
Airline Passenger Satisfaction Prediction And Analysis

第一組

陳韋傑、石子仙、莊啟宏、李宇軒

大綱

- 前言
- 資料探索
- 資料前處理
- 模型架設
- 滿意度預測
- 重要變數分析
- 策略建議

前言 — 動機

- 預測並提升航空服務業中顧客的滿意度，有以下兩個好處：
 1. 培養忠誠顧客，讓顧客再次進行消費
 - 留住一個舊顧客的成本只有開發一個新客戶的 1/5 (Thomas et al., 2004)
 - 國內兩項研究顯示，航空業顧客滿意度高對於購後行為傾向有正面影響
(一、葉晶雯講師, 2003，二、湯玲郎副教授等三人, 2021)
 2. 累積好評，達到正向口碑宣傳，提升企業形象
 - 航空業彼此競爭的產品相似且不易區隔，好的企業形象為顧客重要的決策依據

前言 — 專案目標

- 探究是否能藉由旅客資料與問卷結果準確預測旅客滿意與否
 - 以整體資料建構預測模型，透過各項指標檢視模型成效
- 了解實施問卷調查是否有助於預測旅客滿意與否
 - 比較以旅客資料建構的預測模型與以問卷結果建構的預測模型
- 找出不同客群在意的重要指標，藉此提出經營策略建議
 - 以不同客群的資料建立預測模型、比較重要變數的異同

資料探索 — 來源

- Kaggle: Airline Passenger Satisfaction
- 資料筆數
 - Train : 103904
 - Test : 25976
- 欄位數 : 24
 - 自變數 (x) : 23 個欄位
 - 依變數 (y) : Satisfaction , 爲二元變數



變數介紹 — 旅客資料

自變數

a. 旅客資料

欄位名稱	說明	欄位名稱	說明
id	旅客編號	Class	艙位分級
Gender	性別	Flight Distance	飛行距離
Customer Type	旅客分類	Departure Delay in Minutes	起飛延誤時間
Age	年齡	Arrival Delay in Minutes	抵達延誤時間
Type Of Travel	旅行目的		

變數介紹 — 問卷結果 (5分制)

依變數

欄位名稱

說明

Satisfaction

滿意度

自變數

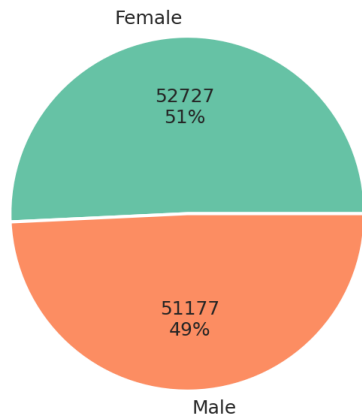
b. 問卷結果 (5分制)

欄位名稱	說明	欄位名稱	說明	欄位名稱	說明
Inflight wifi service	機上Wi-fi網路	Inflight entertainment	機上娛樂	Baggage handling	行李處理
Departure/Arrival time convenient	航班時間的便利性	On-board service	登機服務	Checkin service	登機服務
Ease of Online booking	線上訂票便利性	Leg room service	腿部空間	Inflight service	機上服務
Gate location	登機門位置	Online boarding	線上報到	Cleanliness	整潔度
Food and drink	餐飲	Seat comfort	座位		

