

和雲行動服務 X NTUDAC 合作專案

【期末報告】

2024.05.31



和雲行動服務



專案目標

分為「舊場貼標」及全台不同地區網格單位車格月收入之「新場預測」。

舊場貼標

- 舊場整體前 3 大重要特徵為：附近公園面積、假日費率、附近餐廳數量。
- 建議在合理情況可提高費率至特定金額，並考慮提高部分停車場之最高收費上限。
- 以廣州街、愛國西路停車場為例，推測造成兩場收入差異的原因為停車場入口位置。

新場預測

- 新場整體重要特徵多為附近地點類別數量，建議新場選址時可著重觀察。
- 「住商用建築比例」與網格的預測收入呈現正相關；反之，「超市與交通樞紐之覆蓋率」與網格的預測收入則為負相關。

成果展示

- 舊場貼標：整體、各場重要特徵之圖表、試算表與 Power BI 儀錶板。
- 新場預測：12 個月之新場網格預測收入的 Power BI 儀錶板。

AGENDA

- 1 專案目標
- 2 舊場貼標流程
- 3 舊場貼標結果
- 4 新場預測流程
- 5 新場預測結果
- 6 Q&A
- 7 Appendix

AGENDA

- 1 專案目標
- 2 舊場貼標流程
- 3 舊場貼標結果
- 4 新場預測流程
- 5 新場預測結果
- 6 Q&A
- 7 Appendix

將專案目標分為舊場貼標，及預測全台不同地區新場之 **單位車格月收入**。

舊場貼標

- 整合停車場周邊及內部資訊，如附近餐廳數量、停車場費率等因素對於單位車格月收入影響的重要性，並挖掘 **重要影響變數**以作為標籤。

新場預測收入

- 將台灣拆分為數個網格，利用舊場訓練之模型預測不同網格的單位車格月收入，最後呈現不同地區之預測收入高低。

AGENDA

1

專案目標

2

舊場貼標流程

3

舊場貼標結果

4

新場預測流程

5

新場預測結果

6

Q&A

7

Appendix

訓練模型後，選擇最佳模型並運用
以作為場站標籤。

SHAP 與 PDP 尋找與解讀重要影響特徵



和雲行動服務

數據蒐集

彙整所有內外部資料中，可能影響月收入的變數。

預處理

填補資料中的缺漏值等，預處理過程詳見 Appendix 2。

特徵工程

將不同特徵轉變為適合模型的樣子，如標準化等。

模型訓練

將資料拆分為 Training、Validation、Testing Set，以 Training Set 訓練模型，Validation Set 調整模型參數，Testing Set 進行測試。

模型評估

透過 Testing Set 的預測結果選出表現較佳的模型。

為各停車場貼標

使用 SHAP 和 PDP 找出重要影響特徵作為標籤並作解釋。

舊場訓練資料以場站 x 月作為單筆資料定義，以 Training Set 訓練模型，Validation Set 調整模型參數，Testing Set 測試模型表現。



和雲行動服務

停車場 / 月份	單位車格收入	平日費率	假日費率	土地面積	人口數	人口密度
A 停車場 / 1 月	105,080	50	60	7.6071	146,970	19,320.10885
B 停車場 / 2 月	115,180	50	60	13	147,515	19,391.75244
C 停車場 / 3 月	103,135	50	60	18	148,103	19,469.04865

⋮

D 停車場 / 1 月	48,510	20	20	8.5	147,515	19,391.75244
E 停車場 / 2 月	50,400	20	20	9.2	148,103	19,469.04865
F 停車場 / 3 月	64,830	20	20	10	148,505	19,521.89402

⋮

80%
Train
+
Validation




...

20% Test

註：資料皆已經過標準化、遺漏值處理。(詳見 Appendix 2)

舊場貼標模型以 XGBoost 表現最佳, 預測平均誤差 18.31 %。

- MAPE 18.31% 代表平均而言預測單位車格月收入的實際值會落在預測值的 $\pm 18.31\%$ 區間內。
- RMSE 719.6 代表平均而言模型對於單位車格月收入的預測值和實際值會差 719.6 元/格。
- R^2 0.95 代表模型可以解釋資料 95% 的變化。

指標 \ 模型	Random Forest	XGBoost	LGBM
MAPE 	61.82%	18.31%	20.67%
RMSE 	1467.5	719.6	724.3
R^2 	0.79	0.95	0.95

