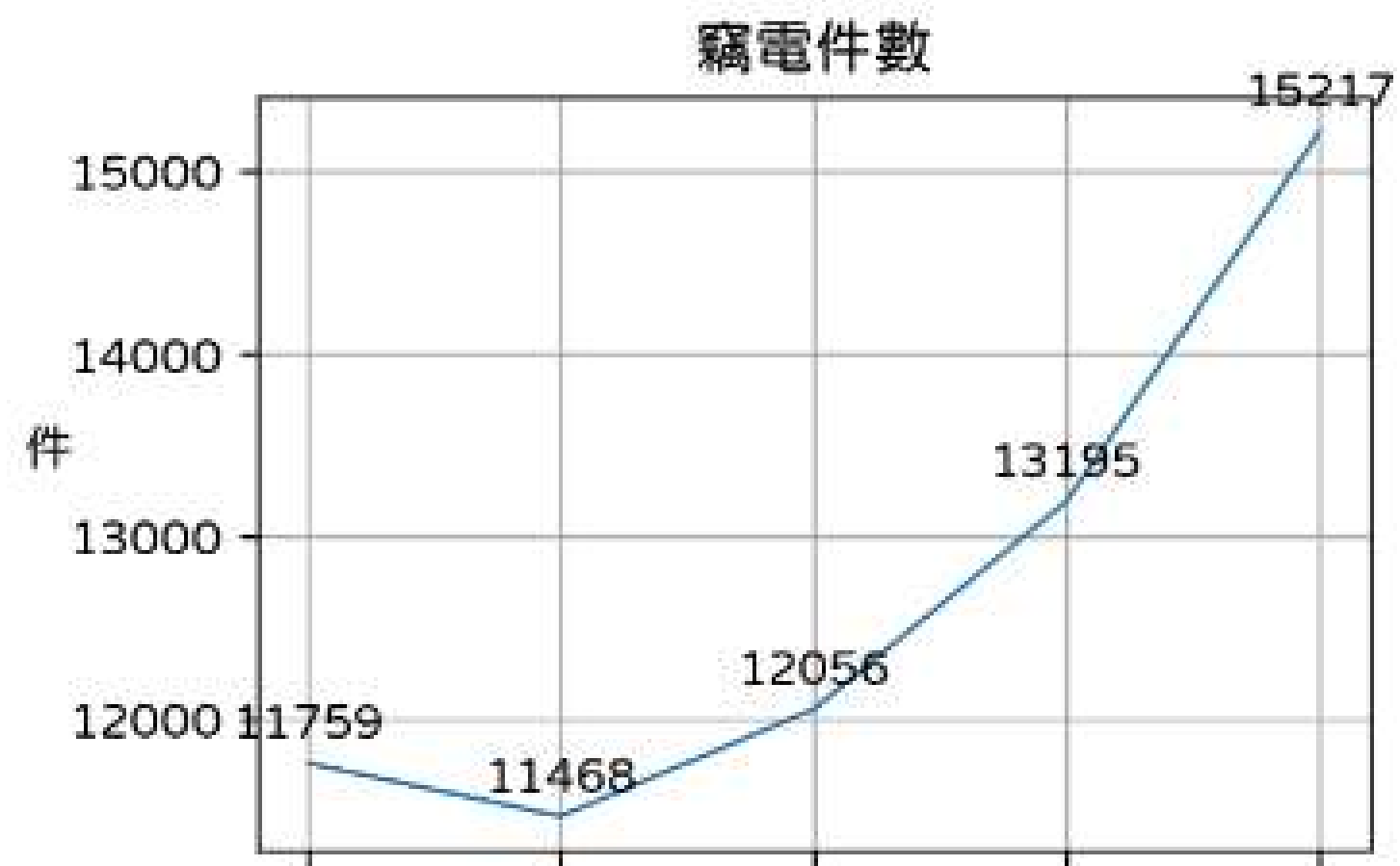




# 背景說明

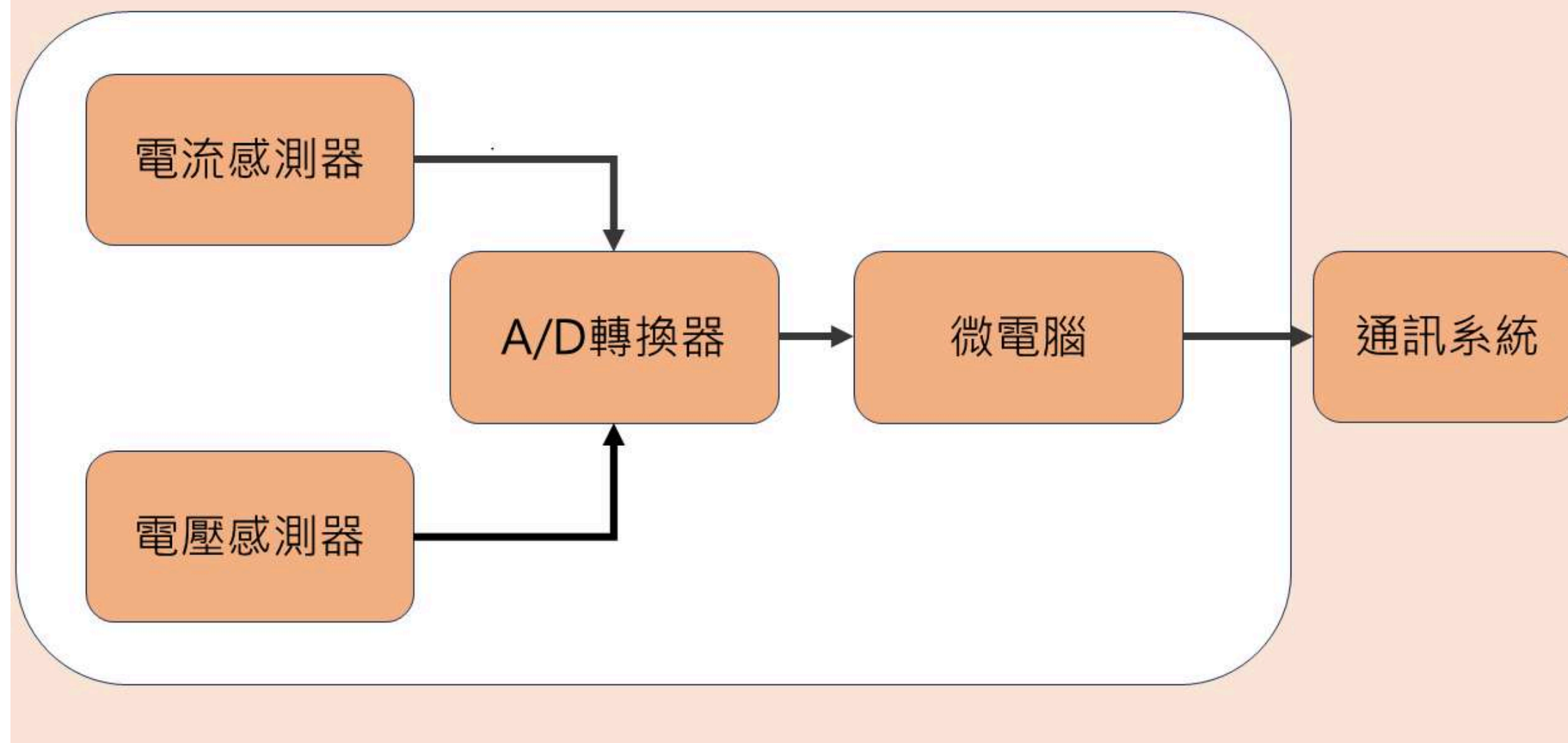


# 102到106年的偷電查獲比率



# 如何偷電

## 智慧型電表內部



跳過電表

測量數據

儲存數據

傳輸數據

# 如何偵測偷電

賽局理論

電網分析

硬體分析

機器學習



A photograph of a modern office interior, overlaid with a semi-transparent green filter. The office features wooden desks, ergonomic chairs, and numerous potted plants, including large indoor trees and hanging greenery. The space is bright and airy, with large windows in the background. The title '資料說明與觀察' is centered in white text.