資料分布情形

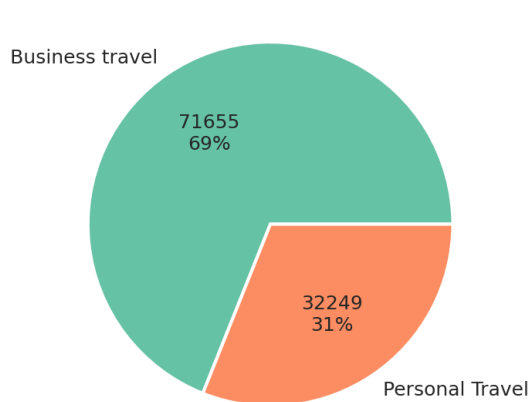
性別

Gender



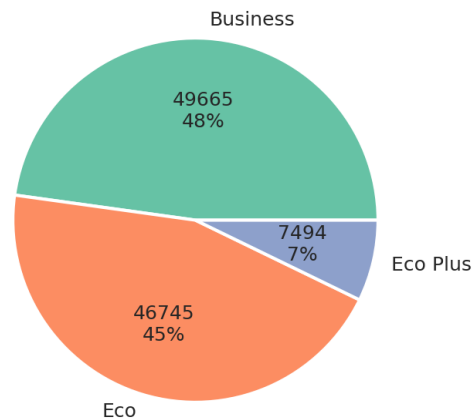
旅行目的

Type Of Travel



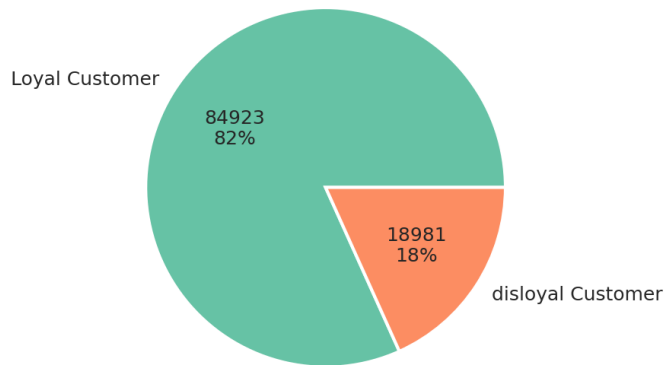
艙位分級

Class

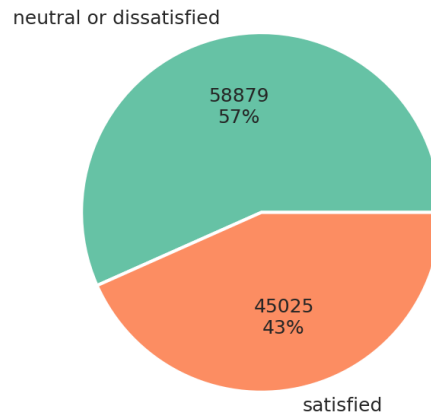


