

保險業務員舞弊風險分析

臺灣大學 許哲榮

臺灣大學 李宸帆

東吳大學 陸珮甄

東吳大學 楊永浚

東吳大學 溫偉成

目錄

- 01 資料描述與違約相關性分析
- 02 資料處理與生成第一次Decision Tree
- 03 第二次Decision Tree及訓練結果
- 04 Decision Tree模型結果實務觀點
- 05 舞弊人員特性分析
- 06 總結

原始資料性質

01

數值資料欄位：384

類別資料欄位：7（包含應變數）

舞弊業務員數：195（0.65%）

非舞弊業務員數：29805（99.35%）

總業務員數：30000（100%）

數值資料

Step 1 檢測X之間的相關性，去除相關性過高的因子

Step 2 以Point-biserial Correlation Coefficient找出處理後資料之相關性

Step 3 以Wilcoxon Rank Sum Test找出數值資料分布在Y=0及Y=1是否不同

類別資料

使用克雷默V係數

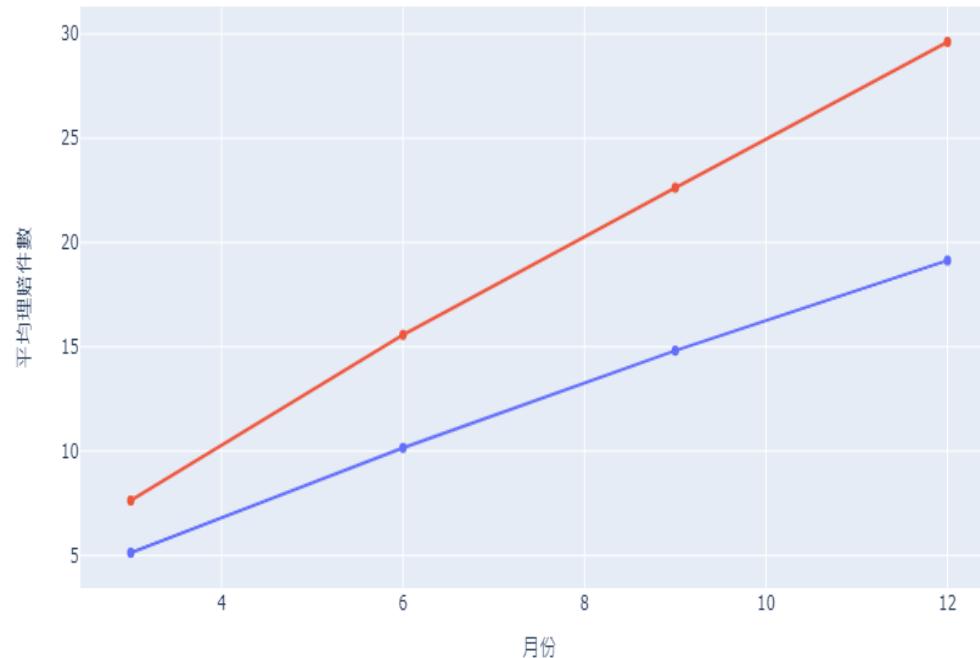
與舞弊前五大相關變數

1. 過往舞弊(含偽簽/代簽)次數
2. 過往不當招攬&未視親簽的次數
3. 持有失效保單之客戶數
4. 墊繳保單數比例
5. 拒賠件數

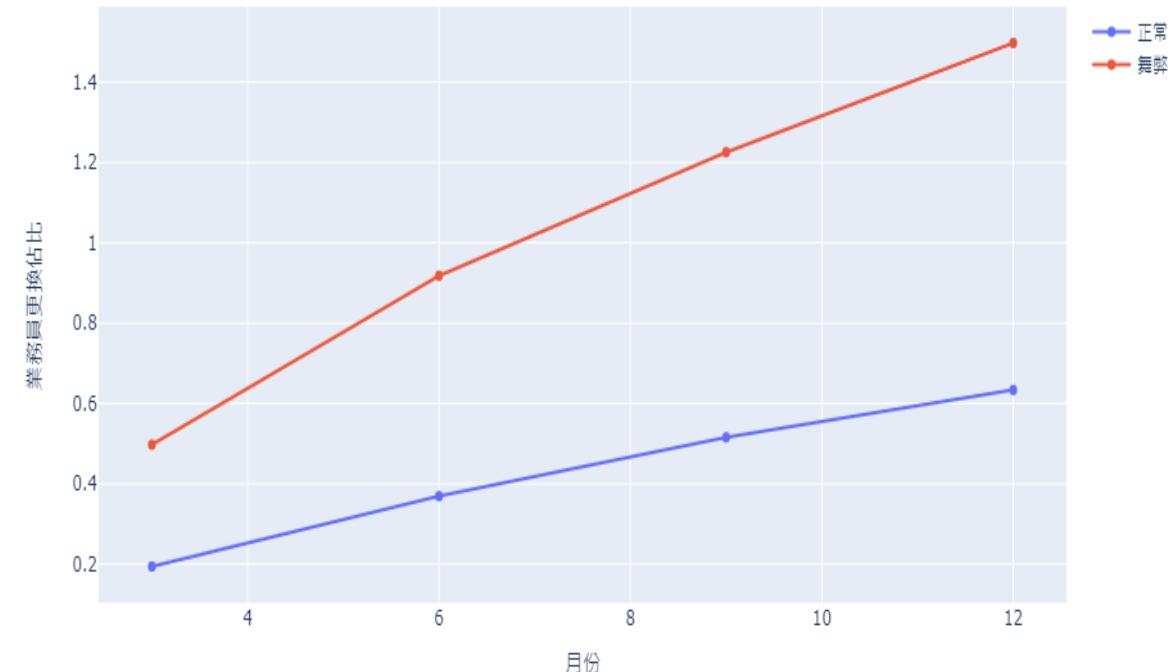
業務表現

- ✓ 短期理賠件數多
- ✓ 業務員更換比高

業務員短期理賠件數



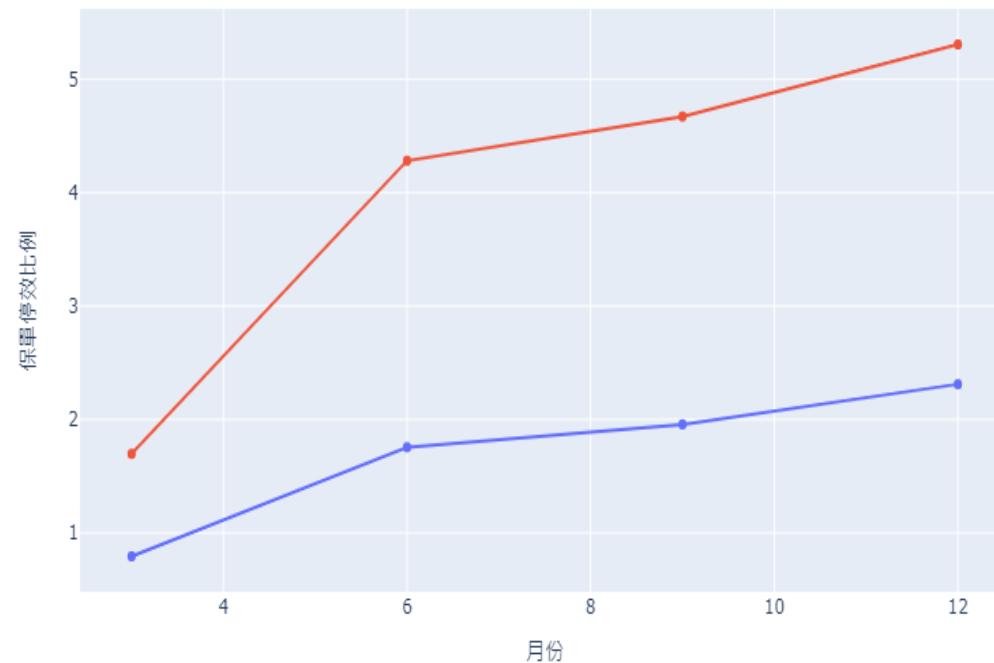
短期業務員更換佔比



