

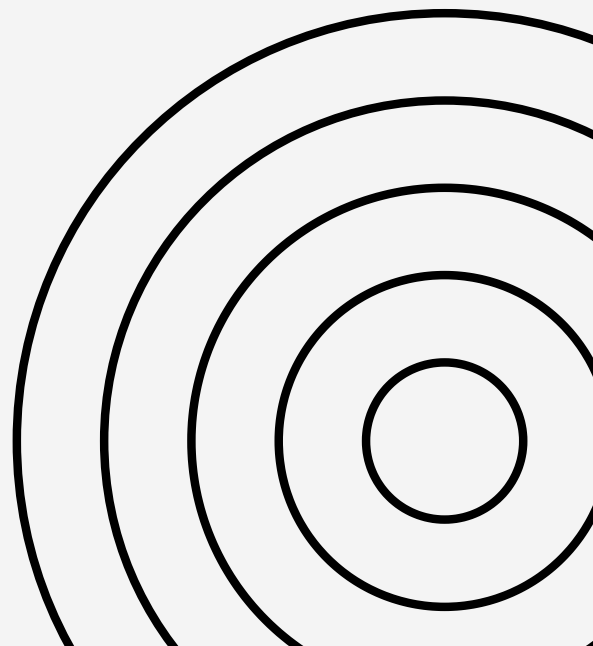


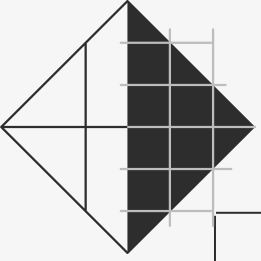
資產活化 AI 應用創新競賽

基於實價登錄資料之土地價值預測

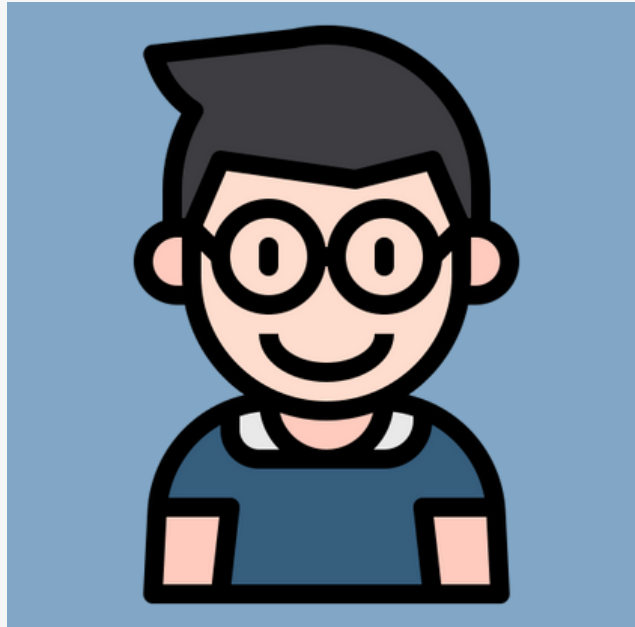
Group Name
老同學

2022 Oct 14





Our Team



何彥南

國立政治大學 | 資訊科學系



莊歲宇

國立政治大學 | 資訊科學系

