

外送平台訂單配送延遲因素分析

製造數據科學 第 N 組

卓庭榛 葉光筠 王皓 孫旖旎

1. Title

外送平台訂單配送延遲因素分析

2. Background and Motivation

2.1 Motivation:

- 外送平台的準時率會直接影響使用者體驗與平台評價，若能預測並改善配送延遲問題，將能有效提升營運效率與顧客滿意度。

2.2 Background:

- 隨著外送平台的擴張，配送準時性成為服務品質的關鍵指標。然而，配送過程中受到多變因影響，如天氣、交通、餐廳備餐時間等，導致準時配送成為一大挑戰。

2.3 Problem Definition:

- 本研究運用機器學習技術預測配送延遲事件，並針對導致延遲的潛在因素進行探討，提出改善策略以供外送平台參考。

3. Methodology

3.1 分析流程(研究框架)

- 資料描述與來源確認: 確認變數含義、瞭解資料結構與特性。
- 資料前處理: 進行缺失值補值、格式統一、距離計算與資料標準化等。
- 特徵工程與標籤定義: 以 Mapbox Api 預估時間推算延遲標籤、時段分類。
- 模型建立與比較: 使用邏輯回歸、Random Forest、XGBoost 進行延遲預測。
- 模型解釋與重要特徵分析: 透過 SHAP 解釋模型與變數關係。
- 洞察與決策建議: 整合模型結果提出平台優化建議。

3.2 分類模型選擇與理由:

我們總共用了三種不同分類模型來預測訂單是否會延遲(延遲定義:實際配送時間>估算配送時間), 選擇理由如下

- Baseline model 為 Logistic Regression, 是解釋性強、可得到關係方程式的模型
- 為提高預測精準度, 我們選集成學習模型當中的兩個樹狀模型: Random Forest / XGBoost, 作為我們的分類模型
 - Random Forest: 屬於Bagging方法的模型, 每棵 decision tree 以隨機樣本、特徵訓練而成, 理論上比起線性模型能提高準確度, 捕捉變項之間的交互作用與非線性關係。
 - XGBoost: 屬於 Boosting 方法, 每一次會提升前一次分類錯誤樣本(如:實際上有延遲但預測為準時、實際上準時但預測為延遲)的權重, 能夠有效將低模型的bias, 達到較高的精確度
- 特別討論 沒用時空模型的原因:

在期中提案時老師提到可以試試進階的時空模型 (Spacial Temporal Model), 在預測外送延遲或時間會更加準確, 但我們研究過時空模型後, 發現這筆資料集雖然有時間日期, 本質上卻不是時間序列。每一筆資料都是一個獨立發生的配送事件, 為某一天某一個時刻, 某位配送員在某個天氣、交通情況下接單、取餐並完成配送。意思是這些訂單之間彼此獨立, 沒有前後的因果關係或是時間的連貫性。

時間序列資料, 要具備時間上的依賴性, 就是過去的狀態會影響未來的結果。比如一位配送員今天配送的情況, 會因為昨天連續高工時而變慢; 或者一個城市的外送需求, 會因為連續幾天的天氣變化而出現趨勢。這些才是時間序列要處理的問題。資料需要按時間排列, 觀測值之間有順序關係和演化過程。

但這份資料沒有配送員的連續配送紀錄, 也沒有區域隨時間變化的外送趨勢。每一筆資料是孤立的, 今天的訂單不會影響明天, 也不會影響同一時刻其他地區的配送。所以即使硬把這份資料丟進時空模型, 例如 TFT, 模型也無法學習時間上的連續模式。

與其用複雜的時空模型增加運算負擔可能因為模型設計錯誤而導致效能下降, 我們判斷這份資料集更適合用其他的模型

4. Data Collection and Analysis Result

4.1 資料來源與內容

本研究使用來自 Kaggle 平台的 Zomato Delivery Operations Analytics Dataset, 共包含 45,584 筆外送訂單資料與 20 個變數, 涵蓋配送員背景、訂單時間與地理位置、天氣與交通等欄位。

下表為變數名稱及說明整理:

變數名稱	說明	變數名稱	說明
ID	每筆配送的唯一編號	Festival	是否為節慶日
Delivery_person_ID	配送員的唯一編號	City	訂單所在城市
Delivery_person_Age	配送員年齡	Time_taken (min)	配送所花時間(分鐘)
Delivery_person_Ratings	配送員平均評分	Type_of_order	訂單類型(外帶、自取)
Restaurant_latitude	餐廳位置(緯度)	Type_of_vehicle	配送車輛種類
Restaurant_longitude	餐廳位置(經度)	Multiple_deliveries	是否同時配送多筆訂單
Delivery_location_latitude	顧客位置(緯度)	Order_Date	訂單日期
Delivery_location_longitude	顧客位置(經度)	Time_Ordered	訂單下單時間
Weather_conditions	天氣狀況	Time_Order_picked	騎手取餐時間
Road_traffic_density	路況壅塞程度	Vehicle_condition	配送車輛狀況