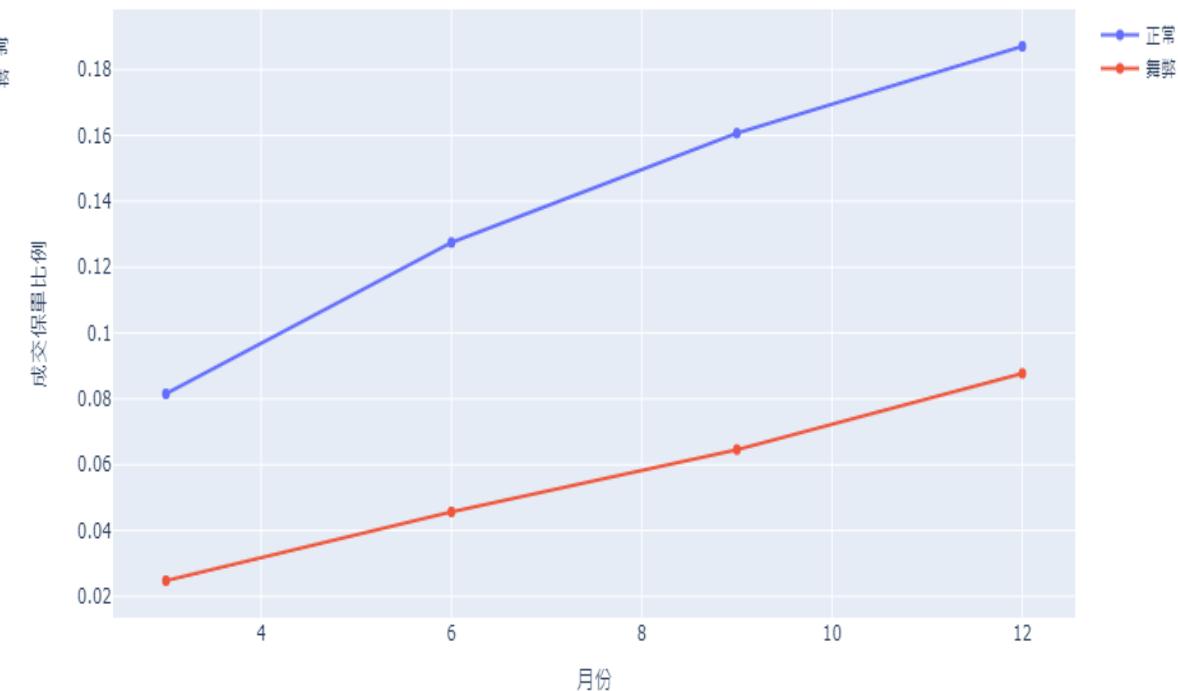
業務表現

- ✓ 保單停效、失效、比例高
- ✓ 保單成交數少

短期保單停效比例



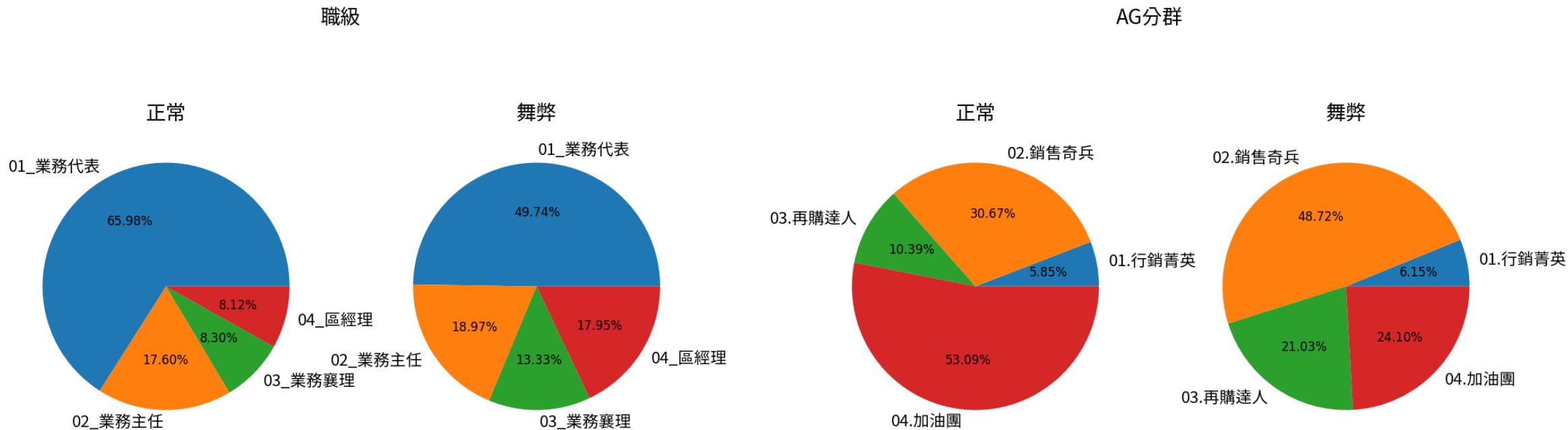
短期成交保單比例



01

身分類別

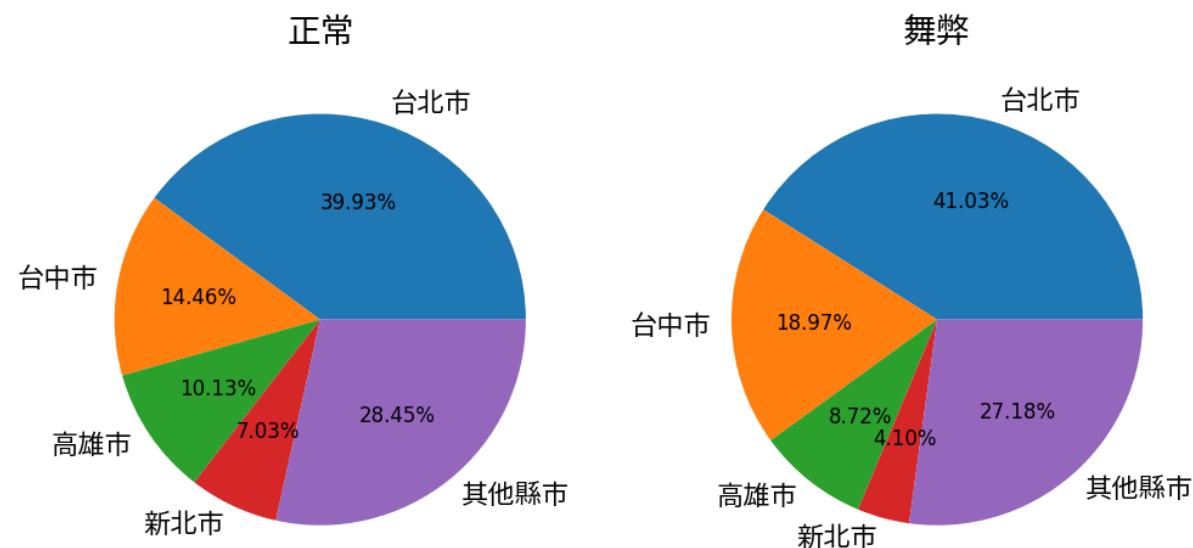
- ✓ 職級：區經理&業務襄理舞弊機率較高
- ✓ AG分群：銷售奇兵 & 再購達人舞弊機率較高



通訊處所在縣市

- ✓ 台中市舞弊機率較高

通訊處所在縣市



各變數與舞弊風險的關係

1. 偽簽 & 代簽

判例指出，曾有業務員冒名簽收理賠支票，並頂替客戶之名開立帳戶，在轉至業務員之銀行帳戶。最後保險公司須負連帶賠償責任。

2. 不當招攬

以誇大不實的說法，誘使客戶誤信關於保單的特質，例如優惠的保證收益，或者使客戶誤信解約所造成的損失不大等未使客戶充分明白合約內容。若不當招攬屬實，則需將保費全數退回給客戶。造成公司財務及信譽上的損失。

3. 未視親簽

無法清楚瞭解消費者投保的動機

無法確認要保人與被保險人及被保險人與受益人之關係

無法瞭解要保人與被保險人之投保目的與需求

財務狀況並考量保單適合度、保險費、保險金額與保障需求間之適當性

要保人、被保險人是否親自填寫要保書之內容

金管會於102年已增加此規定

4. 持有失效保單之客戶數

判例指出，有保戶未繳交保費之理由是業務員未去收費，導致保單失效。而該保戶說其業務員未提到保單有未繳費之情形，最終判決為保險公司未提醒保戶繳費，因此保單有效且保戶須補繳保費。因此若失效保單客戶數高，可能表示業務員未盡其義務。