# 資料說明與觀察



## 資料說明

透過SGCC資料集進行機器學習  
訓練模型以偵測偷電用戶。

製作可以輸入用戶用電量後  
預測用戶偷電嫌疑程度的網頁app。

透過追蹤偷電用戶開始偷電的  
時間點估算應追繳電費。



# 資料說明

原資料筆數：42372 筆

■ 訓練資料集：42222 筆

■ 測試資料集：150 筆

變數

資料說明

電表編號  
CONS\_NO

ID 碼



標記  
FLAG

0:沒偷電

1:有偷電

日期  
MM/DD/YYYY

某天的用電量 (2014/01/01-2016/10/31  
2016/09/18 缺失，共1034天)

# 資料說明

	訓練集資料	測試集資料	總個數
缺失值個數	11174212	42872	11217084
離群值個數	1705795	6105	1711900

原資料遺失率

25.6 %

訓練資料遺失率

25.29 %

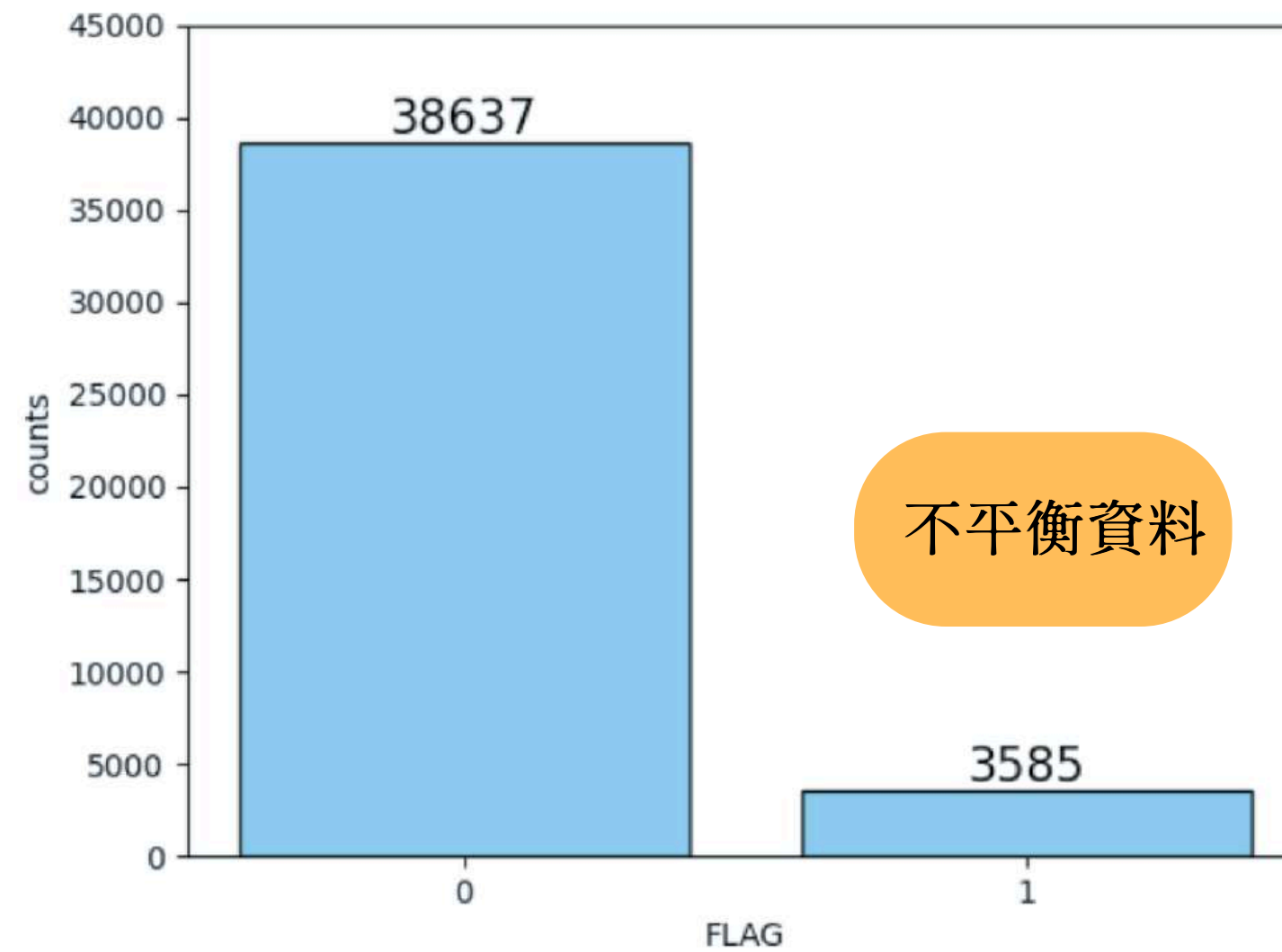
測試資料遺失率

27.64 %



# 資料觀察

## 標記 (FLAG) 累積次數圖



標記 FLAG

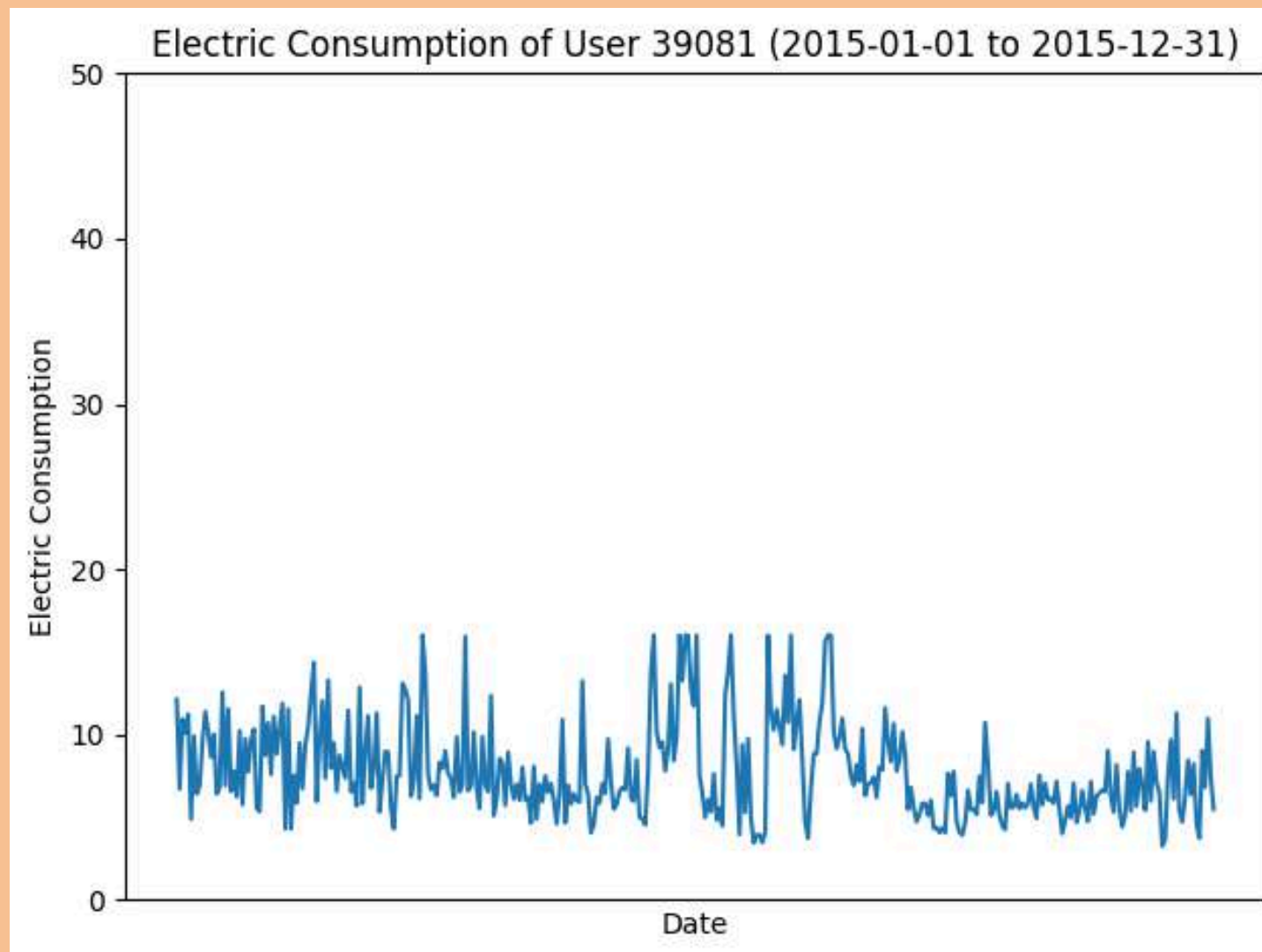
0 : 沒偷電

1 : 有偷電

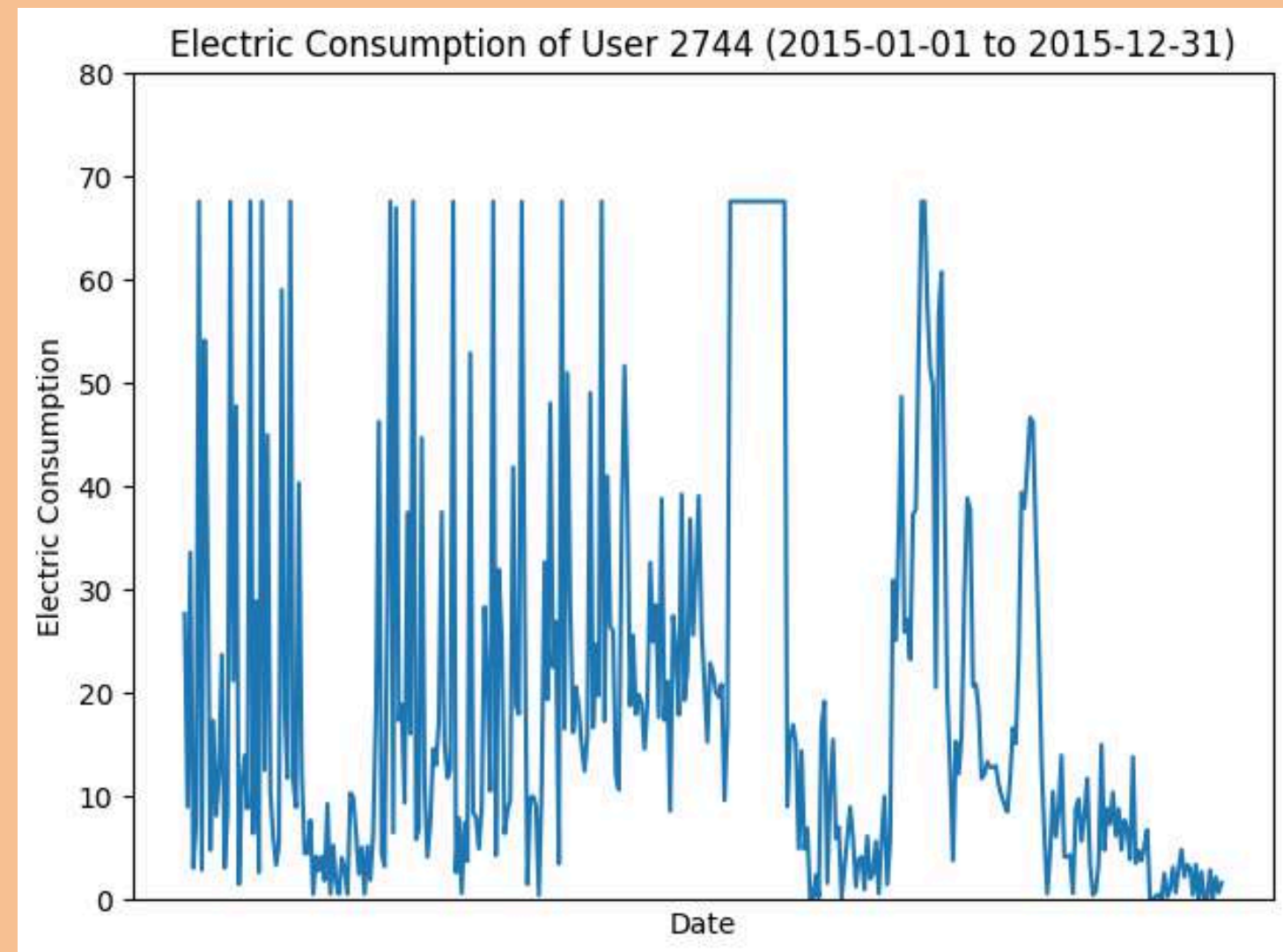
不平衡資料

# 資料觀察

比較一年中每日用電量  
(2015/01/01 ~ 2015/12/31)