備註:Vehicle_condition 變數意義不明, 經討論後於資料前處理階段中排除分析。

4.2 分析結果

1. 資料前處理與特徵工程

a. 資料清洗與缺失補值

欄位刪除與資料修正：

- 移除無分析意義之欄位:ID、Delivery_person_ID、Vehicle_condition。
- 已確認資料來源地區為印度, 修正經緯度為負值的異常情況, 並篩除落於合理經緯度範圍外的資料點。

缺失值補值策略：

- Delivery_person_Ratings、Delivery_person_Age:依據(城市 + 車種)群組計算群組平均或中位數補值。

- ii. Time_Orderd、Time_Order_picked:補齊 Excel 格式轉換錯誤, 無法修復者則估算平均備餐時間回推。
- iii. City:使用 KD-Tree 以地理位置比對最近鄰點補值。
- iv. Weather_conditions 與 Road_traffic_density:以時間 ± 1 小時、距離最近前 5 筆資料中的眾數補值, 結合時間與空間資訊進行推論。

資料篩選與處理:

- 移除 Festival = Yes 或為缺值的資料, 因該類型樣本過少。
- 移除 multiple_deliveries 欄位為缺值之觀測值, 避免無法判定配送特性。

b. 時間與空間特徵工程

時間處理與特徵生成:

- i. 建立 order_datetime 作為時間戳記欄位, 結合 Order_Date 與 Time_Orderd。
- ii. 擷取 hour_ordered(下單小時)、hour_picked(取餐小時)與 day_of_week(星期幾)作為分類特徵。
- iii. 自訂時段分類(time_segment):
 - morning_noon:08–11點
 - afternoon:12–16點
 - evening:17點後
 - 其他時間標記為 other、缺失標記為 unknown

延遲標籤定義:

- i. 根據實際配送時間與估算配送時間(使用 Mapbox API 計算)比較, 若超出預估時間則標記為延遲(delayed = 1), 否則為 0。

欄位型態分類:

- i. 數值型特徵如:年齡、評分、距離等。
- ii. 類別型特徵如:天氣、交通、城市、車種、時段等。

2. 模型訓練

a. 模型準確率與資料量分析(程式碼可見:學習曲線.ipynb)

Logistic Regression

我們首先使用了 Logistic Regression 作為基準模型進行分類預測。由於其屬於線性分類器,可透過統計學習理論中的 Vapnik–Chervonenkis 維度來推估模型的複雜度與所需樣本數量。

對於線性分類模型, VC 維度約可視為其輸入特徵數 d 的線性函數。在本專案中,資料經過前處理(數值標準化、類別 OneHotEncoding、時間特徵工程)後,總共包含 84 個特徵欄位,因此估計 VC 維度約為 84。

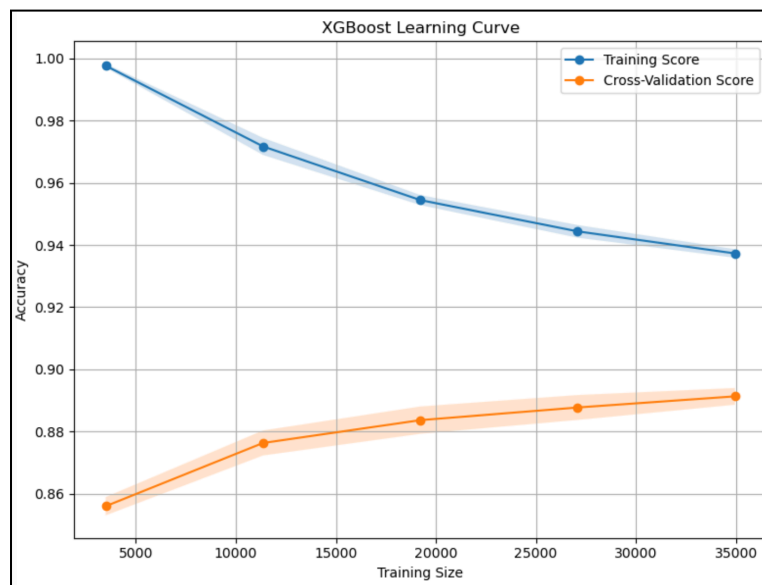
根據學習理論,若希望模型在泛化時的誤差不超過 $\epsilon = 0.05$, 且信心水準達 $1 - \delta = 95\%$, 則理論上至少需滿足以下樣本數條件:

$$n \geq 1/\epsilon(4 \cdot VC \cdot \log(2e/\epsilon) + \log(4/\delta))$$

將參數帶入 ($VC = 84$, $\epsilon = 0.05$, $\delta = 0.05$) 得出所需樣本數 $n \approx 31,597$ 筆。而本資料集經前處理過後,共計 38,835 筆樣本,高於理論下限,顯示模型訓練資料充足,足以支撐有效學習,亦符合統計學習理論對於模型泛化性的需求。

