

製造數據科學期末報告

再生能源建置決策支援

— 透過太陽能日容量因子預測與建置區域辨識

r08725059 張煜柔

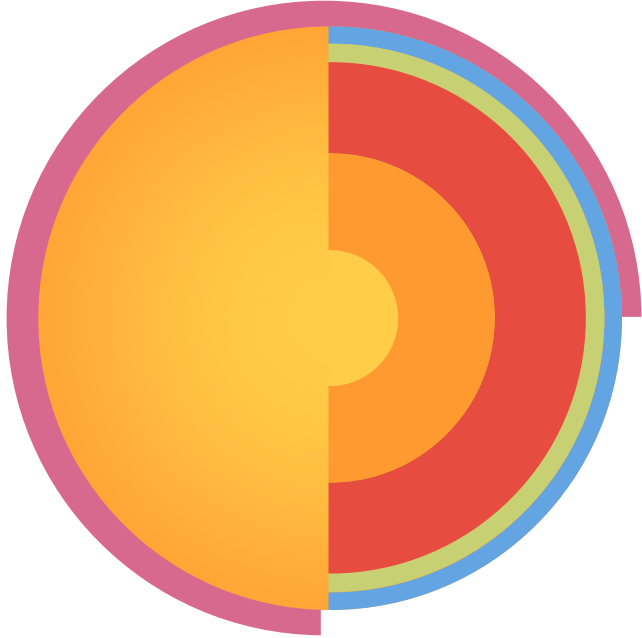
r09722016 劉庭安

r09725006 劉心鈺

r10725014 莊芯瑜



Agenda



**01 Background &
Motivation**

02 Research Process

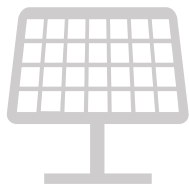
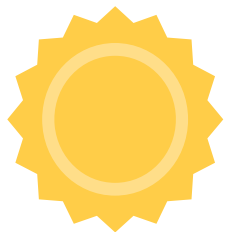
**03 Capacity Factor
Prediction**

**04 Roof Area
Recognition**

05 System

06 Conclusion

01 Background & Motivation



需求：
企業的綠電需求



供給：
家戶可以賣綠電



研究問題：

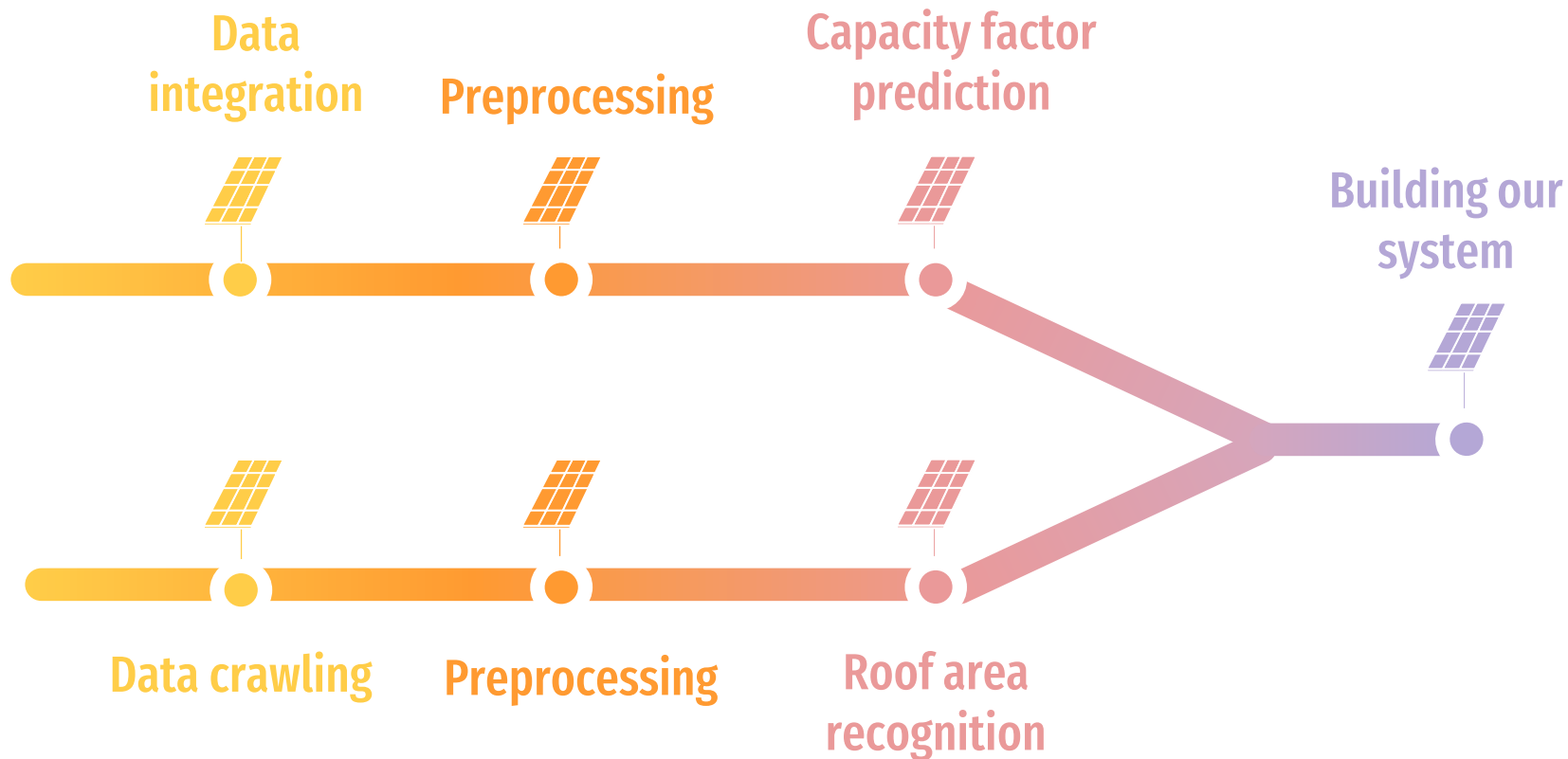
假設有一公司欲收購某位置附近的房子屋頂以建置太陽能板，使企業每年皆可達到使用 P 度太陽能源的永續目標。本研究參考該位置過去一年的天氣狀況，幫助公司了解需設多大的裝置容量，以及收購多少面積、多大範圍的屋頂。