FLAG=0 沒偷電

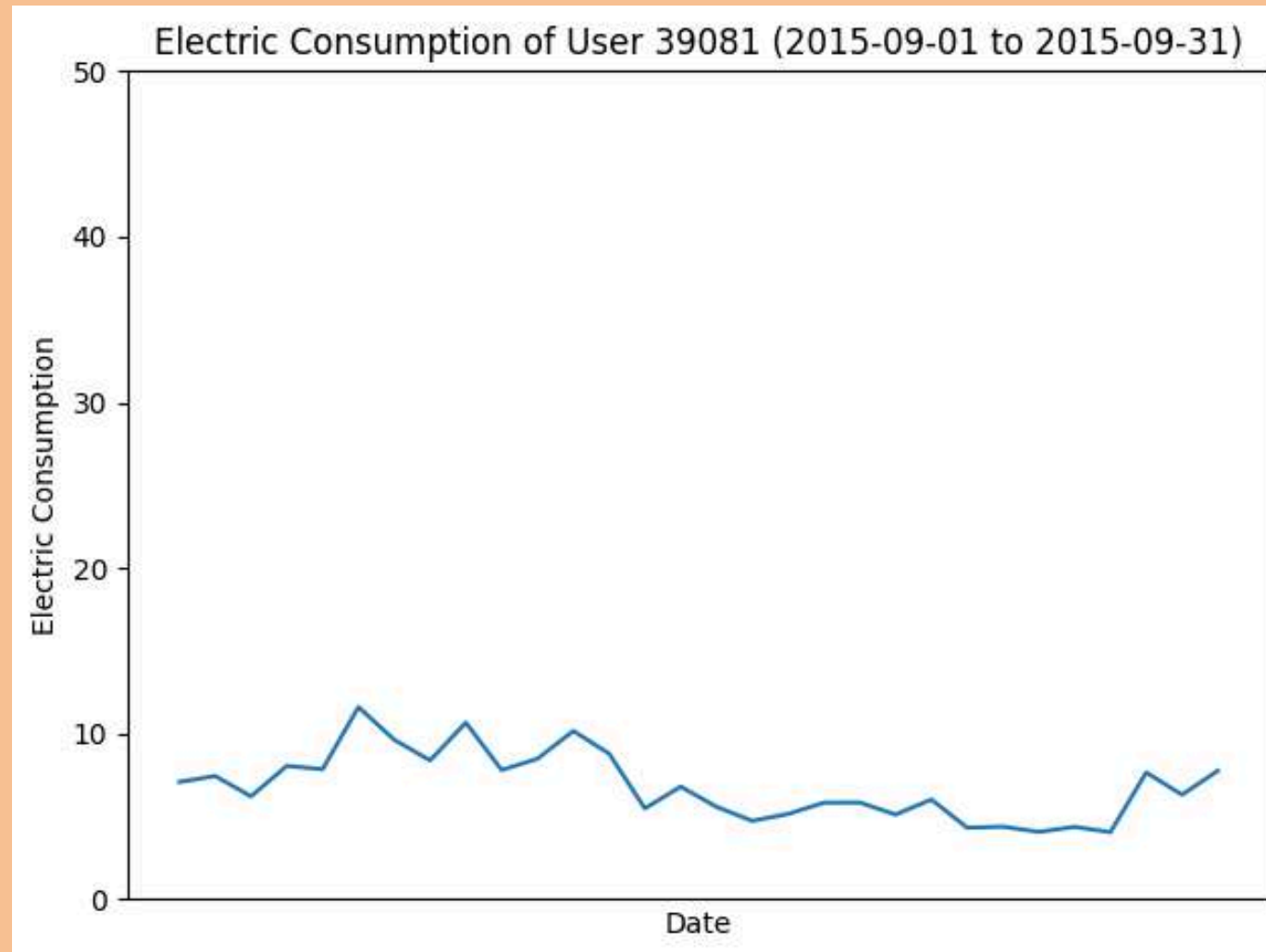


FLAG=1 有偷電

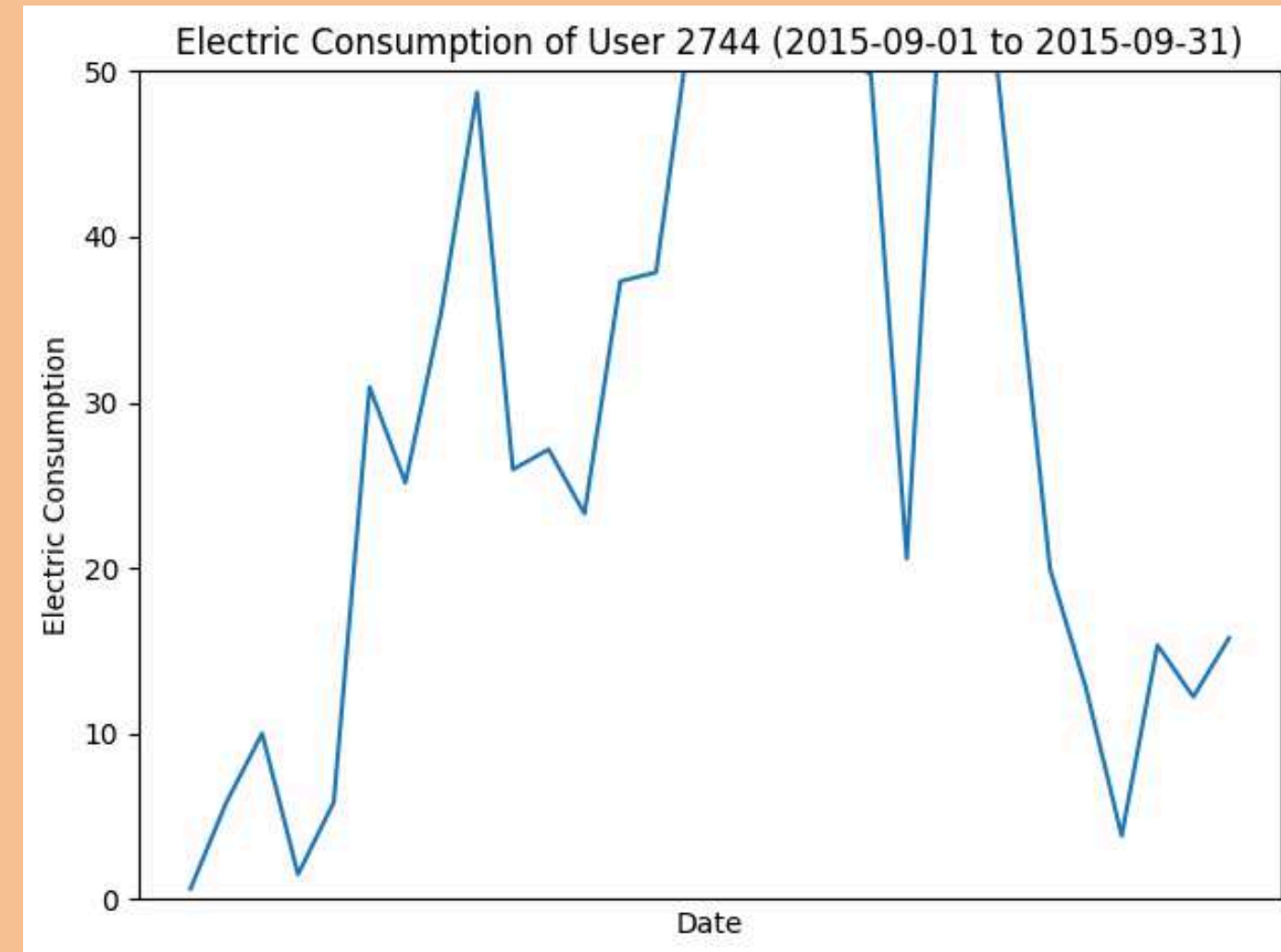


# 資料觀察

比較單月每日用電量  
(選擇用電高峰期：2015 年 9 月 / 夏季)