5. 墊繳保單數比例

保單進入墊繳表示保戶未準時繳交保費，是成為失效保單的前一步，因此與第4點指出的原因類似，其中可能包含業務員未盡其告知義務。

1. 填補缺失值

eg. 將 RECENCY 月份缺失值填入極大值

2. 資料不平衡

非舞弊的資料只有0.65%

使用oversampling，生成 $y=1$ 的資料，使 $y=0$ 及 $y=1$ 的資料比為1:1，共47688筆資料

3. 加入時間維度

將含有時間資訊的變數，前後三個月資料相減，考慮時間變化量的影響，將數量及是否舞弊之二維資訊提升為變化量與舞弊之三維資訊。

→當符合其他條件，停效保單總數變化量及投資型保單總數可能為抓出舞弊業務員的變數

Decision Tree方法說明

方法

1. 先將資料切 training set 和 validation set, 再將 Over Sampling 後的資料倒入模型
2. 不做 Over Sampling, 直接將資料丟入模型，調整舞弊資料的權重

以 Cost Complexity 和 tree 的深度訓練資料，找在 Validation Set 表現最好的模型

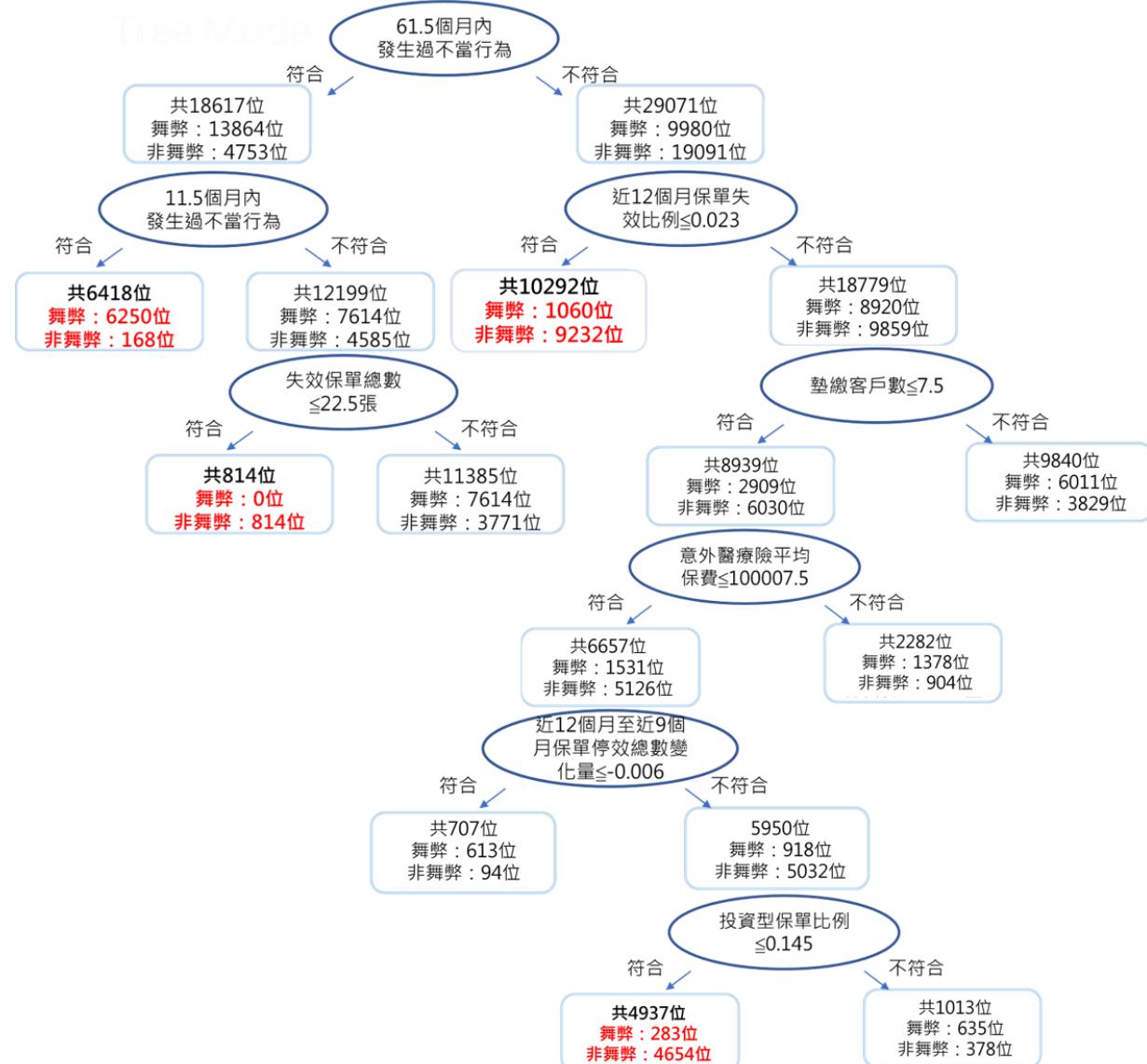
Tree Model 1

02

第一次Decision Tree結果

- 訓練資料前5%捕捉率40.00%
- 測試資料前5%捕捉率29.09%

捕捉變數→



觀察結果

- 發生過往不當行為是很好的預測變數
→ 捕捉率達 25.13%
- 拿掉發生不當行為RECENCY
→ 重新訓練Decision Tree
→ ~~再將11個月內發生不當行為的舞弊發生率 100%~~
- 訓練模型時，為了避免 Overfit，Tree 分支少
→ 模型預測相同機率的 Agent 多

過往不當行為Recency < 12 ←—————

	機率	總人數	舞弊人數	Bad 佔比
0	0.0%	3862	3	1.54%
1	4.53%	11512	11	5.64%
2	15.43%	6269	11	5.64%
3	68.2%	3519	42	21.54%
4	71.74%	3790	60	30.77%
5	73.61%	629	11	5.64%
6	86.7%	119	4	2.05%
7	90.1%	26	1	0.51%
8	92.47%	17	1	0.51%
9	97.38%	252	49	25.13%
10	99.65%	5	2	1.03%

Training 結果

訓練資料前5%捕捉率40.51%

	Bad 滲透度	倍數	Bad 占比	Bad 累積占比
5%	5.27%	8.1	40.51%	40.51%
10%	1.80%	2.77	13.85%	54.36%
20%	1.37%	2.1	21.03%	75.38%
30%	0.63%	0.97	9.74%	85.13%
40%	0.47%	0.72	7.18%	92.31%
50%	0.37%	0.56	5.64%	97.95%
60%	0.10%	0.15	1.54%	99.49%
70%	0.03%	0.05	0.51%	100.00%
80%	0.00%	0	0.00%	100.00%
90%	0.00%	0	0.00%	100.00%
100%	0.00%	0	0.00%	100.00%

