

大數據分析方案說明書

一、基本資料

作品名稱：i 郵箱智慧化郵政佈點的創新與優化

團隊名稱：「科」顆顆小隊

指導教授：葉向原 教授

成員：曾芷葳、賴士瑜、藍言馨、余瑄瑤

二、動機與目的

目前郵政系統面臨許多挑戰與機遇，特別是在數位化浪潮的影響下，傳統郵件的需求逐漸減少，網路購物的普及為郵政服務帶來了新的挑戰，消費者愈加依賴網路平台來滿足其購物需求，這不僅影響了寄件的需求，也改變了物流市場的格局。隨著消費者對便捷和快速配送的期待提高，傳統郵政服務逐漸受到新興物流平台的競爭威脅，需要滿足消費者對於配送時間和服務品質的需求。在這樣的背景下，傳統郵政的服務多樣化不足的問題日益凸顯，相比於現代物流和快遞公司，郵政業務主要依賴信件和包裹的運輸，這使得郵政服務在消費者心中的吸引力降低，顯得功能單一，無法有效滿足現代消費者的需求，這幾年來開始逐步推廣 i 郵箱，但相較於市場上的物流寄送，仍然不是使用者的第一首選。

便利商店不僅是寄件和收件的場所，還同時提供各種生活必需品，包括食品、飲料、日用品等，消費者可以在一個地點解決多項需求。此外，便利商店通常具備 24 小時營業的優勢，方便消費者在任何時候進行購物或取件，並且經常提供促銷活動以吸引顧客。便利商店提供的綜合服務，不僅滿足了消費者對便捷性和多樣性的需求，還強化了消費者在同一個地點的購物體驗。相比之下，i 郵箱雖然是 24 小時自助服務，但取貨設置在便利性上不及超商和蝦皮店到店，便利商店通常位於人流量大的地點，並且提供更便捷的取貨服務，包括與其他購物需求的整合，讓消費者可以一站式完成多項需求。而蝦皮的取貨機制則能夠結合網路購物的整個流程，消費者從下單到取貨均可在同一平台上完成，這使得購物體驗更加流暢。

因此要提升 i 郵箱的市場競爭力，除了強化其寄件和收件的基本功能外，還需要思考如何增加服務機會以及優化使用體驗，從而提升整體的吸引力和實用性。因此，我們的目標在於探索 i 郵箱佈點的創新與優化，將是提升郵政服務競爭力的重要途徑，以提升其市場吸引力。我們將融合公開資料並透過圖資分析的

方式來解決 i 郵箱的佈點問題，，利用環境周遭指標從而掌握 i 郵箱目前的運作狀況及其潛在的使用需求，同時也運用機器學習模型來預測 i 郵箱在這些不同位置的使用率，期望能透過有效選址來提升 i 郵箱的使用性，促進其在激烈競爭中的生存和發展。

三、創新解方的整體流程與步驟

為了達到此目標，我們擷取 i 郵箱與蝦皮店到店的地點資料，運用多元外部資料對 i 郵箱、蝦皮店到店進行比較，並進行 GIS 環域分析來監測其周遭的環境設施，將其劃分為不同指標，包括醫療設施、便利商店、教育資源、交通運輸等，採用圖資套疊分析技術以視覺化方式呈現資料結果，並透過機學習的回歸模型來預測使用率並分析點位周遭狀況間的關聯，此創新方法將幫助快速識別出最適合設立新 i 郵箱的位置或是撤銷使用不佳的 i 郵箱位址。這種以數據為基礎的決策過程能有效減少傳統選址方法中主觀判斷的偏差找到最佳的佈點策略，從而找到最佳的佈點策。此創新方法也建立一個動態的監控系統，隨著環境的變化及需求的演變，適時調整 i 郵箱的佈點策略，保持其在市場上物流市場的競爭力，持續適應並滿足消費者的需求，期待能夠有效提升其市場佔有率和使用率。執流程與架構如圖一。



