

# 氣候相關財務揭露報告書 的文字分析

指導教授：余清祥 博士

研究生：徐韶汶



# 目錄

1 研究目標

4 寫作風格分析

2 文獻與研究素材

5 結論與建議

3 評分模型

6 限制與展望





## 研究目標 | 輔助專家進行評鑑



建立文字探勘評分模型，作為檢核人工計分的輔助工具

區隔四大要素寫作風格，檢視要素是否分工明確揭露資訊，提供指引訂定建議

準備台灣高質量氣候變遷數據資料集

# 文獻回顧 | TCFD框架缺乏一致的評分方法

## Amar等人（2020）

- 過去40年，通常用綜合指數衡量企業環境揭露程度
- 提出綜合指標Climate Risk and Opportunity Reporting Index （ CRORI ）
- 以延伸TCFD揭露事項的方式設計虛擬變數

→問題：綜合指標無法說明程度、權重設置缺乏科學依據

→解決：我們改良評分方法並驗證其有效性

# 文獻回顧 | TCFD建議的揭露事項缺乏具體定義

Auzepy 等（2023）

- BART零樣本文本分類（Zero-Shot Text Classification）自動多種標籤標記，檢視各要素揭露的主題
- 銀行選擇性揭露氣候風險資訊，比如在碳密集型行業的投融資揭露率低

→問題：因為規範模糊，企業只揭露對其有利資訊等漂綠行為

→解決：我們初步檢視各要素的揭露內容是分工明確

## 文獻回顧 | 氣候變遷開放數據集有限

Webersinke等（2022）、Diggelmann等（2020）、Morio等（2023）

- 提出氣候相關數據集

→問題：相較於大語言模型需要的資料量，氣候開放數據集有限

→解決：我們人工標記82家本國銀行+保險業的永續報告書數據集

# 研究素材

38家銀行、21家壽險及23家產險2022年的TCFD或ESG永續報告書

## 評分模型 (單位) 整本報告書

| 樣本數  | 銀行 | 壽險 | 產險 |
|------|----|----|----|
| TCFD | 31 | 15 | 11 |
| ESG  | 7  | 6  | 12 |
| 總樣本數 | 38 | 21 | 23 |

## 寫作風格分析 (單位) 要素

| 四大要素 | 銀行  | 壽險 | 產險 |
|------|-----|----|----|
| 治理   | 31  | 16 | 11 |
| 策略   | 31  | 16 | 11 |
| 風險   | 31  | 16 | 11 |
| 指標   | 30  | 16 | 11 |
| 總樣本數 | 123 | 64 | 44 |



# 研究素材

FSB “2021 Recommendations of the Task Force on Climate-related Financial Disclosures”  
金管會“本國銀行業氣候風險財務揭露指引”、“保險業氣候相關風險財務揭露指引”

| TCFD 指標對照表 |   |                             |
|------------|---|-----------------------------|
| 層級         | 一般行業揭露建議                                  | 章節                          |
| 治理         | A. 描述董事會監督氣候相關風險與機會                       | 1.1 董事會對氣候相關風險與機會的監督情況      |
|            | B. 描述管理階層在評估與管理氣候相關風險與機會之作用               | 1.2 管理階層在評估與管理氣候相關風險與機會的情況  |
| 策略         | A. 描述組織已鑑別出之短、中、長期的氣候相關風險與機會              | 3 氣候風險與機會                   |
|            | B. 描述會對組織業務、策略與財務規劃有產生重大衝擊的氣候相關風險與機會      | 3 氣候風險與機會                   |
|            | C. 描述組織的策略韌性，將氣候變遷不同的情境納入考量，包括 2°C 或更低的情境 | 3.3 實體風險與風險分析<br>3.4 轉型風險分析 |
| 風險管理       | A. 描述組織鑑別和評估氣候相關風險的流程                     | 3.1 氣候風險鑑別與評估               |
|            | B. 描述組織管理氣候相關風險的流程                        | 2 氣候風險管理                    |
|            | C. 描述組織在鑑別、評估和管理氣候相關風險的流程，如何整合納入整體的風險管理   | 2 氣候風險管理                    |
| 指標及目標      | A. 揭露組織在符合策略與風險管理流程下，使用於評估氣候相關風險與機會的指標    | 5 指標與目標                     |
|            | B. 揭露範疇一、二、三（若適用）的排放量與相關風險                | 4.1 低碳營運                    |
|            | C. 描述組織在管理氣候相關風險與機會之目標，以及相關目標之表現績效        | 5 指標與目標                     |