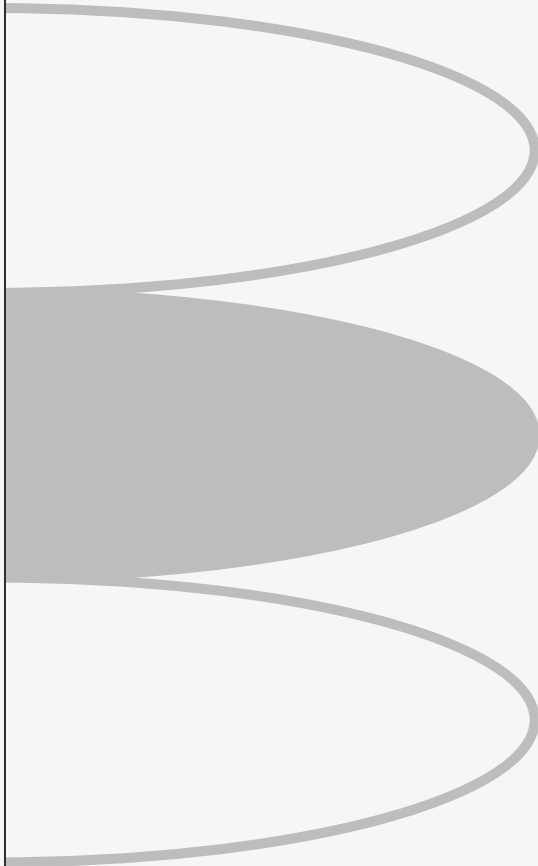


周健因

國立台灣科技大學 | 資訊管理系



Table of Content



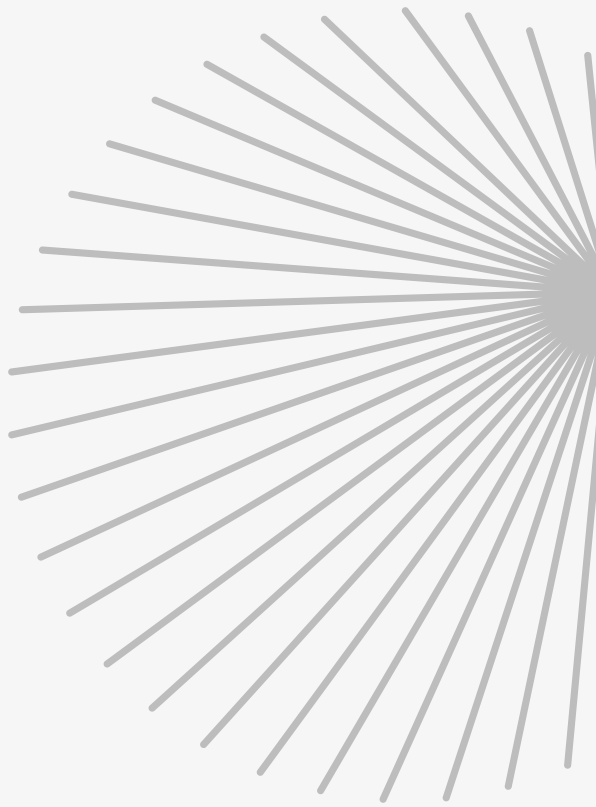
01. 摘要與研究目的

02. 模型介紹

03. 資料集/資料處理/方法

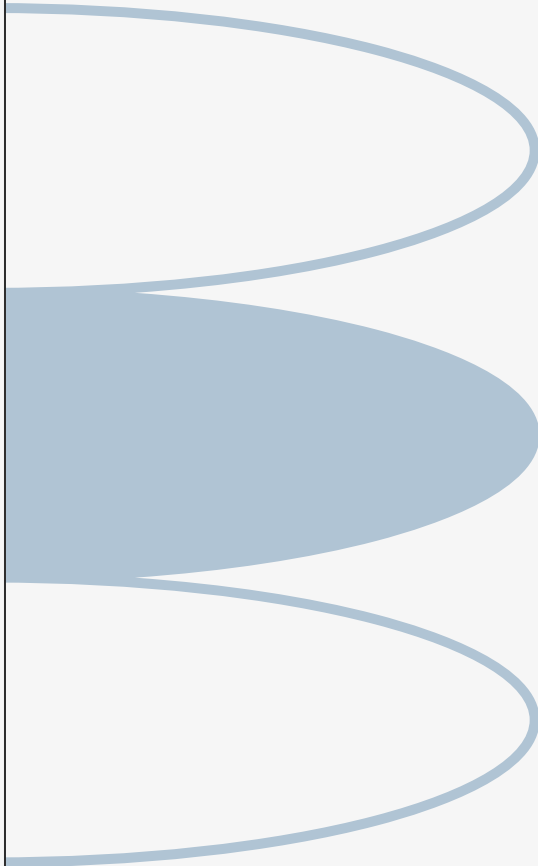
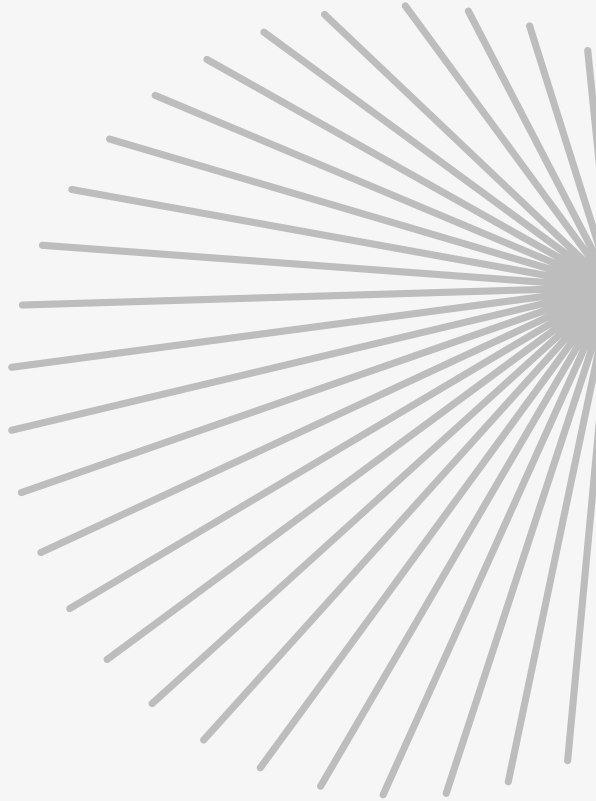
04. 成果展示