圖一：流程架構圖

(1) 蝦皮店到店與 i 郵箱的位址與其區域數目

首先，我們擷取蝦皮購物官方網頁之服務據點公告中的服務地區說明，將該段文字內容解析出所有服務據點的城市名稱及其對應的區域，透過 Google Maps API 尋找每個服務據點區域方圓 10,000 公尺內的蝦皮店到店點位資訊，去除重複的內容整理匯出為 csv 檔案，方便後續處理及作業。我們採用了 Pandas 和 Geopandas 資料處理套件，Pandas 提供了強大的資料操作

和分析功能，使我們能夠輕鬆處理大型數據集，包括數據清洗、篩選和聚合等。Geopandas 則允許我們根據一個或多個地理特徵(例如位置、區域、社區等)對資料進行分組，然後應用各種聚合函數，以便深入分析和理解數據的內在結構和分布特徵。將蝦皮店到店點位資料與〈提供資料-1-1. i 郵箱資訊.csv〉欄位名稱統一後，利用 Geopandas 將坐標資料轉換為地理資料格式 GeoDataFrame，從檔案中讀取〈台灣里界圖_TWD97〉，將其轉為相同的座標系統，透過空間連接功能，將蝦皮店到店點位資料與台灣地圖資料合併，為每個 i 郵箱和蝦皮店點加上對應行政區的多邊形幾何，視覺化處理蝦皮店點和 i 郵箱在各區的分布情況，以利分析其佈點的比例原則。這樣的方式可以明確識別出目前佈點的集中或不足情況。這種評估可以幫助我們發現哪些地區的 i 郵箱使用頻率過高，可能導致過度擁擠，以及哪些地區的使用率低，甚至可能需要考慮廢除或重新佈點。讓後續的資源分配提供重要依據，確保每個 i 郵箱的設立都能最大限度地滿足消費者的需求。

(2) i 郵箱位址的環域分析

環域分析是一種常見的地理空間分析技術，透過設置特定的距離範圍，我們能夠評估某一地理現象(如 i 郵箱或蝦皮智取櫃)對其周邊環境的影響。我們明確定義一個影響範圍(250 公尺、500 公尺、750 公尺)，以圓形的形式來界定，從而探討該範圍內的其他特徵(如醫療設施、便利商店、教育資源、交通運輸等)的點之間的關聯性。透過對這些空間資料的深度分析，可以更清楚地了解如何優化 i 郵箱的佈點策略，提升其在市場中的競爭力和使用便利性。我們採用了環域分析的方式來深入了解蝦皮智取櫃的座落位置與周邊特性，並將這些資訊與目前 i 郵箱的特性進行綜合評估，不僅讓我們能夠直觀地看到兩者可能的優勢和劣勢，進而做出更為精準的決策，這一分析將為這些特性能夠幫助我們判斷哪些地點適合增設 i 郵箱，從而補足現有的佈點缺口。