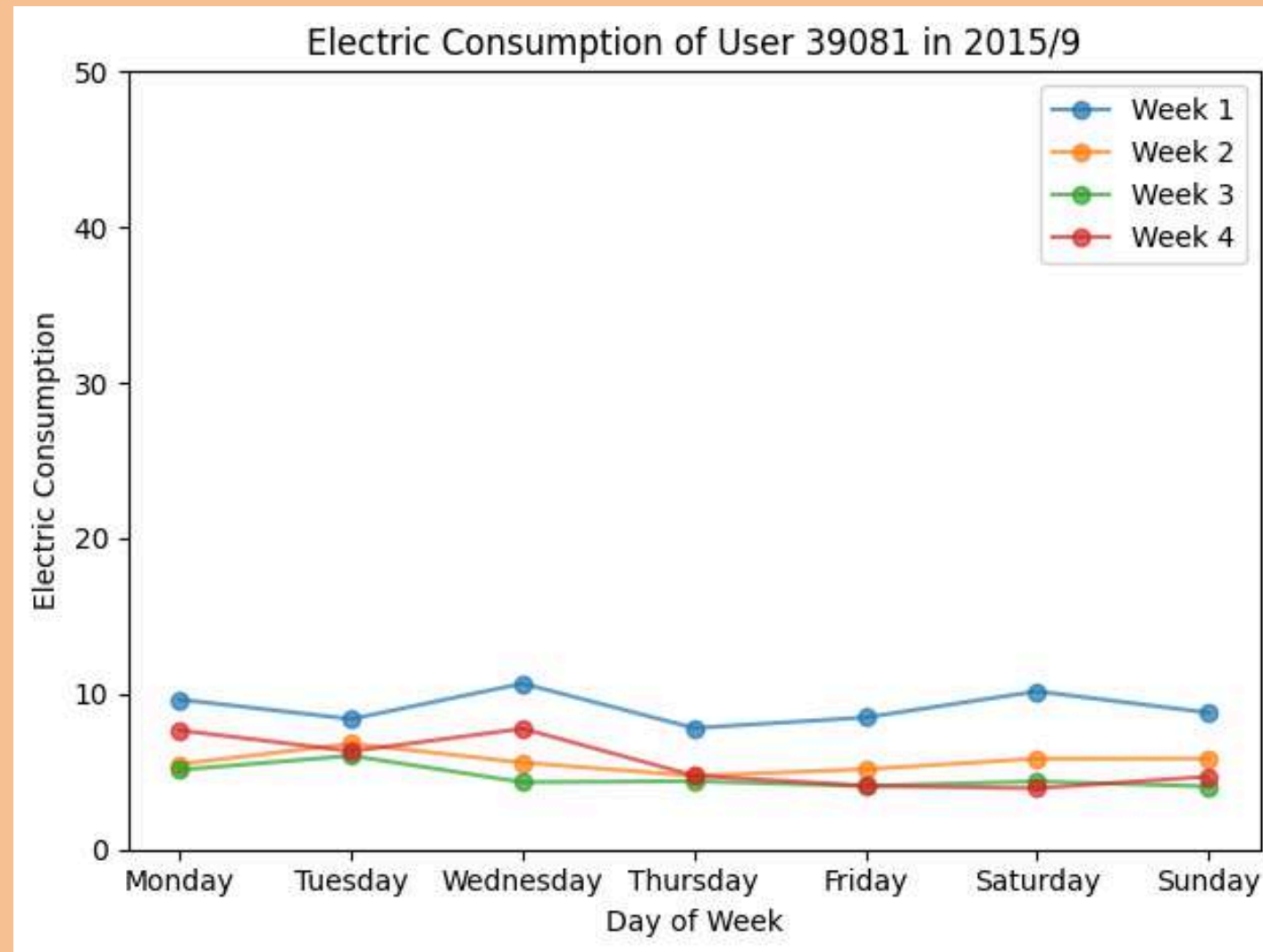
FLAG=0 沒偷電



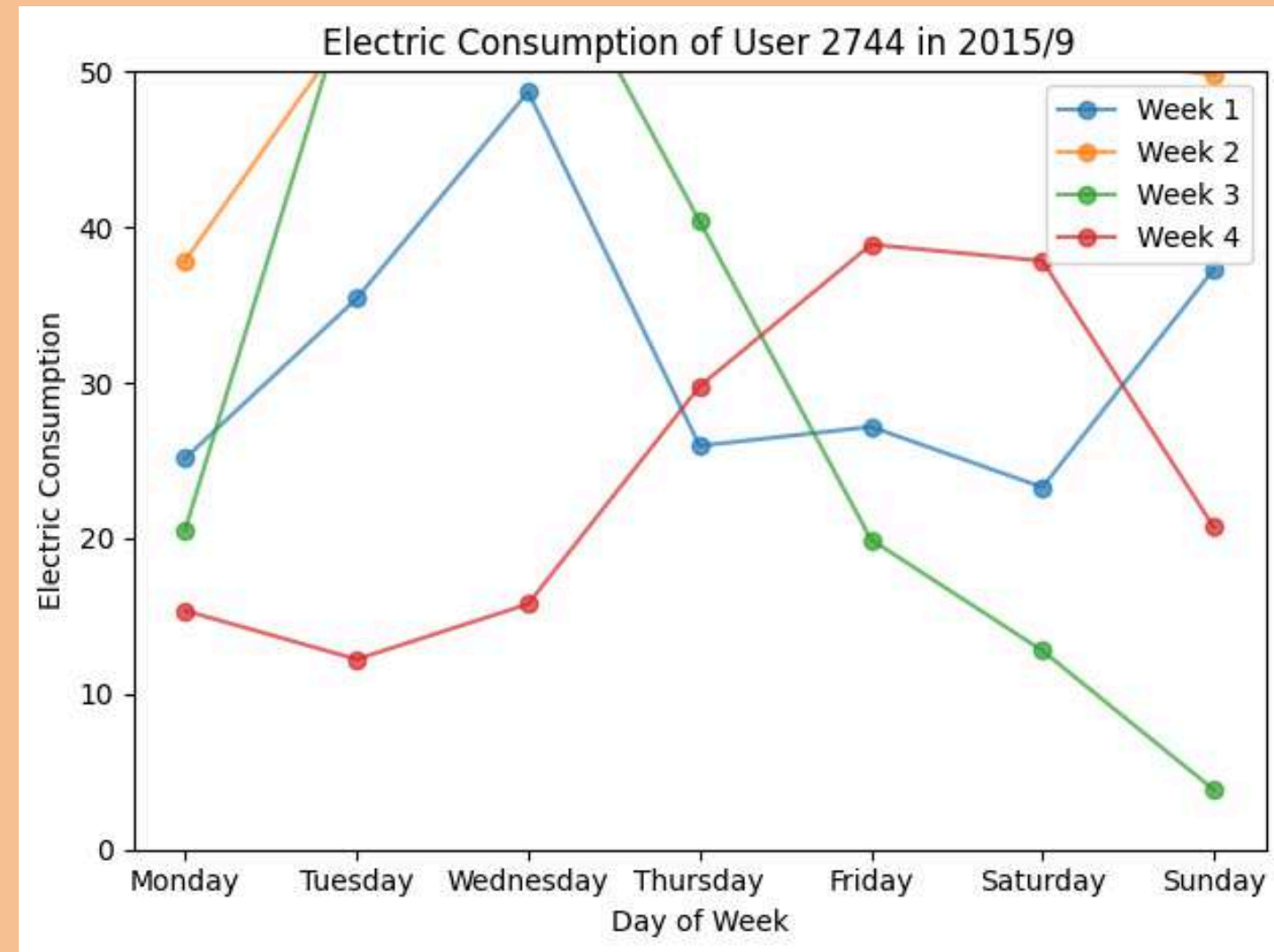
FLAG=1 有偷電

# 資料觀察

比較單月每周用電量  
(選擇用電高峰期：2015 年 9 月 / 夏季)