註: 指標為該模型在 Testing Set 的表現

AGENDA

1

專案目標

2

舊場貼標流程

3

舊場貼標結果

4

新場預測流程

5

新場預測結果

6

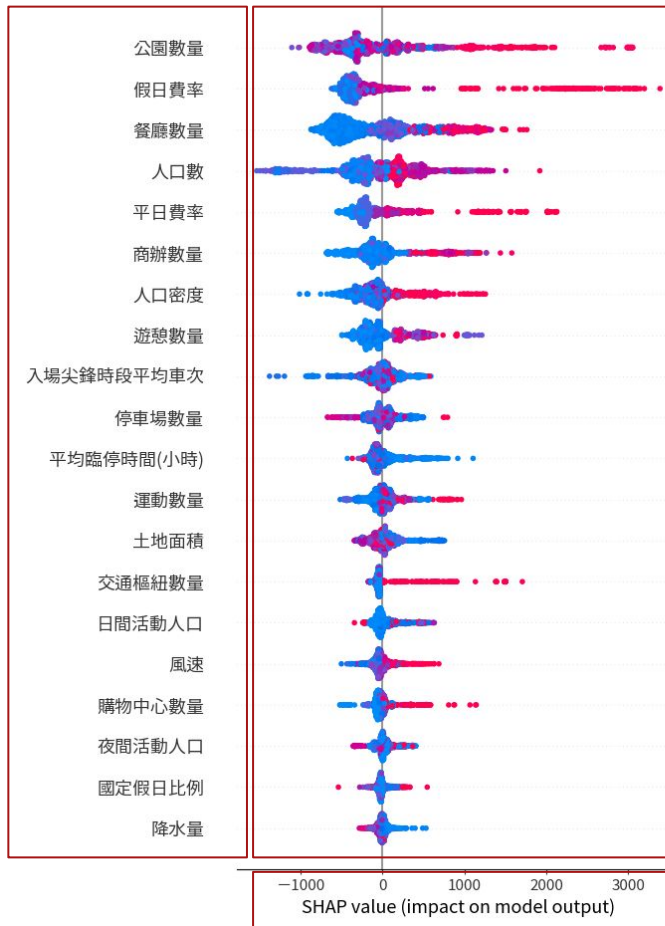
Q & A

7

Appendix

運用 SHAP 相關圖表解讀 整體停車場的重要影響特徵。

由上到下為
前 20 大重要特徵排序



資料點的顏色代表各個特徵
的數值大小：

- 愈紅 = 特徵數值愈大
- 愈藍 = 特徵數值愈小

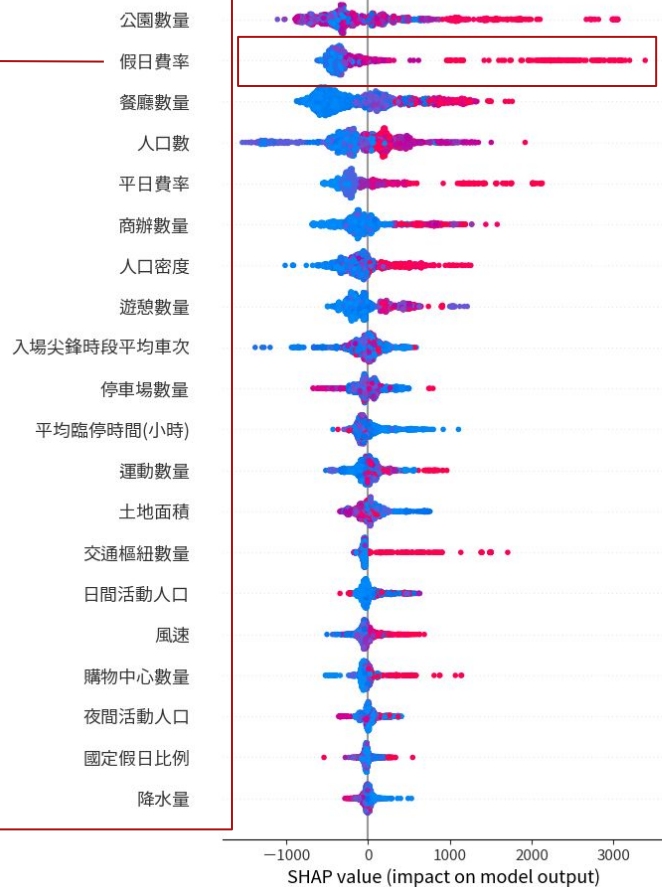
SHAP Value 代表特徵的重要性：

- 正值 = 對預測收入有正向影響
- 負值 = 對預測收入有負向影響

運用 SHAP 相關圖表解讀 整體停車場的重要影響特徵。

以假日費率為例

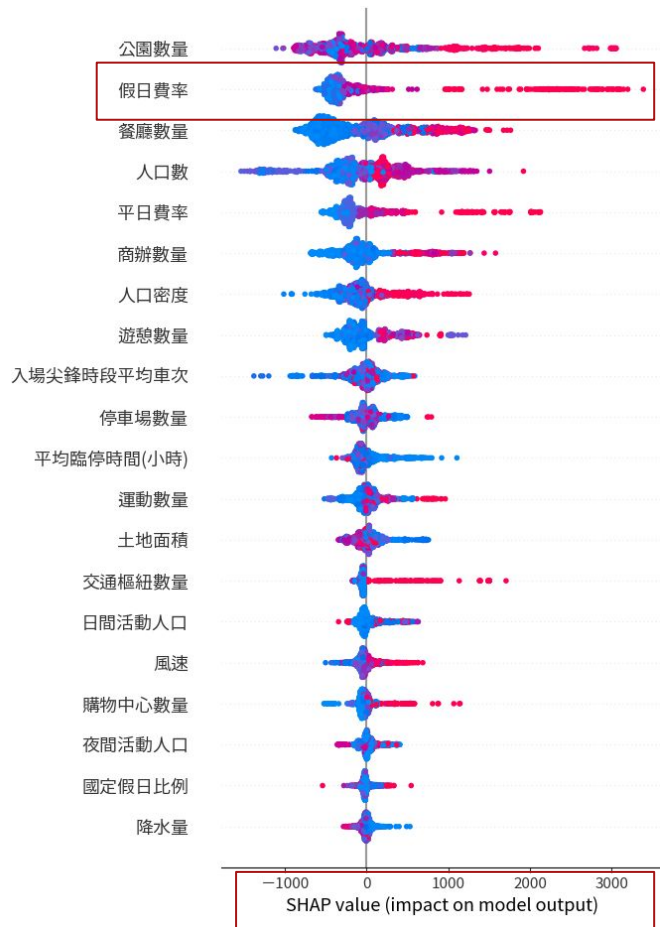