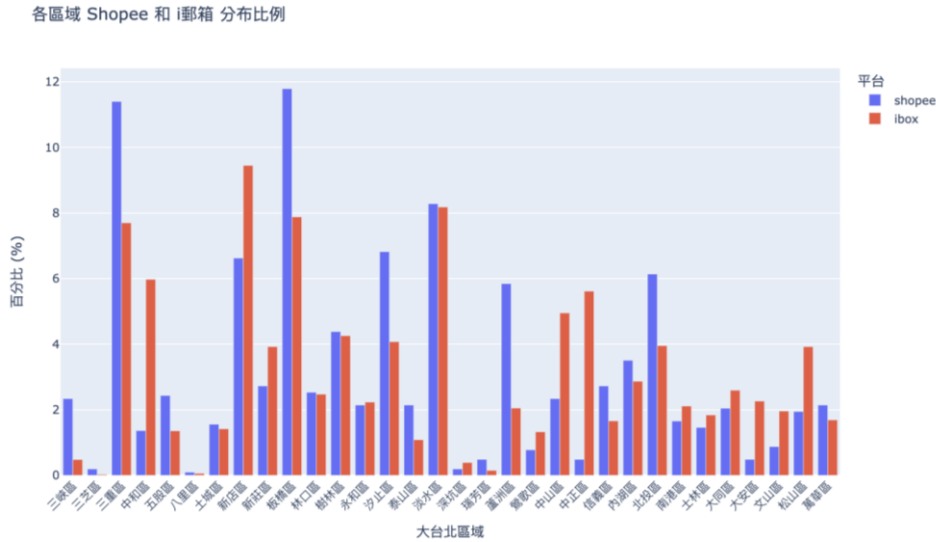
(3) i 郵箱環域分析特徵與使用率模型預測分析

為了將來對於新的選址能夠預測可能的使用率，我們利用目前的 i 郵箱的貨態資訊與環域分析特徵來建立使用率預測模型，從而支持選址決策。我們使用 Scikit-learn 套件，其為一個廣泛使用的機器學習庫，提供多種分類、回歸和聚類算法。我們採用了類神經網路和隨機森林這兩種機器學習模

型進行訓練(80%)與測試(20%)。為了評估這些模型的表現，我們將使用了均方根誤差(RMSE)和命中率(HitRate)作為主要的評估指標。RMSE 是一種衡量預測模型準確度的統計指標，它的計算方式為先求出每個預測值與真實值之間的差異，將這些差異平方後取平均，再取平方根。RMSE 的值越小，表示模型的預測效果越精準。其次，如果模型的預測值在實際值的 10%、20%以及 30% 範圍內，我們就將其視為“命中”，可以直觀地反映模型在一定容忍範圍內的準確性，幫助我們了解模型在實際應用中的有效性。透過這兩項指標的綜合運用，我們能夠全面評估模型的表現，進而調整和優化預測策略。

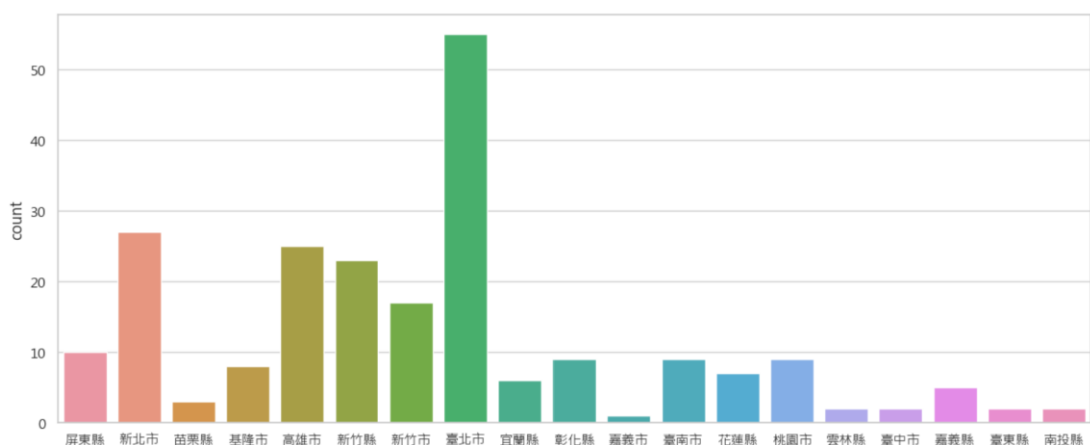
四、創新解方實驗結果：數據可視化

我們採用資料來源包括政府開放資料(DATA.GOV.TW)，其提供各類政府公開數據，包括交通、人口、經濟等多方面的資訊，數據將有助於了解當前的社會經濟環境。社會經濟資料服務平台(SEGIS.MOI.GOV.TW)，該平台提供有關各種社會經濟指標的資料，這對分析 i 郵箱使用情況及周邊特性至關重要。