資料分布情形

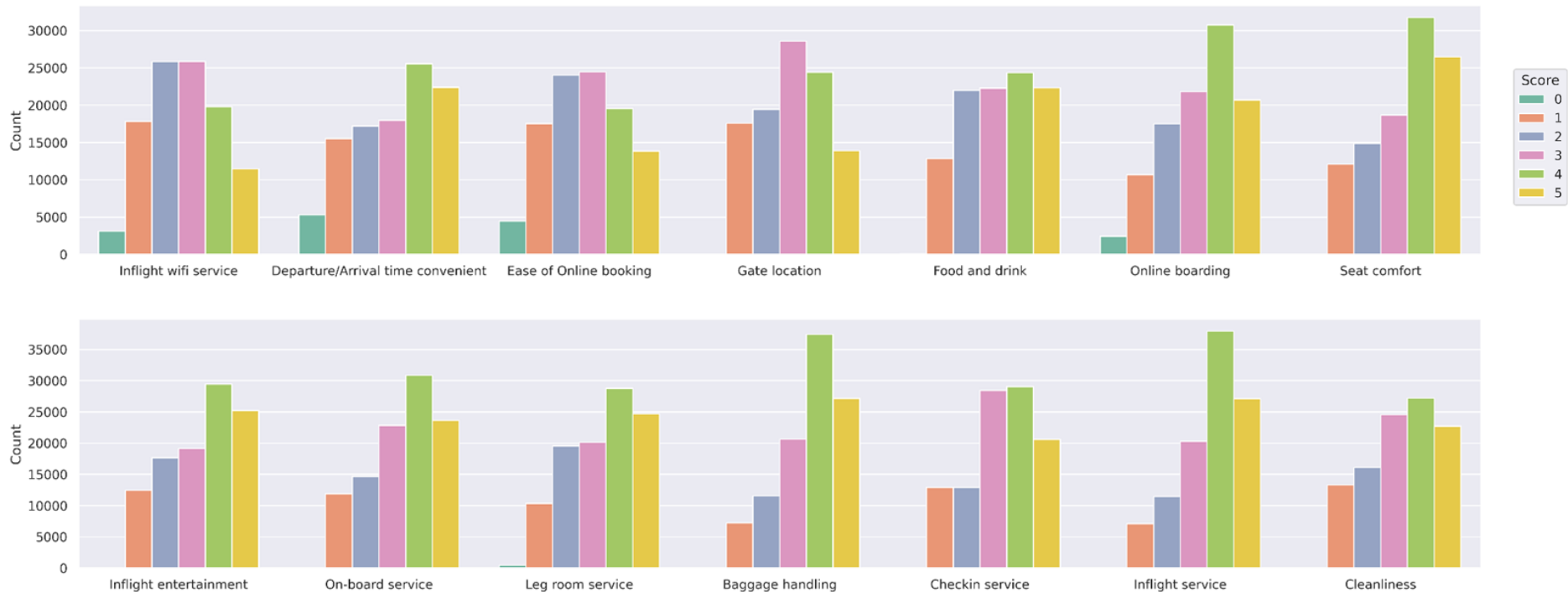
顧客種類
Customer Type



滿意度
Satisfaction

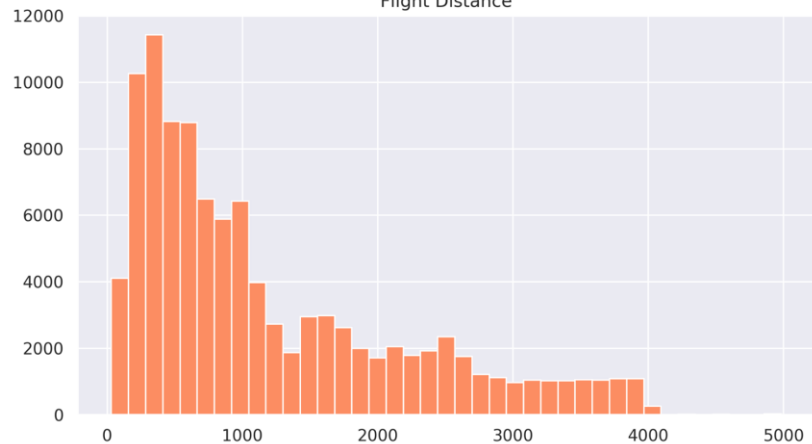


資料分布情形 — 問卷各項欄位評分



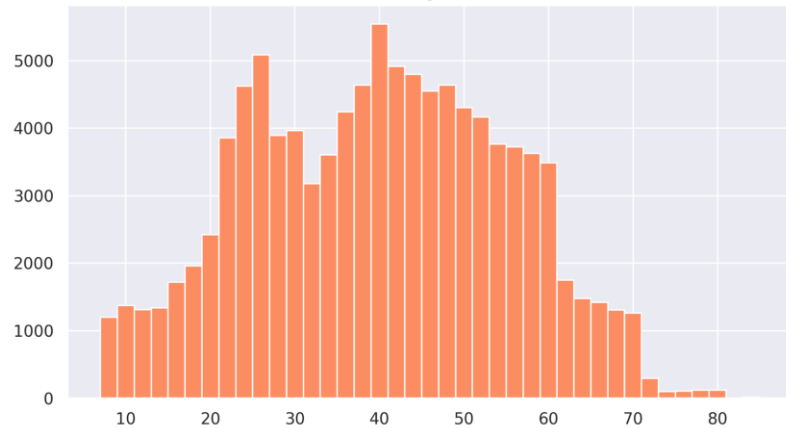
飛行距離

Flight Distance