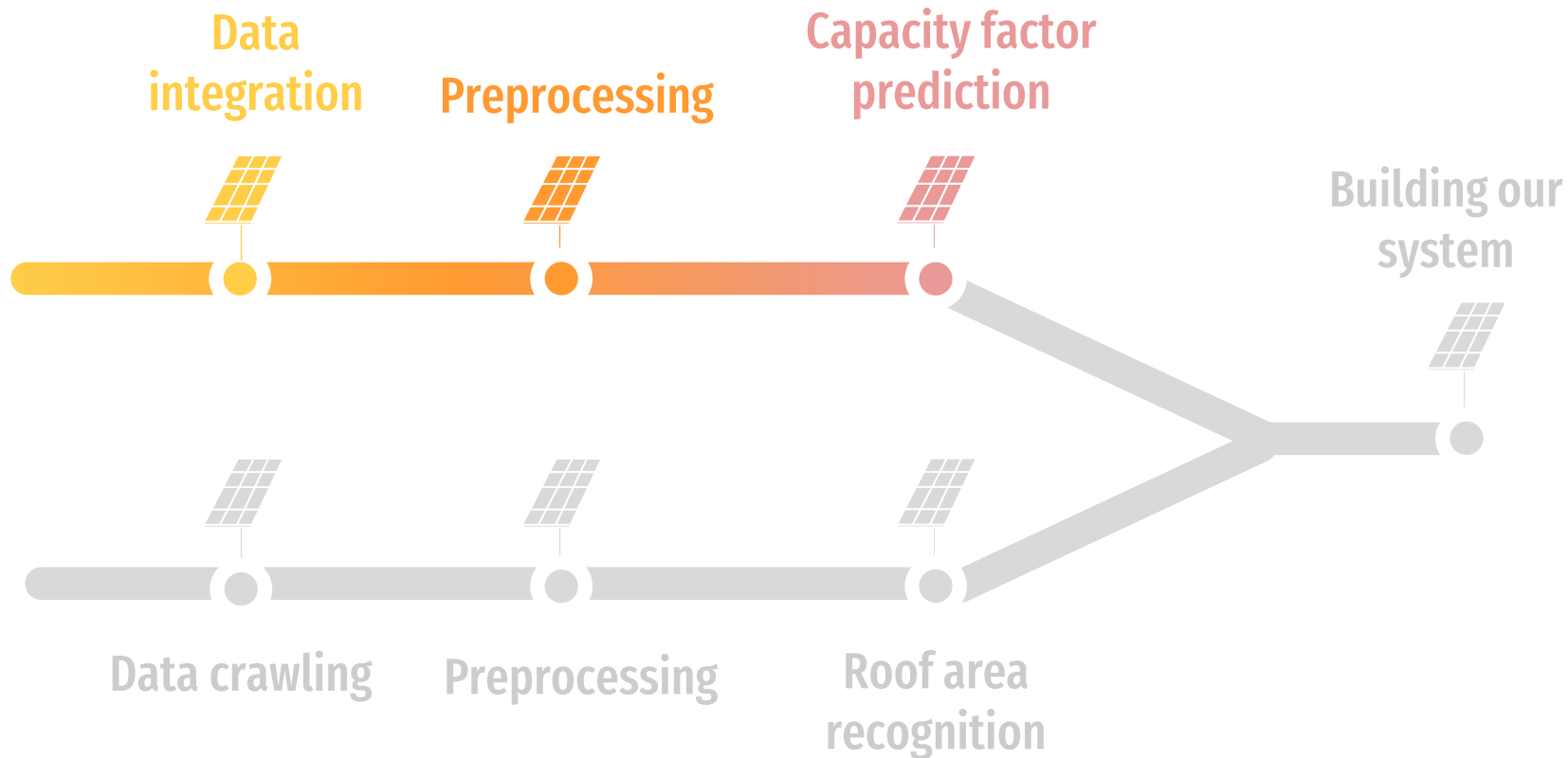
名詞解釋：

- **裝置容量**：設備出廠時，所設計滿載（百分之百全力發電）時的最大值，可視為發電廠的 Capacity
- 1 瓩裝置容量的太陽光電系統約需 10 平方公尺（約 3 坪）的設置面積
- **容量因子（Capacity factor）**：可視為發電效率
容量因子 = 設備全日總發電量 ÷（裝置容量 × 日時數）
- 台灣太陽能發電平均容量因子為 14 %

為方便理解，以下統稱容量因子為「發電效率」



Capacity Factor Prediction



預測值

太陽能日發電效率

- 總共 44 座太陽能發電廠
- 2017/01/01 - 2021/09/30 每日發電量
- 透過發電廠裝置容量，將日發電量轉換成日發電效率

地理資料

- Google Earth API 爬取發電廠與天氣測站經緯度
- 計算發電廠與最近的天氣測站距離

發電廠



天氣測站



距離

特徵值

天氣數值

- 資料來源：觀測資料查詢平台
- 爬取離發電廠最近的天氣測站之相對應時間區段天氣資料
- 天氣欄位包含氣溫、降水量、日照、氣壓等共 36 個欄位

1. 類別變數處理

- 針對類別變數進行 get dummy

2. 缺失值處理

- 刪除過多空值欄位
- KNN 補值：以中央氣象站的滿欄位資料為補值對象，將其餘有缺值氣象測站資料補齊

3. 檢視是否有重複性高的特徵值

- 為降低共線性問題
- 找出相關係數高於 0.9 或低於 -0.9 的自變數組合，並從各組合中各挑出一個變數

1. 8:2 隨機切分成 training dataset 與 testing dataset
2. 以 training dataset 資料分布對 training dataset 與 testing dataset 的特徵值做標準化
3. 使用 training dataset 做 cross validation tune hyperparameters

Linear models

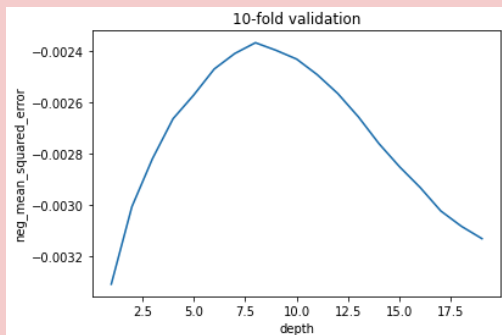
- OLS regression
- LASSO regression
- Ridge regression
- Elastic Net

Tree-based models

- CART regression tree (tune depth)
- Random Forest (tune depth)
- Gradient Boosting Decision Tree (tune the number of boosting stages)