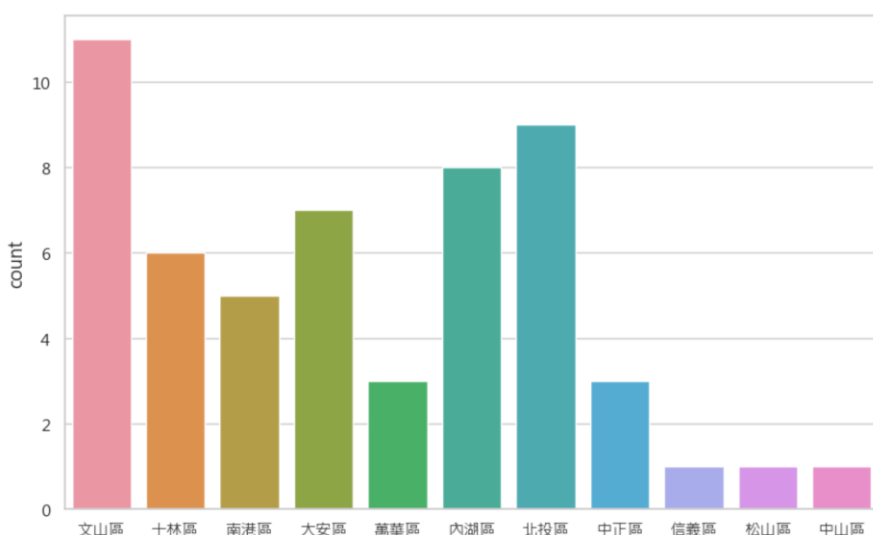
圖二：雙北各區域中蝦皮店到店(藍色)和 i 郵箱(紅色)的分布比例

從圖二可以看出，不同區域的蝦皮店到店和 i 郵箱的分布比例存在顯著差異，像三重區和板橋區，蝦皮店到店的比例遠高於其他地區，而在中和區和新店區，i 郵箱的比例明顯較蝦皮店高。由於蝦皮店到店的選址模式與人口密度相關，因此我們可以以這樣的模式來對照 i 郵箱實際設點狀況，觀察是否有閒置以及供不應求的狀況發生，或是是否符合分配的比例原則，近一步提出增加與刪除點位建議。



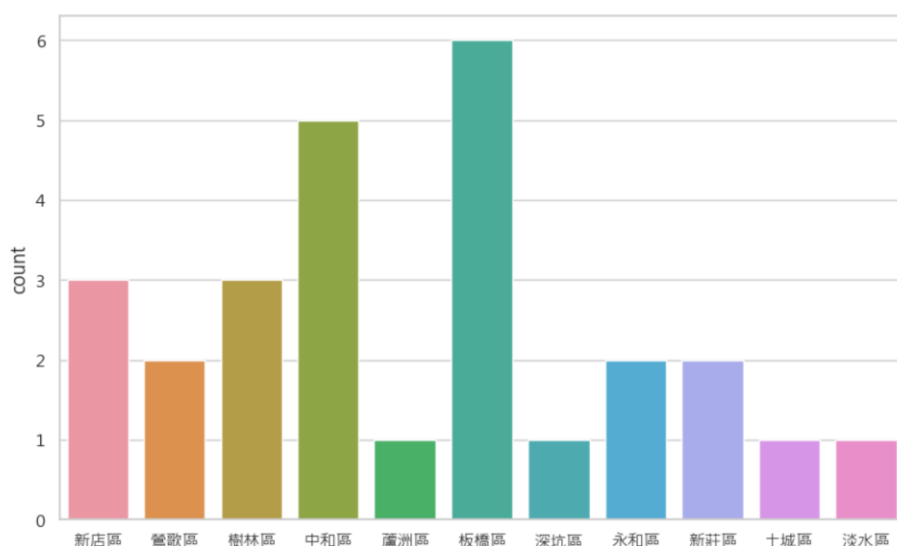
圖三：各縣市使用率低於 50%的 i 郵箱數量直條圖

另一方面，我們也從 i 郵箱的貨態狀況選出使用率低於 50%的縣市，從圖三中可以看出，各縣市的使用率低於 50%的 i 郵箱數量，以台北市和新北市為最高，接下來為高雄市與新竹縣。