05. 研究限制與未來展望





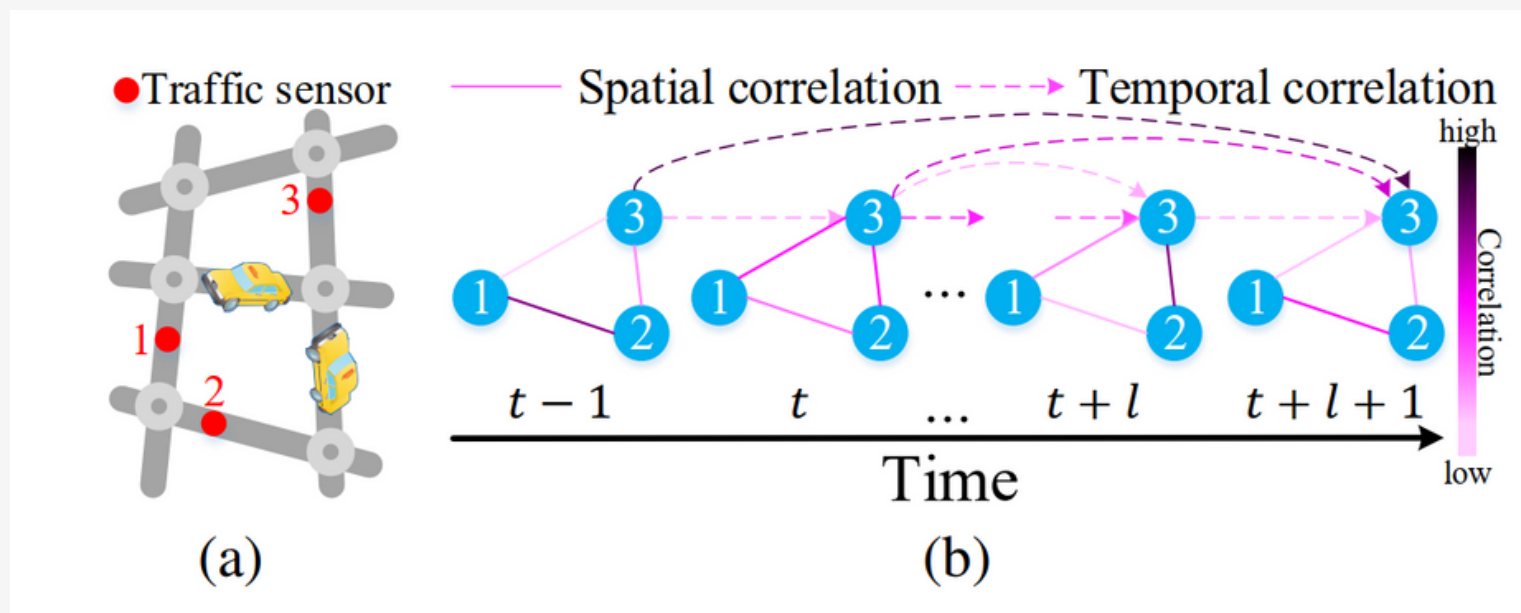
摘要與研究目的

- 採用穩定且公開的實價登陸資料，以其中的土地交易資料與客製化的需求設定預測目標，並以時間序列角度對未來土地價值預測，藉此應用在後續的決策上。
 - 選擇在交通流量預測上表現很好的 GMAN 模型，使用模型特性達到 Multi-Step 時間序列的預測。
 - 本研究以土地價值指標為範例，此方法還可以應用在其他客製化的區域指標上，像是下個月目標點附近的交易數量等資訊。
 - 目的是能透過此模型與流程，將各地區於不同時段之預測土地價值或是其他客製化指標，以此預測結果去結合其他現有資訊去輔助決策，以產出於不同區域、時段下，最適合該土地的利用選擇與開發順序；並輔助資產活化決策，期望能將利益達到最大化。
- 
- 

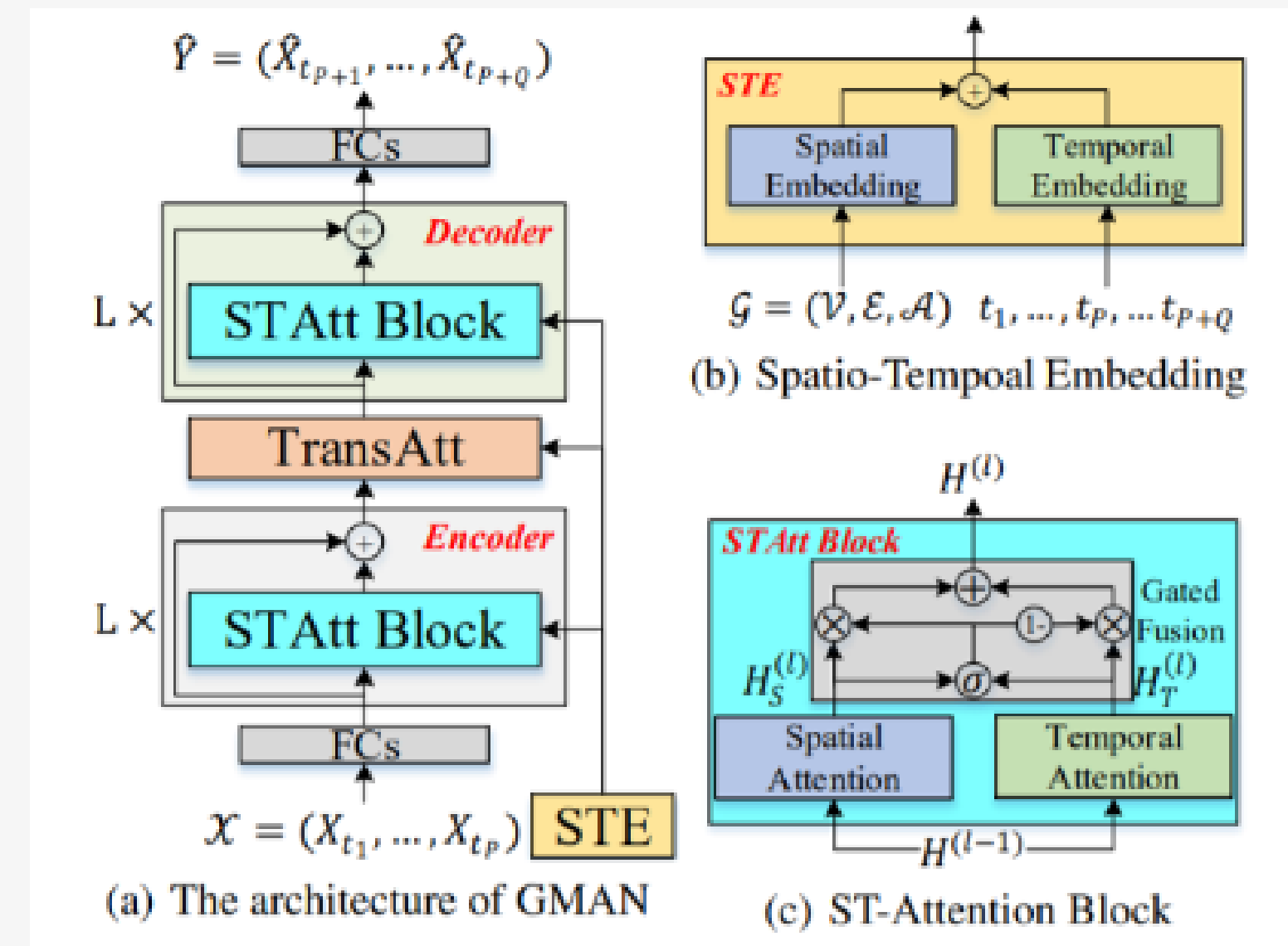
模型介紹

GMAN 模型

- Sensors in a road network
- Encoder – Decoder 技術
→ Multi-Step Time Series Forecasting
- Spatio-Temporal Embedding
→ SE(地理資訊)、TE(時間資訊)
- Attention 機制
→ 讓模型可以學習將注意力放在那些特徵上