XGBoost

XGBoost 模型在不同訓練樣本數下的表現如下圖:



- Training Score(藍線)

在樣本數較少時(如 4000 筆), 準確率接近 100%, 顯示模型具高度擬合能力, 容易記住少量資料 → 存在明顯過擬合風險。但隨著訓練資料增加, 訓練準確率逐漸下降, 約降至 94% 左右, 表示模型逐漸轉向泛化學習, 不再死記資料。

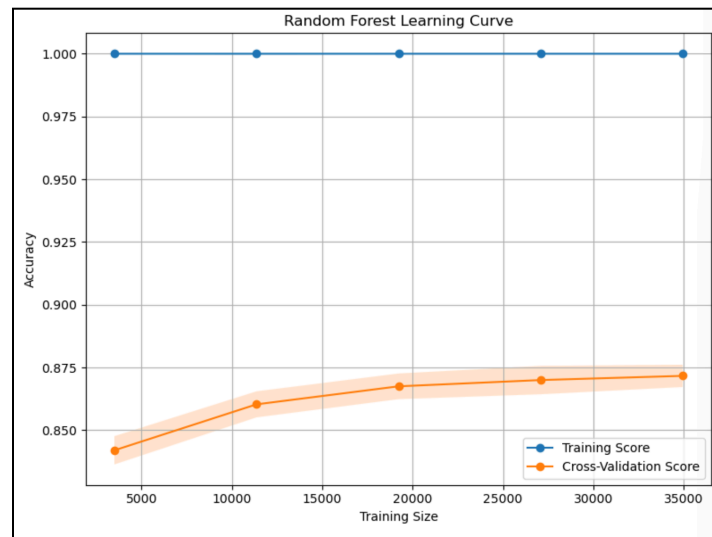
- Cross-Validation Score(橘線)

起始約為 85.6% 隨著資料量增加穩定提升至 88.6%, 收斂趨勢明顯, 且變異範圍逐漸縮小, 顯示模型學習穩定、泛化能力提升。

從學習曲線觀察可見, XGBoost 模型隨著訓練資料增加, 訓練準確率略降, 交叉驗證準確率穩定上升, 兩者間差距逐漸縮小, 反映出模型在 Bias-Variance 之間取得良好平衡, 具備穩定且可信的泛化能力。

Random Forest

Random Forest 模型在不同訓練樣本數下的表現如下圖:



- Training Score(藍線)

模型在所有訓練樣本規模下皆達近乎 100% 準確率, 顯示其對訓練資料具極高擬合能力。然而, 這也代表模型可能過度記憶訓練集特徵, 未能有效學習資料中的普遍規律。

- Cross-Validation Score(橘線)

模型在驗證集上的準確率由 84.2% 穩定上升至 87.2%, 雖呈現學習趨勢, 但始終與訓練準確率存在明顯落差。

從學習曲線中可見, 此現象屬於典型高變異(high variance)表現, 顯示模型對訓練資料過擬合, 泛化能力不足。可能原因包含:

- 森林中樹木過深、參數未受限制
- 類別特徵經 One Hot Encoding 導致特徵維度上升
- 資料集可能包含雜訊或不具代表性分布

4.3 解釋與洞察 (IMPORTANT)

- 預測效果分析: 各模型表現差異、優缺點比較

模型	Accuracy	Precision	Recall	F1-Score	AUC
Logistic Regression	0.8349	0.7955	0.7468	0.7703	0.9090
Random Forest	0.8730	0.8543	0.7930	0.8225	0.9466
XGBoost	0.8902	0.8621	0.8380	0.8498	0.9611

說明: 從研究結果可見, XGBoost 模型在所有評估指標中皆表現最佳, 特別是在 Recall、F1-Score 與 AUC 等關鍵指標上明顯優於 Logistic Regression 與 Random Forest。這樣的結果與本研究資料特性高度相關。研究中所使用的資料涵蓋多種型別變數與非線性特徵, 例如地理距離、天氣與交通狀況、外送員特質與訂單類型等, 其間存在高度複雜的交互作用與非線性關係。XGBoost 採用梯度提升演算法, 能夠逐步修正前一棵樹的錯誤, 並透過內建的正則化項控制模型複雜度, 避免過擬合, 使其能在高維且特徵間交互性強的情境下有效學習潛在模式。