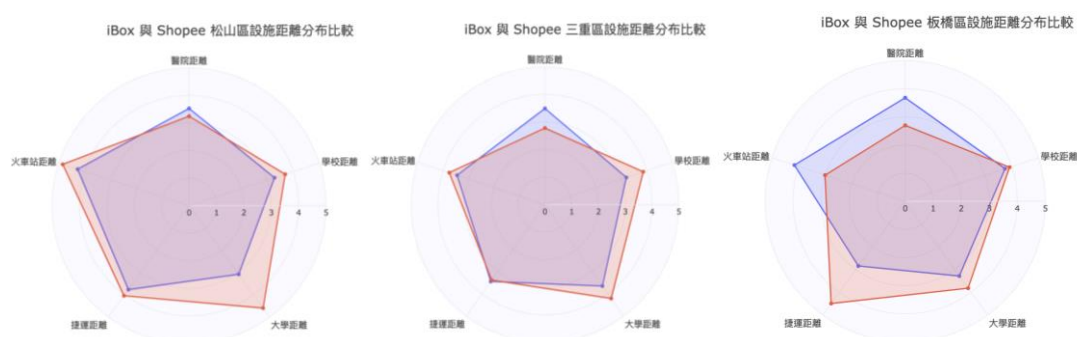
圖四：台北市使用率低於 50%的 i 郵箱數量直條圖

台北市行政區中，使用率低於 50%的 i 郵箱數量以長條圖呈現如圖四。在各區中，文山區的使用率最低，對照圖二可見，文山區的蝦皮點位數量比例少於 i 郵箱，導致閒置有 i 郵箱閒置的問題，且可以考慮將部分據點遷移至其他地區。



圖五：新北市使用率低於 50% 的 i 郵箱數量直條圖

新北市的板橋區、中和區、樹林區、新店區等地區，i 郵箱的使用量較低，其中新店區以及中和區對照圖二，i 郵箱數量亦遠遠多於蝦皮店到店的數量比例，可知此兩處可能更適合減少 i 郵箱的設置數量以避免資源浪費。



圖六：i 郵箱、蝦皮店到店其周邊設施距離比較雷達圖

從圖六由左至右分別為松山區、三重區、板橋區表達周邊設施距離比較雷達圖，紅色為蝦皮店到店與相關設施之距離，藍色為 i 郵箱與相關設施之距離，從十二點鐘方向順時針的圖例為醫院距離、學校距離、大學距離、捷運距離、火車站距離。於此圖六中我們可以看見蝦皮的選址策略與 i 郵箱不同處，為其離醫院距離較近，學校、大學、捷運距離較遠，松山區及三重區的蝦皮店到店離火車站距離較遠。



圖七：i 郵箱、蝦皮店到店與周邊設施數量比較雷達圖(II)

從圖七由左至右分別為松山區、三重區、板橋區；紅色為蝦皮店到店與周邊設施數量，藍色為 i 郵箱與周邊設施之數量。從十二點鐘方向順時針的圖例為速食店數量、公車站數量、便利商店數量、診所數量、圖書館數量。於此圖七中我們可以看見蝦皮店到店的周遭速食店及診所的數量較多。