GMAN 模型中地理與時間上的注意力機制示意圖



GMAN 模型架構圖

本研究透過導入參考點，將問題轉換為 Spatio-Temporal 問題，以增加資訊的採樣的廣度，並套用到 GMAN 模型上。



資料集

#1 主辦單位提供之土地資料

- 主要使用欄位：鄉鎮市區、地段、地號、使用分區
- 因研究需爬取大量實價登錄資料，因此將先以桃園市資料作為實作範例進行展示

#2 歷年實價登錄資料

- 來源：「內政部不動產成交案件實際資訊資料供應系統」
- 時間：2012年7月 ~ 2022年04月。
- 主要使用欄位：鄉鎮市區、交易標的、土地位置建物門牌、使用分區(相關)、交易年月日、單價元平方公尺
- 本研究主要使用不動產買賣中只有土地交易資料，經過前處理後，約 78000 筆。

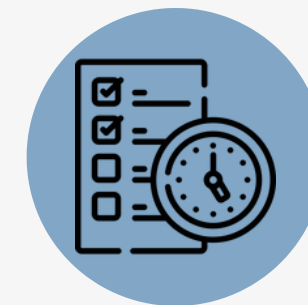
#3 各鄉鎮市區人口密度

- 使用「政府資料開放平台」中「各鄉鎮市區人口之密度」資料集
- 涵蓋 102 年至 110 年之人口資料，其中包含：統計年、區域別、年底人口數、土地面積、人口密度等欄位
- 本研究使用此資料輔助判斷土地開發決策

資料前處理

1 - 資料前處理

- 取用「交易標的」為土地的資料
- 透過 regex (正規表達式, Regular Expression) 切分出地段與地號
- 亂碼、重複值、輸入錯誤與缺失值編碼、移除等處理



2 - 土地爬取經緯度

- 資料來源:「地號 GeoJSON API」、「地籍圖資網路便民服務」
- 大約 68000 筆土地資料，成功爬取到約 54000 筆經緯

3 - 舊地號轉新地號

- 使用「桃園地政資訊服務網」中的「新舊地號查詢」服務
- 使用 python 的 selenium 等工具進行爬蟲，可以自動處理批量資料

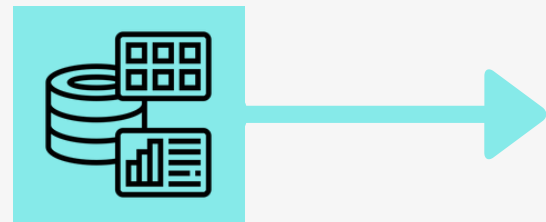
5 - 資料量與時間區間選擇

- 時價登入資料中的前幾個月，資料品質較不穩定，所以不採用。
- 本研究會 101 年 6 月至 111 年 5 月的資料為主

4 - 土地使用的分類與篩選

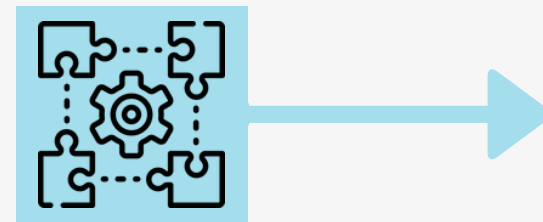
- 主要針對商業區和住宅區之資料，其中商業相關約 1,375 筆，住宅相關約 13,654 筆。

