

運用機器學習預測台灣上市公司 財務風險之研究

以財報與市場數據為基礎



412580096 智慧資安二 曾雁宸

412580149 智慧資安二 楊晏丞

目錄

1. 研究背景與動機
2. 研究目的與問題定義
3. 研究流程與方法架構
4. 資料來源與樣本說明
5. 特徵建構與資料處理
6. 財務風險定義與標記方式
7. 模型建構與評估指標
8. 實驗結果與分析討論
9. 應用場景與研究貢獻
10. 結論與未來方向



研究背景與動機

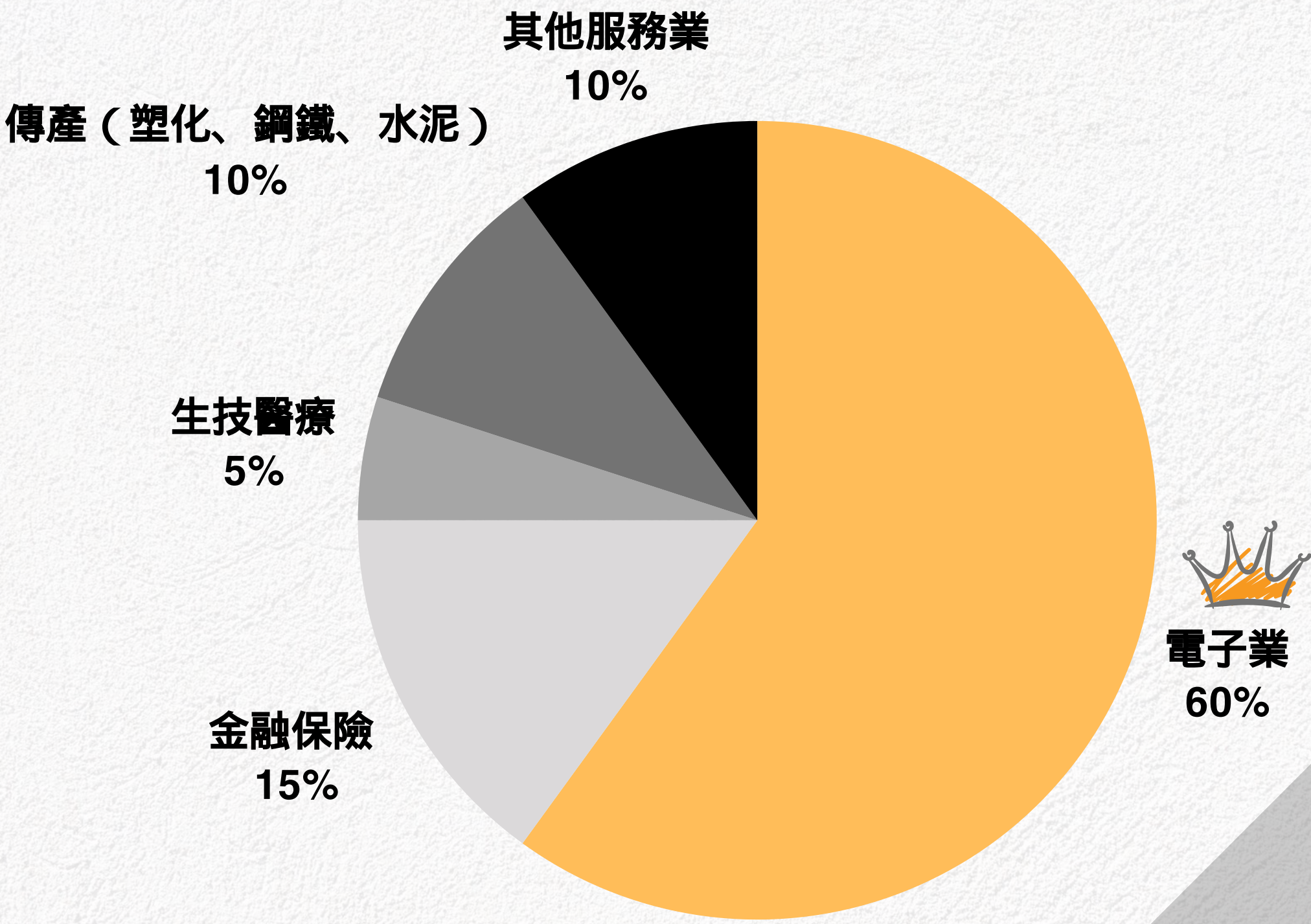
台灣電子業為資本市場主力，其財務穩健與否直接影響股市波動與經濟發展。然而，傳統如 Z-score、Altman 模型難以掌握即時市場動態，忽略非線性與高維度資料的潛在關聯性。