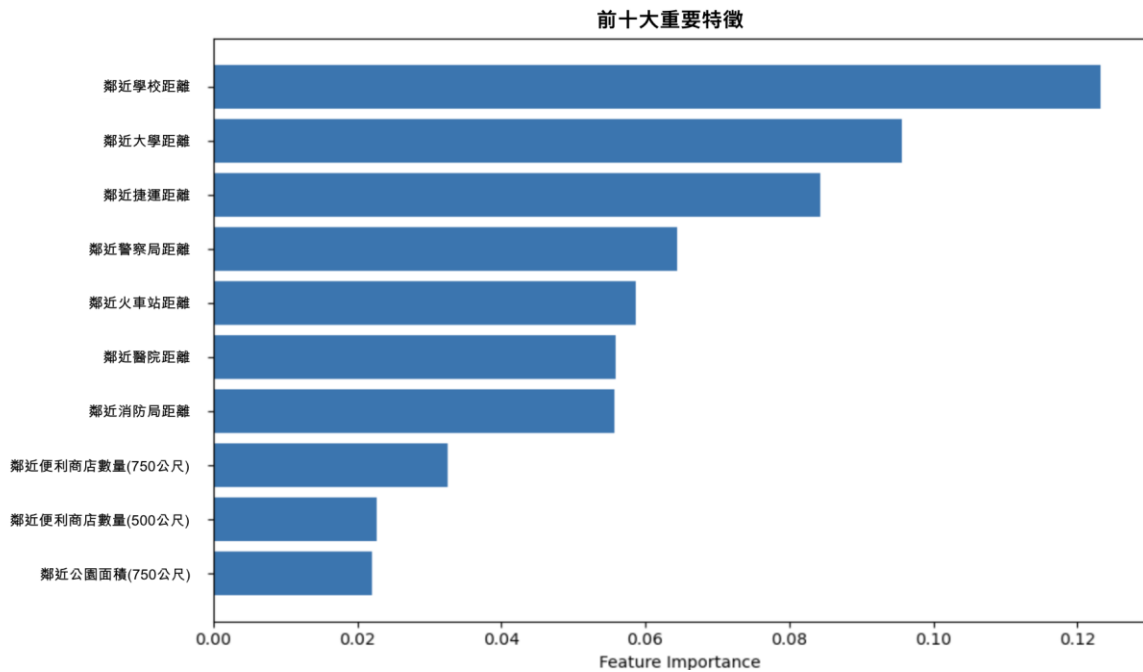
除了之周遭的環域分析外，我們更針對環域特徵進而來推估使用率，因此我們以環域圖資特徵當作輸入，以 i 郵箱的貨態使用率當作輸出，去建構一個回歸預測模型，以評估周邊設施數量及距離對於 i 郵箱使用量的關係。我們使用類神經網路、隨機森林進行訓練與測試，以 RMSE 及 HitRate 作為判定模型好壞的評估指標，RMSE 的結果為 14.9590，預測值落在實際值正負 10% 的比例為 45.61%，正負 20% 者為 75.73%，正負 30% 者為 85.36%，如圖八所表示。

```

=== RandomForest 回歸模型評估結果 ===
RMSE: 14.9590
預測值落在實際值±10%內的比例: 45.61%
預測值落在實際值±20%內的比例: 75.73%
預測值落在實際值±30%內的比例: 85.36%
  
```

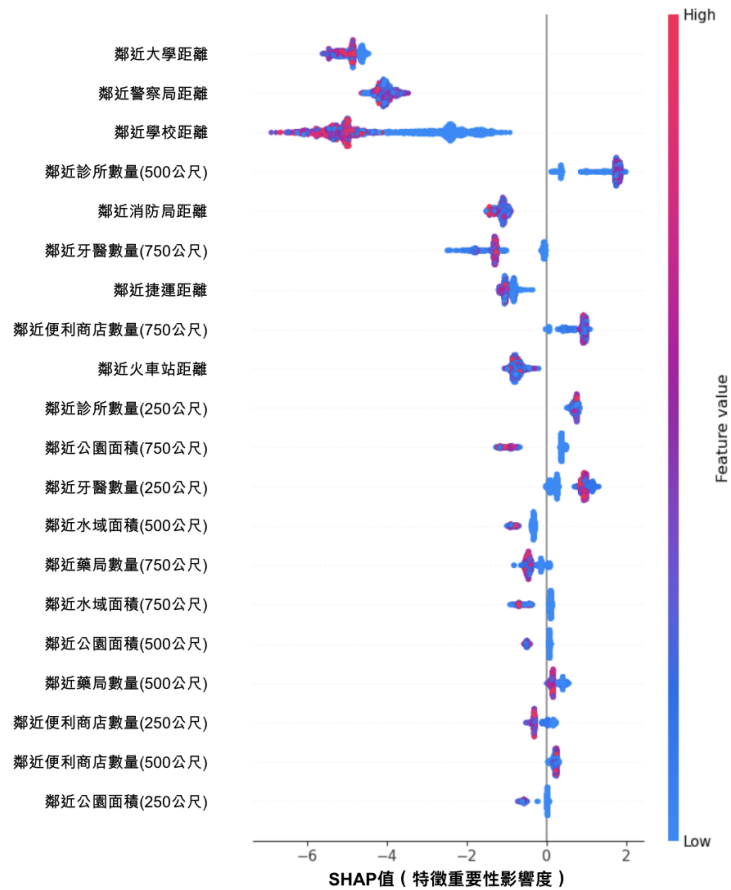
圖八：回歸模型評估結果

在預測模型方面，圖九顯示出各特徵的重要性排序，數值越大表示該特徵對模型預測的影響越大。圖中鄰近學校距離和鄰近大學距離為影響力最大的兩個特徵，代表這些距離變量對模型預測具特徵顯著性，再來則以交通與醫療院所具相關性。