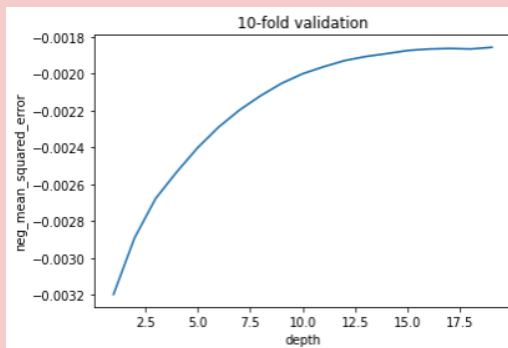
Tree-based models – tuning results

CART regression tree



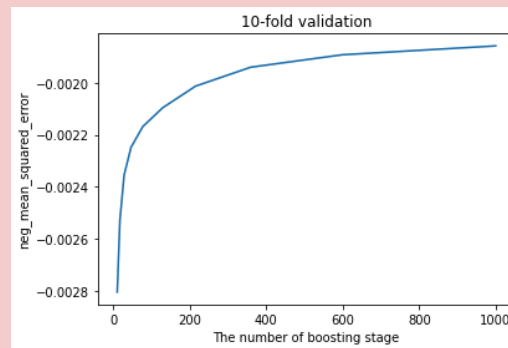
取 depth = 8

Random Forest



取 max_depth = 19

GBDT