相較之下, Logistic Regression 為線性模型, 僅能學習變數與目標變數之間的線性關係, 對於非線性或交互特徵難以捕捉, 導致其在 Recall 與 F1-Score 表現上相對較差, 顯示其在預測實際發生延遲的訂單方面具有較高的錯失風險。雖然 Logistic Regression 在模型解釋性方面具有優勢, 適合於初步變數分析與風險評估, 但其預測效能無法與其他進階模型相比。

至於 Random Forest, 雖然同樣具備處理非線性與多樣特徵的能力, 其在研究中之表現亦優於 Logistic Regression, 但仍不及 XGBoost。原因在於 Random Forest 採用的是隨機抽樣與平均策略, 較不具備針對「難分類樣本」的學習機制, 也缺乏正則化來抑制過度擬合, 因此在極度不平衡或特徵訊號分佈不均的資料中, 難以達到 XGBoost 那樣的高度精確與穩定。

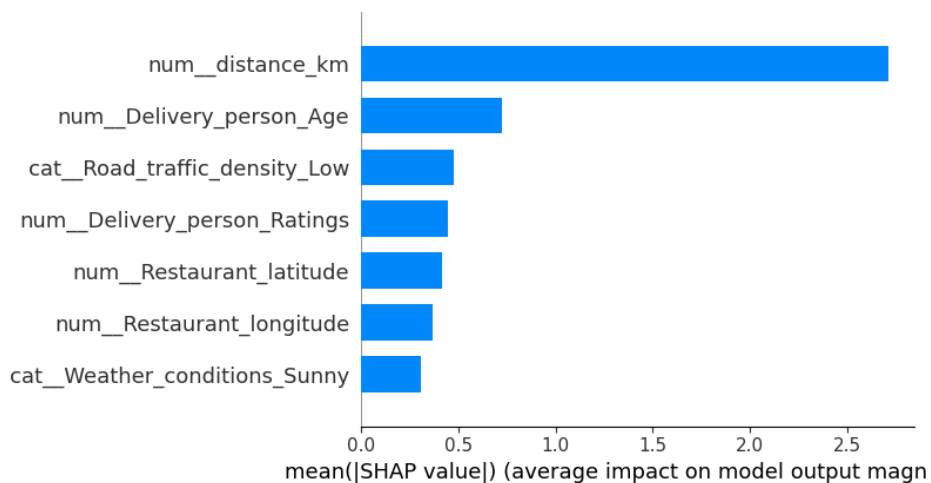
- 風險推論: 錯誤預測會帶來什麼風險?
 - 高風險訂單誤判為低風險 (False Negative): 可能對平台營運帶來實質影響。此類誤判將導致平台無法即時察覺潛在延遲風險, 錯失調派資源與預先調整路線的時機, 最終可能出現訂單堆積與延遲爆單的情形, 進一步造成顧客不滿、差評增加, 甚至顧客流失與品牌信任受損。

- 低風險訂單誤判為高風險 (**False Positive**): 可能導致資源錯置, 例如將優質外送員過度集中於實際風險不高的訂單, 造成整體配送效率下降, 進而影響營運成本與調度效能。

- 模型解釋性分析 (**SHAP** 分析)

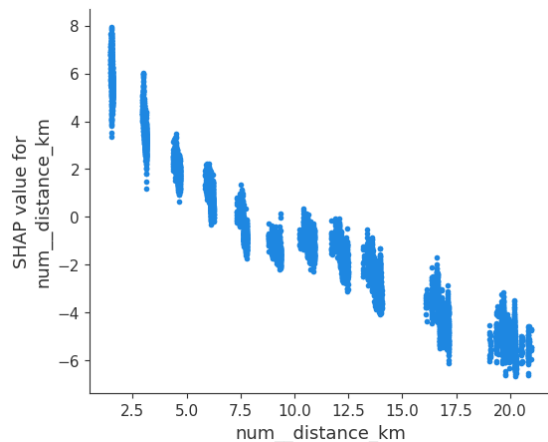
由於三個分類模型中, XGBoost的分類accuracy最高, 因此我們以XGBoost作為我們的最佳模型, 做進一步模型的解釋性分析, 希望能了解哪些因素最容易造成延遲。

首先是模型整體的分析, 下圖為整體模型的summary plot, 由上而下是前 7 大影響模型預測結果的變數、橫軸代表該變數對延遲預測的「影響力大小」, 由圖可見, 使模型容易預測訂單為延遲的前三大影響因子為: 配送距離、外送員年齡、車流密度是否為低。以第一大影響因子「距離」舉例, 其SHAP值大約是2.7, 代表會讓模型預測的 log-odds 增加約 2.7, 等於是讓預測「延遲」的機率提升至93.7%。