FLAG=0 沒偷電

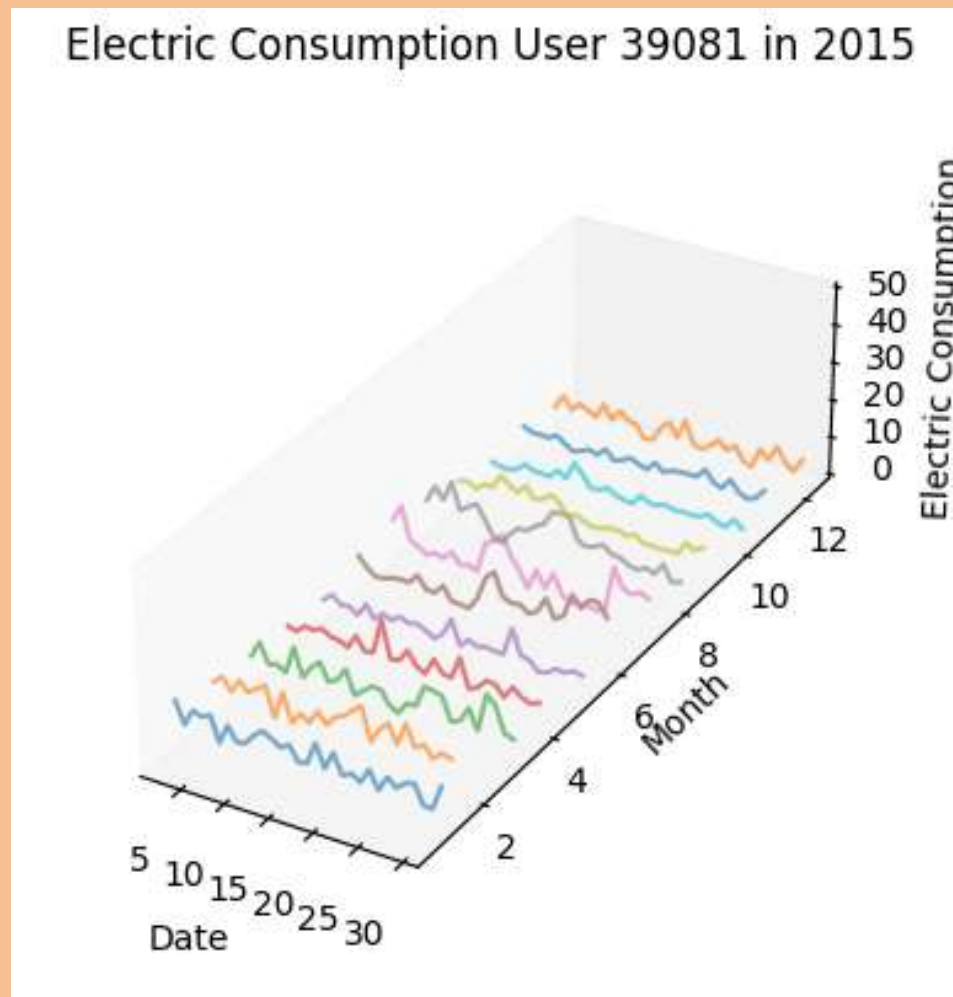


FLAG=1 有偷電

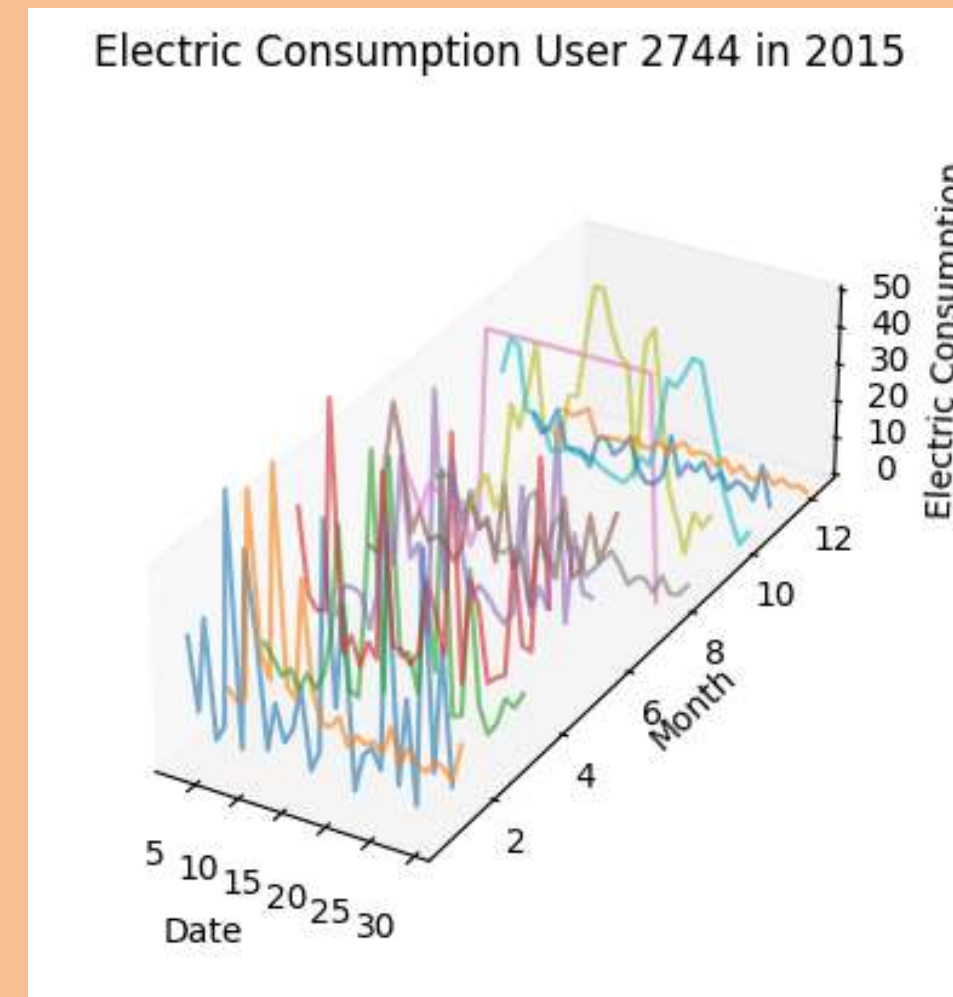


# 資料觀察

按月份比較一年中的每日用電量  
(2015/01/01 ~ 2015/12/31)



FLAG=0 沒偷電



FLAG=1 有偷電



# 資料前處理方法



# 缺失值處理

**寬深 CNN**

$$f(x_i) = \begin{cases} \frac{x_{i-1} + x_{i+1}}{2} & \text{if } x_i \in \text{NaN}, x_{i-1}, x_{i+1} \notin \text{NaN} \\ 0 & \text{if } x_i \in \text{NaN}, x_{i-1} \text{ or } x_{i+1} \in \text{NaN} \\ x_i & \text{if } x_i \notin \text{NaN}, \end{cases}$$

**CLOF**

$$G(x_i) = \begin{cases} \text{mean}(x) , & x_i \in \text{NaN} \\ x_i , & \text{otherwise} \end{cases}$$

# 異常值處理

寬深 CNN

$$f(x_i) = \begin{cases} \text{avg}(x) + 2 \cdot \text{std}(x) & \text{if } x_i > \text{avg}(x) + 2 \cdot \text{std}(x) \\ x_i & \text{otherwise,} \end{cases}$$

CLOF

$$G(x_i) = \begin{cases} \frac{x_{i-1} + x_{i+1}}{2}, & \text{if } x_i > \text{avg}(x) + 3 \cdot \text{std}(x), \\ & x_{i+1}, x_{i-1} \neq \text{NaN} \\ x_i, & \text{otherwise} \end{cases}$$