方法



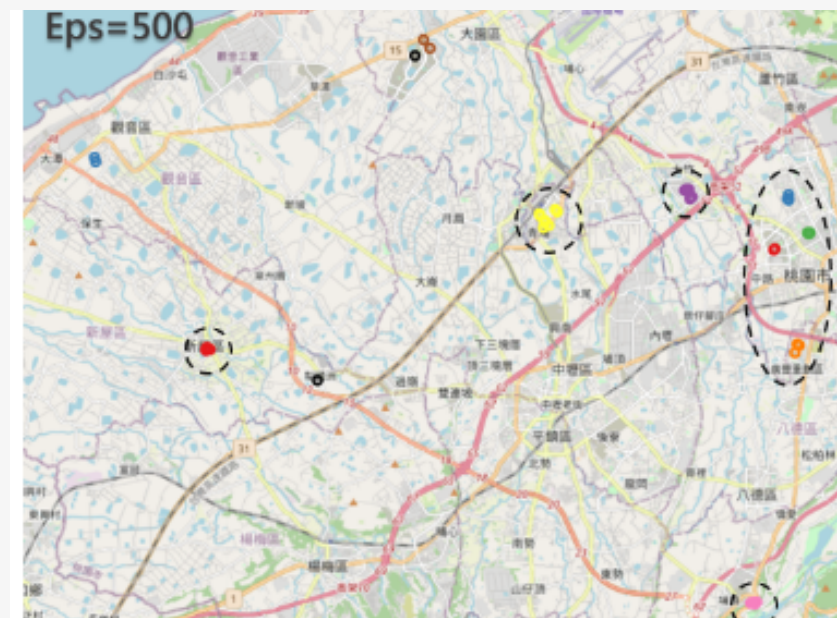
1 - 目標區域範圍設定

- 計算各自範圍內，在各月中的有效點數量
- 本研究以區域範圍 **3000 公尺** 為範例

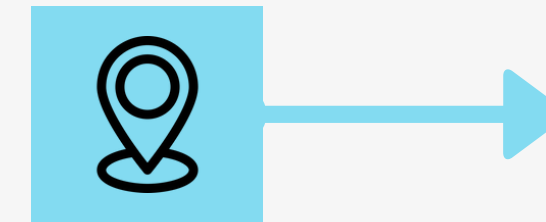


2 - 目標土地分群

- 將目標土地使用 **DBSCAN**，以經位度間的實際距離作為分群依據($esp=500$)，並以**該群中心經緯度**為代表**目標點**

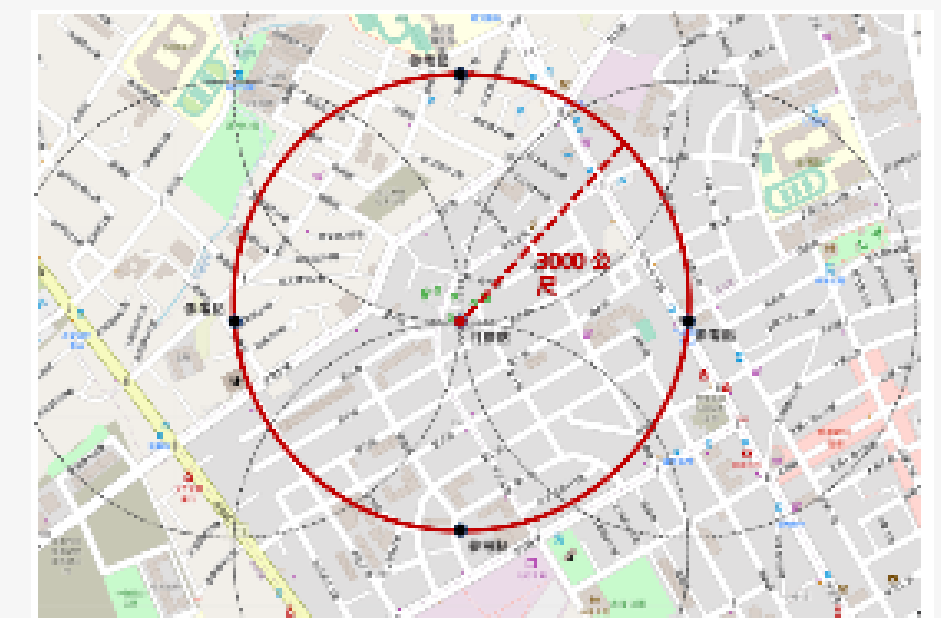


DBSCAN 分群結果



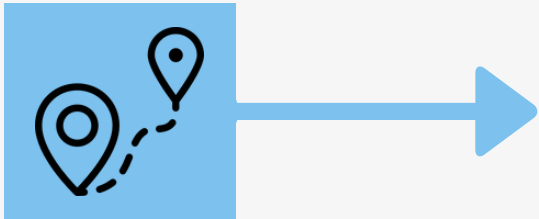
3 - 設置各群參考點

- 採取**上下左右四個方向向外推 3000 公尺**的距離作為**參考點**



參考點示意圖

方法



4 - 計算目標點、參考點與交易土地之距離

- 獲取目標點、參考點到所有交易土地經緯距離，並使用 geopy 套件計算

目標點、參考點到各土地距離矩陣

	目標點	參考點 1	~	參考點 n
交易土地 1	20	50		100
交易土地 2	280	350		500
...				

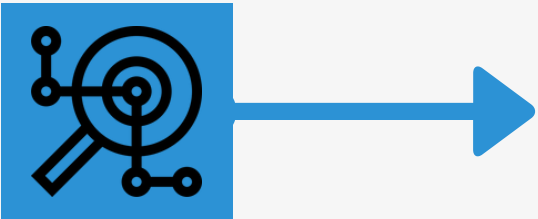


5 - 篩選目標經緯度一定範圍內之交易點

- 從交易資料中篩選出 3000 公尺內所有交易土地點

篩選 3000 公尺內的距離

	目標點	參考點 1	~	參考點 n
交易土地 1	✓	✓		✓
交易土地 2	✓			
...				