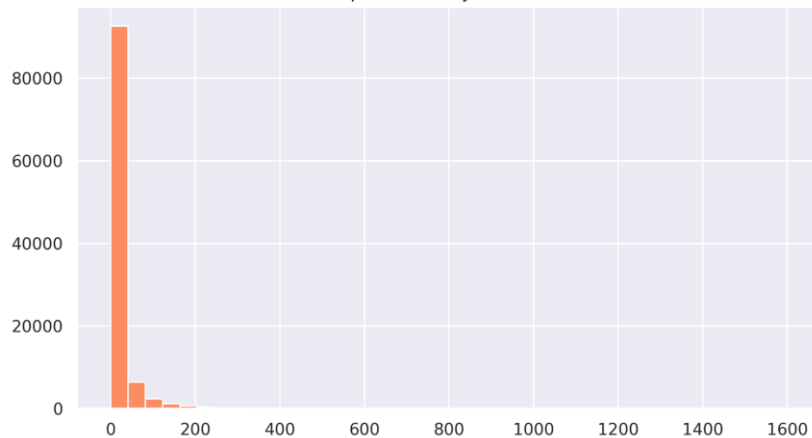
年齡

Age



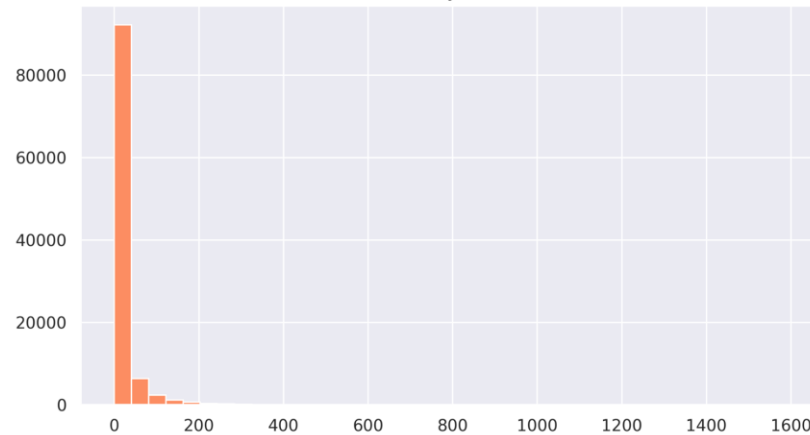
起飛延誤時間

Departure Delay in Minutes



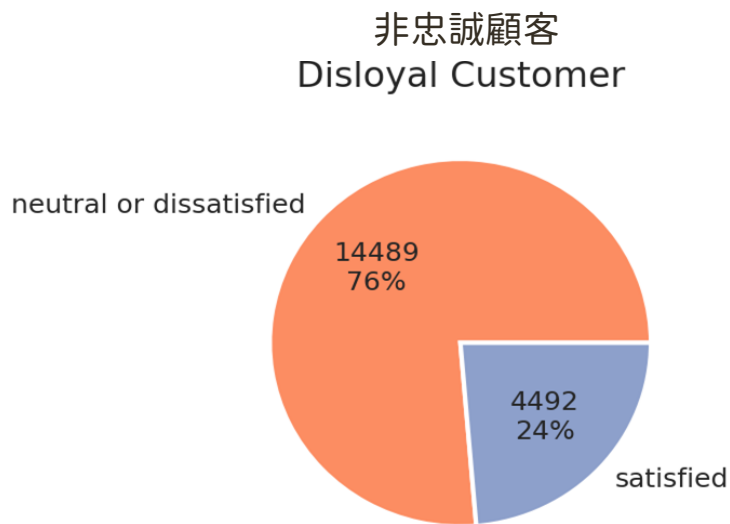
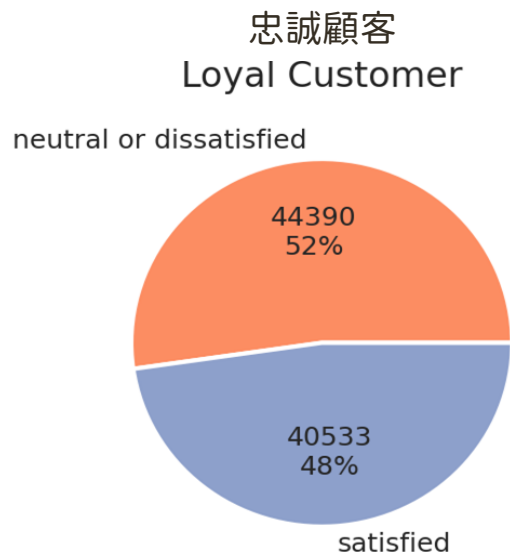
抵達延誤時間

