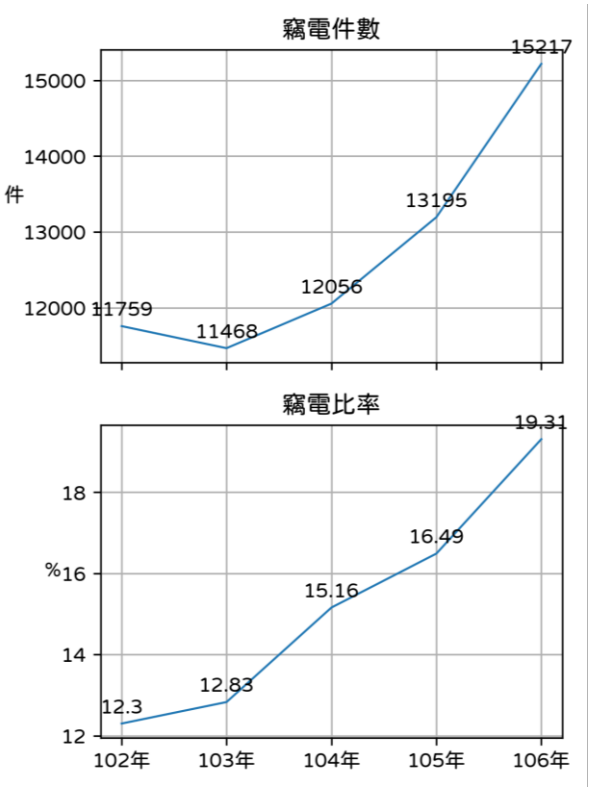
摘要

本次研究旨在找尋有效偵測竊電的方法，並且以這些方法製作應用程式以期達到協助判斷竊電的可能性，我們系統地評估了現有的檢測方法，最終選擇了三種模型，寬深 CNN、CLOF、CBLOF 做為判斷依據，他們各自適合不同的分析情境，結果顯示，新的檢測方法在防止電力盜竊方面具有潛力。

前言

全球每年因電力盜竊造成的經濟損失高達近千億美元。電力盜竊問題不僅對經濟造成巨大損失，還可能影響電力系統的穩定性和安全性。在台灣，電力盜竊問題同樣嚴重且逐年加劇。台電公司 106 年度全系統發購電電量達 2,310.80 億度，扣除抽蓄用電、公司自用電及售電量後，全年線路損失量為 88.27 億度，線路損失率為 3.82%。102 年度至 106 年度查獲竊電情況如圖(1)，反映電力被竊用的情形有日趨嚴重之勢。



現行偵測竊電的研究大致分為四個方向，賽局理論、電網分析、硬體分析、機器學習，由於機器學習以外都有相當程度的限制，我們會主要探討機器學習相關的方式。

專案背景與目的

本研究的背景在於應對現代化電力系統中日益嚴重的電力盜竊問題。隨著篡改電表讀數的方法變得更加多樣和隱秘，現有的檢測方法面臨新的挑戰。本研究的主要目的是基於 SGCC(STATE GRID Corporation of China)資料集，系統地調查和評估各種電力盜竊檢測方法。包括基於機器學習和測量不匹配的方法。通過這些分析，我們期望能夠提出檢測方法，以方便我們利用應用程式檢測竊電者，以達到追回欠繳電費的目的。

資料觀察

資料集描述

本研究所使用的資料集涵蓋了 42372 個電力用戶在 1035 天內的用電數據，資料記錄期間從 2014 年 1 月 1 日至 2016 年 10 月 31 日。以下是資料集的主要特徵：

- 日期格式（MM/DD/YYYY）：每筆記錄表示該日的電力消耗量。
- 電網編號（CONS_NO）：電網編號為字符串類型，代表每個電網的唯一身份。
- 標誌欄位（FLAG）：此欄位用於標示是否存在竊電行為，其中 0 表示無竊電，1 表示有竊電。在此資料集中，標示為 0 的記錄有 38637 筆，而標示為 1 的記錄則有 3585 筆。