6 - 計算各月之目標點和參考點的區域指標

- 依需求，計算每個月的區域指標

計算區域指標

	目標點	參考點 1	~	參考點 n
2012.7	50	100		80
2012.8	30	60		70
...				
2022.4	100	50		120

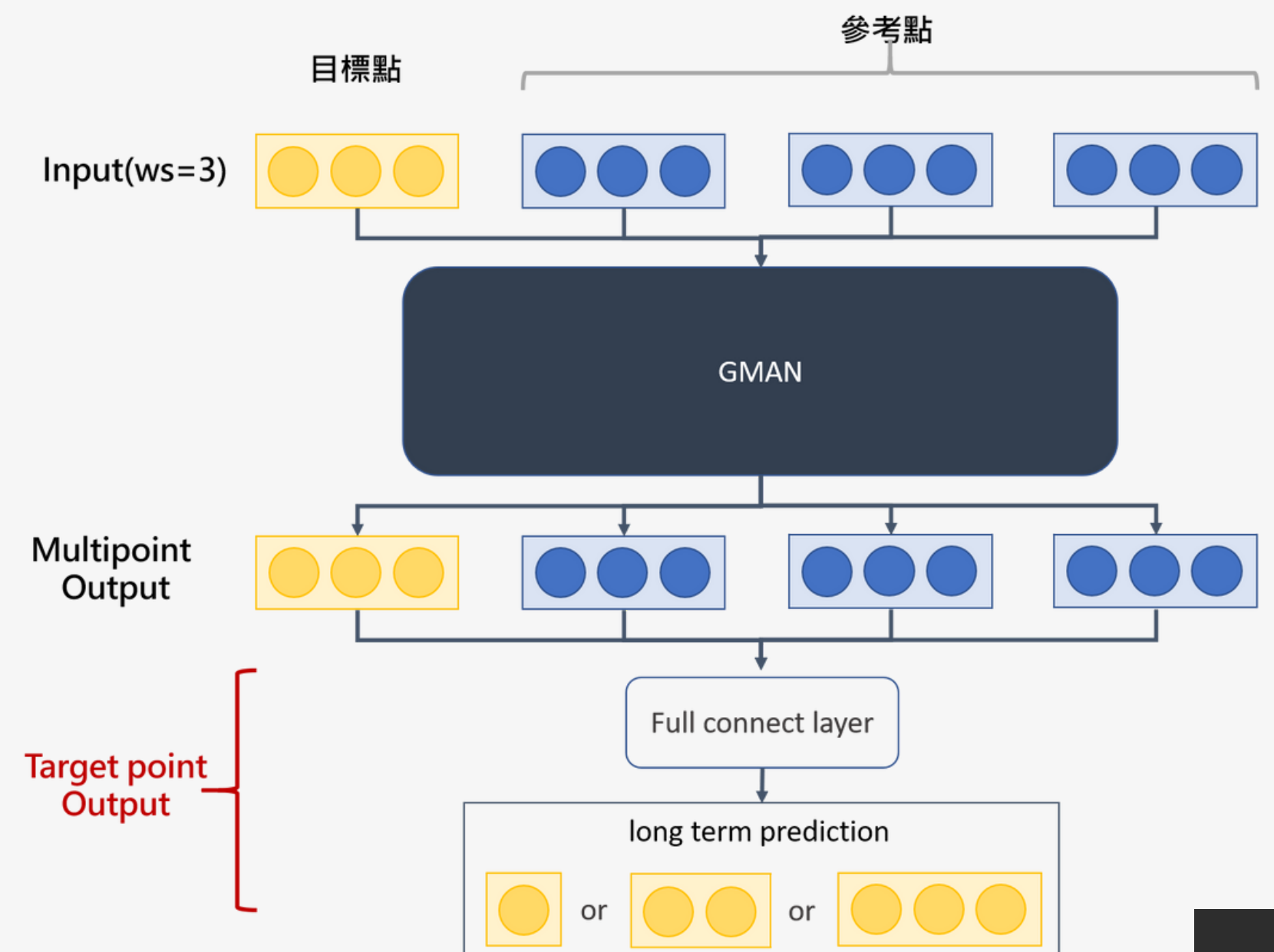
方法 - 資料建模與預測

訓練資料

Input(ws=3)				Label
				Target
History1 History2 History3				
Train	t1	t2	t3	t4
	t2	t3	t4	t5
	t3	t4	t5	t6
Test	t4	t5	t6	t7
	t5	t6	t7	t8
Predict	t6	t7	t8	unknown

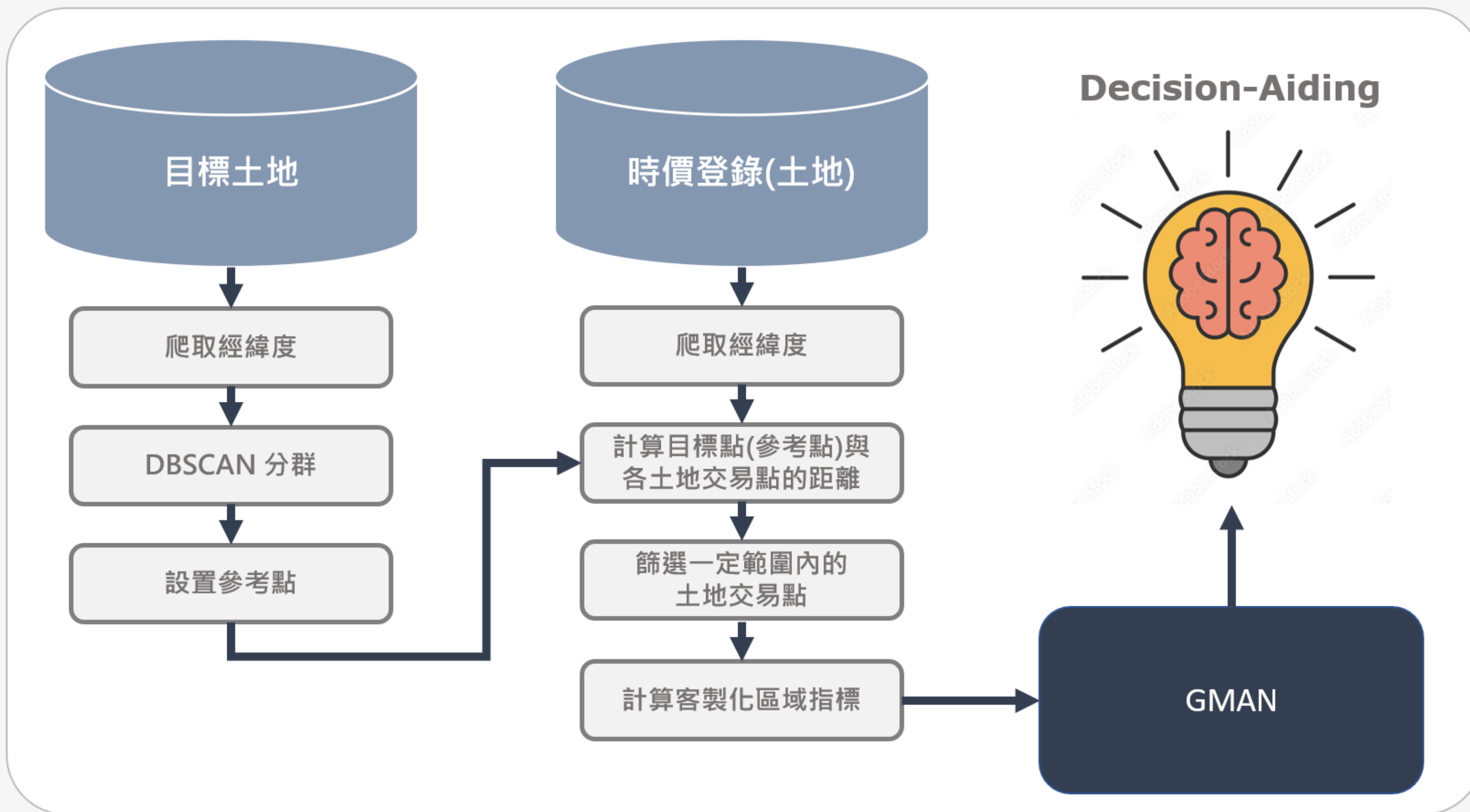
訓練資料格式 (window_size=3)