Arrival Delay in Minutes



資料分布情形

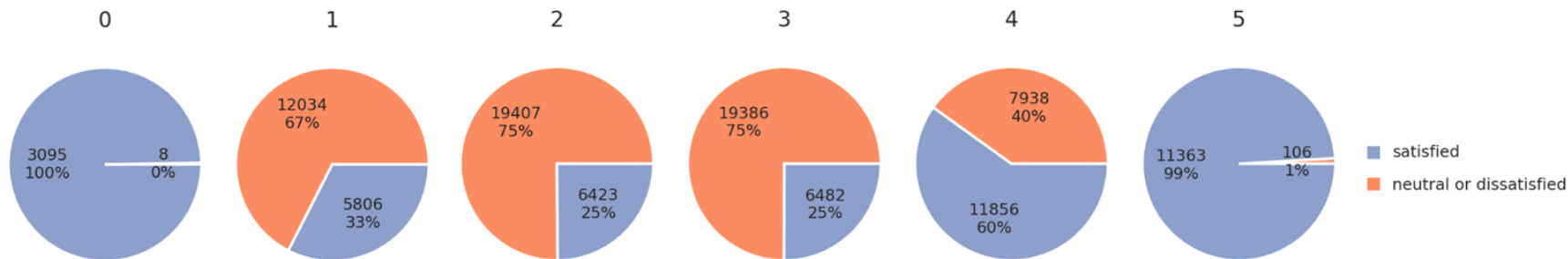
- 顧客種類對滿意度的分布



資料前處理

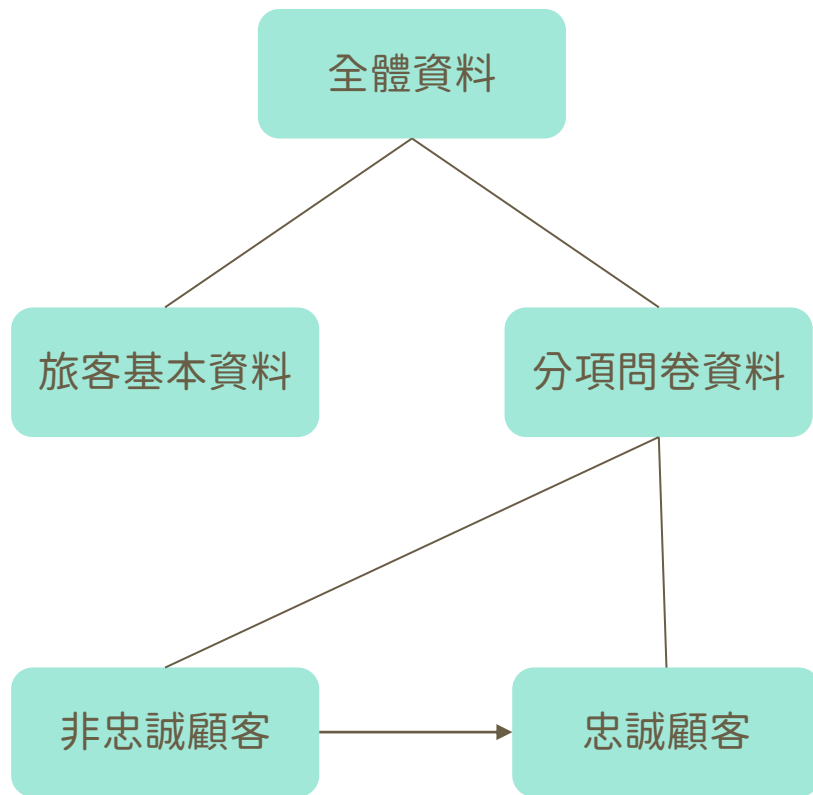
- 將問卷服務滿意度，值為 0 的部分轉換為NA
- 將含有 NA 值的資料點刪除
 - 包含 "Arrival Delay in Minutes"（僅占0.3%）、服務滿意度欄位...
- 對各欄位進行 Normalization

Inflight Wifi Service to Satisfaction



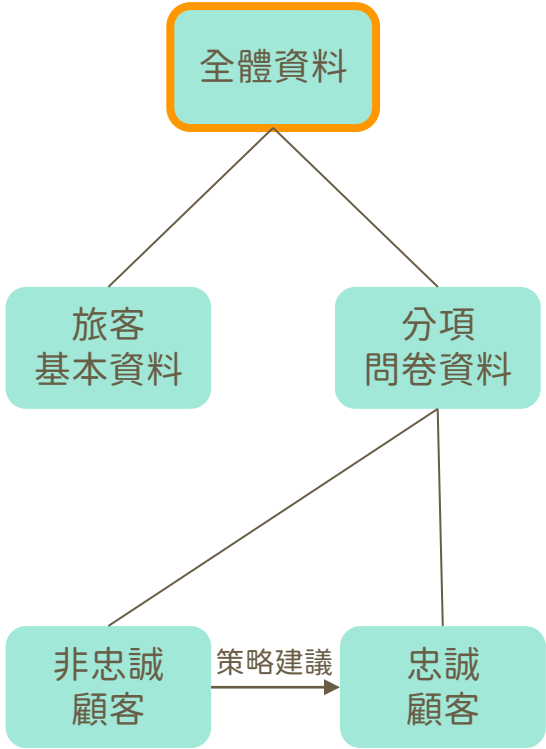
模型架設

- 使用多種模型進行預測與分析
 - KNN
 - Gradient Boosting Machine
 - Random Forest
 - XGBoost
 - Decision Tree
 - Ensemble of Above Models
- 預測顧客對於整體服務是否滿意
- 分析重要變數，給出具體策略建議
 - 使用投票法找出共同重要變數