假日費率為整體第2重要特徵



紅色的資料點代表較高的
假日費率，藍色則較低。

運用 SHAP 相關圖表解讀 整體停車場的重要影響特徵。

以假日費率為例



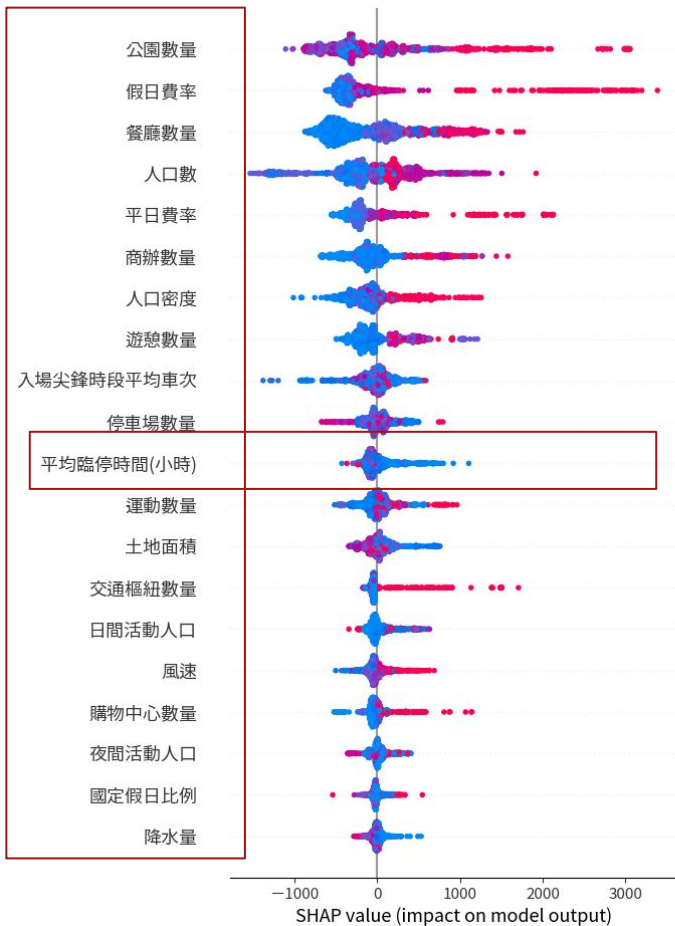
假日費率愈高(顏色愈紅),
SHAP Value 愈大,
即對收入的正向影響愈大。

假日費率愈高, 單位車格收入愈高。

運用 SHAP 相關圖表解讀 整體停車場的重要影響特徵。

以平均臨停時間為
例

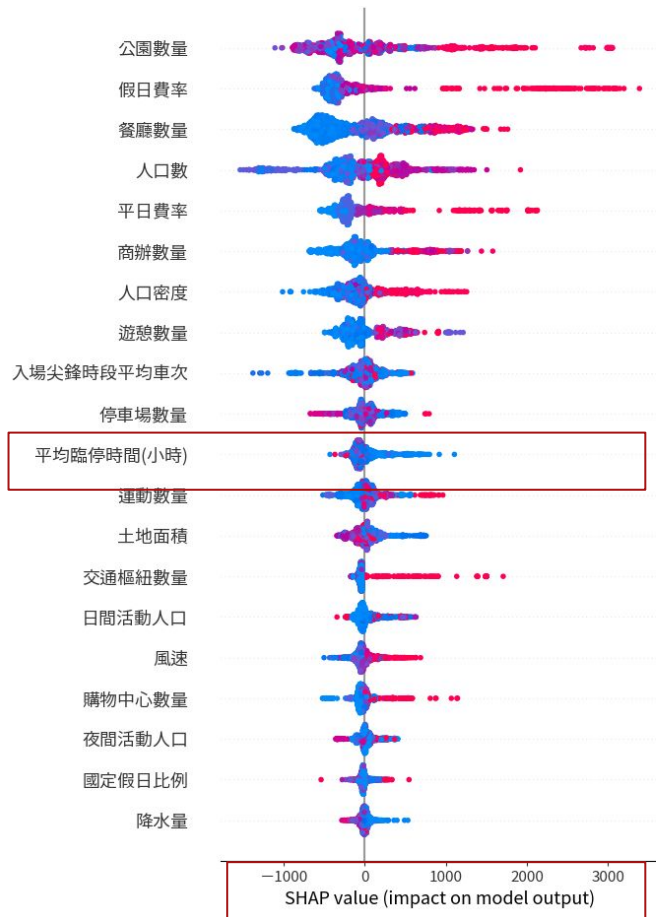
平均臨停時間為整體第11重要特徵



紅色的資料點代表較長的
平均臨停時間，藍色則較短。

運用 SHAP 相關圖表解讀 整體停車場的重要影響特徵。

以平均臨停時間為
例



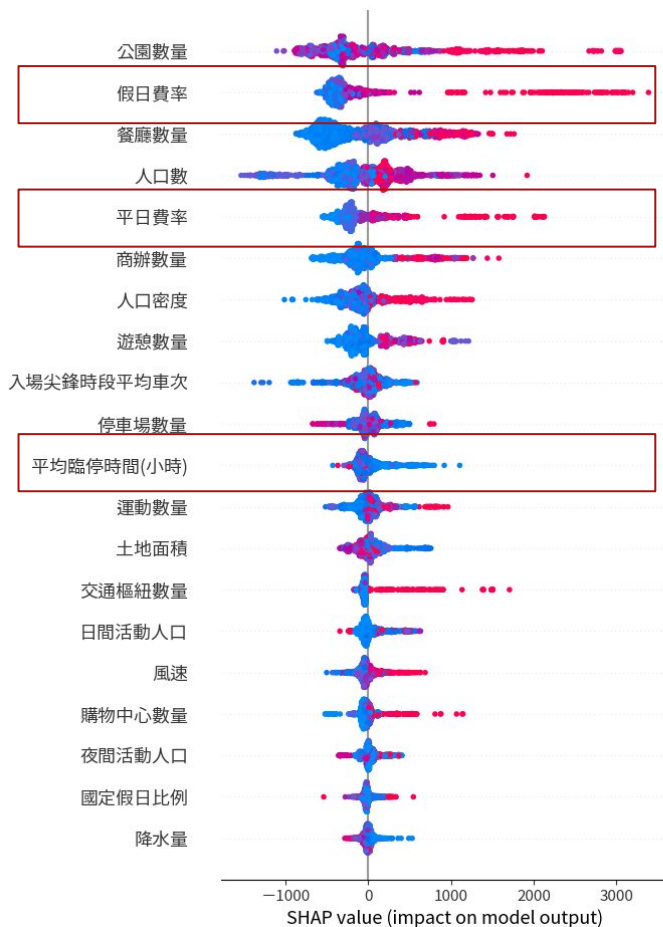
平均臨停時間愈短(顏色愈藍),
SHAP Value 愈大,
即對收入的正向影響愈大。