（2016 年 9 月 18 日的數據缺失，因此後續我們當作缺失值處理。）

資料觀察後發現的問題

1. 資料缺失值與離群值多，若不處理將會影響後續的分析及預測

	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

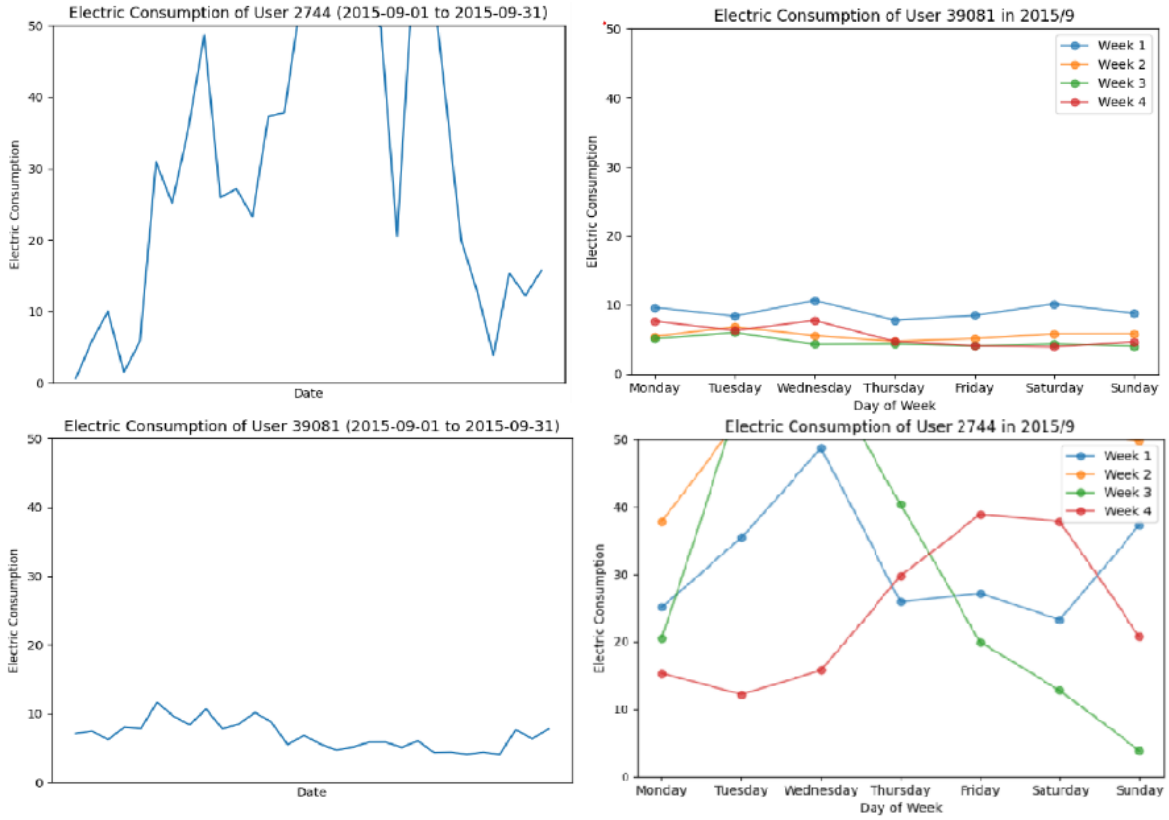
資料集	**遺失率**

原始資料集	25.60%
訓練資料集	25.29%
測試資料集	27.64%

2. 資料極度不平衡資料，若不處理，可能會導致準確度高，但精確低十分低的問題
3. 不同用戶的用電量差異很大，需要進行資料標準化或正規化處理，以提高模型的效果
4. 除了缺失值，即便是正常用戶，可能因為出國或其他原因，導致用電量在幾個月內為零，這也會對預測結果造成影響。
5. 資料集的日期並未照順序排序，在視覺化或是分析可能會遇到問題，因此將資料按時間排序也是前處理的重點

資料視覺化的發現

我們分別畫出了單月每週用電量、單月每日用電量以及按月份比較一年中的每日用電量。觀察結果顯示，涉嫌竊電者的用電起伏較大，而正常用電者的用電則較為平穩或具有週期性。基於這一發現，在處理缺失值時，我們除了使用常見的前後平均值法外，還採用了每個月該星期的平均值來補全缺失數據。在報告中，我們找出了能夠明顯以視覺方式呈現正常用電與竊電用戶差異的圖表。然而，仍有許多差異是肉眼難以辨識的，因此需要進一步透過模型訓練、預測和排序來精確識別。下圖中左測為其中一筆偷電者的用電量視覺畫圖、右測為其中一筆正常用戶的用電量圖。



(圖片使用簡報中的即可：

https://www.canva.com/design/DAGBpzYSkuI/fQdINNTQhVU8pxJfJ_N_9w/edit）

(程式碼：

https://colab.research.google.com/drive/1ol2G0Sn0bN7BcygiR_8jV17heJq3mj1g?usp=drive_link)

資料前處理與分析方法原理

寬深 CNN(Wind and Deep CNN)

資料前處理

（此模型的架構與前處理方式均參考《Wide and Deep Convolutional Neural Networks for Electricity-Theft Detectionto Secure Smart Grids》這篇論文）

1.修正資料排序

因為此為時間序列資料，若時間沒有按照正常順序排序的話會很大程度的影響模型判斷。因此先將原先雜亂的日期整理成按照時間先後排序的資料。

2.修正離群值

如果某個資料點超過均值的三個標準差，則認為該點是錯誤的，使用"三西格瑪準則"來修正。公式如下：

$$G(\tilde{u}_{i,t}) = \begin{cases} \frac{\tilde{u}_{i,t-1}+\tilde{u}_{i,t+1}}{2} & \text{if } \tilde{u}_{i,t} > 3\sigma(\tilde{u}_i) \text{ and } \tilde{u}_{i,t-1}, \tilde{u}_{i,t+1} \neq \text{NaN} \\ \tilde{u}_{i,t} & \text{otherwise} \end{cases}$$

其中 x 是由每天的 xi 組成的向量，avg(x) 是 x 的平均值，std(x) 是 x 的標準差。

3.處理缺失值

當該筆中有資料缺失時，使用該筆資料中的其他可用資料均值來替換缺失值。公式如下：

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

其中 xi 代表一段時間（例如一天）內電力消費數據中的值。如果 xi 是空值或非數字字符，

我們將其表示為 NaN（NaN 是一個集合）。

4.**正規化**

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

其中 min(x) 是 x 中的最小值，max(x) 是 x 中的最大值。

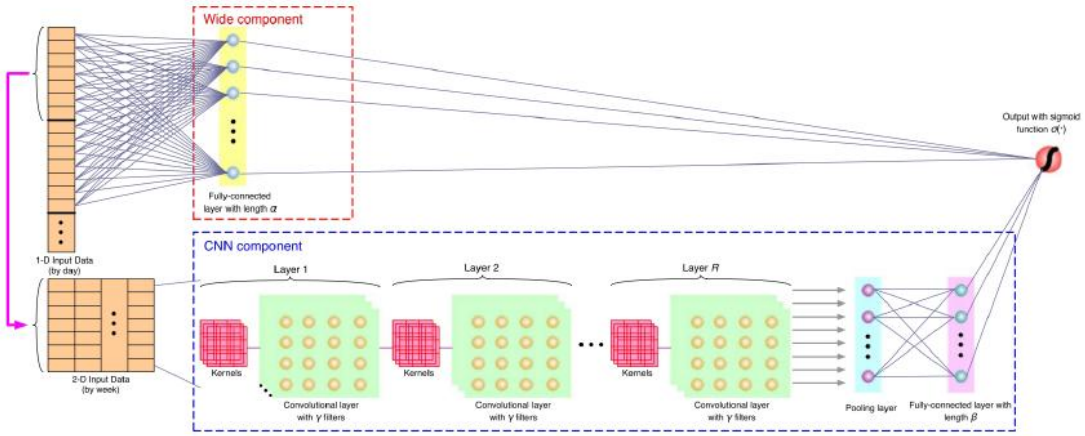
分析方法原理

方法介紹

根據「資料視覺化的發現」，我們得知了能從是否具有週期性來去判斷是否有偷電。為了能有效的應用這個發現，我們使用了寬深 CNN 來去分析電力消費數據

寬深 CNN 的框架主要由寬組件(Wind Component)跟深度卷積神經網絡組件(CNN Component)兩個部分組成， 寬組件可以學習全局知識，而深度卷積神經網絡組可以捕捉電力消費數據的週期性。這個模型整合了寬組件