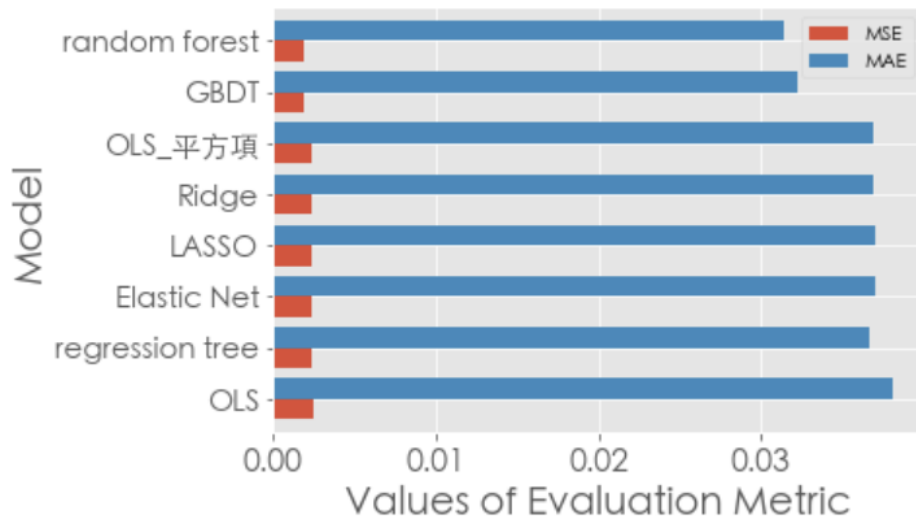


取 the number of stages = 1000

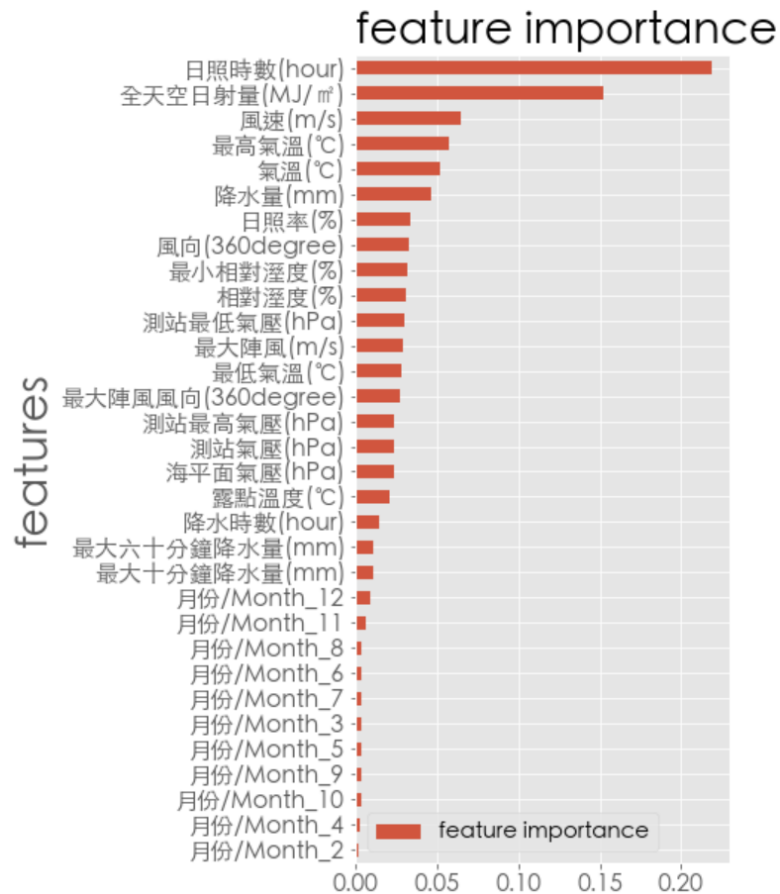
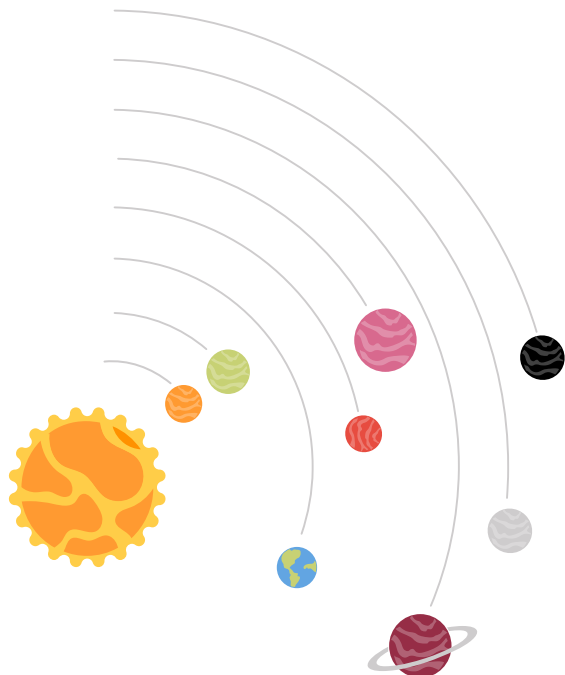
Capacity Factor Prediction / Result