本研究動機源於：

- 台灣電子業為資本市場主力
- 財務風險事件影響重大
- 傳統模型侷限性明顯
- 金融科技開啟新機會



研究背景與動機



▲ 台灣產業市值分布

資料來源：TWSE 市值週報



研究背景與動機

比較項目	傳統財務預警模型 (如 Z-score)	機器學習模型 (如 XGBoost、RF)
📌 資料來源	僅使用財報數據	財報＋市場數據（多元來源）
🕒 資料更新頻率	季報/年報（低頻）	可納入每日股價變化（高頻）
📈 模型彈性	固定公式，彈性低	高度可調整參數、適應新資料
🔍 變數間關係	須由人為設定（線性假設）	自動學習非線性、交互關係
🧠 特徵選擇	依研究者主觀選擇	可自動計算特徵重要性並選擇
🎯 預測準確性	適用於特定產業或時期	可跨產業應用，具泛化能力
🚫 資料不平衡處理	不易處理	有多種技術（如SMOTE、加權學習）支援



▲ 預警能力比較表

研究目的與問題定義

本研究目標如下：

- 建構一套針對台灣上市電子業公司之財務風險預測模型
- 探討財報資料與市場指標在模型中的貢獻度
- 評估不同機器學習模型之效能表現與應用可行性

「如何透過多源資料與機器學習演算法，預測上市公司財務風險事件？」



研究流程與方法架構

研究流程分五大階段：

1. 資料範圍與樣本公司確定
2. 資料清洗與指標建構
3. 模型建立與訓練
4. 結果分析與評估
5. 撰寫結論與提出未來應用方向



研究流程與方法架構

1

財務資料

- 財報（資產負債表、損益表、現金流量表）
- 財務比率（如流動比率、負債比率、營收成長率等）

2

市場資料

- 股票價格、成交量
- 波動率指標（如VIX）

3

產業與宏觀經濟資料

- 利率、匯率、GDP、產業景氣指數
- 同業競爭者表現

資料來源與樣本說明

這份測試用資料是GPT根據常見的財務風險研究設計原則，構造出來的小型模擬資料集目的是可以先驗證機器學習模型流程，之後再擴展應用到實際財報數據。

📌 特別說明(From GPT)：

- 我刻意設計了一些「正常」公司和一些「風險」公司，使得機器學習模型能從中學到某些變數的預測價值。
- 雖然這是模擬資料，但設計邏輯是仿真真實財報分析中可能出現的關聯與特徵。