Test 結果

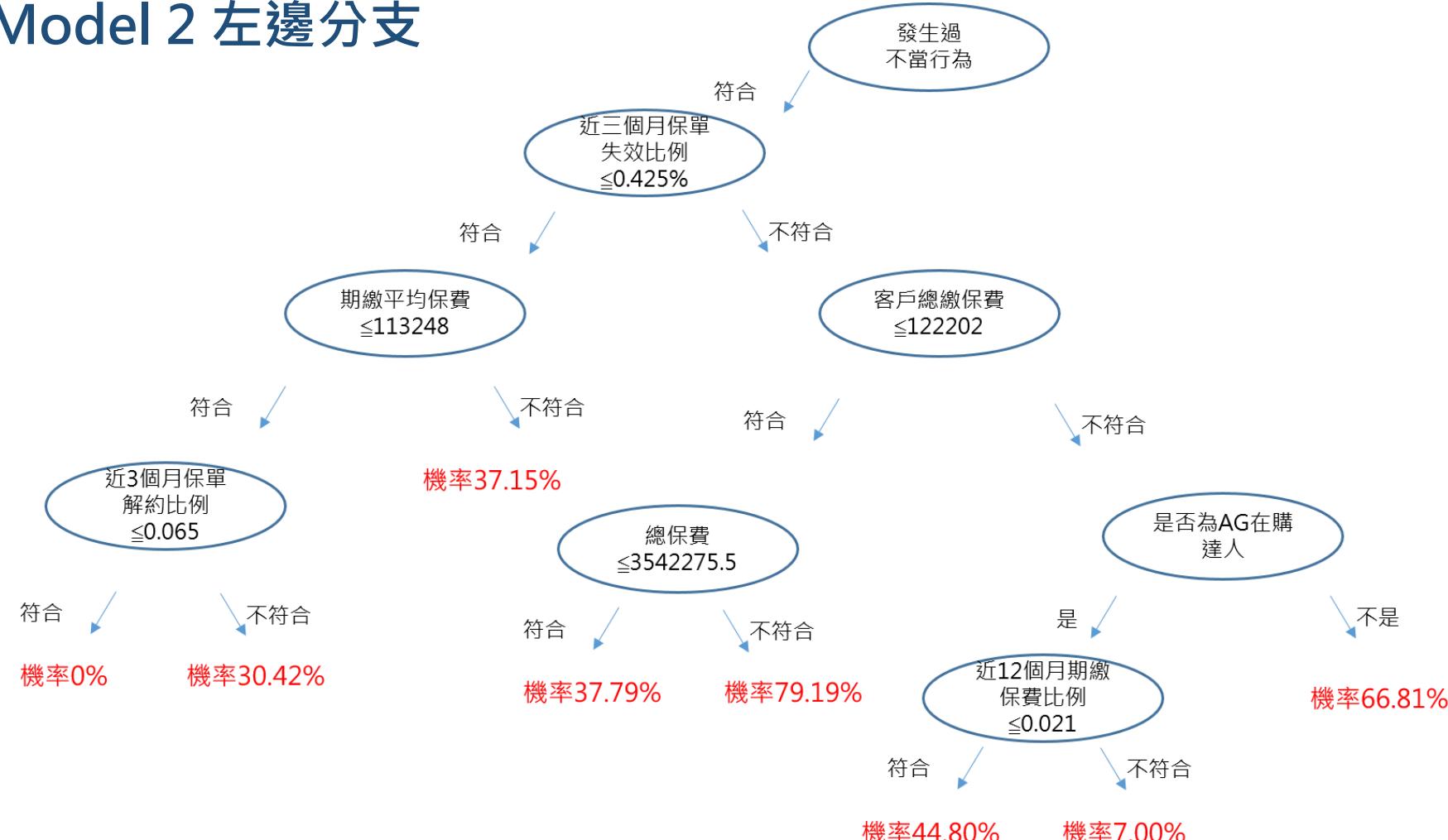
測試資料前5%捕捉率32.73%
增加3.64%捕捉率

	Bad 滲透度	倍數	Bad 占比	Bad 累積占比
5%	1.20%	6.55	32.73%	32.73%
10%	0.20%	1.09	5.45%	38.18%
20%	0.10%	0.55	5.45%	43.64%
30%	0.20%	1.09	10.91%	54.55%
40%	0.23%	1.27	12.73%	67.27%
50%	0.07%	0.36	3.64%	70.91%
60%	0.13%	0.73	7.27%	78.18%
70%	0.20%	1.09	10.91%	89.09%
80%	0.07%	0.36	3.64%	92.73%
90%	0.07%	0.36	3.64%	96.36%
100%	0.07%	0.36	3.64%	100.00%

Tree Model 2

03

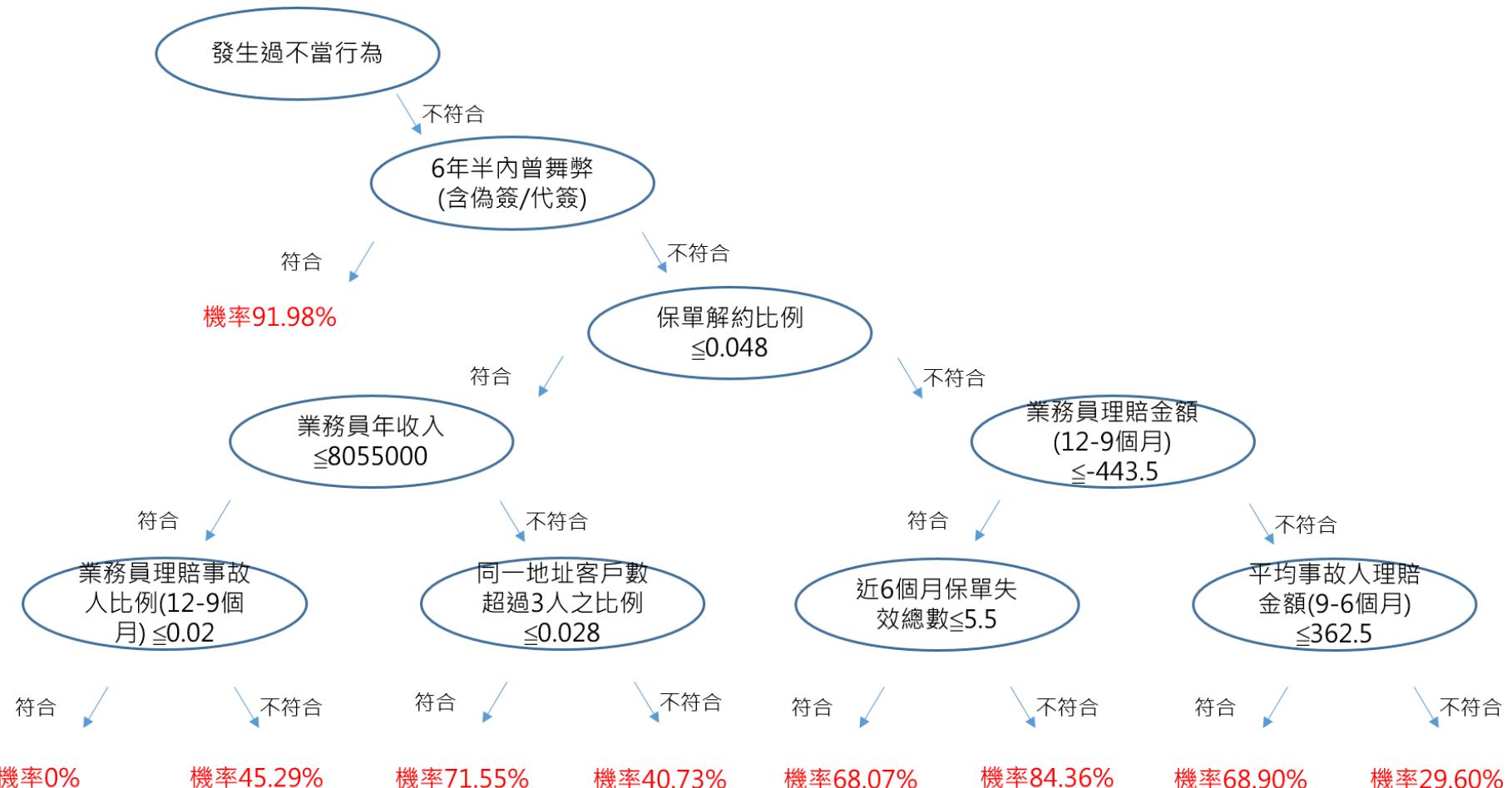
Tree Model 2 左邊分支



Tree Model 2

03

Tree Model 2 右邊分支



Tree Model 條件統整

03

✓ 條件 1

過往沒有不當行為, 近3個月保單失效比例 > 0.4%, Total AUM < 122202, 總保費 > 3542275.5

✓ 條件 2

過往有不當行為 → 過往舞弊行為Recency > 77 個月 → 保單解約比例 > 4.8% → 業務員近12 - 9個月平均(件數)理賠金額 <= -443.5
→ 近6個月保單失效總數 > 5.5

✓ 條件 3

過往有不當行為 → 過往舞弊行為Recency <= 77.5

✓ 條件 4

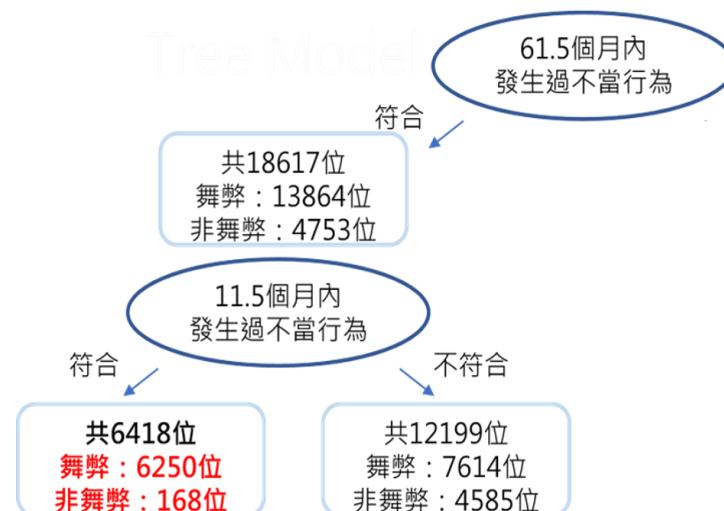
持有失效保單之客戶數占比 > 8.3% → 過往舞弊行為Recency <= 78.5