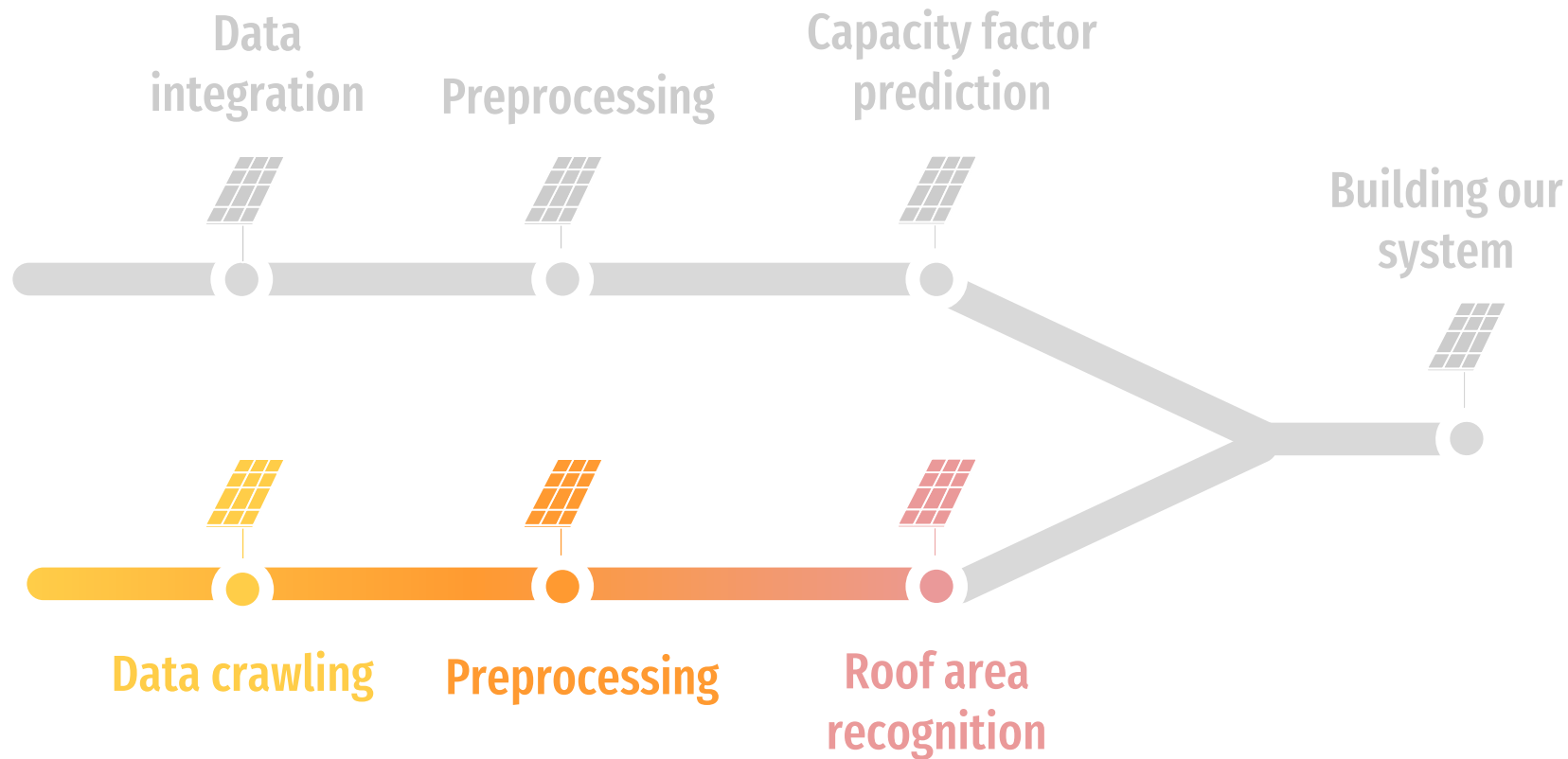
- Tree-based models 預測結果優於 linear models
- Linear models 中有考慮 regularization term，預測結果較好
- Random Forest 預測結果最好，取 random forest 作為系統預測使用