↓
平均臨停時間愈短, 單位車格收入愈高。
(反之, 平均臨停時間愈長, 單位車格收入愈低。)

分析可透過營運策略改變的變數， 以提升停車場單位車格月收入。

可透過營運策略改變的變數：

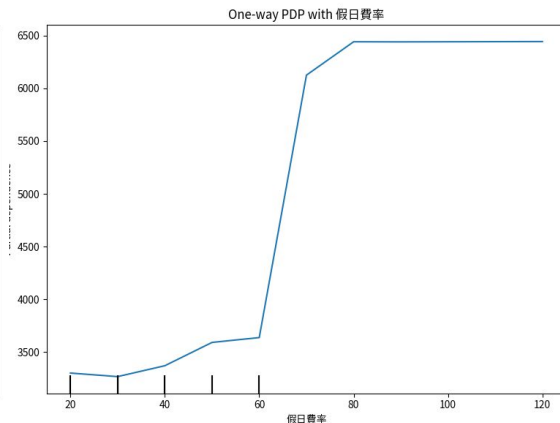
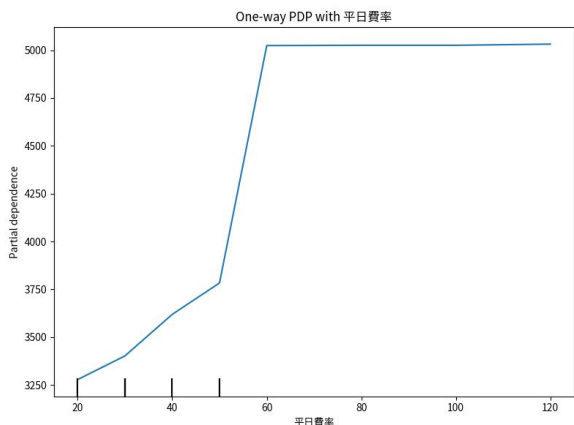
- 平 / 假日費率
- 平均臨停時間



平 / 假日費率 與 平均臨停時間 為重要影響特徵，且可透過內部策略改動對收入帶來正向影響。

平 / 假日費率

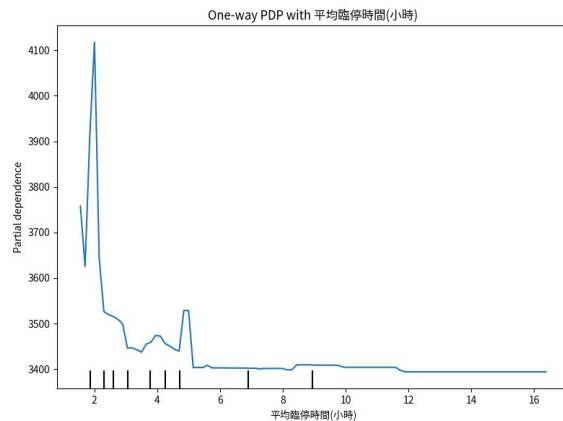
- 平 / 假日費率的資料點分別集中在20~50 及 20~60。
- 此範圍內假日費率愈高，預測單位車格月收入愈高。