圖九：重要度分析橫條圖

為了瞭解每個特徵的正向與負向影響，我們利用 SHAP 值表示每個特徵對模型預測結果的貢獻方向與程度。SHAP 值越高，代表該特徵對模型預測的正向影響越強，SHAP 值越低則表示負向影響。圖十中每個點代表一筆數據，根據數據的特徵值以顏色呈現，高特徵值為紅色，低特徵值為藍色整體上。其中鄰近大學距離、鄰近學校距離該特徵的高低值對模型接有負向影響；而鄰近牙醫與診所數量、鄰近便利商店數量(750 公尺)的數據則集中在右側，且該特徵的高值對模型有正向影響。對於便利商店數量(250 公尺)越多，越具有負向影響。



圖十：重要度分析 SHAP

五、創新解方實驗結果與分析

我們透過以上的實驗比較與分析，提出以下幾點進行深入探討：

問題一：i 郵箱設置在大學或學校附近可能不如設置在診所或牙醫附近具有優勢。

1. 郵件需求量和頻率不同

學校附近的使用者以學生為主，他們的郵寄需求多偏向日常生活中的小件收發，且通常依賴便利商店進行寄取，既方便又可搭配購物需求。相較之下，診所或牙醫診所附近的客群通常為來自周邊社區的居民，這些人可能會定期有穩定的郵寄需求，且對 i 郵箱的使用頻率較高。

2. 便捷性和即時性需求差異

在學校附近，由於學生日常活動集中在校園內，出入頻率不固定，對郵寄需求的即時性和便捷性要求不高，以至於 i 郵箱的設置難以達到預期效果。診所和牙醫診所附近的成人客群通常具備穩定的作息，且診所附近的民眾對文

件或快遞的接收需求可能具有較高的即時性，因此 i 郵箱在這些地區的設置更能滿足客戶的實際需求。

3. 穩定客群特徵

學校區域的學生流動性較高，畢業或實習期會導致常駐人數波動，而診所和牙醫診所周邊的居民通常為社區固定人口，構成相對穩定的客群，這種穩定性有助於維持 i 郵箱的使用率。

問題二：為什麼 i 郵箱不要設便利商店附近？i 郵箱設置在便利商店附近可能效果不佳？

1. 競爭壓力大，便利商店服務覆蓋全面

便利商店擁有強大的寄取件服務網，且通常與多家物流公司合作，提供包裹寄件和取件、付款、加值等多樣服務。這些店面通常距離住宅區、商業區較近，且二十四小時營業，使用率極高。i 郵箱設在這些地方，容易因競爭激烈而難以吸引客戶，因為多數民眾會選擇更為便捷的便利商店作為收發件首選。

2. 缺乏全天候服務和人員協助

便利商店能提供全天候服務，且有店員在場協助寄取件程序，特別適合有需要即時協助的顧客。而 i 郵箱多數沒有專人服務，僅能依賴機器自助操作，因此使用者的體驗會較受限，尤其對不熟悉操作的人來說，便利商店的服務顯得更具吸引力。

3. 便利商店流量遠高於 i 郵箱

便利商店因兼具購物、繳費、加值等多功能服務，每天吸引大量人流。這些人流同時也會帶動寄取件的需求，形成自然的高使用率。而 i 郵箱若設在便利商店旁，因便利性較弱和缺乏附加功能，較難吸引人群前往使用，可能導致使用率偏低。

4. 服務費用相近但便利性不足

多數便利商店提供的寄取件服務收費與 i 郵箱相差不大，甚至有時會因為合作活動而更優惠。當服務費用差異不明顯時，民眾更傾向選擇服務內容豐富且更便捷的便利商店，而非僅提供簡單郵寄功能的 i 郵箱。