	MAE	MSE
Random Forest (max_depth = 8)	0.0314	0.0018
GBDT (n_estimator = 1000)	0.0323	0.0018
CART regression tree (depth = 8)	0.0367	0.0024
OLS_平方項	0.0369	0.0024
Ridge	0.0369	0.0024
LASSO	0.037	0.0024
Elastic Net	0.037	0.0024
OLS	0.0381	0.0025



- 最好模型 Random Forest 重要變數





01

讀取 google
road map



02

將圖片轉為灰階
並增強圖像



03

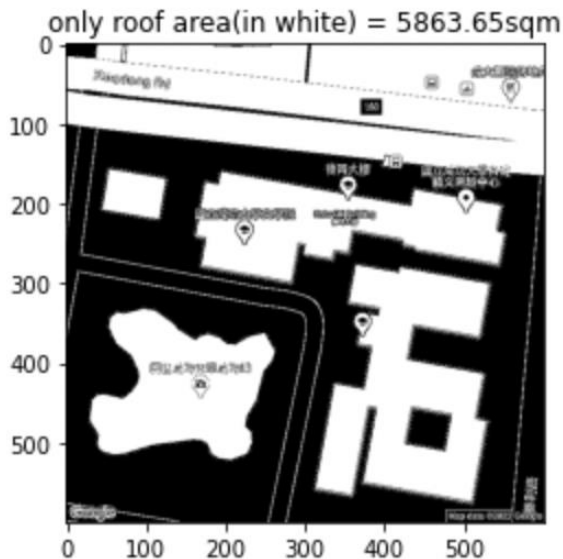
影像二值化



04

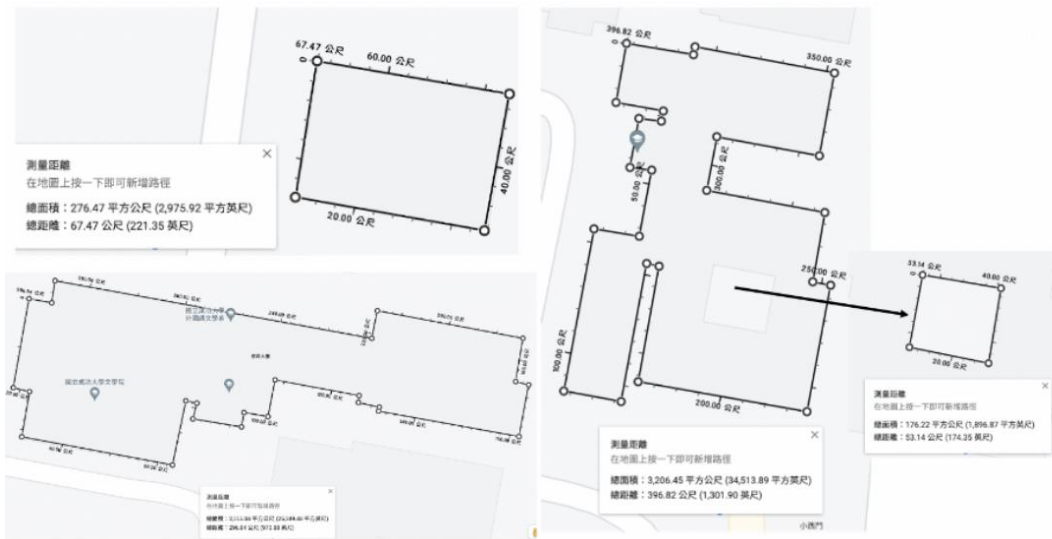
計算屋頂面積

我們計算出屋頂面積



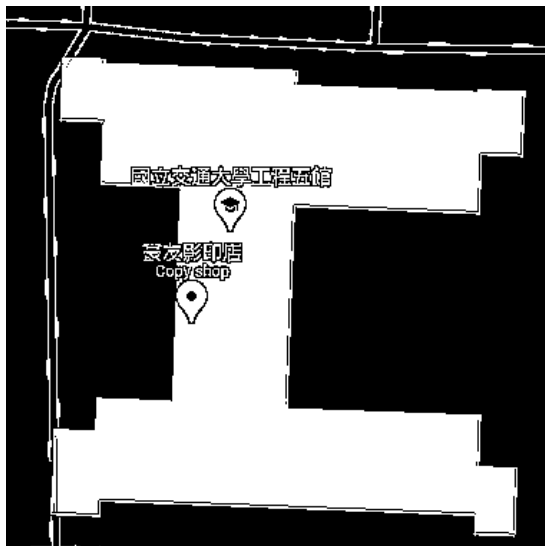
屋頂面積為 5863.65 m²

人工計算的屋頂面積