模擬資料庫

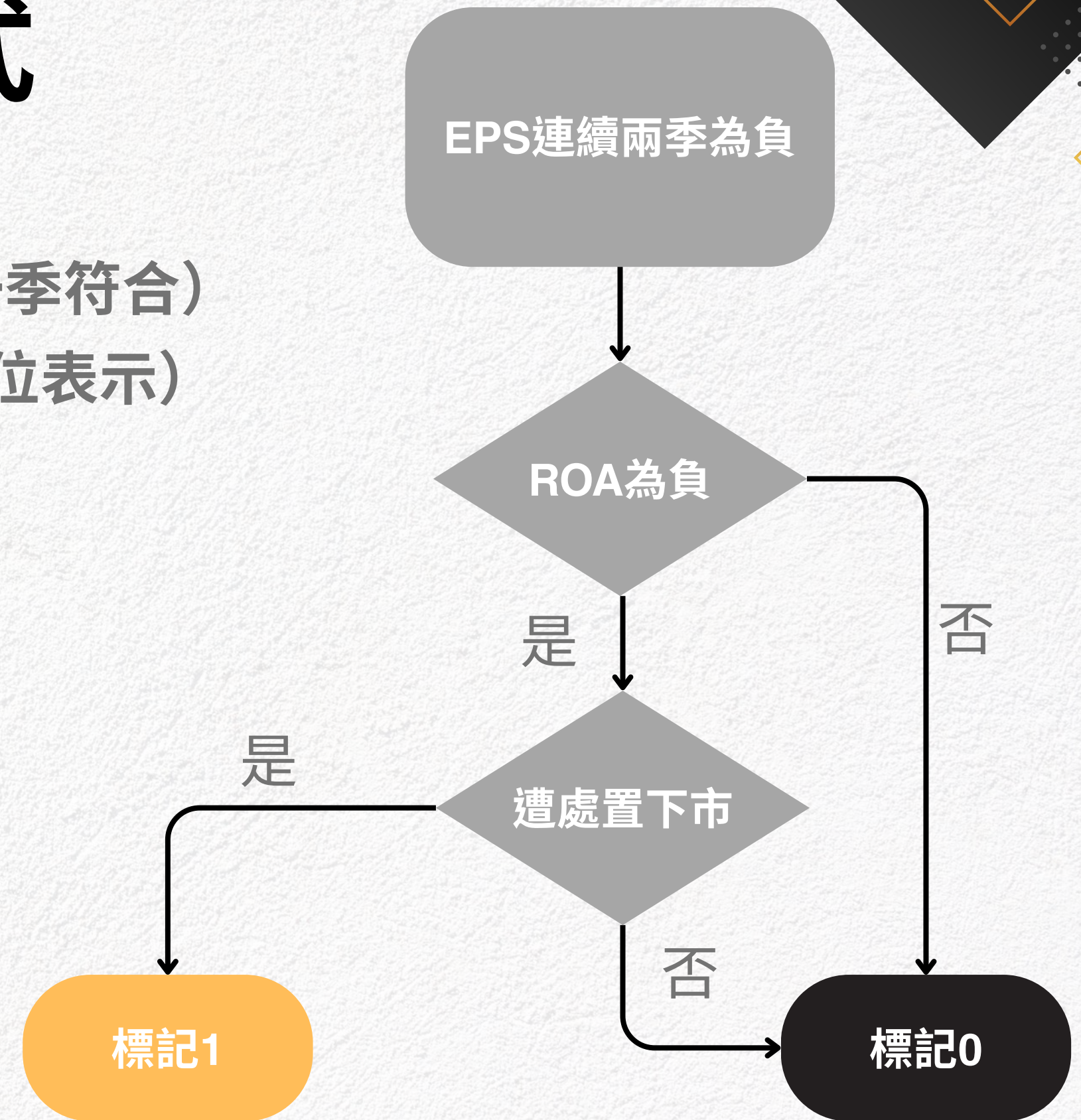
特徵建構與資料處理

- ◆ 財報指標（來自財報欄位）：
 - 流動比率：衡量公司短期償債能力
 - 負債比率：衡量資本結構風險
 - ROA、EPS：反映獲利與資產使用效率
- ◆ 市場資料（來自即時市場與新聞）：
 - 營收年增率：代表營運成長趨勢
 - 媒體情緒分數：自動分析新聞情感性
- ◆ 資料處理流程：




財務風險定義與標記方式

- ◆ 風險定義準則（程式中實作）：
 - 連續兩季 $EPS < 0$ 且 $ROA < 0$ （簡化為任一季符合）
 - 被處置、下市等風險事件（模擬中以標記欄位表示）
- ◆ 資料標記方式（是否風險事件欄位）：
 - 若公司該季資料符合風險定義，即標記為 1
 - 否則標記為 0
 - 實務上應考量連續季觀察與事件公告資訊