# 正規化處理

**寬深 CNN**

$$f(x_i) = \frac{x_i - \min(x)}{\max(x) - \min(x)}$$

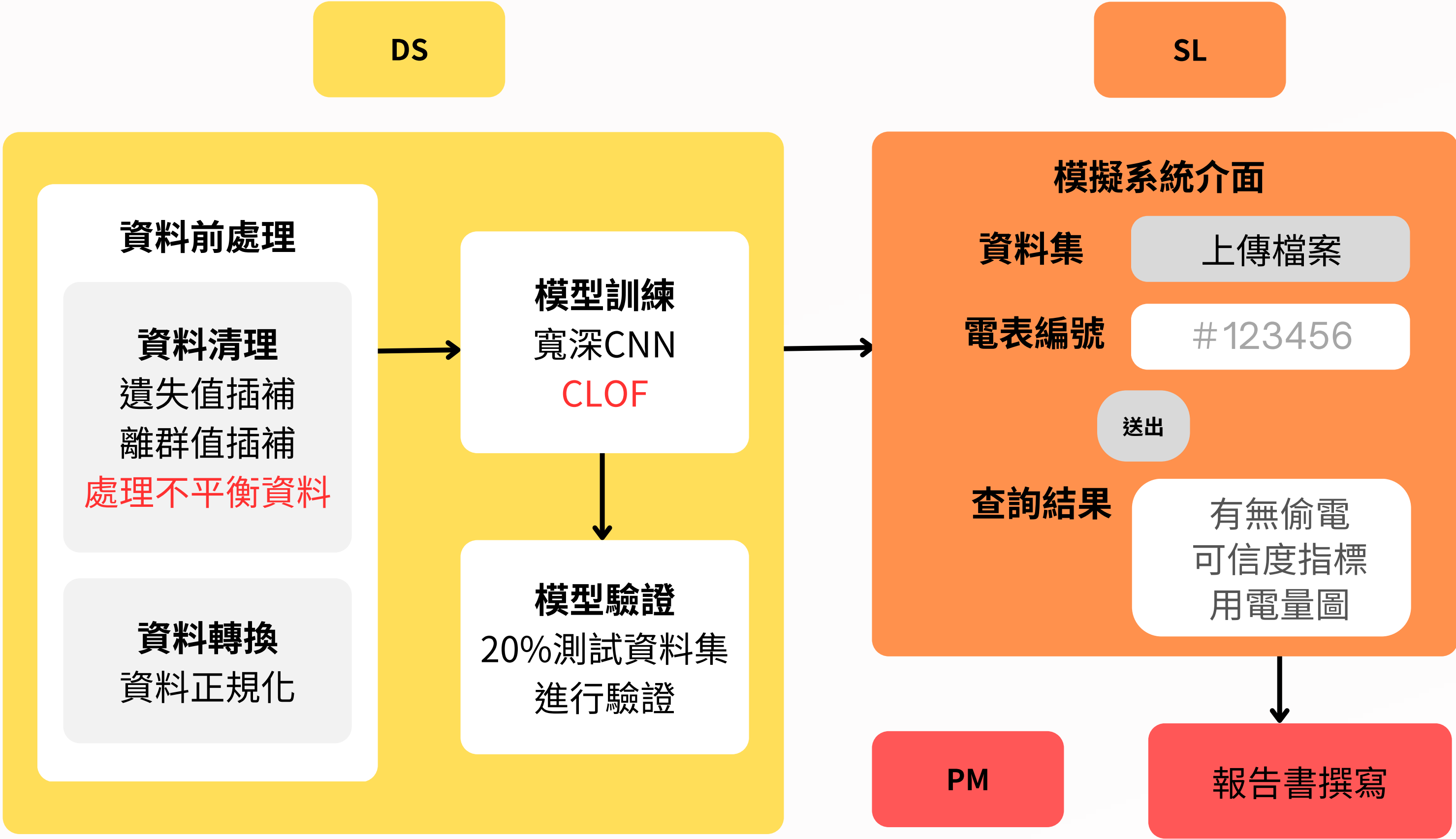
**CLOF**

$$G(x_i) = \frac{x_i}{\max(x)}$$



# 流程介紹

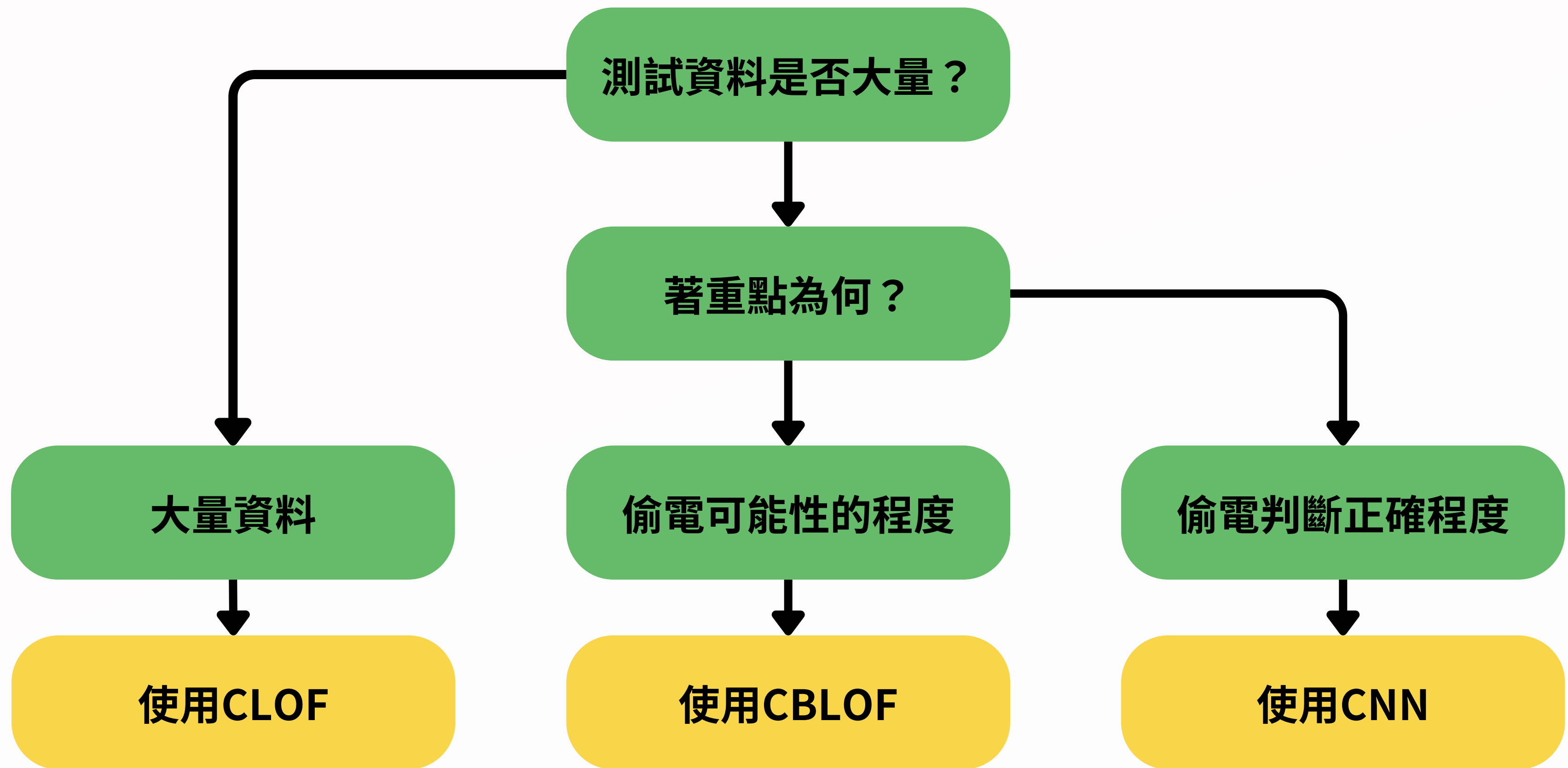




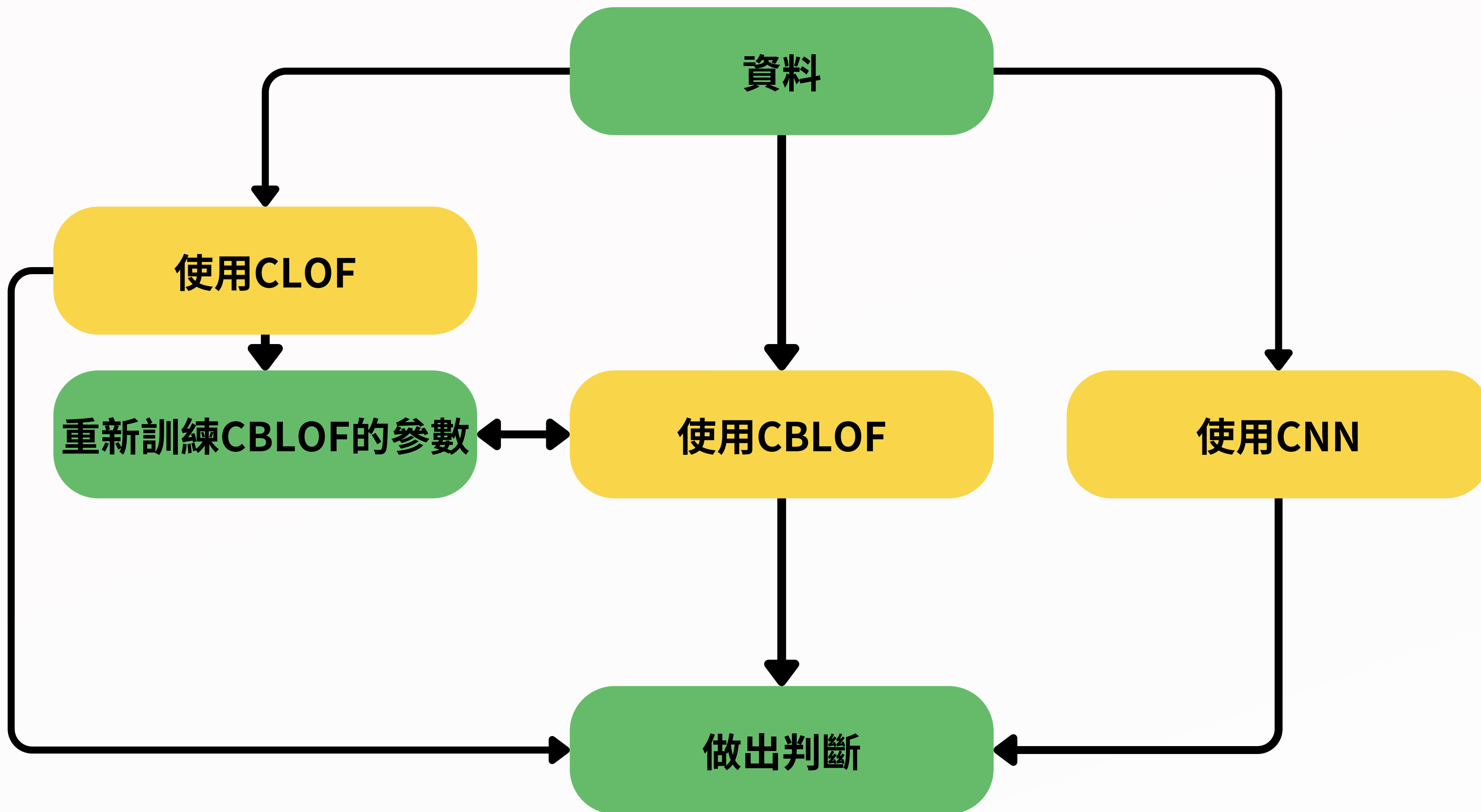


# 建議之決策方式











# 驗證結果



# 模型比較與說明

模型	評斷模型指標	百分比	異常臨界值	特點
Clof	precision,F1 score, recall,accuracy	11.1%,13.4%, 17.1%,81.3%	$\geq 1.27$ (lof 值)	運算速度快、利用密度 計算出異常得分，透過 分數去預測竊電用戶
寬深 CNN	AUC,F1 score, MAP@200,MAP@100	81.2%, 39.8%, 89.5%,90.9%	$\geq 0.3$ (可能偷電的機率)	模型效果最佳，較準確 找出潛在竊電用戶
Cblof	recall,F1 score, AUC,precision, MAP@50	16.6%,27.66%, 66.87%,82.98% 98.0%	$\geq 0.25$ (異常分數)	對數據進行排序，識別 出異常數據，能優先 鎖定潛在竊電用戶

Clof : Clustering and Local Outlier Factor  
寬深CNN : Wide and Deep Convolutional Neural Networks  
Cblof : Cluster-Based Local Outlier Factor  
AUC : Area Under Curve  
MAP : mean average precision