若能在合理情況下，將平日費率提升至 60、假日費率提升至 70，預計可對收入造成顯著的正向影響。

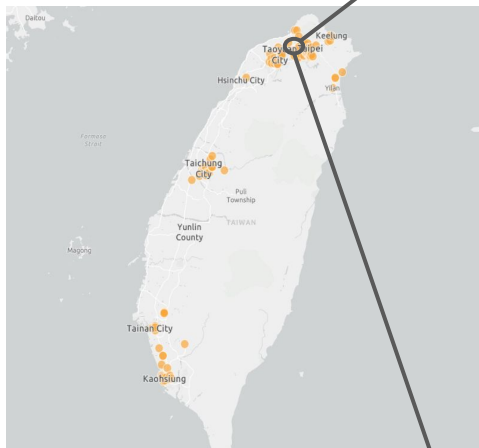
平均臨停時間

- 平均臨停時間 ≤ 2 小時：對收入有正向影響。
- 平均臨停時間 > 2 小時：預測收入有明顯跌幅，且此後呈現負向影響。



推測可能原因：臨停時間過久，造成停車場持續滿位。 ➡
對於平均臨停時間長，且有收費上限的停車場，
若提升最高收費上限，可能避免損失潛在收入。

單一停車場的重要特徵解釋： 透過儀錶板可了解特徵數 值高低對於各場收入的影響程度。



各場之重要特徵不同, 影響機制也不同: 以廣州街及愛國西路停車場為例。

廣州街停車場

餐廳、商辦與遊憩場所數量多，
對收入有正向影響。



台北愛國西路停車場

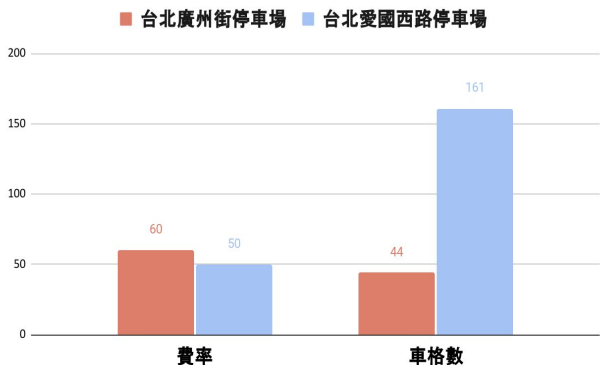
入場尖峰時段平均車次低、鄰近停車場數量多，
對收入的影響力較大, 且為負向影響。



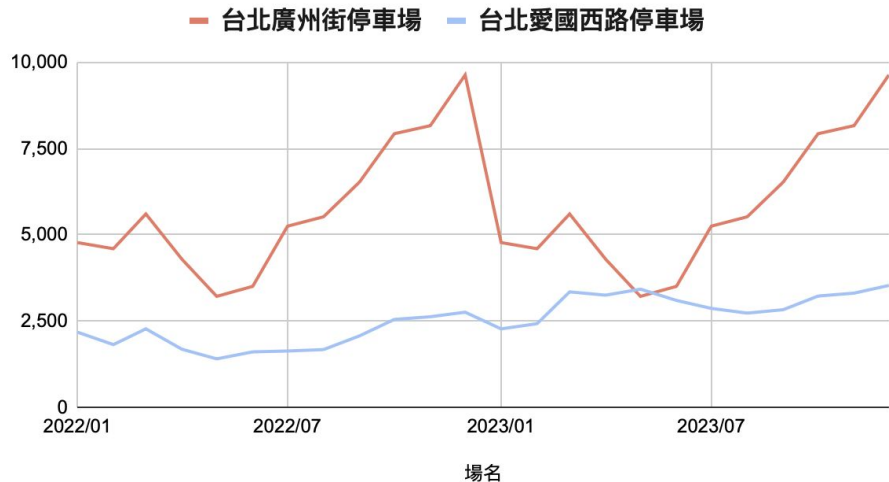
廣州街停車場的費率較高、車格數較少，
相比僅隔一條巷弄的愛國西路停車場，單位車格月收入卻較高。

停車場	費率	車格數
廣州街	高	少
愛國西路	低	多

車格數及費率比較

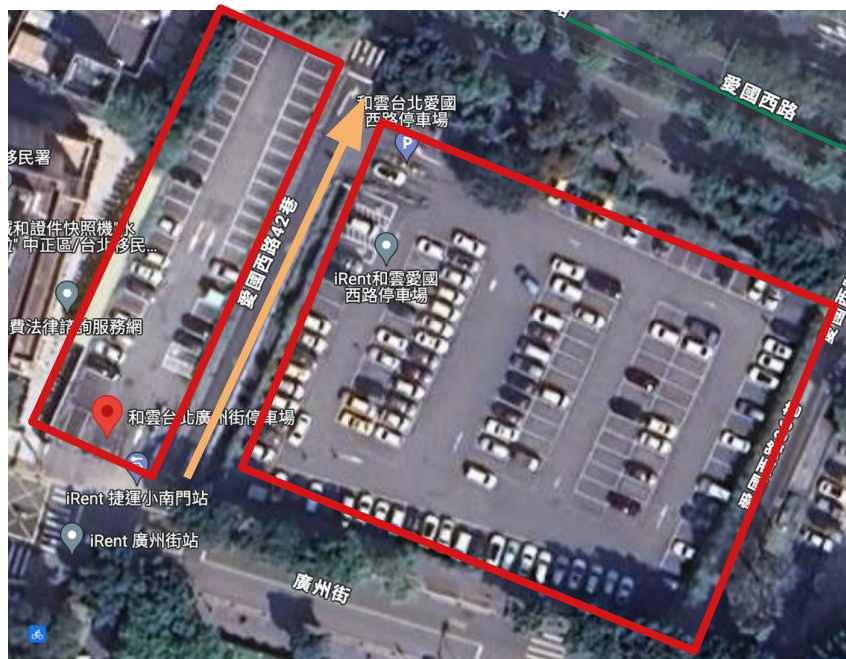


單位車格月收入比較



受惠於**停車場入口位置**，廣州街停車場的單位車格收入較愛國西路停車場高出許多。

愛國西路 42 巷為**單行道**，進入後將首先遇到 **廣州街停車場**，
因此雖廣州街停車場的費率較高、車格數較少，
然受惠**停車場入口位置**，單位車格月收入仍比愛國西路停車場高。



受到**停車場入口位置**影響，廣州街及愛國西路停車場的重要特徵不同。

廣州街停車場

受到**附近地點數量**影響較大，與入場尖峰時段平均車次、鄰近停車場較無關。



台北愛國西路停車場

受到**入場尖峰時段平均車次、鄰近停車場數量**影響較大。



AGENDA

1

專案目標

2

舊場貼標流程

3

舊場貼標結果

4

新場預測流程

5

新場預測結果

6

Q&A

7

Appendix

訓練模型後，選擇最佳模型並用以預測各網格 在未來 2025 年 12 個月的單位車格月收入。

數據蒐集

以網格為依據蒐集資料，蒐集方式詳見 Appendix 4。

預處理

填補資料中的缺漏值等，預處理過程詳見 Appendix 2。

特徵工程

將不同特徵轉變為適合模型的樣子，如標準化等。

模型訓練

將資料拆分為 Training、Validation、Testing Set，以 Training Set 訓練模型，Validation Set 調整模型參數，Testing Set 進行測試。