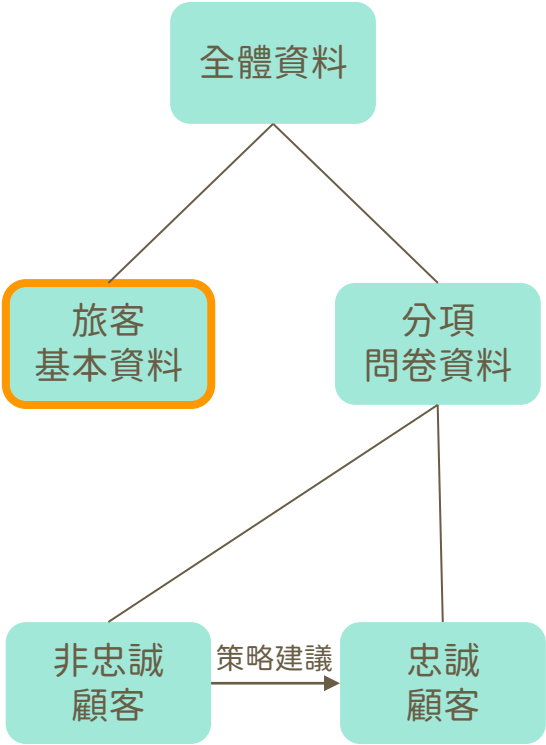
滿意度預測（全體資料）

Model / Metric	Accuracy	Precision	Recall	F1-score
KNN	0.9383	0.9637	0.8901	0.9254
Gradient Boosting	0.9604	0.9695	0.9373	0.9532
Random Forest	0.9623	0.9721	0.9394	0.9555
XGBoost	0.9649	0.9730	0.9425	0.9575
Decision Tree	0.9432	0.9412	0.9258	0.9334
Ensemble	0.9627	0.9729	0.9396	0.9560
Average	0.9553	0.9654	0.9291	0.9468



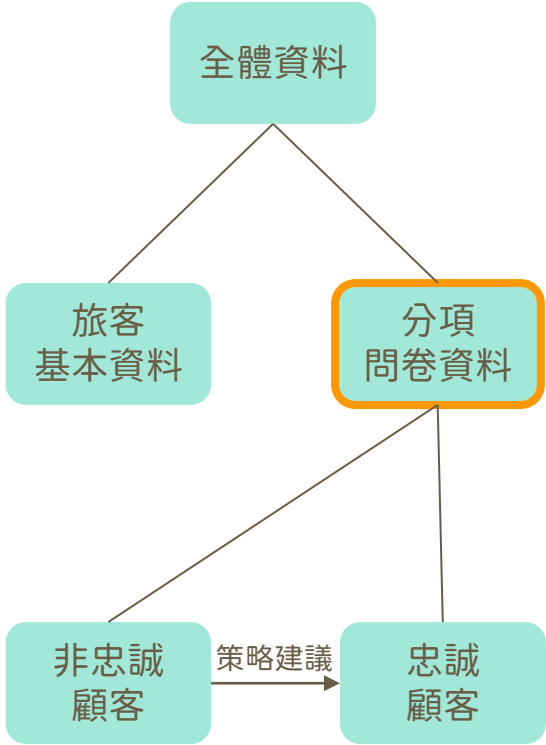
滿意度預測（旅客基本資料）

Model / Metric	Accuracy	Precision	Recall	F1-score
KNN	0.7905	0.7614	0.7473	0.7543
Gradient Boosting	0.8065	0.7614	0.8013	0.7809
Random Forest	0.7730	0.7370	0.7348	0.7359
XGBoost	0.8061	0.7627	0.7975	0.7797
Decision Tree	0.8039	0.7555	0.8049	0.7794
Ensemble	0.8055	0.7629	0.7950	0.7786
Average	0.7976	0.7568	0.7801	0.7681