模型建構與評估指標

- ◆ 使用模型（程式中建構）：
 -  Random Forest：擅長處理非線性與變數間交互作用
- ◆ 評估指標（以 sklearn 顯示）：
 - Accuracy：整體正確率
 - Precision / Recall：針對風險事件預測的精準與完整性
 - 混淆矩陣 + 分類報告：模型效能完整呈現
- ◆ 模型效果亮點：

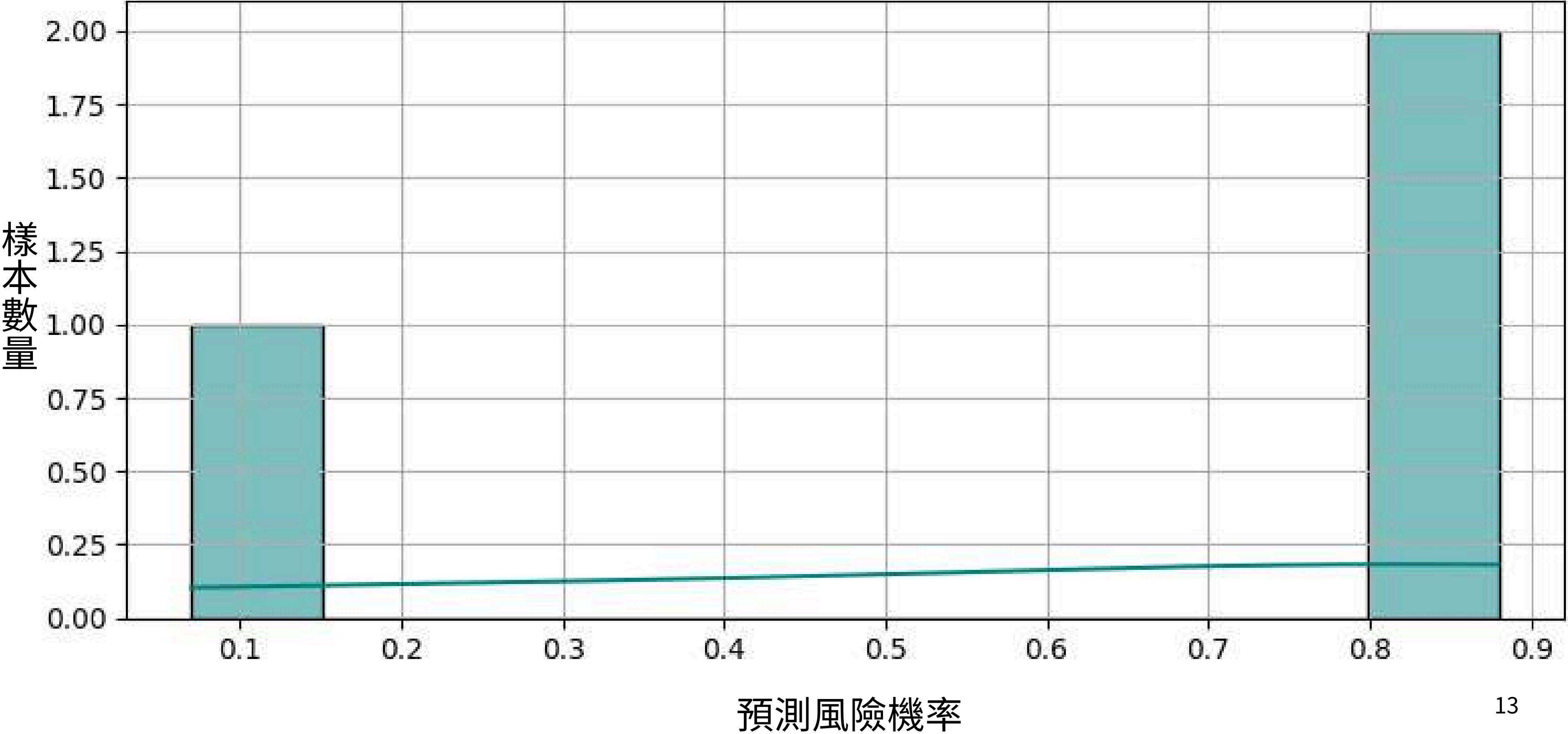


自動學習變數間非線性關係

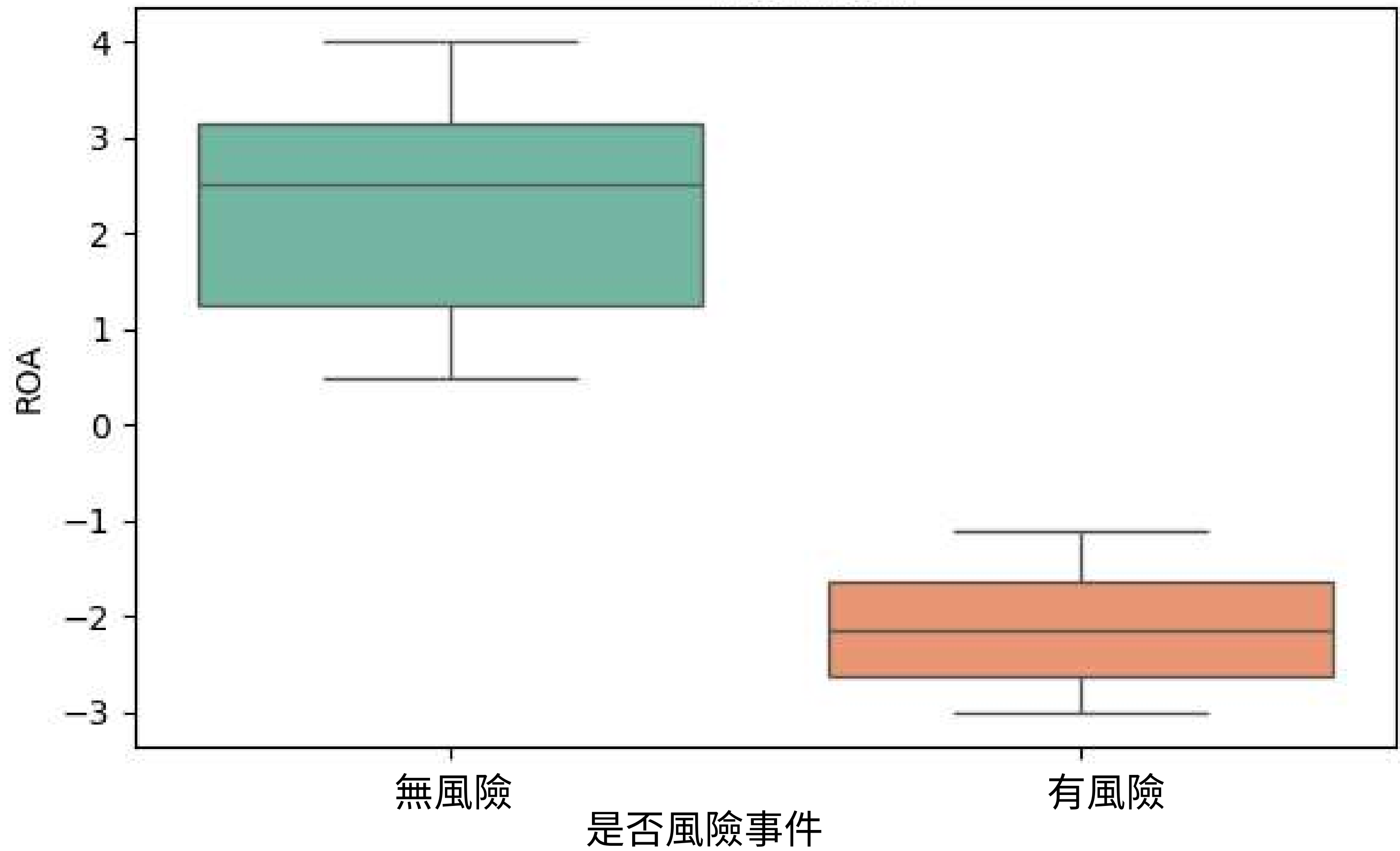


擅長處理多元異質特徵
(如 EPS + 媒體情緒)