和深 CNN 組件的優點，因此在電力盜竊檢測中能有優秀的表現。其模型的運作方式如下圖所示。



1. **寬組件(Wind Component)**

寬組件是一個全連接層，主要用於從一維電力消費數據中學習全局知識。電力消費數據通常是隨時間變化的時間序列數據，而這些數據可能顯示出周期性模式或非周期性模式。寬組件的設計目的是捕捉這些全局模式。全連接層中的每個神經元根據以下方程使用一維電力消費數據計算其自身的得分：

$$y_j := \sum_{i=1}^{n} w_{i,j} x_i + b_1$$

其中 yj 是第 j 個神經元在全連接層的輸出，n 是一維輸入數據 x 的長度，w{i,j} 表示第 i 個輸入值和第 j 個神經元之間的權重，b1 是偏差。計算後，它將通過激活函數將此值發送到更高層的連接單元，以確定它對下一步預測的貢獻程度。激活函數如下所示：

$$u_j := f(y_j) = \max(0, y_j)$$

其中 uj 是激活計算後的輸出，f(·)表示激活函數。在本文中，我們使用修正線性單元（ReLU）作為激活函數，這將僅激活正值。這個函數可以有效地防止過擬合。

2. 深度卷積神經網絡組件(CNN Component)

深度卷積神經網絡組件由多個卷積層組成，這些卷積層用於提取輸入數據的局部特徵。這部分的設計目的是從電力消費數據中自動學習更複雜的特徵表示。

- **卷積層**：
    - 每個卷積層包含多個卷積過濾器，這些過濾器專門設計來捕捉電力消費數據中的特定模式，如周期性或非周期性模式。
    - 卷積過程將輸入數據轉換為特徵圖，通過多個卷積層的處理，逐漸提取更高層次的特徵。
- **技術細節**：
    - 深度 CNN 組件的訓練過程中，輸入的二維電力消費數據將通過多個卷積層，這些層能夠有效地捕捉二維數據的特徵。
    - 在實驗中選擇了不同數量的卷積層（例如，γ 個過濾器），以調整模型性能。

寬深卷積神經網絡的整合

寬組件和深度 CNN 組件分別捕捉全局特徵和局部特徵，通過整合這兩部分的輸出，最終形成一個具有強大特徵學習能力的模型。

分析與驗證結果

驗證指標

- 不錯的區分能力（AUC: 0.81）：模型的 AUC 值為 0.81，顯示出它有較好的整體分類性能，能夠有效區分大部分的盜竊行為和正常行為。

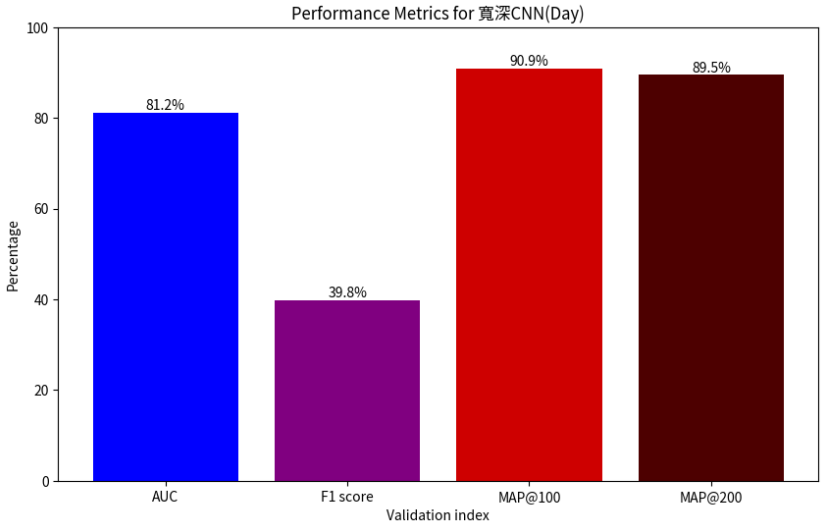
- 中等的 F1 分數（F1 Score: 0.39）：儘管模型在高優先級位置上的精確度很高，但整體的精確度和召回率之間的權衡不夠理想，這表明模型在全面檢測所有盜竊行為時可能有一定的不足。

- 不錯的前 100 名和前 200 名預測精確率（ MAP@100, MAP@200: 0.91,0.9）：這兩個指標表明，模型在前 100 和 200 個排序位置上的平均精確度非常高，能夠有效地將最可能的盜竊行為排在前面，這對於實際應用中需要優先處理的場景非常有用。