滿意度預測（分項問卷資料）

Model / Metric	Accuracy	Precision	Recall	F1-score
KNN	0.9268	0.9463	0.8799	0.9119
Gradient Boosting	0.9420	0.9477	0.9157	0.9314
Random Forest	0.9464	0.9520	0.9219	0.9367
XGBoost	0.9485	0.9562	0.9227	0.9391
Decision Tree	0.9249	0.9314	0.8910	0.9108
Ensemble	0.9470	0.9550	0.9202	0.9373
Average	0.9393	0.9481	0.9086	0.9279

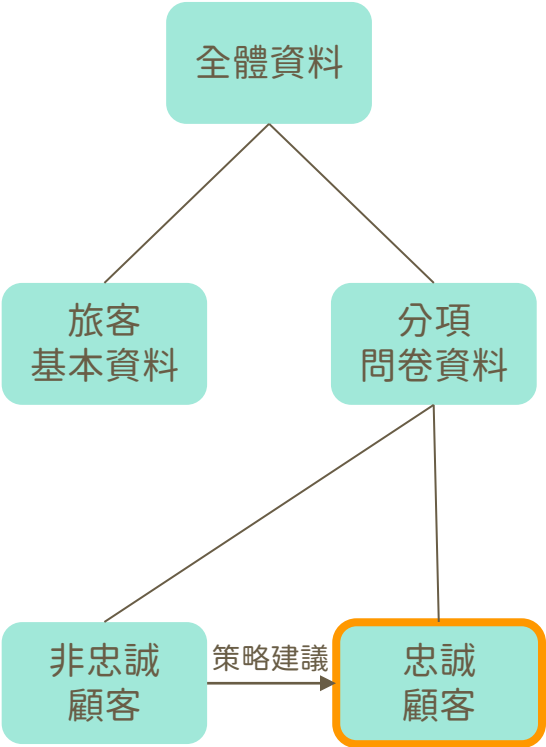


滿意度預測 — 小結

- 全體資料滿意度預測之 Accuracy, Precision, Recall, F1 皆達到 0.9 以上
- KNN 表現多較差、Boosting 與 Ensemble 表現多不錯
- 問卷資訊對了解顧客滿意度的確有幫助

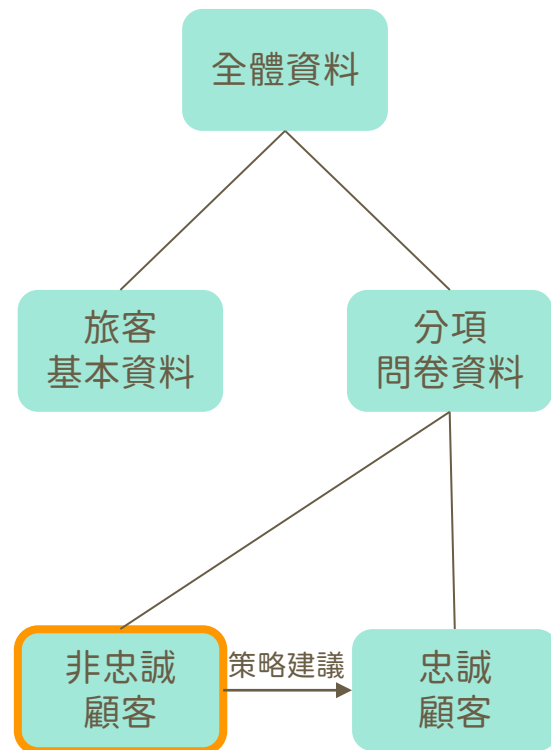
重要變數分析（忠誠顧客）

Features (Not in order)	Votes	Description
Leg room service	4	Satisfaction level of Leg room service
Inflight wifi service	4	Satisfaction level of the inflight wifi service
Departure/Arrival time convenient	4	Satisfaction level of Departure/Arrival time convenient
Gate location	4	Satisfaction level of Gate location
On-board service	4	Satisfaction level of On-board service
Inflight entertainment	4	Satisfaction level of inflight entertainment
Online boarding	4	Satisfaction level of online boarding



重要變數分析（非忠誠顧客）

Features (Not in order)	Votes	Description
Leg room service	4	Satisfaction level of Leg room service
Inflight wifi service	4	Satisfaction level of the inflight wifi service
Departure/Arrival time convenient	4	Satisfaction level of Departure/Arrival time convenient
Gate location	4	Satisfaction level of Gate location
On-board service	4	Satisfaction level of On-board service
Baggage handling	4	Satisfaction level of baggage handling
Checkin service	4	Satisfaction level of Check-in service
Inflight service	4	Satisfaction level of inflight service