A modern office interior with a warm, natural aesthetic. The space features light-colored wooden beams on the ceiling, with various green plants hanging from them. In the foreground, there's a long wooden table with several black office chairs. A large planter box with various green plants sits on the table. In the background, there are more wooden tables, chairs, and a large sofa. The floor is made of light-colored wood. The overall atmosphere is bright and airy, with a focus on natural elements.

# 輔助決策工具 DEMO



# 模擬系統介面

模擬系統介面

資料集

上傳檔案

電表編號

# 123456

送出

查詢結果

資料區間、缺失值  
平均用電量、最大值  
有無偷電、用電量圖  
偷電時間、偷電行為  
偷電地點、偷電度數  
偷電金額

# 模擬系統介面

使用說明

結果展示(潛在偷電名單)

使用者查詢

偷電人人有責!

Upload CSV file

↑

拖放檔案至此處

- 或 -

點擊上傳

輸入ID

選擇使用方法

Wide\_CNN

Clof

Cblof

選擇起始日期

2024/06/17

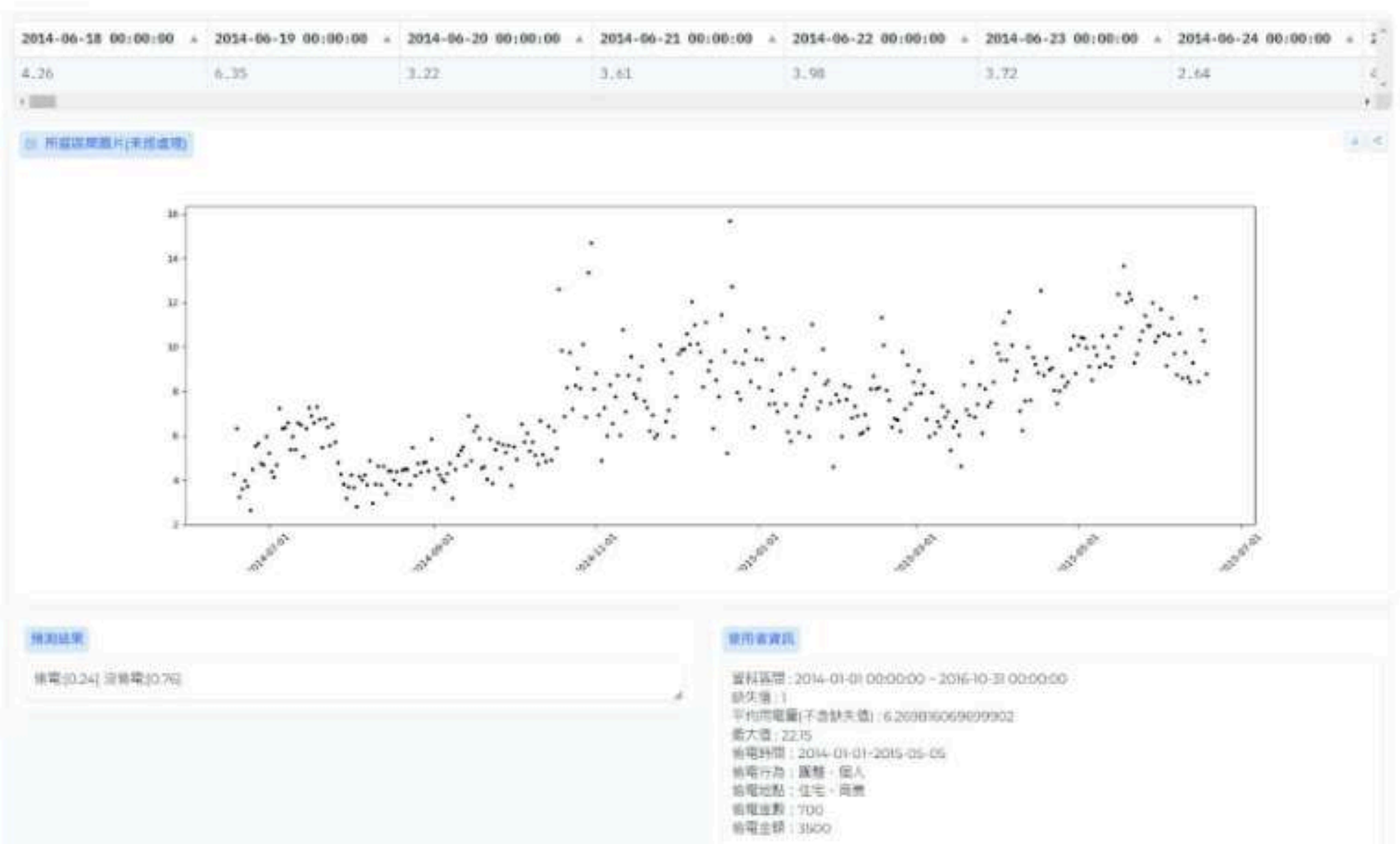
選擇結束日期

2024/06/17

Run



# 模擬系統介面





**THANK  
YOU**