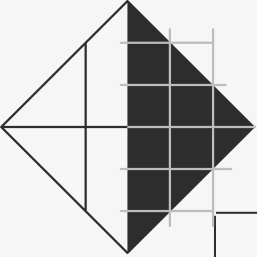
建模預測



訓練過程與調整設計 (window_size = 3)

方法 - Overview





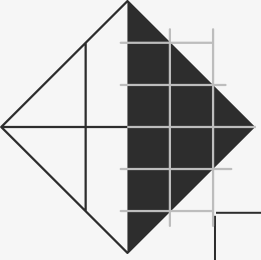
模型指標 - 各群的預測表現

每月單位土地價

Group	MAE	RMSE	MAPE
0	8559	11221	0.1200
1	13177	15033	0.1714
2	4725	6197	0.1551
3	12471	15810	0.1817
4	36140	43242	0.5367
5	36388	37937	0.6182
6	18018	20783	0.3152
7	4718	6186	0.2148
8	3126	4391	0.1766
9	27635	32195	0.4216
10	8805	10027	0.3290
		avg.	0.29

每月交易量

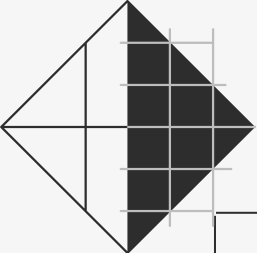
Group	MAE	RMSE	MAPE
0	16.3	22.4	0.4673
1	16.3	20.5	0.4492
2	8.6	10.5	0.4751
3	10.3	12.3	0.3869
4	8.3	10.2	0.7115
5	7.5	8.9	1.1352
6	4.3	4.4	0.7005
7	7.5	9.5	0.4964
8	7.5	8.6	0.4324
9	10.8	13.2	1.8828
10	7.9	10.5	0.6284
		avg.	0.706



模型指標 - Multi-Step Time Series Forecasting

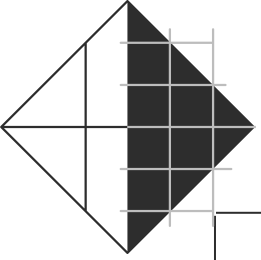
Group1, 3 step 預測結果

	MAE	RMSE	MAPE
train	4912.06	6281.67	7.18%
val	12454.65	13950.57	16.33%
test	14835.74	16654.49	19.32%
performance in each prediction step			
step: 01	15731.80	16938.54	20.56%
step: 02	14642.03	16327.90	19.21%
step: 03	14133.37	16691.37	18.19%
average:	14835.74	16652.60	19.32%



模型成果展示 - 以桃園區、新屋區為例

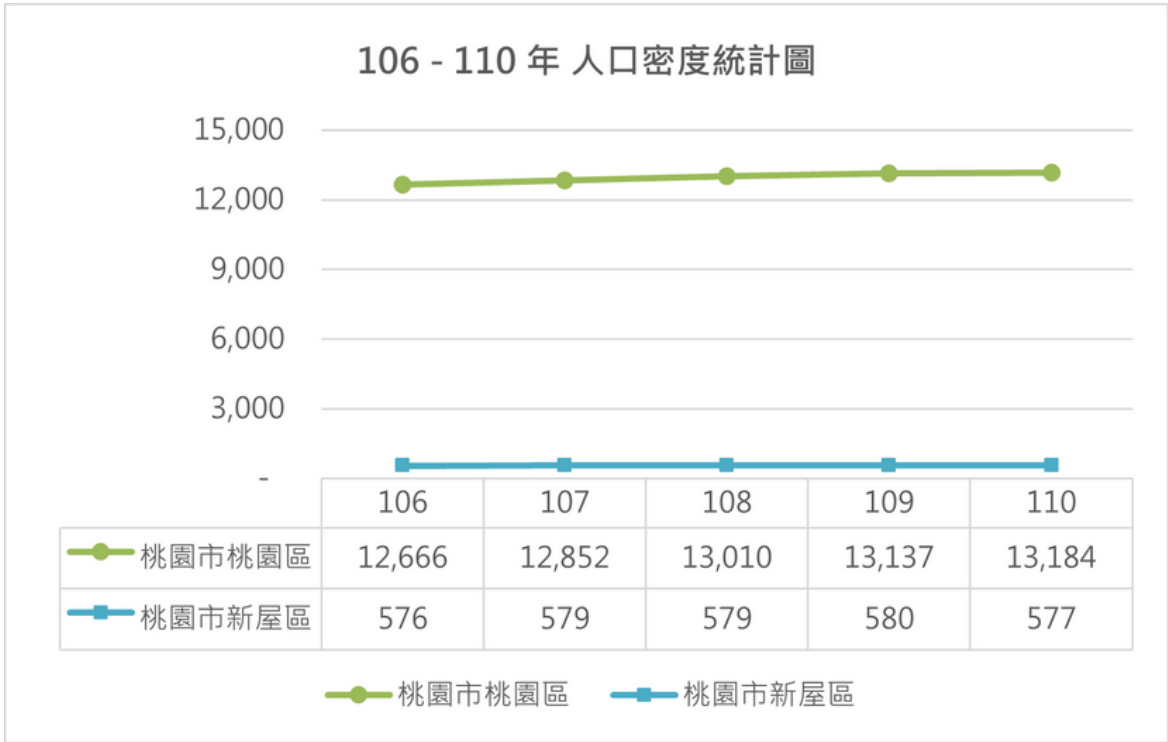
縣市	重劃區名稱	鄉鎮市區	地段	地號	土地面積(m2)	使用分區	DBSCAN
桃園市	中路重劃區	桃園區	中路二段	101	4316.15	住宅區	0
桃園市	中路重劃區	桃園區	中路二段	103	1532.84	住宅區	0
桃園市	中正五街重劃區	桃園區	長安段	1	1126	住宅區	1
桃園市	中正五街重劃區	桃園區	長安段	10	121	住宅區	1
桃園市	中正五街重劃區	桃園區	長安段	13	890	住宅區	1
桃園市	中正五街重劃區	桃園區	埔子段埔子小段	1913-1	740	住宅區	1
桃園市	中正五街重劃區	桃園區	埔子段埔子小段	1913-199	46	住宅區	1
桃園市	中正五街重劃區	桃園區	埔子段埔子小段	1913-200	108	住宅區	1
桃園市	中正五街重劃區	桃園區	埔子段埔子小段	1913-151	178	住宅區	1
桃園市	中正五街重劃區	桃園區	埔子段埔子小段	1913-177	128	道路	1
桃園市	新屋中華市地重劃區	新屋區	中華段	353	4296.47	農業區、住宅區	7
桃園市	新屋中華市地重劃區	新屋區	中華段	363	1210.07	住宅區	7
桃園市	新屋中華市地重劃區	新屋區	中華段	364	2249.62	住宅區	7
桃園市	新屋中華市地重劃區	新屋區	中華段	388	1536.94	住宅區	7
桃園市	新屋中華市地重劃區	新屋區	中華段	389	1909.66	住宅區	7