重要變數分析 — 小結

- 對所有顧客而言
 - 座位空間、機上 Wi-fi、航程時間方便性、登機門位置、機上服務皆是影響滿意度的重要因子
- 對忠誠顧客而言
 - 重視機上娛樂與線上登機方便性
 - 推測因常搭乘，故較注重搭乘體驗與訂位方便性
- 對非忠誠顧客而言
 - 重視行李托運、報到服務與航程中服務
 - 推測多是新客，較注重整體服務品質與流程順暢度






策略建議

- 策略對象：美國各大航空公司
- 策略目標：透過提升顧客滿意度，將非忠誠顧客轉為忠誠顧客
- 根據模型篩選出在忠誠與非忠誠顧客資料中相同的重要因子，進行策略建議
 - 機上服務 (On board service)
 - 機上網路 (Inflight wi-fi service)
 - 登機門位置 (Gate location)
 - 座位空間 (Leg room service)

策略建議 — 機上服務 (On-board service)

問題描述	服務種類過少
策略建議	<p>提升服務多樣性：</p> <ul style="list-style-type: none">● 機上 Wi-fi● 機上娛樂：電視、電影、遊戲、音樂● 機上餐食：正餐、點心、飲品、酒水● 購物產品：周邊商品、Outlet● 個人化服務：依照旅客特殊需求提供

策略建議 — 機上網路 (Inflight wi-fi service)

<p>問題描述</p>	<p>根據 2021 TPG 美國最佳航空公司的報告，前 7 名的公司都有 Wi-fi 服務，8 至 10 名則無 (Spirit、Frontier、Allegiant)</p>	<p>ABOUT FREE MESSAGING</p> <p>Message friends and family throughout your flight, courtesy of Delta. Both you and the recipient must have one of the compatible apps, which include: iMessage, Facebook Messenger, and WhatsApp.</p> <div> iMessage</div> <div> Facebook Messenger</div> <div> WhatsApp</div> <p>Supports text messaging only using these apps. Photos and videos are not supported.</p> <div></div>
<p>策略建議</p>	<p>規劃不同方案給不同旅客，且保障基本網路使用</p> <ul style="list-style-type: none">● 基本方案：僅限傳送文字訊息，限速限流量● 進階方案：不限訊息格式，提升速度● 訂閱制：一次購買一個月，可無限量使用	

策略建議 — 登機門位置 (Gate Location)

問題描述	不同登機門常有顧客流量不平衡的情形，造成某些登機門堵塞
策略建議	利用 Airport Gate Assignment (AGA 模型)，最大限度的減少每個登機口的乘客數量差異 (資料來源：Airport Gate Assignment for Improving Terminals' Internal Gate Efficiency, Lee et al., 2016)
目標函式	$\min \sum_{k \in K} \sum_l (G_{lkt})^2 + M^2 \sum_{ij} y_{ij}$
變數定義	G ：在時間點 t 時，使用第 k 個內部登機門前往第 i 個登機門的乘客數量 y ：當航班 i 使用登機門 j 時， y 為 1，否則為 0 M ：一個很大的數字

策略建議 — 座椅空間 (Leg Room Service)

問題描述	座椅空間過小，導致腳無法自然伸展，降低旅客航程舒適度				
策略建議	<p>參考美國各大航空公司，座椅空間需至少 30 英吋</p> <ul style="list-style-type: none">根據 2021 TPG 美國最佳航空公司的報告，Delta 為第 1 名、United 為第 3 名、Spirit 為第 8 名				
					
Jet Blue		Virgin America, Southwest	Delta, American	United	Spirit
34"		32"	31"	30"	28"

結論

- 使用多種模型針對顧客滿意與否進行預測
 - 準確度等指標均在 0.9 以上
- 問卷的實施對預測顧客滿意與否確實有幫助
 - 僅使用問卷資料即可達到不錯的預測
- 針對重要指標，提出多項策略建議
 - 提供多樣機上服務
 - 提供方案制機上網路
 - 以 AGA 模型最佳化登機門的分配
 - 確立座椅空間的最佳標準

Thanks for Listening