模型評估

透過 Testing Set 的預測結果選出表現較佳的模型。

預測單位車格月收入

使用最佳模型預測各網格在未來 2025 年 12 個月的單位車格月收入。

新場訓練資料以網格 x 月作為單筆資料定義，以 Training Set 訓練模型，Validation Set 調整模型參數，Testing Set 測試模型表現。

網格索引	月份 1	雨量	國民住宅	一般假日	國定假日	單位車格收入
0	0	5	0.007236	6	3	2,569.880952
1	0	50	0.007236	5	2	3,182.559524
2	1	24	0.007236	3	4	2,201.369048

80%
Train
+
Validation

網格索引	月份 1	雨量	國民住宅	一般假日	國定假日	單位車格收入
39XXX	1	23	0.00157	3	4	2,569.880952
39XXX	0	34	0.00157	6	2	3,182.559524

20% Test

註：資料皆已經過遺漏值處理。

將台灣劃分成等面積 (1km * 1km) 的網格
每列資料都是某網格某月份的資料

以網格 x 月作為單筆資料定義, 預測 2025的12個月 各網格的單位車格月收入

網格索引	月份 1	雨量	國民住宅	一般假日	國定假日
0	0	24	0.012	4	4
1	1	27	0.013	2	8
2	1	100	0.0172	2	8

⋮

39XXX	1	127	0.007	2	8
39XXX	0	29	0.02	4	3
39XXX	0	39	0.01	3	2

⋮

單位車格收入
?
?
?
⋮
?
?
?

使用
訓練完畢
的模型
進行預測

註 1: 資料皆已經過遺漏值處理。

註 2: 天氣等資料由歷史資料取平均而得, 平假日天數則為實際值。

新場預測模型以 XGBoost 表現最佳，預測平均誤差為 16.65%。

- MAPE 16.65% 代表平均而言預測單位車格月收入的實際值會落在預測值的 $\pm 16.65\%$ 區間內，儀錶板中將呈現此預測範圍。
- RMSE 837.98 代表平均而言模型對於單位車格月收入的預測值和實際值會差 837.98 元/格。
- R^2 0.93 代表模型可以解釋資料 93% 的變化。

指標 \ 模型	Random Forest	GBM	XGBoost	LGBM
MAPE ↓	15.77%	19.40%	16.65%	24.64%
RMSE ↓	879.36	890.23	837.98	878.61
R^2 ↑	0.92	0.92	0.93	0.92