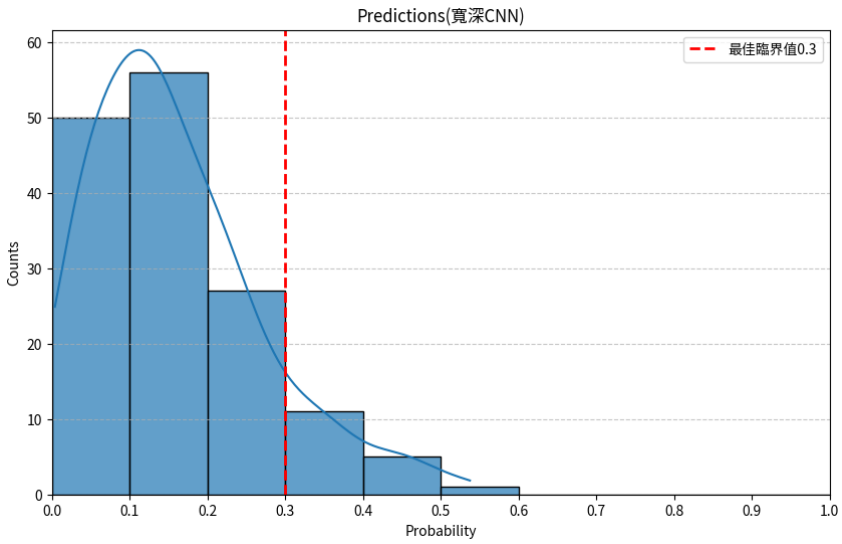
改進方向

- **提高召回率:** 通過調整模型或引入更多特徵來提高召回率，從而提升 F1 Score。
- **平衡精確率和召回率**：考慮使用其他方法，如調整閾值或引入代價敏感學習，以平衡精確率和召回率。

- **進一步分析特徵**：對於模型的特徵進行進一步分析，找出哪些特徵對於檢測盜竊行為更為關鍵，並對模型進行優化。



（圖：寬深 CNN 的指標）



（圖：在測試資料集中，定義偷電用戶的閥值）

補充:平均精確率均值 (MAP) 指標

MAP（Mean Average Precision）是常用於信息檢索和排名模型的一個評估指標。它主要衡量模型在不同位置的預測準確性，特別適用於排序任務。

Cluster-based Local Outlier Factor(CLOF)

資料前處理

1.處理缺失資料

當負載向量中的資料缺失時，使用該向量中其他可用資料的均值來替換缺失值。公式如下：

$$G(\tilde{u}_{i,t}) = \begin{cases} \text{mean}(\tilde{u}_i) & \text{if } \tilde{u}_{i,t} \in \text{NaN} \\ \tilde{u}_{i,t} & \text{otherwise} \end{cases}$$

2.修正錯誤資料

如果某個資料點超過均值的三個標準差，則認為該點是錯誤的，使用"三西格瑪準則"來修正。公式如下：

$$G(\tilde{u}_{i,t}) = \begin{cases} \frac{\tilde{u}_{i,t-1}+\tilde{u}_{i,t+1}}{2} & \text{if } \tilde{u}_{i,t} > 3\sigma(\tilde{u}_i) \text{ and } \tilde{u}_{i,t-1}, \tilde{u}_{i,t+1} \neq \text{NaN} \\ \tilde{u}_{i,t} & \text{otherwise} \end{cases}$$

3.正規化

每個負載向量的每個元素會被該向量的最大值進行歸一化處理。

分析方法原理

CLOF（Cluster-based Local Outlier Factor）方法將聚類技術與 LOF 方法結合起來，用於識別電力消費數據中的異常行為，特別是電力偷竊。這種方法的主要步驟包括：

1. 聚類分析（Clustering）：
    - 使用 k-means 聚類算法對用戶的電力消費模式進行分組。這一步的目的是根據消費特徵將用戶劃分為若干類別，使得每個類別中的用戶具有相似的消費行為。
    - 對於每個聚類結果，計算聚類中心並衡量每個用戶與其聚類中心的距離。距離較遠的用戶可能是潛在的異常值。
2. 局部離群因子（Local Outlier Factor, LOF）：
    - 在聚類分析的基礎上，對每個聚類內的用戶應用 LOF 算法。LOF 算法通過評估每個數據點的局部密度，來衡量該數據點相對於其鄰域內其他數據點的異常程度。
    - 具體來說，LOF 計算每個數據點的局部可達密度（Local Reachability Density, LRD），並與其鄰域內其他數據點的 LRD 進行比較。如果一個數據點的 LRD 顯著低於其鄰域內的其他點，則該點被認為是潛在的異常值。
3. 異常檢測框架（Detection Framework）：
    - 最終，CLOF 方法結合聚類分析和 LOF 的結果，識別並標記那些可能存在異常行為的用戶。
    - 該方法能夠有效地處理數據集中的不平衡問題，並能檢測到各種類型的偷電攻擊，從而提高電力偷竊行為的檢測準確性。

Local Outlier Factor(LOF) 方法介紹

LOF（Local Outlier Factor）是一種基於密度的異常檢測方法，它通過比較每個數據點與其鄰域內其他數據點的密度來評估該數據點的異常程度。LOF 方法的主要步驟包括：

1. 鄰域定義：
   - 對於每個數據點 p，找到其 k 個最近鄰居（k-nearest neighbors）。這些鄰居構成該數據點的鄰域。
2. 可達距離（Reachability Distance）：
   - 對於數據點 p 和其鄰居 o，計算可達距離 reach-dist(k,p,o)。可達距離是 o 和 p 之間的實際距離以及 o 的 k 距離中較大的一個，即：

$$\text{reach-dist}(k, p, o) = \max(\text{k-distance}(o), \text{dist}(p, o))$$

其中，k-distance(o) 是 o 的第 k 個最近鄰的距離。

3.局部可達密度（Local Reachability Density, LRD）：

- 計算數據點 p 的局部可達密度，即 p 的鄰居的平均可達距離的倒數：

$$\text{LRD}(p) = \left(\frac{\sum_{o \in N_k(p)} \text{reach-dist}(k, p, o)}{|N_k(p)|} \right)^{-1}$$

其中， N_k(p)是 p 的 k 個最近鄰居的集合。

3. 局部離群因子（Local Outlier Factor, LOF）：

- 計算數據點 p 的局部離群因子：

$$\mathrm{LOF}(p) = \frac{\sum_{o \in N_k(p)} \frac{\mathrm{LRD}(o)}{\mathrm{LRD}(p)}}{|N_k(p)|}$$