條件	符合人數	舞弊人數	舞弊發生率	捕捉率
過往不當行為Recency < 12個月	252	49	19.4%	25.13%
1	700	17	2.43%	8.72%
2	704	24	3.41%	12.31%
3	706	52	7.37%	26.67%
4	711	52	7.31%	26.67%

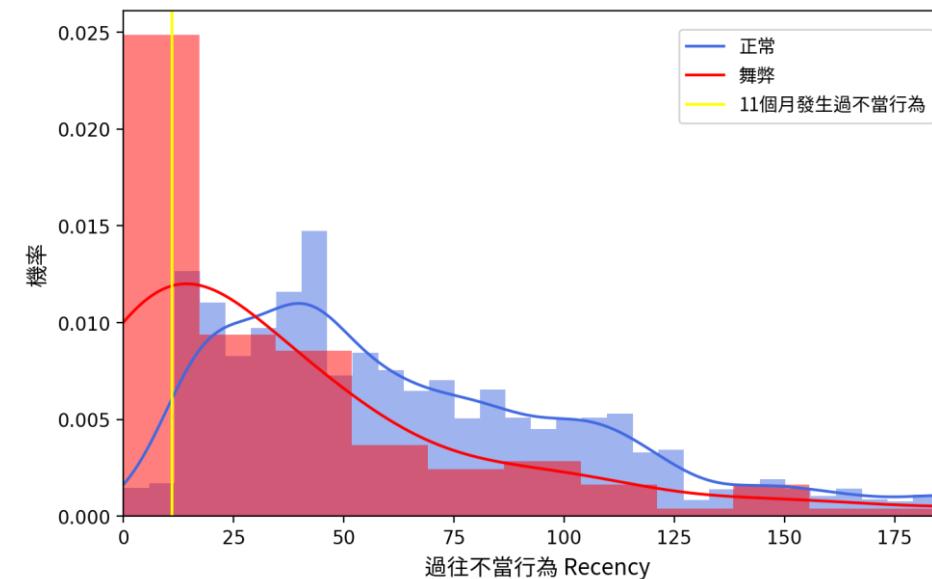
觀點一：不當行為發生時間越接近，越可能發生舞弊

以近期是否發生不當行為為出發點，過往不當行為的發生時間越近，越能抓出舞弊，若近期有發生不當行為，則**再犯可能性高**。

下圖可看見，當11個月內有不當行為，則能抓出大部分的舞弊業務員。



圖一

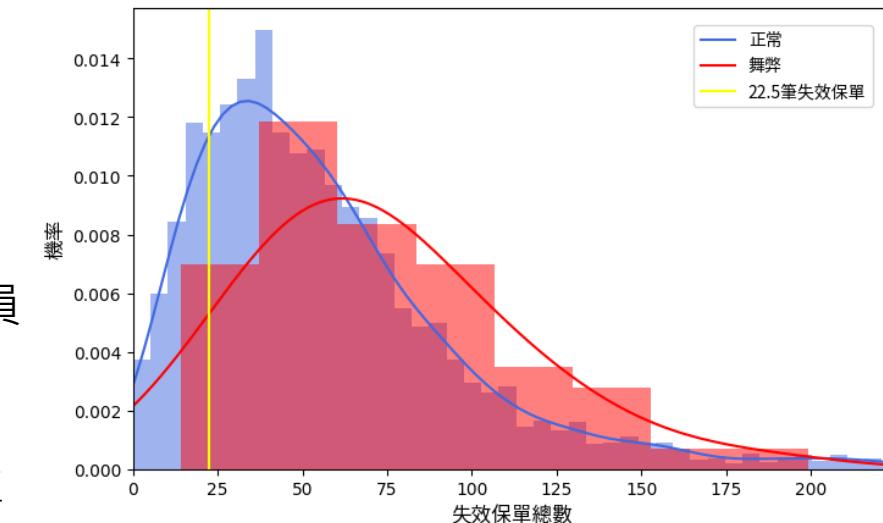


圖二

觀點二：

失效保單總數越高，舞弊可能性越高（須滿足觀點一條件）

過往不當行為在近期內發生，加上保單失效總數高，可能代表業務員態度不佳，或者未妥善處理客戶保單，導致失效保單增加。

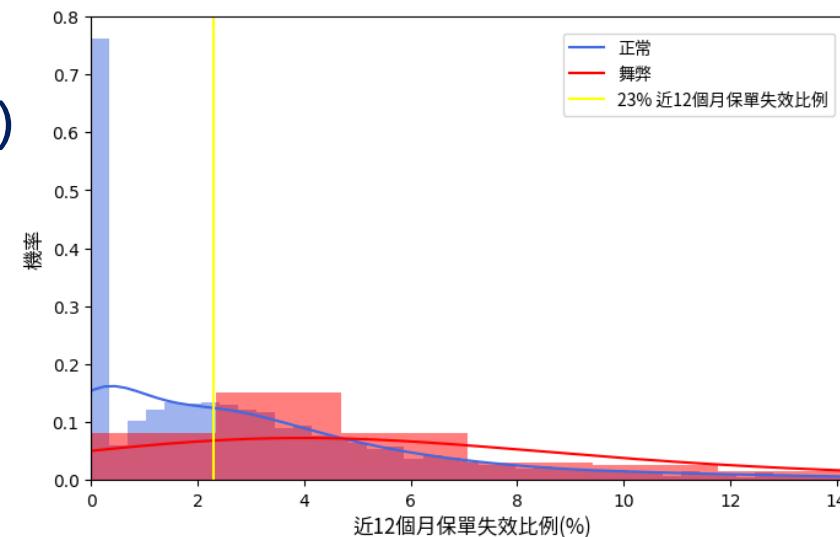


圖三

觀點三：

近期保單失效比例越高，舞弊可能性越高（須滿足觀點二條件）

若近12個月失效保單比例高，可能代表業務員最近出了狀況，更有可能從事舞弊行為。

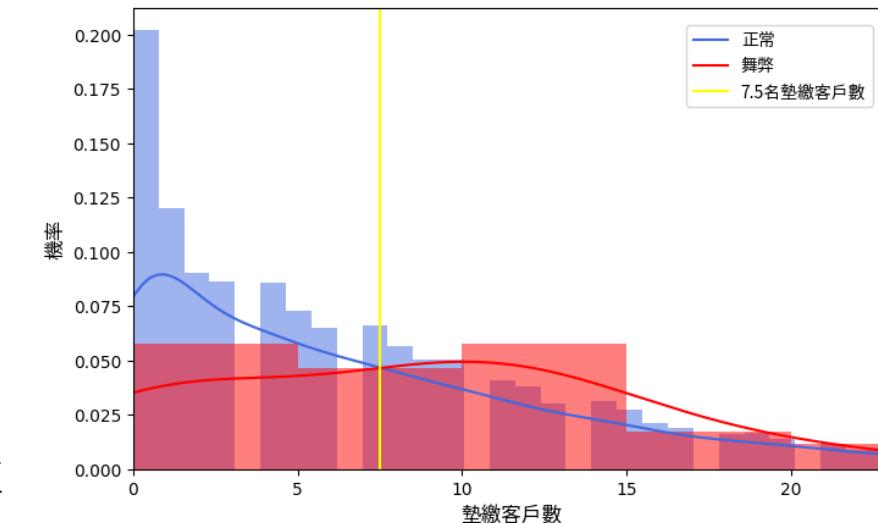


圖四

觀點四：

墊繳比例高，較可能出現舞弊（須滿足觀點三條件）

墊繳情況發生，表示客戶未準時繳費。與舞弊相關可能原因有業務員未按時催繳，沒有用心經營客戶；業務員主要經營短期關係，勸說親朋好友購買人情保單，以達成業績。

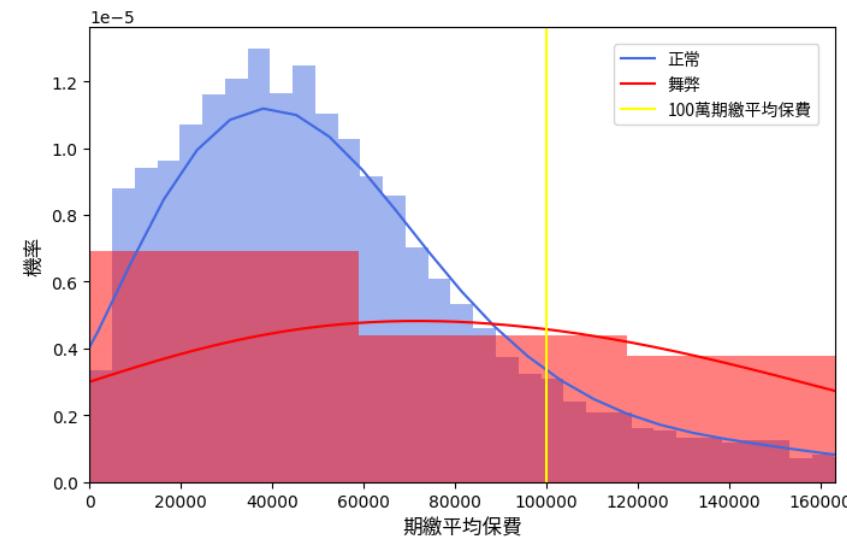


圖五

觀點五：

期繳平均保費高，較可能出現舞弊（須滿足觀點四條件）

客戶為每期繳交保費，比起一次繳清，業務員有更多的機會從中舞弊，控制金錢流動。



圖六

觀點七：

沒有舞弊行為者，業務表現可能影響舞弊機率

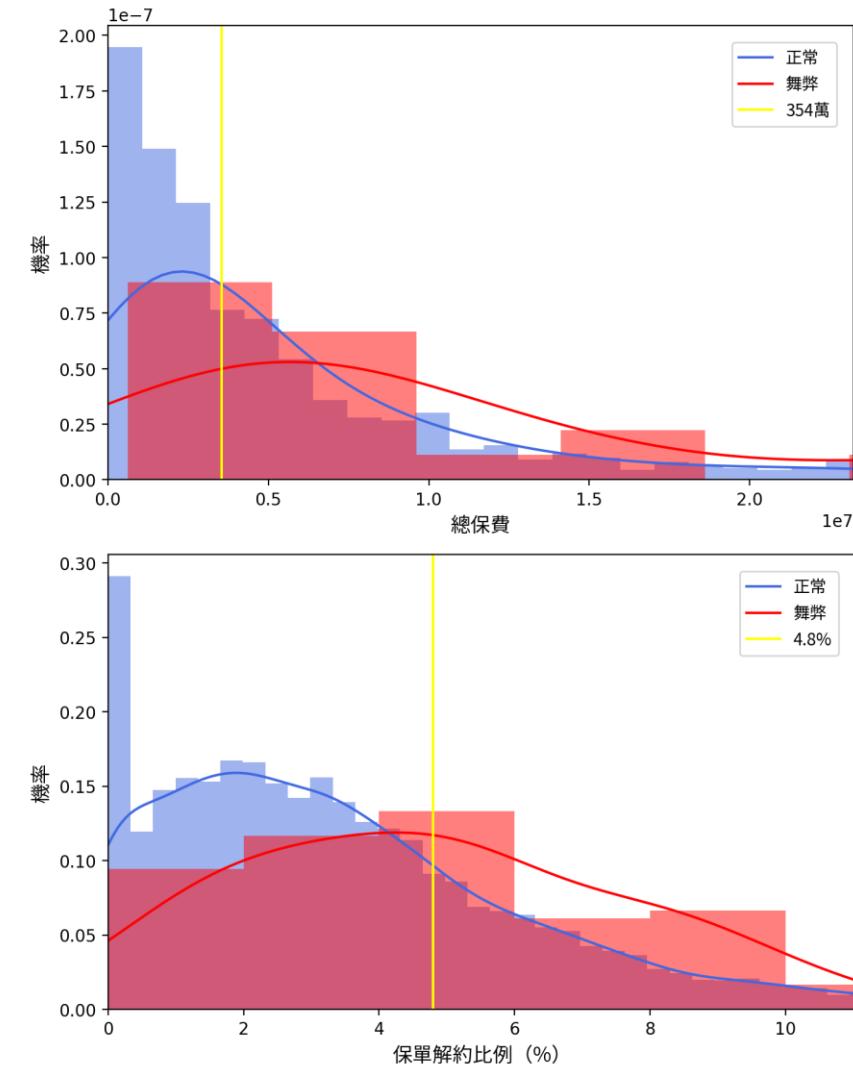
業務員在保單解約比例較高、客戶總繳保費低的前提下，短期總保費較高有較高的舞弊機率。