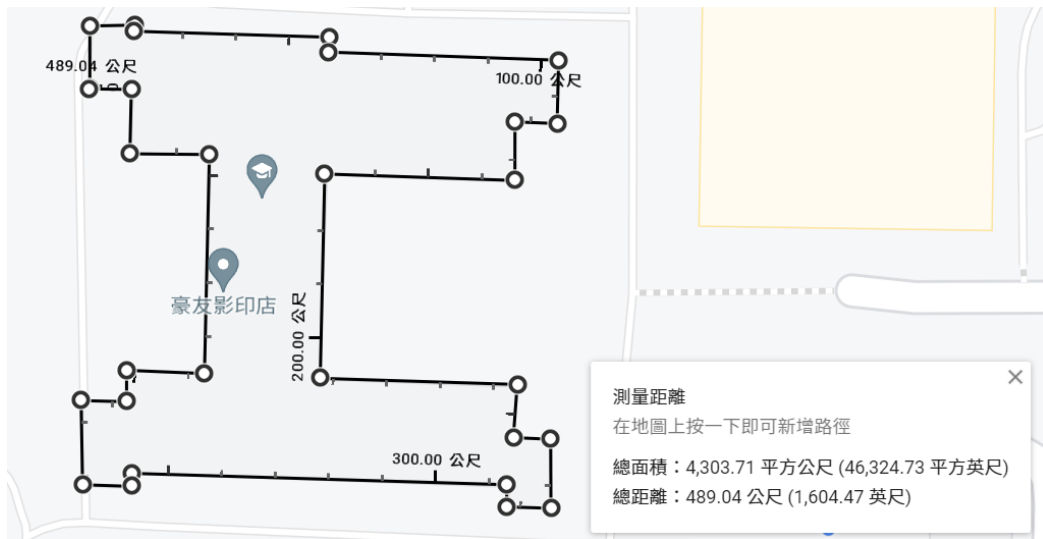
屋頂面積為 5661.74 m²

我們計算出屋頂面積

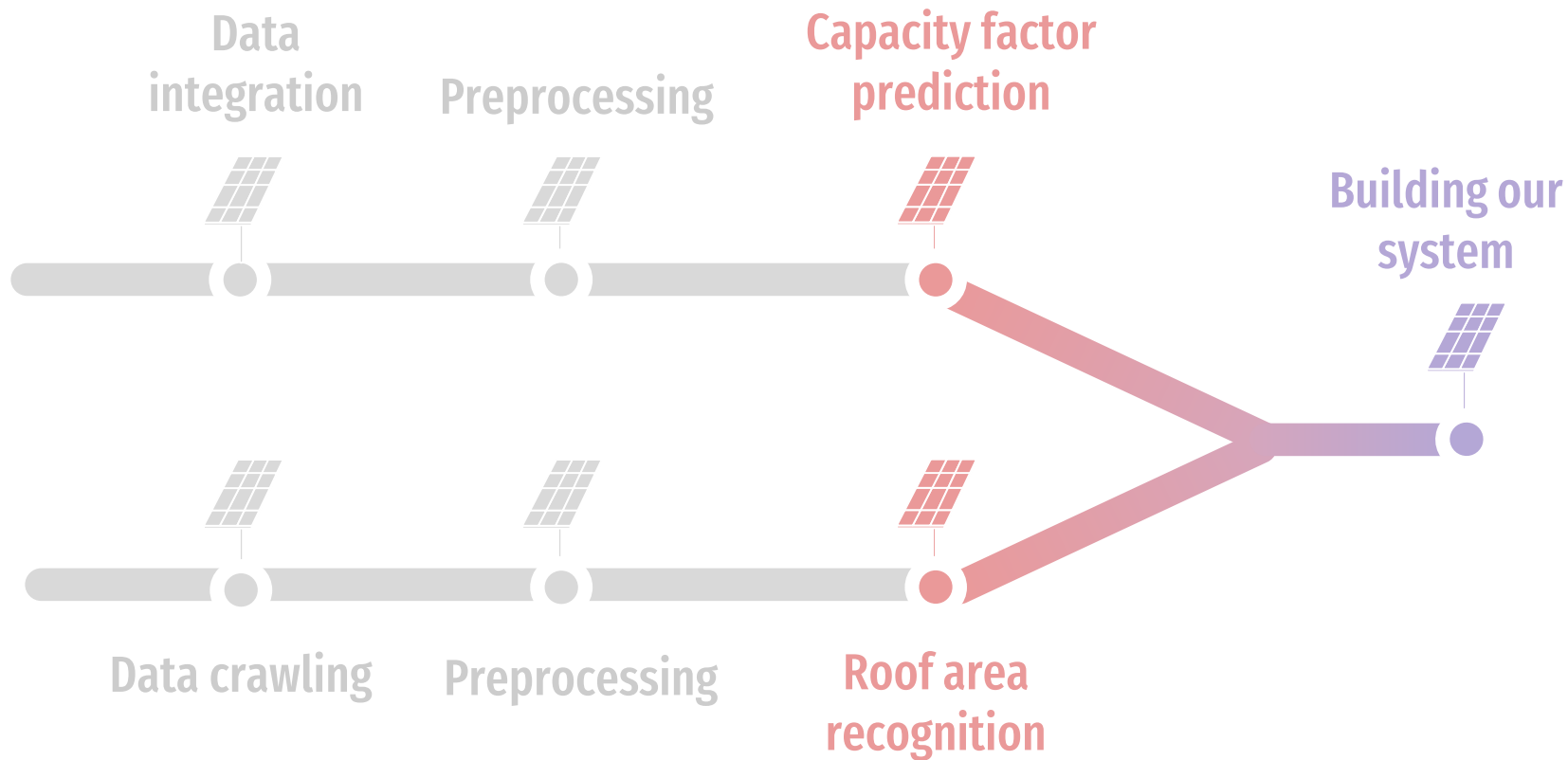


屋頂面積為 4315.88 m^2

人工計算的屋頂面積



屋頂面積為 4303.71 m^2



系統流程

Capacity Factor Prediction : 預測特定地點之日發電效率

Data Collection

合併發電廠發電資料與最近天氣測站之天氣資料

Data Preprocessing

類別變數處理、缺失值處理、高相關自變數處理

Training and Testing

以天氣因子預測日發電效率

Model Comparison

比較各模型，取預測結果最好的模型作為最終使用模型

開始

獲取使用者輸入之預計建置地址與預期達成之年發電量

找離建置地址最近的天氣測站之天氣資料

預測太陽能日發電效率

計算所需裝置容量

計算所需屋頂面積

輸出所需屋頂範圍

結束

Roof Area Recognition : 計算屋頂面積

Data Retrieval

從 Google Map API 獲取 roadmap 圖像

Image Preprocessing

將圖片轉為灰階並增強圖像，最後將影像二值化

Calculate the rooftop areas

final_xin.ipynb - Colaboratory frontend

127.0.0.1:5000/#/

太陽能預測與建廠選址

參數輸入

預計設廠位置
24.803647, 120.965585

預計達成之年發電量
度

**輸入預計設置太陽能板位置
(經緯度、地址或移動圖標)**

計算結果

可收購屋頂面積
0

至少所需裝置容量
0

以下範圍可收購之屋頂面積
0

下載圖片



屋頂面積計算

- **Capacity Factor Prediction**
 - Tree-based models 的預測結果比 linear models 好
 - Random Forest 預測效果最佳
- **Roof Area Recognition**
 - 使用 Google Road Map 辨識與計算屋頂面積
- **Building our system**
 - 建立人性化介面，協助公司做屋頂購買決策



感謝聆聽 (๑´ㅂ`๑)