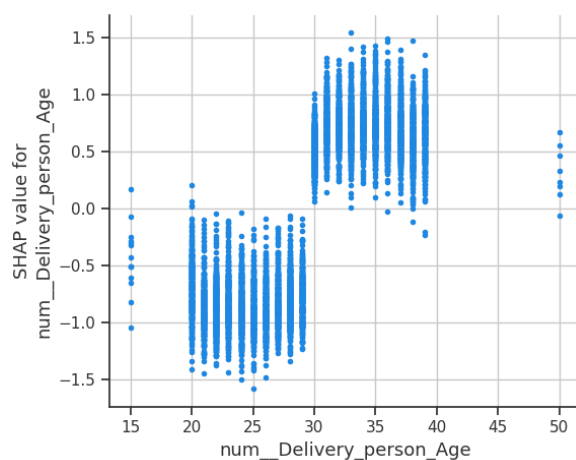
接著我們取出前三大影響因子製作dependence plot, 來觀察這些因子對於模型影響的方向性。

1. 配送距離對訂單延遲的影響



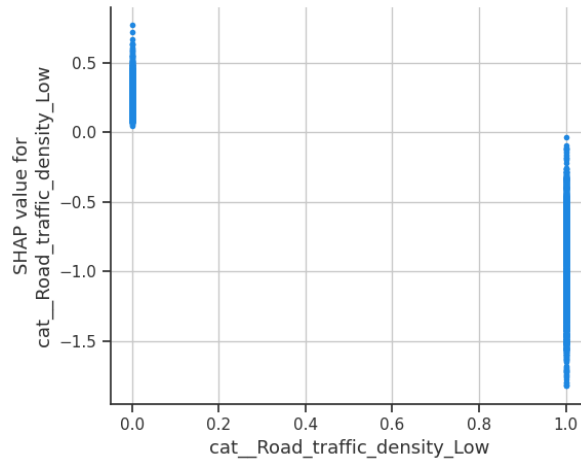
由圖可見，配送距離越短，SHAP 值越大；配送距離越長，SHAP 值越小。代表配送距離越短，模型越認為會延遲；配送距離越長，模型越傾向預測準時。大約在 6 公里處 SHAP 值由正轉負，模型從傾向預測延遲轉為傾向預測準時。「短距離訂單容易延遲」這個結論較不符合我們直覺，稍後在洞察與建議的部分，我們將繼續探討為什麼短距離訂單會是使模型容易預測訂單為延遲的第一要素。

2. 外送員年齡對訂單延遲的影響



從此圖中可以看到在 30 歲的地方有明顯的分界，外送員年齡 < 30 歲，SHAP 值為負；外送員年齡 > 30 歲，SHAP 值為正。代表模型認為年紀較大的配送員，較可能造成延遲。

3. 車流密度對訂單延遲的影響



此圖中顯示了明顯的趨勢：當 Road_traffic_density_Low = 1 (交通很順)，SHAP 值為負，當其為 0 (交通阻塞)，SHAP 值為正、接近 0。代表了，交通順暢的區域，模型認為越不可能延遲，和常理相符。

- 洞察與建議：

- 短距離訂單延遲因素

接續SHAP分析的結論：短距離訂單是讓模型預測訂單延遲的第一大影響因子，以下我們將分析短距離訂單是否真的比較容易延遲，以及短距離延遲訂單的特性。

- 事前定義：

- 短距離訂單定義：

在派送距離與SHAP值的dependence plot當中，大約在 6 公里處 SHAP值由正轉負，因此我們將短距離訂單定義為：配送距離 \leq 6 公里的訂單。

- 延遲時長定義：

為了深入了解短距離訂單和延遲的數量關係，我們不再採用「是否延遲」的二元標籤，而是計算延遲時長作為連續的結果變項，延遲時長的定義為：實際配送時間 - API預測時間。