問題三：為什麼 i 郵箱不要設在位於交通要道的建築裡面？須顧及便利性也不宜離交通要道太遠？

1. 人流特性與使用需求不符

交通要道的建築內人流量雖然大，但大多數人主要是匆忙穿行或前往其他目的地，停留時間短暫，並非前往郵寄或取件的目標客群。設置 i 郵箱於此處，可能無法吸引穩定的使用者，導致設備閒置。相反地，距離交通要道稍近但非主幹線位置的地點，居民或通勤者會更傾向於利用 i 郵箱寄取件，增加使用率。

2. 便利性與易到達性

設置於交通要道附近的建築外圍，能讓使用者在進出交通樞紐的過程中方便到達，無需走進繁忙的建築物內。這樣的選址更符合寄取件需求者的便利性期待，讓他們能夠快速完成操作後立即前往其他地點。與交通要道內部相比，這樣的位置可以更好地滿足使用者的快速、便捷需求。

3. 隱私性和使用體驗更佳

交通要道建築內部通常人流密集且較開放，這會降低用戶的隱私性。郵件寄取過程中包含個人資訊，使用者可能會更偏好在相對隱蔽的地方進行操作。選擇距離交通要道較近的戶外或商業區域位置，能夠提升隱私性，提供更良好的使用體驗，吸引穩定客群。

4. 使用環境的舒適度

交通要道內部建築往往人潮擁擠且環境較吵雜，這樣的環境容易讓寄取件的操作變得不便或壓力增大，甚至會影響人們對 i 郵箱的使用意願。在交通要道旁的相對安靜區域設置 i 郵箱，則可以提高使用者的舒適度，讓寄取件過程更輕鬆自在。

六、創新解方預測效益與可行性

我們的創新解方透過整合公開數據資源並透過 GIS 環域分析，結合相關環境指標進行評估，不僅有助於更全面地瞭解選址周遭，還能降低佈點風險，進而進行精準的佈點決策，提高運營的可持續性。對蝦皮智取櫃和 i 郵箱的環域分析，評估目前的佈點情況，確定哪些地區存在取貨需求不足或過於集中的問題，進而做出針對性的調整與優化，我們可以更有效地識別出使用率不佳的 i 郵箱，並評估其是否適合轉換位置或進行功能調整，以提升其服務效益。本創新思維充分利

用政府的公開資料以及蝦皮店到店的佈點策略作為參考，除了相互視覺化比較外，更利用建構模型來因應未來的選址使用率預測，達到更完善的功能與落地實踐。此舉不僅有助於優化現有的選址策略，還能確保未來選址更符合實際需求，進一步強化選址決策的重要性並提高 i 郵箱的使用效率。

七、結論與建議

i 郵箱的選址需找出自身的差異化路線和優勢，才能達成最大效益。以下是我們的建議：

1. 人群結構至關重要選址前應調查周邊人群密度，以確定 i 郵箱的輻射範圍和目標客群的數量，人口密度越高則越有利於發展。如果 i 郵箱設置於人流較少但常住居民多的區域，可以累積穩定的客群，提供長期便捷的服務。此外，因此選址時可優先考量商業區或住宅區，這些地區往往能吸引較多的穩定使用者。
2. 競爭情況影響選址效果，考量當地其他類似服務的數量和分布十分重要。在競爭對手眾多的區域，i 郵箱可能較難取得優勢，建議避免直接與大量超商據點競爭。
3. 環境因素為考量，選址時距離和便利性是影響使用率的因素，建議優先考量靠近公車站、捷運站或火車站等公共交通設施，以方便民眾前往。不過，i 郵箱的設置若位於交通要道內，可能因人流過於匆忙且缺乏隱私而不適合。另一方面，鄰近學校的區域也未必為理想選址，因學生的郵寄需求有限且偏向非商業性的頻繁使用。診所與牙醫的坐落更可能具有優勢。因此，應尋找具有適度人流、相對私密且交通便利的地點。

八、資料來源

1. 政府開放資料(<https://data.gov.tw/>)
2. 社會經濟資料服務平台(<https://segis.moi.gov.tw/>)
3. 內政部實價登錄網站(<https://lvr.land.moi.gov.tw/>)
4. 蝦皮購物官方網頁服務據點公告(<https://shopee.tw/m/spxservicepoint>)