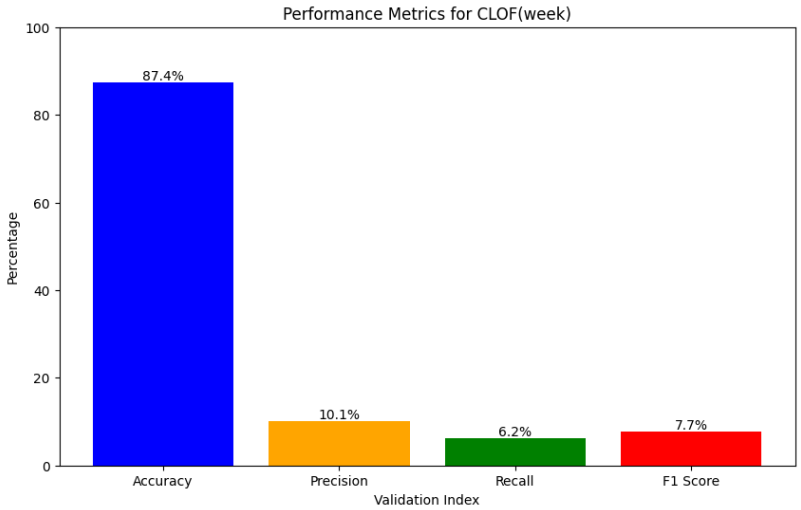
LOF 值衡量了 p 的局部密度與其鄰居的局部密度的比率。如果 p 的 LOF 值顯著大於 1，則表示 p 是異常點。

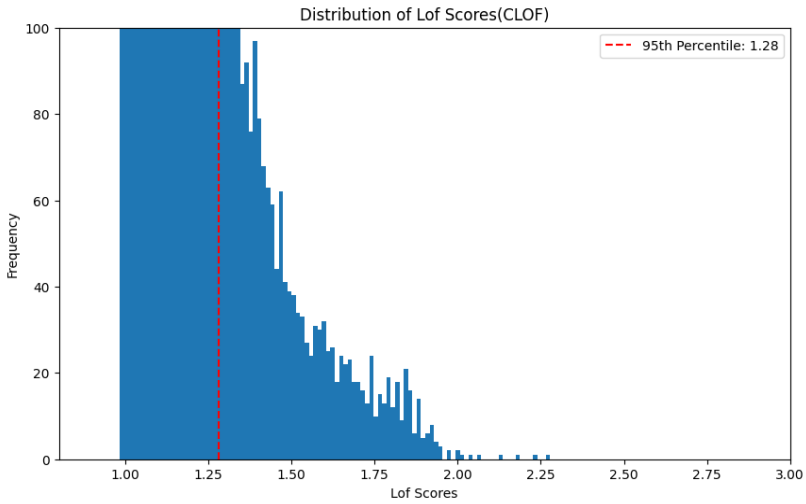
分析與驗證結果

驗證指標

1. 一定的區分能力（Accuracy: 0.87）：模型在整體上能夠比較準確地區分出竊電用戶和非竊電用戶。
2. 低精確率（Precision: 0.10）：模型預測的竊電用戶中，只有極少數真正是竊電者，這表示模型有很多錯誤的警報（誤報）。
3. 低召回率（Recall: 0.06）：模型僅檢測到極少部分的竊電用戶，大部分的竊電行為沒有被模型捕捉到。這可能意味著模型過於保守，只在非常明顯的情況下才會做出預測。
4. 低的 F1 分數（F1 Score: 0.08）：表示模型在精確率和召回率之間的平衡較差，需要大幅改進以提升整體性能。

這些指標表明，儘管模型的準確率較高，但在檢測竊電用戶時存在很大的不足。特別是精確率和召回率都非常低，導致 F1 分數也非常低，說明模型在竊電行為的檢測上並不理想，需要進一步優化。





Cluster-Based Local Outlier Factor(CBLOF)

資料前處理

（這裡的前處理均參考《An Ensemble Deep Convolutional Neural Network Model for Electricity Theft Detection in Smart Grids》此篇論文）

1. **處理缺失值**

   我們參考了論文提出的填補缺失值方法，考慮了不同季節、月份、工作日和週末的用電模式差異。具體方法如下：

對於特定天（如星期一）的用電量，使用下列公式來填補缺失值：

$$f(CM_i) = \begin{cases} \frac{\sum_{n=1}^{x} CM_n}{n} & \text{若 } 1 \leq n \leq 4, n \in \mathbb{N} \\ CM_i & \text{若 } M_i \notin NaN \\ -1 & \text{若 } n = 0 \end{cases}$$

其中，CMi 是特定月份中的第 □ 個星期一的用電量， CMn 表示該月份中其他星期一的用電量。

2. 處理離群值

我們使用了論文中的 Winsorization 方法來替換離群值。具體操作是使用 Least-Winsorized-Square 方法來將離群值替換為最接近的可接受值。

3. 資料正規化

$$f(c_i) = \frac{c_i - \mathrm{Min}(C)}{\mathrm{Max}(C) - \mathrm{Min}(C)}$$

其中，ci 表是該用電量，Min(C) 和 Max(C) 分別表示用電量的最小值和最大值。

4. 處理資料不平衡問題

我們採用了 Random Under Bagging 的方法

- 建立平衡的子集：

  首先，先把所有竊電的資料（少數類別）都放進每個子集裡。

  接著，從正常用戶（多數類別）的資料裡，隨機選取和竊電資

料數量相同的樣本加入子集中。這樣每個子集裡，竊電和正常用戶的數量都是一樣的。

- 重複這個過程：重複上述步驟，將原始的不平衡資料集轉換成多個平衡的子集。(在實際操作中，我們總共建立了 15 個子集，並取其平均來判斷其指標）

分析方法原理

1. **CBLOF 簡介**

   CBLOF（Cluster-Based Local Outlier Factor）是一種用於異常檢測的無監督學習方法，特別適合處理不平衡資料和異常值的問題。該方法通過對資料進行聚類分析，識別出離群點，進而提升異常檢測的準確性。