註：指標為該模型在 Testing Set 的表現，新場模型處理詳見Appendix 6。

AGENDA

1

專案目標

2

舊場貼標流程

3

舊場貼標結果

4

新場預測流程

5

新場預測結果

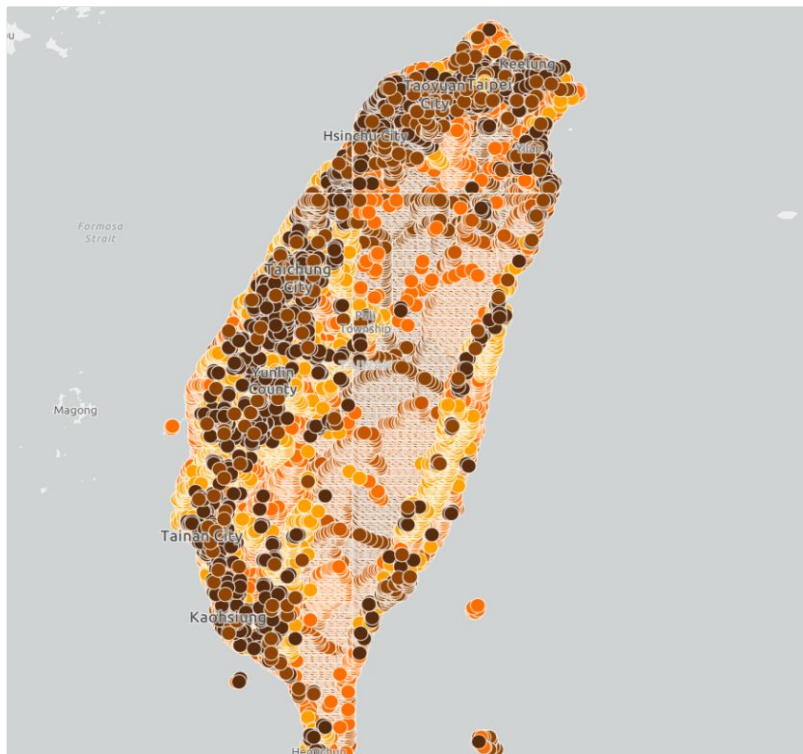
6

Q&A

7

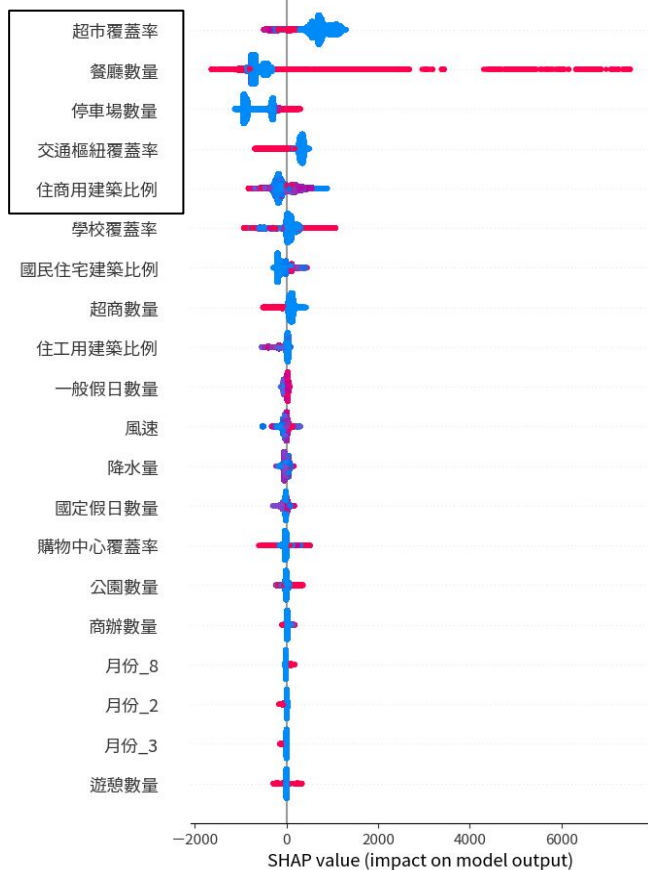
Appendix

預測結果 Power BI 儀錶板展示



註：準確度計算方式詳見Appendix 7。

整體以網格 內附近地點類別數量、住商用建築比例 為重要特徵。

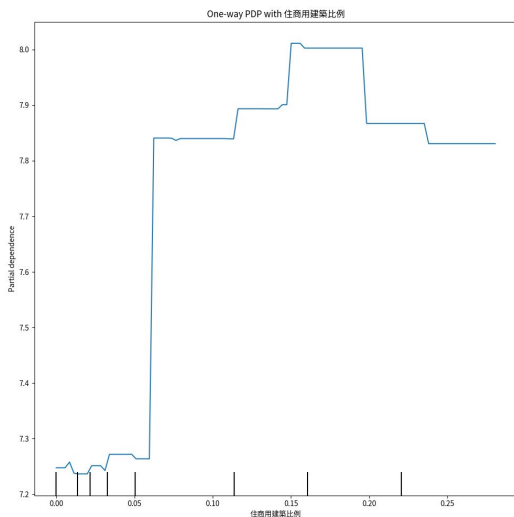


建議新場選址時
可著重觀察上述的地理特徵。

整體網格重要特徵之 PDP 解釋

住商用建築比例越高，
網格的預測單位車格月收入愈高。

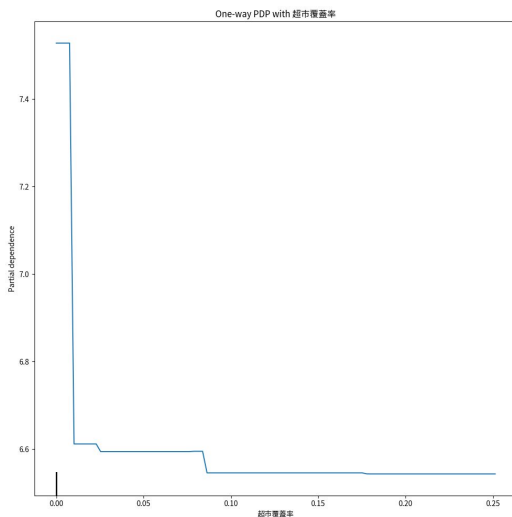
住商用建築比例



住商用建築比例越高，則單位車格收入越高，故建議將新停車場設置在商業區。

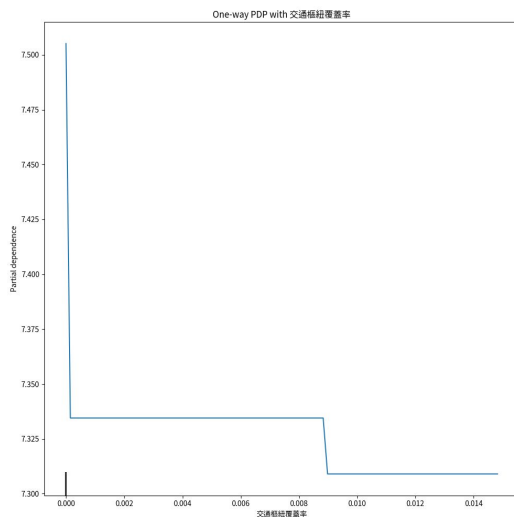
超市與交通樞紐之覆蓋率愈高，
網格的預測單位車格月收入愈低。

超市覆蓋率



由於超市可能附設停車場，排擠其他停車場的營利空間，故超市覆蓋率愈高，單位車格收入愈低。

交通樞紐覆蓋率



推測當網格附近公共運輸選擇較多時，開車的誘因降低，進而減少停車需求。

AGENDA

1

專案目標

2

舊場貼標流程

3

舊場貼標結果

4

新場預測流程

5

新場預測結果

6

Q & A

7

Appendix

The background is a stylized, light-colored illustration of a city street scene. On the left, a person wearing a white helmet and a dark vest is riding a white scooter. The scooter has a red 'i Rent' logo on its side. In the center, a yellow car is visible, also featuring a red 'i Rent' logo. To the right of the car, a person is sitting at a desk, working on a laptop. Another person is standing next to the desk, looking at the laptop. The scene is set against a backdrop of stylized buildings and a clear sky.