預測風險機率分布

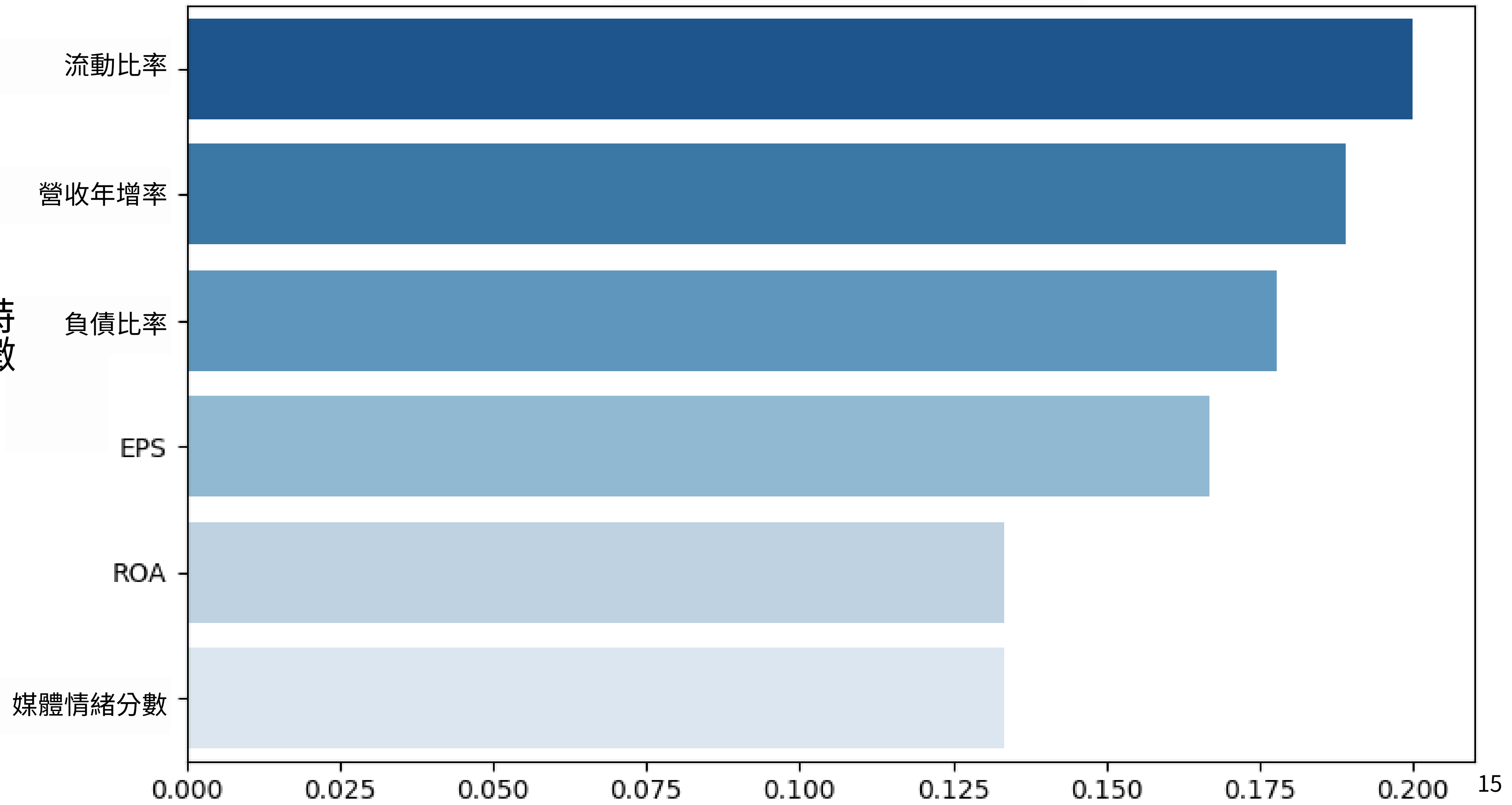


ROA與風險事件的關聯



模型特徵重要性模型特徵重要性

特徵



重要性

🇺🇸 所有公司風險預測結果：

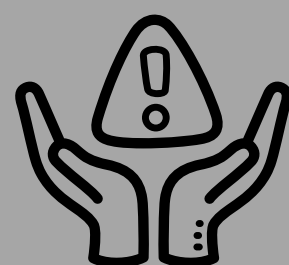
公司代號	預測風險機率	實際是否風險事件
1108	0.90	1
1106	0.88	1
1102	0.87	1
1104	0.86	1
1109	0.07	0
1101	0.00	0
1105	0.00	0
1103	0.00	0
1107	0.00	0
1110	0.00	0