2. **步驟與公式**

   他的核心思想是基於資料的聚類結構來計算離群因子。它將資料點劃分到不同的聚類中，並根據每個資料點與其所屬聚類的距離來評估其異常程度。

   主要包括以下步驟：

- 聚類分析：使用 K-means 或其他聚類算法，將資料集劃分成多個聚類。(在實際操作中，我們選用 K-means ）
- 離群因子計算：對於每個資料點，計算其與所屬聚類中心的距離，並結合聚類的大小，評估該資料點的離群因子。

具體的計算公式如下：

$$CBLOF(p) = \begin{cases} |C_i| \cdot \min(distance(p, C_j)) & \text{where } p \in C_i, C_i \in SC \text{ and } \\ & C_j \in LC \text{ for } j = 1 \text{ to } b \\ |C_i| \cdot distance(p, C_i) & \text{where } p \in C_i \text{ and } C_i \in LC \end{cases}$$

- 對於屬於大聚類的資料點，其 CBLOF 值為該點到聚類中心的距離乘以聚類的大小。
- 對於屬於小聚類的資料點，其 CBLOF 值為該點到最近的大聚類中心的距離乘以小聚類的大小。

3. 選擇此模型原因：

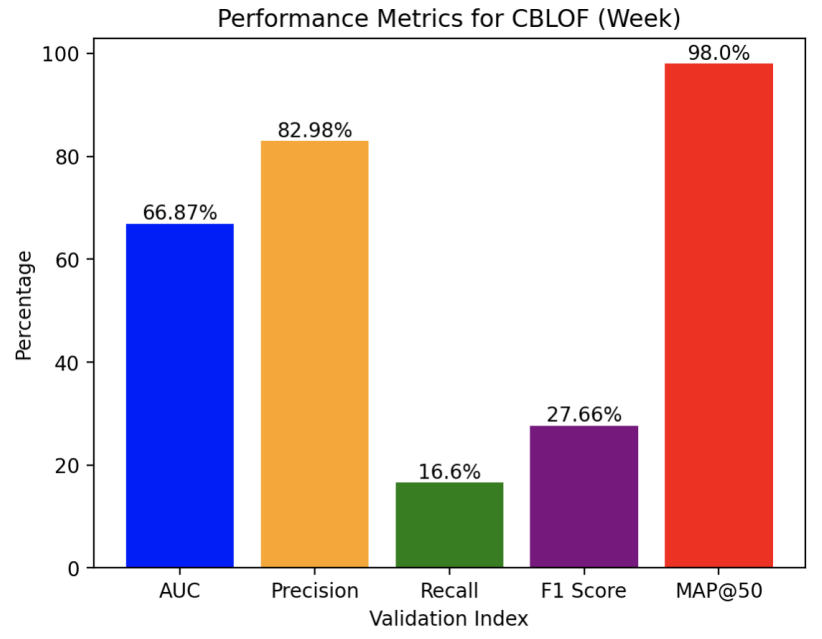
   CBLOF 值越高，表示該資料點越可能是異常值。我們可以根據這個值進行排序，並設定要抓出竊電者最異常的前 5 %（使用者可自訂）。由於電力公司的資源有限，無法逐一檢查所有用戶，因此這個方法能夠生成異常排名，讓電力公司的人員可以優先查詢最異常的用戶。

分析與驗證結果

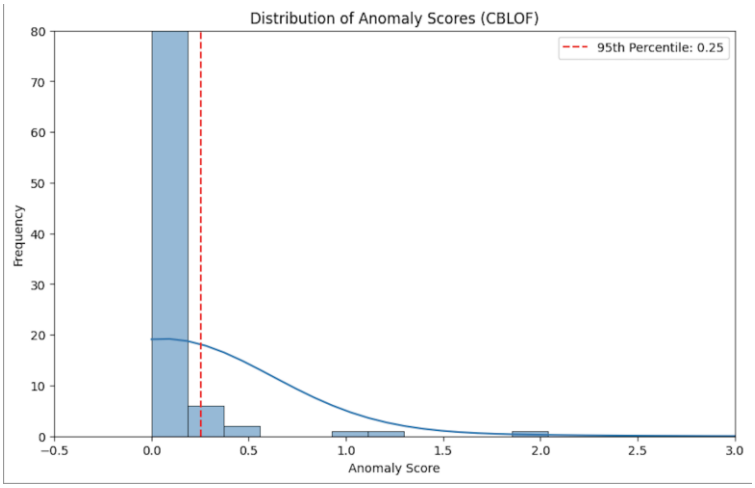
驗證指標

- 一定的區分能力（Average AUC: 0.67）：模型能夠比隨便猜測更好地區分竊電用戶和非竊電用戶。
- 高精確率（Average Precision: 0.83）：模型預測的竊電用戶中，大多數確實在竊電。

- 低召回率（Average Recall: 0.17）：模型僅檢測到部分竊電用戶，說明還有改進空間。這可能是因為在使用 CBLOF 時將 beta 設定得較大，讓模型可以更專注於極端異常值。此外，儘管有做不平衡處理，但資料不平衡的問題仍然可能影響召回率。
- 中等的 F1 分數（Average F1 Score: 0.28）：表示模型在精確率和召回率之間的平衡還需要改進。
- 不錯的前 50 名預測精確率（Average MAP@50: 0.98）：模型在前 50 個預測中表現很不錯，適合用於抓取最異常的竊電行為。而因為我們主要是想要看異常偷電的排名，所以 MAP@50 是我們看的主要指標。



（圖：CBLOF 的指標）



（圖：在測試資料集中，找出前 5 % 最異常用戶的閥值）

（程式碼：https://colab.research.google.com/drive/14q_Yi6FcwT2aGlNRbvSl9VC-6HxZM0FT?usp=drive_link）