可能表示業務員為了挪用客戶保費，增加更多保單但總繳保費並未同時上升。客戶單方面發現後逕行解約，未進一步投訴。

觀點八：

77個月前曾舞弊者，保單解約比例高，較可能出現舞弊

保單解約比例高，可能為短線保單或未具備足夠專業，亦可能為客戶單方面發現舞弊行為而解約。

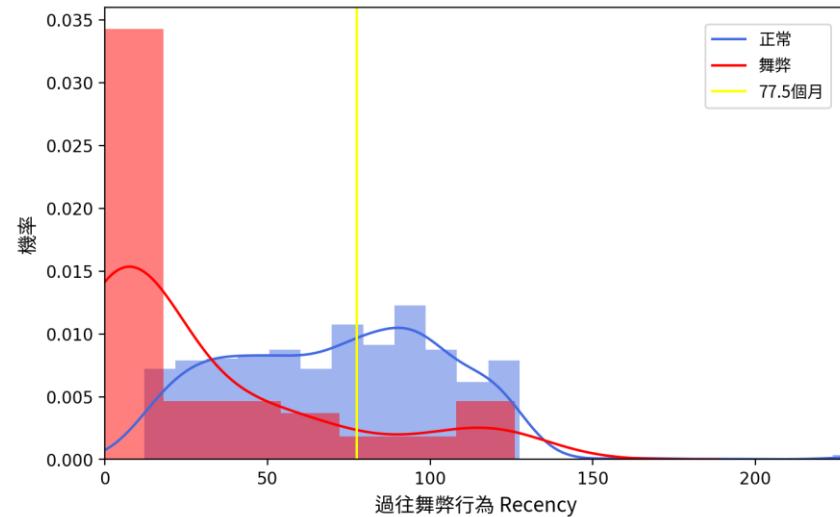


圖八

觀點九：

77個月內曾舞弊者，近期發生過舞弊，可能再犯率高

若近期內發生過舞弊，其舞弊機率明顯較高，可能是因為業務員的經濟狀況不佳，導致需要一再舞弊以獲取不當利益。

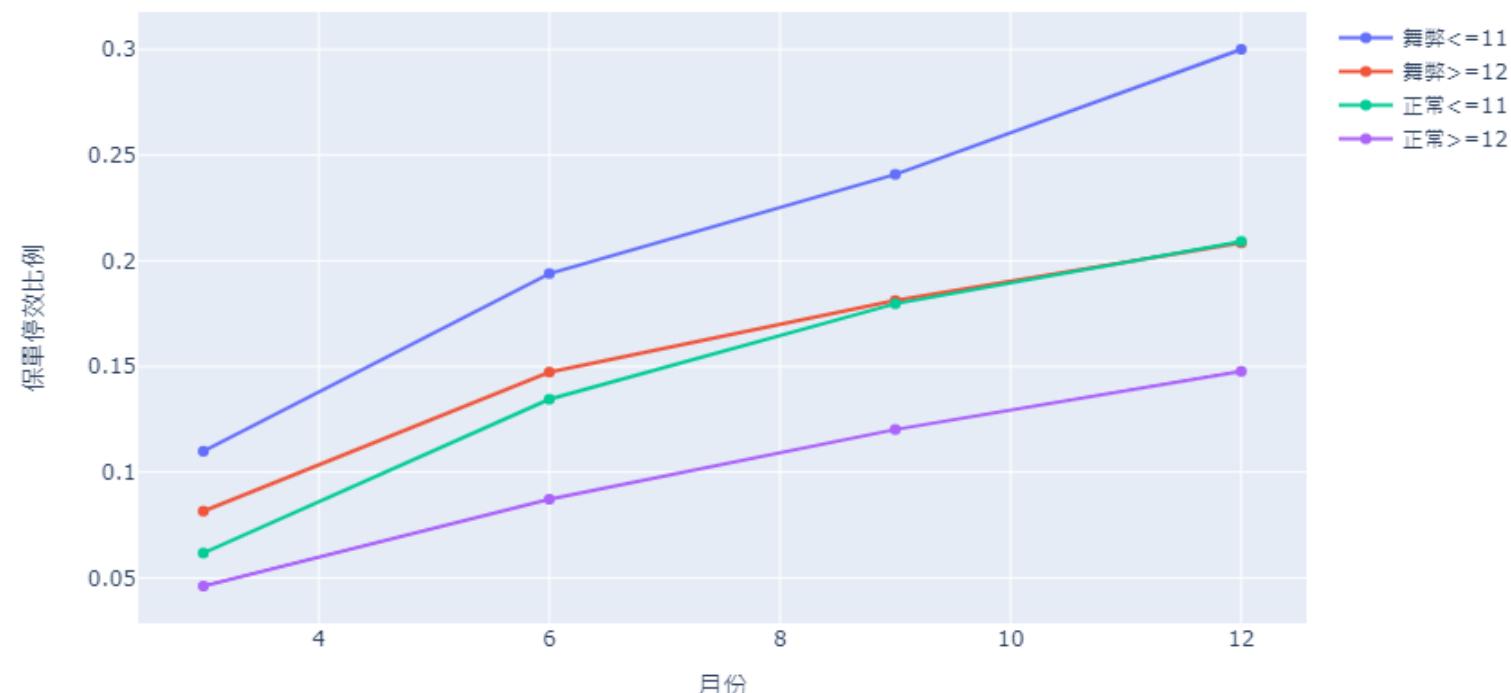


圖十

業務表現