應用場景與研究貢獻



投資決策輔助系統

- 根據預測結果輔助法人、散戶做出投資選股判斷
- 將風險分數納入多因子選股模型中



金融監理單位風控工具

- 協助主管機關或交易所，提前發現潛在財務異常企業
- 強化市場穩定機制與預警能力



公司內部財務預警系統

- 可監控關鍵指標、自行預測財務危機風險
- 適用企業風險控管部門或稽核系統



應用場景與研究貢獻

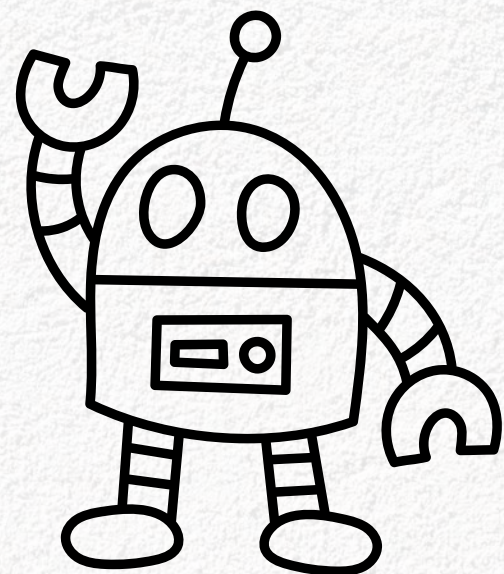


財報資料 + 市場行為
(如媒體情緒) 整合模型

VS

優於單一資料來源
(如僅用財報)

機器學習模型提升預測效能
捕捉非線性特徵與變數交互作用



結論與未來方向



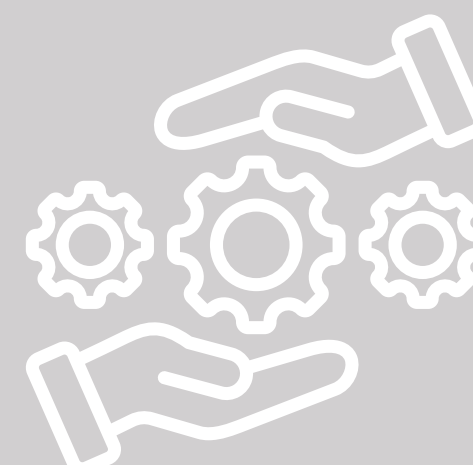
納入
非結構化資料



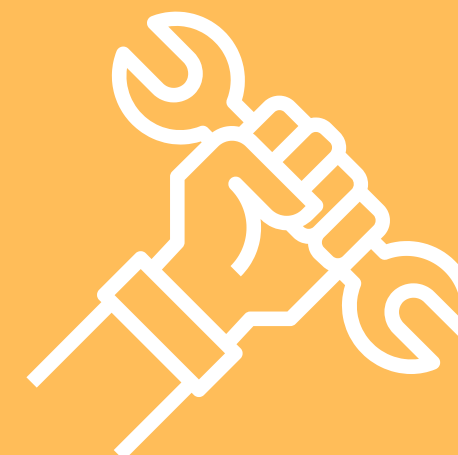
即時預警系統



擴展至其他產業



多模型融合



實務應用工具

參與競賽

2025第13屆資訊應用服務創新創業新秀選拔

初選、複選評分項目：

創新與創意	30%
可實現性	30%
技術可行性	20%
市場可行性	10%
團隊執行力	10%

THANK YOU