輔助決策工具設計稿

系統使用說明書

系統連結

https://huggingface.co/spaces/peter572210355/demo_steal_electricity_detect

使用說明介面

本系統利用不同的統計方法來檢測和預測可疑的偷電行為。提供了結果展示和使用者查詢介面介面，分別針對整體用電資料分析及特定用戶分析。

使用注意事項

資料格式：確保匯入的電力資料格式正確，請使用如同 electricitytheft_test 格式之資料，以避免影響分析結果的準確性。

資料形式參考：

電表編號/日期	2014/1/1	2014/1/10	2014/1/11	2014/1/12	2014/1/13	2014/1/14
F52CFF361D1F87ACF5E1A9BD5D255A3C						
6F4919E1A9FEEF57C26BFF9DCEF97B1E	0.0	0.0	0.0	0.0	0.0	0.0
1FA9C27D9BE77C43A919891A6DCB40D0	0.0	0.0	0.0	0.0	0.0	0.0
065573157D11407F877066A70A25CFA8	5.5	0.0	0.0	0.0	6.02	3.6
3A8E03953795361C00A49C0D93434424	0.0	9.0	0.0	6.07	4.38	2.68

方法選擇：不同的預測方法適用於不同的情境，建議根據實際需求選擇最合適的方法進行分析。

結果解讀：預測結果僅供參考，建議結合其他實際數據和情況進行綜合分析。

預測方法簡介

提供三種的預測方法，分別是 Wide_CNN、Clof 和 Cblof ，每種方法適用於不同的數據分析，如表

預測方法	
Wide_CNN	
Clof	
Cblof	

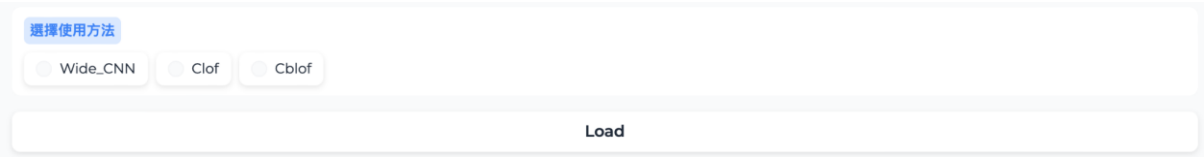
結果展示（潛在偷電名單）介面：整體用電資料分析

1.匯入資料，如圖。

2.點擊「匯入資料」按鈕，選擇您要上傳的電力資料文件，文件支援格式包括 CSV 和 Excel。

3.資料上傳成功後，系統會自動在界面上顯示數據樣本，讓您檢查是否匯入正確。

4.選擇預測方法:從三種預測方法（Wide_CNN、Clof 和 Cblof）中，選擇一種進行分析，如圖 ，選擇預測方法後，點擊「Load」按鈕，系統將運行選定的模型進行分析。