成果展示 - 人口密度 以桃園區、新屋區為例

- 根據右圖中的資料，可以得知:

	類型	人口密度	人口變化
桃園區	都市	高	不穩定
新屋區	鄉村	低	穩定



結合每月單位土地價預測趨勢

桃園區

- 當價格趨勢平穩 or 上升: 可以考慮較靈活的方案像是短期租約、停車場等設施。等待更好的時機做操作。
- 當價格波動大 or 下降: 可以選擇長期租約或外包，減少風險。

新屋區

- 價格平穩地區: 以長期發展的項目為主，像是發展觀光建設。
- 當價格波動大地區: 該區域可以做動態租約調整。像是趨勢為上升短期出租為主，而趨勢下降則以長期出租為準



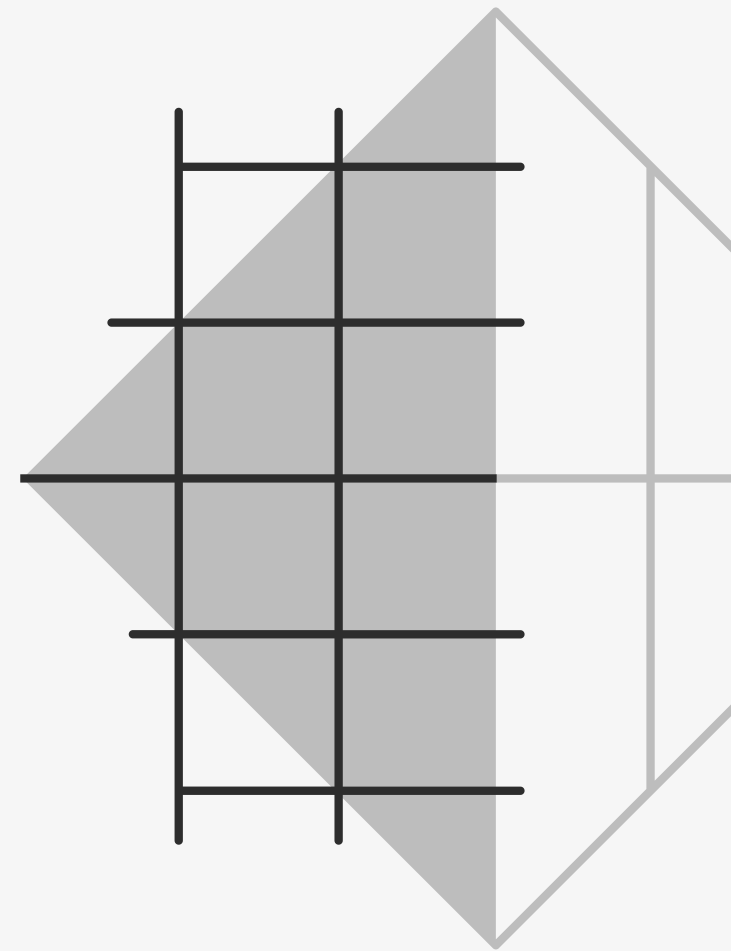
研究限制與未來展望

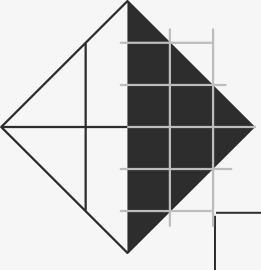
研究限制

- 資料量不足(每月來看10年只有117筆)
- 區域指標資料不足(區域範圍、時間跨度限制)
- 某些目標區的預測結果較差，其預測目標本質上較為複雜。

未來展望

- 模型優化、預測目標的調整和增加模型泛化能力。
- 指標優化(是否可找到更專業一點的判斷依據)。
- 將模型實際運用到產業、決策中，進行更實務面的應用。





Thank you for listening

2022 Oct 14

基於實價登錄資料之土地價值預測