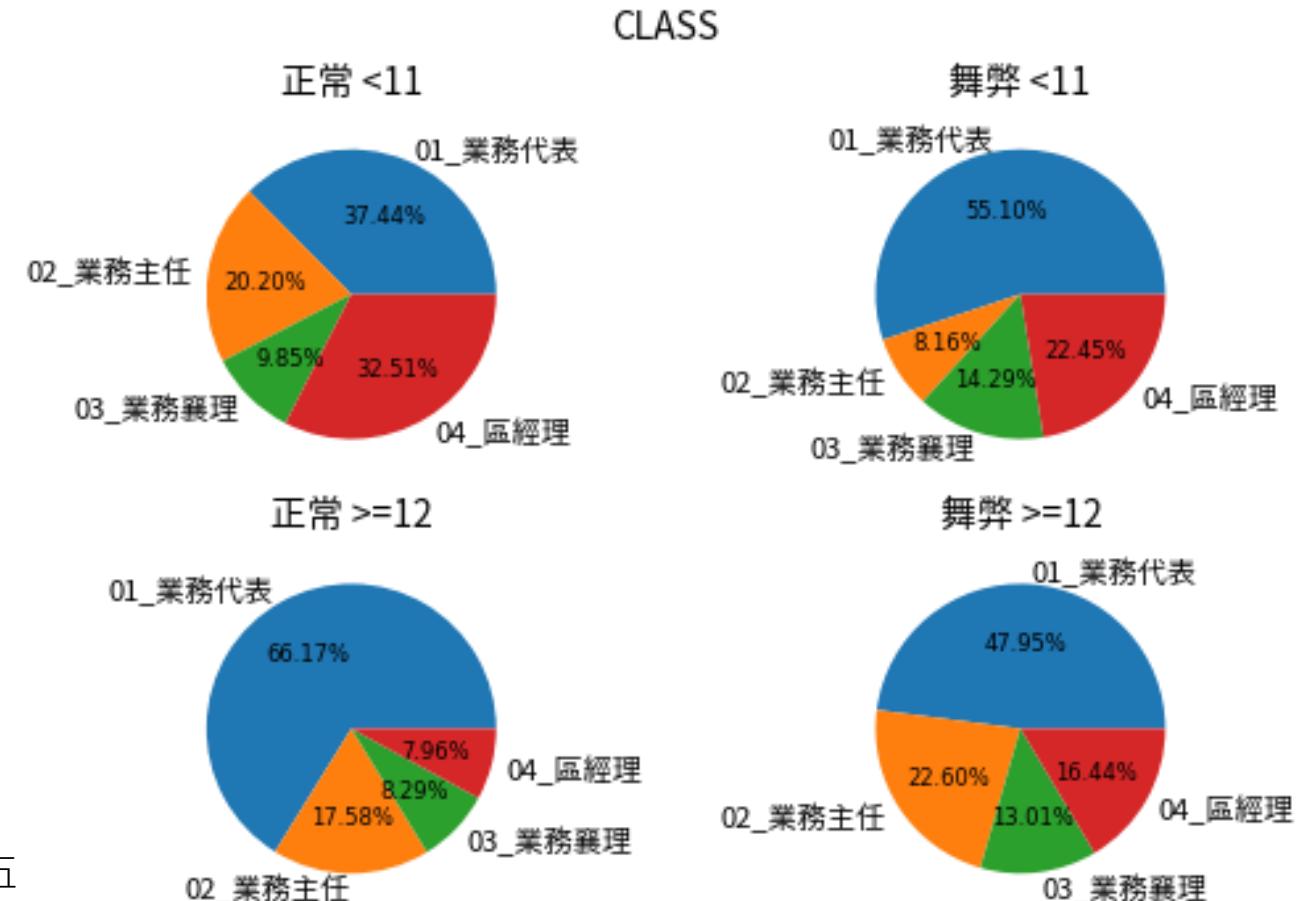
- ✓ 近期發生不當行為且發生舞弊行為者，短期保單停效比例皆明顯大於其他分類

短期保單停效比例



職級

- ✓ 近期有不當行為且發生舞弊者中，較多業務代表、較少業務主任

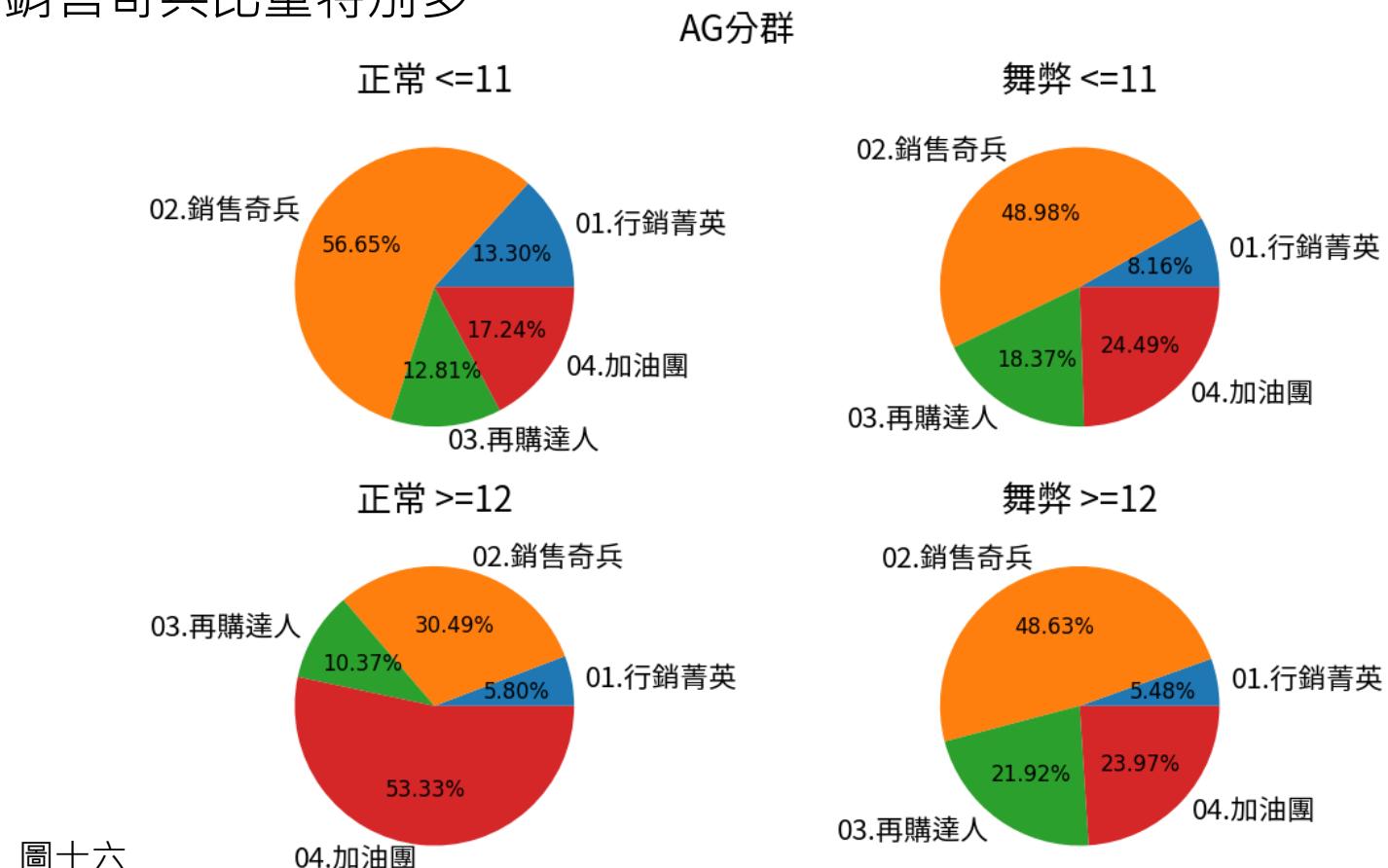


圖十五

05

AG 分群

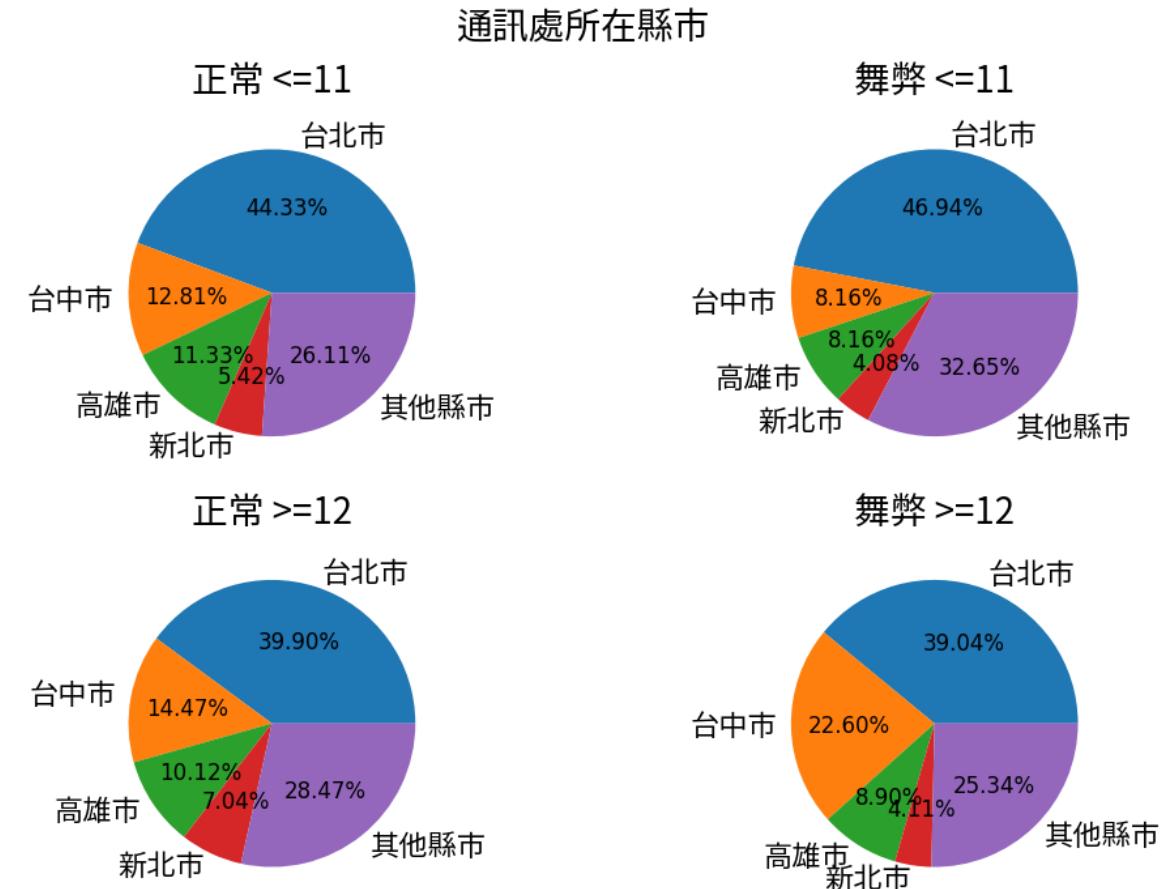
- ✓ 近期發生不當行為者，銷售奇兵比重特別多



圖十六

通訊處所在縣市

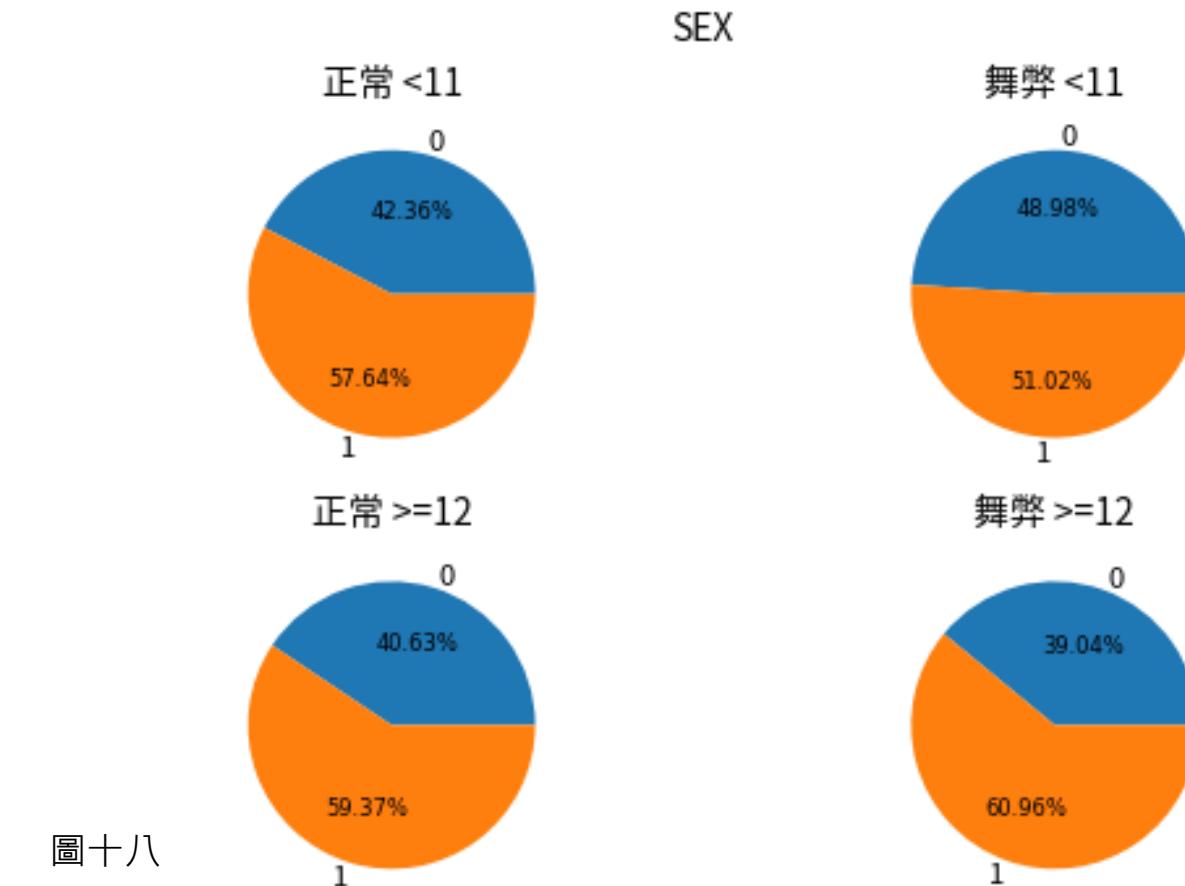
- ✓ 近期發生不當行為且舞弊者，以台北市為大宗
- ✓ 近期未發生不當行為且舞弊者，台中市比例大幅上升



圖十七

生理性別

- ✓ 近期有不當行為時，發生舞弊業務員男性佔比增加



圖十八

- 使用Decision Tree 建立舞弊風險評估模型所抓出的變數大致上都有實務上經驗可解釋之處
- 近期不當行為發生時間與之後是否舞弊有很密切的關聯
- 保單的失效、停效比例增加，都是偵測業務員風險的警訊
- 業務員的近期業務表現不佳可能成為業務員舞弊的誘因
- 業務員進行的業務類型可能為判斷舞弊風險的變數之一



感謝聆聽

小組分工

許哲榮：資料處理、建立及訓練模型、各變數圖表繪製

李宸帆：資料統整、舞弊變數分析、簡報製作

陸珮甄、楊永浚：繪製數值資料圖表

溫偉成：類別資料分析、3分鐘廣宣製作