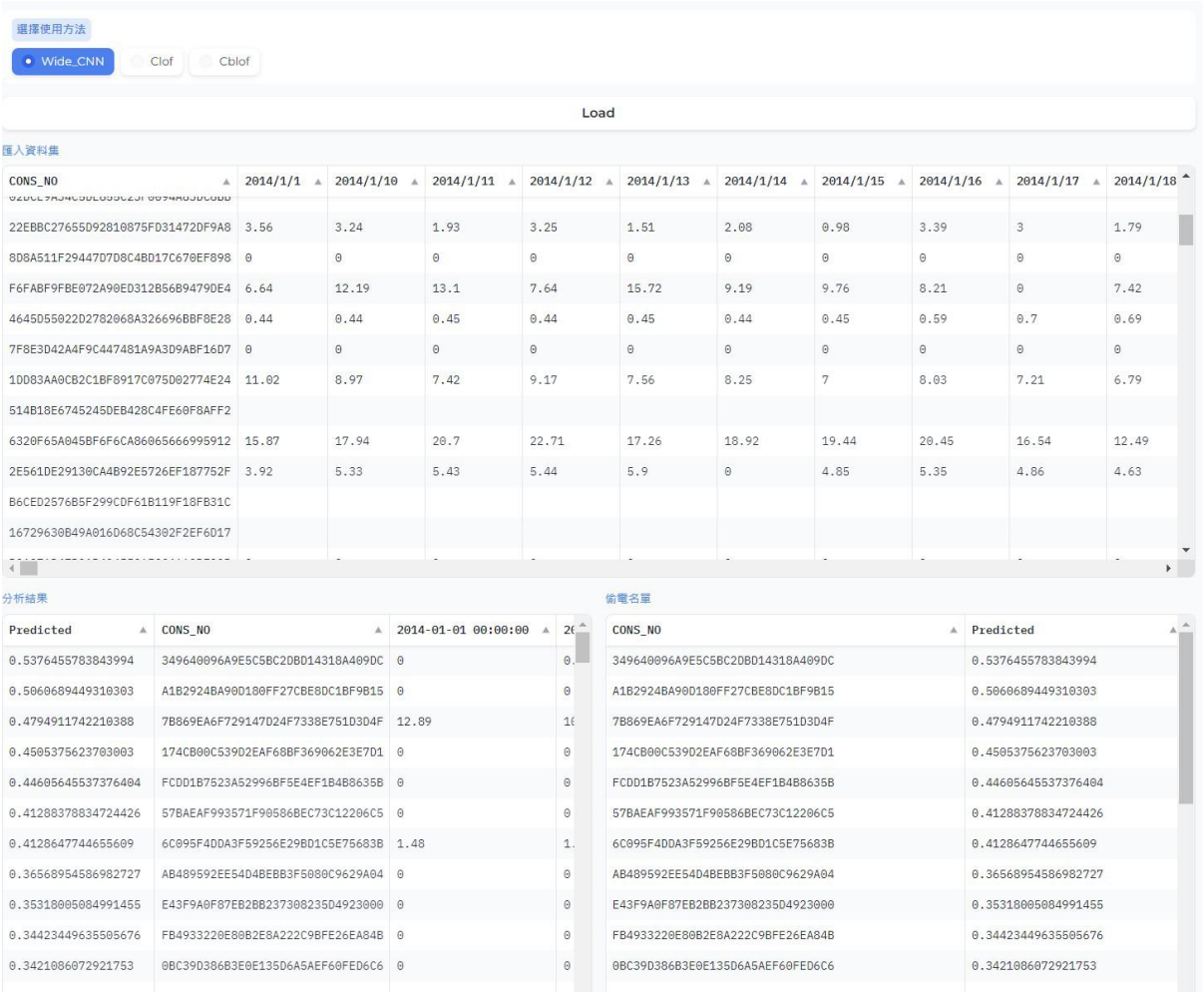


5.分析完成後，系統會生成相關的指標值，如圖 (cnn 取 >= 0.25, Clof 取>= 1.003, Cblof 取>= 0.4 就也是大概 5%-7%上下)和預測結果。

預測結果將依照指標值進行排序，生成潛在的偷電用戶名單。

點擊某一用戶可查看詳細的預測指標和用電行為分析報告。

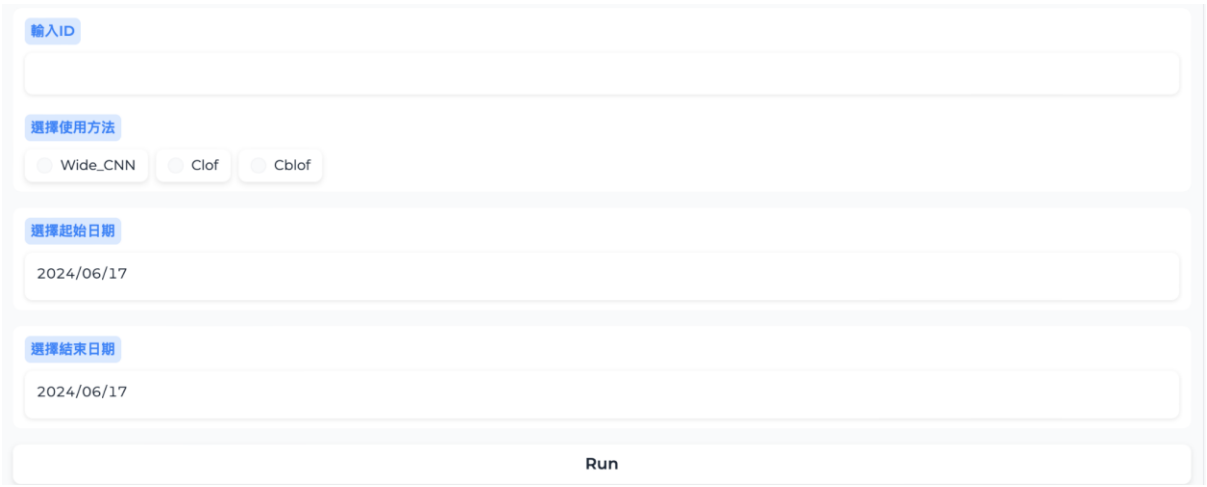
結果示



使用者查詢介面：特定用戶用電行為分析

1.匯入資料:點擊「匯入資料」按鈕，選擇並上傳電力資料文件（支援 CSV 和 Excel 格式），如圖

2.選擇用戶與時間段，如圖



從顯示的資料中選擇您要分析的特定用戶。

選擇需要分析的時間段，確保選擇的時間段內有足夠的數據供分析使用。

3.選擇預測方法:同樣提供三種預測方法（Wide_CNN、Clof 和 Cblof ）作為選擇

4.選擇欲查詢的起始和結束日期

5.點擊「Run」按鈕，系統將針對選定的用戶和時間段進行分析，系統會顯示該段時間的用電行為分析結果，預測結果會指示該用戶是否存在偷電行為，並提供詳細的指標及相關資訊，如圖。

結果展示

提供客服支援

如在使用過程中遇到任何問題或需技術支援，請聯絡我們客服團隊，將竭誠為您提供幫助和支持。

感謝您使用我們的電力偷竊預測應用程式。我們致力於為您提供高效且準確的預測工具，幫助您更好地管理和監控用電情況。希望您在使用過程中能夠獲得最佳的體驗。

結論

電力盜竊是一個全球性問題，每年造成的經濟損失達到數十億美元。這不僅對經濟造成巨大影響，還威脅到電力系統的穩定性和安全性。隨著竊電手段越來越多樣和隱秘，現有的偵測方法面臨巨大挑戰。

在本研究中，我們建立了三種不同的模型來應對竊電問題，這些模型各自有其優缺點，並適用於不同的分析情境。這些模型的建立，旨在透過更精確和有效的方法來偵測竊電行為，從而減少因竊電帶來的經濟損失。

除此，也開發了一個簡潔的系統，將這些模型應用於實際數據中，幫助電力公司更快速、更準確地識別潛在的竊電行為。此系統已經顯示出在實際應用中的巨大潛力，為電力公司提供了一個強有力的工具來打擊竊電。

為了能增加本研究的實用性和使用上的便利性，我們下一代的產品將往以下三點的方向做改進，請客戶敬請期待。

1. **提高模型的準確性和穩定性**：通過引入更多樣的數據和進一步優化模型參數，提升預測的準確性和穩定性。
2. **強化系統的實用性**：改進系統的使用流程和界面，使其更加友好和高效，方便電力公司在實際操作中使用。
3. **綜合多種偵測方法**：探索將更多不同的偵測方法綜合應用，以提高整體偵測效果，應對更加多樣化和隱秘的竊電手段。

總之，通過我們的研究和系統開發，為電力公司提供了一種有效的工具來應對日益嚴重的竊電問題，並為未來的進一步研究和改進打下了堅實的基礎。

附錄

組員分工名單

組別	姓名
Project Manager	劉家維、陳雅柔
Data Scientist	邱奕銓、楊家瑋、羅然莉
System Developer	張子恩、韓明澄