- 延遲訂單來源結構

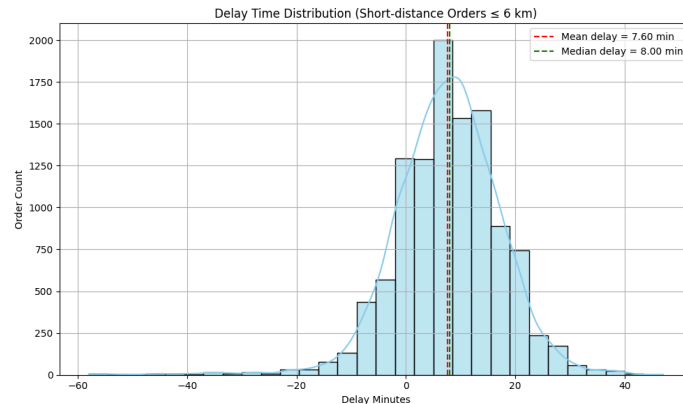
我們將各項訂單數的結果統計如下：

	所有訂單	所有延遲訂單
訂單數	38835	14403

當中短距離訂單數	11218	8904
短距離訂單佔比	28.89%	61.82%

由統計數字可知，短距離訂單被延遲的比例(28.89%)遠高於其在總體中的比例(61.82%)，這支持了我們的猜想：短距離是延遲發生的重要風險因子

■ 短距離訂單與訂單延長時長的關係

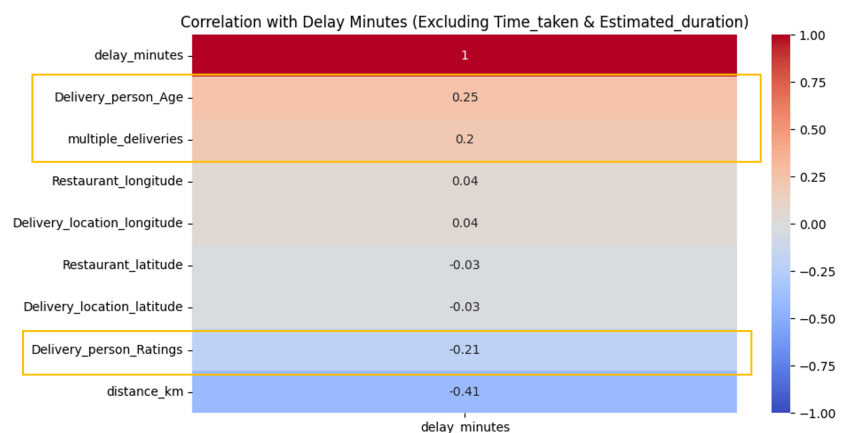


從長條圖中我們發現，短距離訂單中確實有高達 79.4% 的訂單有延遲，而其中甚至有延遲超過 20 分鐘，達到業界「嚴重延遲」的標準（參考Uber Eats標準）。