**Any
Question?**

AGENDA

1

專案目標

2

舊場貼標流程

3

舊場貼標結果

4

新場預測流程

5

新場預測結果

6

Q & A

7

Appendix

Appendix 1 | 舊場資料蒐集

目標變數 (Y)	停車場臨停月營收	各停車場單位車格月收入
解釋變數 (X)	停車場費率表+場站資訊	優惠與否、平 / 假日費率、車格數
	附近停車場資料	附近停車場最低平 / 假日白天費率
	人口資料	人口數、人口密度、土地面積、信令人口
	氣象資料	氣溫、風速、降水量、相對濕度
	附近各地點類別數量	交通樞紐 / 學校 / 運動 / 商辦大樓 / 購物 / 遊憩 / 停車場 / 餐廳 / 公園

欄位

- 每月平均氣溫
- 每月平均風速
- 每月相對溼度
- 附近停車場費率

問題 & 處理方式

問題說明：有缺漏值

處理方式：

- 停車場費率使用眾數補值。
- 其他欄位使用平均數補值。

處理後示意圖

停車場 / 月份	氣溫	風速	停車場費率
A 停車場 / 1 月	13	400	40*
A 停車場 / 2 月	15	390	40
A 停車場 / 3 月	16	380	40

*紅字代表原為缺漏值，經補值的資料。

- 停車場附近各地點特徵數量

處理方式：

- 方法一：直接紀錄地點數量。
- 方法二：根據各地點類別數量由少至多分為 1、2、3 三個等級。

方法一	學校	商辦	遊憩
A 停車場 / 1 月	78	152	5
A 停車場 / 2 月	43	43	31
A 停車場 / 3 月	68	16	18

方法二	學校	商辦	遊憩
A 停車場 / 2 月	3	3	1
A 停車場 / 3 月	1	2	3
A 停車場 / 1 月	2	1	2

欄位

類別型資料

- 月份
- 優惠與否

問題 & 處理方式

使用 One-hot Encoding:

- 月份: 是該月份為1; 否則為0。
- 優惠與否: 有優惠為1; 否則為0。

處理後示意圖

停車場 / 月份	優惠與否	1月	2月
A 停車場 / 1 月	1	1	0
A 停車場 / 2 月	1	0	1
A 停車場 / 3 月	0	0	0

數值型資料

- 類別型資料外的
所有欄位

使用標準化 (Standard Scaler):

- 將每個欄位標準化, 使平均為0、標準差為1, 統一各欄位的數值範圍。

原因:

- 消除不同數值單位的影響。

停車場 / 月份	人口數	雨量	車格數
A 停車場 / 1 月	1.23846	0.9463	0.36518
A 停車場 / 2 月	0.91548	0.0561	-0.4698
A 停車場 / 3 月	-2.4658	1.4681	-1.6498

Appendix 3 | 刪除共線性過高的變數並拆分資料集後，訓練模型。

特徵篩選	<ul style="list-style-type: none">根據 VIF 值將共線性過高的變數刪掉，以避免模型預測結果失準。
異常值處理	<ul style="list-style-type: none">資料中 2022 年 4 ~ 6 月的收入明顯下跌，推測為 COVID-19 影響訓練模型時將其刪除
資料集拆分	<ul style="list-style-type: none">將資料拆分成訓練集、測試集，訓練集用於放入模型進行訓練，而測試集則用來測試模型泛化性和準確度。

Appendix 4 | 新場資料蒐集方式

目標變數(Y)	單位車格月收入	預測在每個網格中新場之單位車格月收入
解釋變數(X)	經緯度	網格中心之經緯度
	人口資料	網格所屬行政區之人口資料
	住宅建築資料	網格所屬行政區之建築型態資料
	氣象資料	距離網格中心點最近的氣象測站資料
	附近各地點類別數量	<ul style="list-style-type: none">● 人流受服務範圍影響較小<ul style="list-style-type: none">○ 直接計算網格內數量，如餐廳、商辦● 人流受服務範圍影響較大<ul style="list-style-type: none">○ 計算網格「屬於該地點類型之服務範圍」的佔比，如交通樞紐。

Appendix 5 | 被刪除網格



問題 & 原因

- 訓練資料中 **單位車格月收入** 呈現偏態
- 人口密度欄位的預測結果與模型訓練結果分布不同, 且較不合理
- 推測失準原因: 人口密度是以行政區計算, 與網格尺度相差較大, 無法代表網格實際人口密度

處理方式

- 將單位車格月收入做對數轉換, 預測對數值, 再轉回原尺度
- 計算模型表現指標時, 以轉回原尺度之單位車格月收入計算
- 將人口密度欄位移除, 並重新訓練模型

Appendix 7 | 各指標介紹

MAPE	<ul style="list-style-type: none">用於想知道實際值和預測值的平均百分誤差表示誤差的相對大小
RMSE	<ul style="list-style-type: none">絕對誤差分析，適用於關心誤差較大的場合適用於實際值較大的場合
R ²	<ul style="list-style-type: none">用於評估模型對變數的解釋能力
準確度	<ul style="list-style-type: none">用於評估預測結果愈靠近 1 愈好

$$MAPE = \frac{1}{n} \sum_{t=1}^n \left| \frac{A_t - F_t}{A_t} \right|$$

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2}$$

$$R^2 = \frac{SS_{res}}{SS_{total}} = 1 - \frac{\sum_{i=1}^n (\tilde{y}_i - y_i)^2}{\sum_{i=1}^n (\tilde{y}_i - \bar{y})^2}$$

與最近和雲停車場距離之倒數

Appendix 8 | 各場重要特徵不同，影響機制也不同： 以廣州街及愛國西路停車場為例

台北廣州街停車場

附近地點類別數量
(餐廳、商辦、遊憩)
對廣州街停車場
有較大的 正面影響。



台北愛國西路停車場

入場尖峰車次、
附近停車場數量
對愛國西路停車場
有較大的 負面影響。



**Appendix 9 | 針對人流受服務範圍影響較大之地點，計算網格「屬於該地點
類型之服務範圍」的佔比(覆蓋率)，如交通樞紐。**