■ 短距離訂單中和延遲時長有關的潛在因素

最後，我們做了一些簡單的視覺化分析，找出短距離訂單中幾個和延遲時長較相關的變項，供未來研究參考。

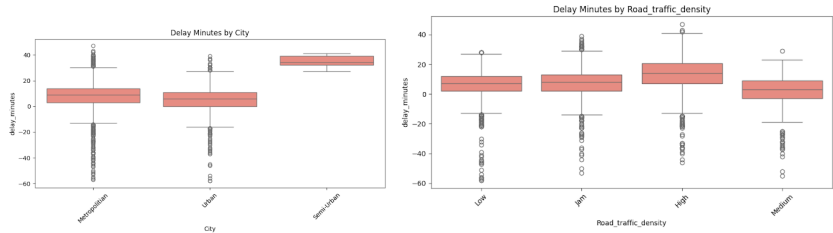
● 連續變項



從其他連續變項和延遲時長的相關係數圖可發現：

年齡越高，延遲越長；同時處理多張訂單，延遲越長；評分高的外送員延遲較短（可反映服務品質）。

- 類別變項



從類別變項的和延遲時長的箱型圖中我們挑出了兩個肉眼上看似有組間差異的變項，可以發現：

在半郊區的訂單，延遲越多；經過高車流密度的訂單，延遲越多。

- 建議平台改善策略

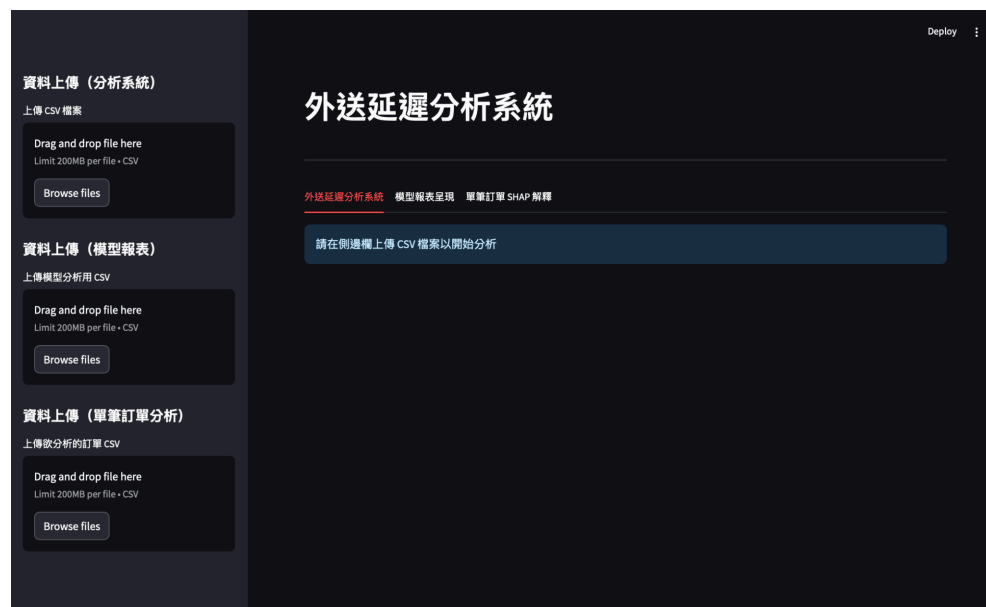
- 延遲風險即時預警機制

- 將模型預測出的高風險訂單進行即時標記，並優先處理。此預測結果可作為外送員指派邏輯的重要參考，將經驗豐富或評價較高的外送員指派至延遲風險較高的訂單，有助於降低實際延遲發生率。

- 平台之優化派單與路線規劃演算法

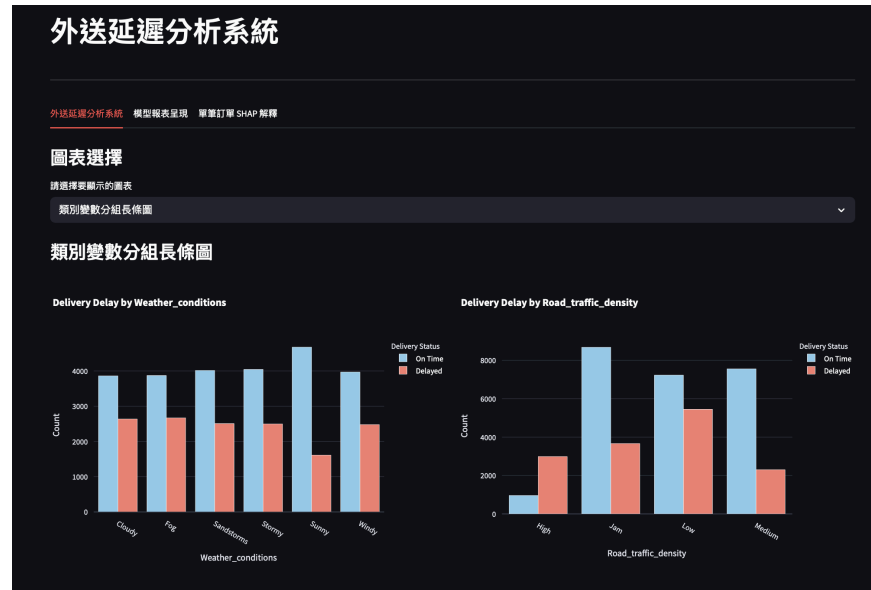
- 針對 SHAP 分析中顯示影響延遲最顯著的變數，設計最佳之路線規劃，並動態調度人力及派單順序。

- 系統平台應用與視覺化介面功能的用途和價值



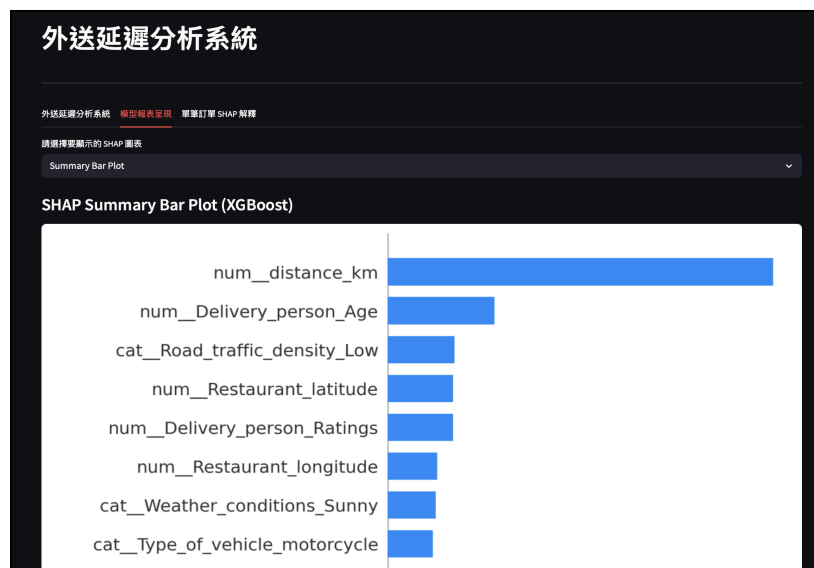
說明: 使用者可於右側上傳透過前處理之 Python 檔處理過之資料集, 且系統包含三大功能, 其個別之功能與用途和價值如下:

1. 外送延遲分析系統: 透過多種統計圖表視覺化潛在具關聯性的變數間關係



用途與價值: 可幫助使用者直觀辨識具有潛在關聯性的變數組合, 助於在特徵工程階段發掘可能影響延遲的關鍵因素, 也可作為管理者監控環境條件與配送表現的依據。

2. 模型報表呈現: 由 4.3 之預設效果分析可知 XGBoost 之表現最好, 故此部分以此模型為基準, 針對我們研究推論之可能影響外送延遲因子做 Shapley value 之視覺化, 藉此看出影響的程度。



用途與價值: 提供模型整體預測效果的評估與 SHAP 值的視覺化解釋, 使用者可明確看出哪些特徵對於模型預測延遲風險的貢獻程度, 更有助於管理者從資料中歸納出具體可行的策略建議(如針對高風險變數進行優化或排程調整)。

3. 單筆訂單 Shap 解釋: 此功能在餵入資料後, 會為使用者根據我們訂出之延遲風險閥值, 將訂單資料區分為延遲風險等級全部、高、中、低, 並有下拉選單可選擇欲查看的風險等級之訂單, 且可於下方搜尋欄輸入訂單資料之 index 細看延遲機率及可能造成延遲之五大因子

外送延遲分析系統

外送延遲分析系統

模型報表呈現

單筆訂單 SHAP 解釋

選擇要查看的延遲風險等級:

高

符合條件的訂單資料:

	Delivery_person_Age	Delivery_person_Ratings	Restaurant_latitude	Restaurant_longitude	Delivery_location_latitude	Delivery_location_longitude	Order_Date	Weather_condition
0	25	4.8	17.4509	78.3793	17.4609	78.3893	2022-04-01 00:00:00	Cloudy
1	35	4.9	19.2546	72.8489	19.2846	72.8789	2022-04-05 00:00:00	Windy
2	37	4.7	22.3081	73.1678	22.3581	73.2178	2022-03-19 00:00:00	Windy
3	34	4.5	12.9357	77.6141	12.9657	77.6441	2022-03-19 00:00:00	Sunny
4	30	4.6358	22.727	75.8842	22.737	75.8942	2022-03-19 00:00:00	Cloudy
5	28	4.6623	27.1609	78.0441	27.1709	78.0541	2022-02-13 00:00:00	Fog
6	31	4.7	22.3103	73.1691	22.3503	73.2091	2022-03-03 00:00:00	Stormy
7	35	4.6	12.324	76.628	12.394	76.698	2022-03-23 00:00:00	Cloudy
8	39	3.9	27.2017	78.0076	27.2817	78.0876	2022-02-16 00:00:00	Windy

請輸入要查詢的 index (0 ~ 27):

4

訂單延遲風險: 高

18

32

4.9

12.334

76.6182

12.354

76.6382

2022-03-11 00:00:00

Sunny

請輸入要查詢的 index (0 ~ 27):

6

訂單延遲風險: 高

延遲因子

第1大延遲因子: num_distance_km

第2大延遲因子: num_Delivery_person_Age

第3大延遲因子: cat_Road_traffic_density_High

第4大延遲因子: cat_Weather_conditions_Sunny

第5大延遲因子: cat_Road_traffic_density_Low

用途與價值: 可提供了個案層級的風險預警能力, 使平台得以在訂單尚未延遲前即進行風險控制(例如調整指派外送員或變更配送路徑), 進一步提升整體服務準時率。

5. Conclusion

本研究透過邏輯回歸、Random Forest 與 XGBoost 三種分類模型建構預測框架, 針對外送平台中訂單配送是否延遲進行預測分析。研究結果顯示, XGBoost 模型在各項評估指標中表現最佳, 特別是在 Recall(83.8%)、F1-Score(84.98%)與 AUC(0.9611)方面顯著優於其他模型。

從 SHAP 解釋模型可知，影響配送延遲最顯著的三個因素依序為：配送距離、配送員年齡、與交通順暢度。其中，短距離訂單延遲的風險最高，與大眾直覺（距離越短應越準時）相悖。進一步分析發現，短距離訂單中高達 79.4% 有延遲現象，其中不乏嚴重延遲(>20分鐘)的案例，推測可能與備餐延遲、等待指派時間過長、或平台排程策略失衡等因素相關。

此外，年長配送員與高車流密度區域亦與延遲顯著相關，凸顯人員經驗與環境因素的重要性。平台若能即時辨識高風險訂單並採取預防措施（如優先派單、動態調度、路線優化），將有助於提升整體服務準時率與顧客滿意度。

總體而言，本研究建立了一個具可行性之預測模型，解釋性分析挖掘出數個具策略價值的洞察。未來建議平台導入即時資料（如即時天氣與交通資訊）、擴充歷史配送行為資料，進一步提升模型精準度與預警能力。透過此類數據科學應用，外送平台可更有效地優化營運流程，降低延遲風險與成本支出。