|    |    |                                 |    |                   |
|----|----|---------------------------------|----|-------------------|
| 目錄 | 2  | 總覽                              | 22 | 2.2.3   低碳營運      |
|    | 4  | 2022 執行成果                       | 24 | 2.2.4   低碳轉型      |
|    | 5  | 董事長的話                           | 34 | 2.2.5   溝通與倡議     |
|    | 6  | 關於本報告書                          |    | 第三章 風險管理          |
|    |    | 第一章 治理                          | 38 | 3.1   辨識、衡量與管理流程  |
|    | 8  | 1.1   職能與職責 - 董事會對氣候相關風險與機會監督   | 42 | 3.2   氣候變遷財務量化分析  |
|    | 10 | 1.2   職能與職責 - 管理階層對氣候相關風險與機會的管理 | 58 | 3.3   氣候變遷情境分析    |
|    | 13 | 1.3   歷年氣候治理之重大決策與執行成果          |    | 第四章 指標與目標         |
|    |    | 第二章 策略                          | 63 | 4.1   低碳營運管理指標與目標 |
|    | 16 | 2.1   氣候風險與機會鑑別與衡量              | 67 | 4.2   低碳轉型管理指標與目標 |
|    | 19 | 2.2   氣候策略制定                    |    | 第五章 未來展望          |
|    | 19 | 2.2.1   科學基礎減碳目標                | 73 | 附錄                |
|    | 21 | 2.2.2   內部碳定價                   |    |                   |

元大銀行 2022 年 氣候相關財務揭露報告書 | 目錄

# 研究素材

每一個樣本是多章節、多頁、多段的文字組合

氣候策略

2-2 氣候議題鑑別及評估流程

為因應氣候變遷相關的風險與機會，由TCFD小組考量氣候風險類型<sup>註1</sup>、時間區間、影響程度、發生可能性、金融業既有風險之關聯性及各項業務性質，確立氣候風險辨識流程，每年辨識或評估氣候風險與機會，定期追蹤監控氣候相關執行情形。

圖3 氣候風險與機會辨識流程

STEP 1. 氣候相關議題鑑別

STEP 2. 辨識影響商業策略與目標的氣候相關風險與機會

STEP 3. 評估氣候相關風險與機會影響程度

STEP 4. 評估風險與機會之重大性

STEP 5. 擬訂合適的因應措施

STEP 6. 氣候風險與機會辨識流程整合

表3 氣候風險與機會－時間範圍

| 時間區間 | 短期   | 中期   | 長期   |
|------|------|------|------|
| 年度   | 1年以內 | 1至7年 | 7年以上 |

註1：主要劃分為兩大類：與氣候變遷相關的「實體風險」及與低碳經濟相關的「轉型風險」，實體風險依據氣候模式區分為立即性極端天氣事件或長期性氣候模式變化，轉型風險依據不同影響因子再細分為政策和法規風險、技術創新風險、市場結構改變、聲譽風險等主要風險因子。

表4 氣候風險與機會－財務影響程度

| 影響程度 | 定義                           |
|------|------------------------------|
| 輕度   | 財務影響或效益低於新臺幣1,000萬元          |
| 中度   | 財務影響或效益介於新臺幣約1,000萬元~2,500萬元 |
| 高度   | 財務影響或效益介於新臺幣約2,500萬元~3.8億元   |
| 重大   | 財務影響或效益介於新臺幣約3.8億~5億元        |
| 總大   | 財務影響或效益超過新臺幣5億元              |

TCFD小組每年定期執行氣候風險與機會辨識流程，並進行質化/量化分析，依據結果對各項風險與機會研擬因應策略，每月進行監控與報告，提報風險管理委員會及董事會。

2-3 以情境分析評估策略韌性

情境分析的範圍及邊界

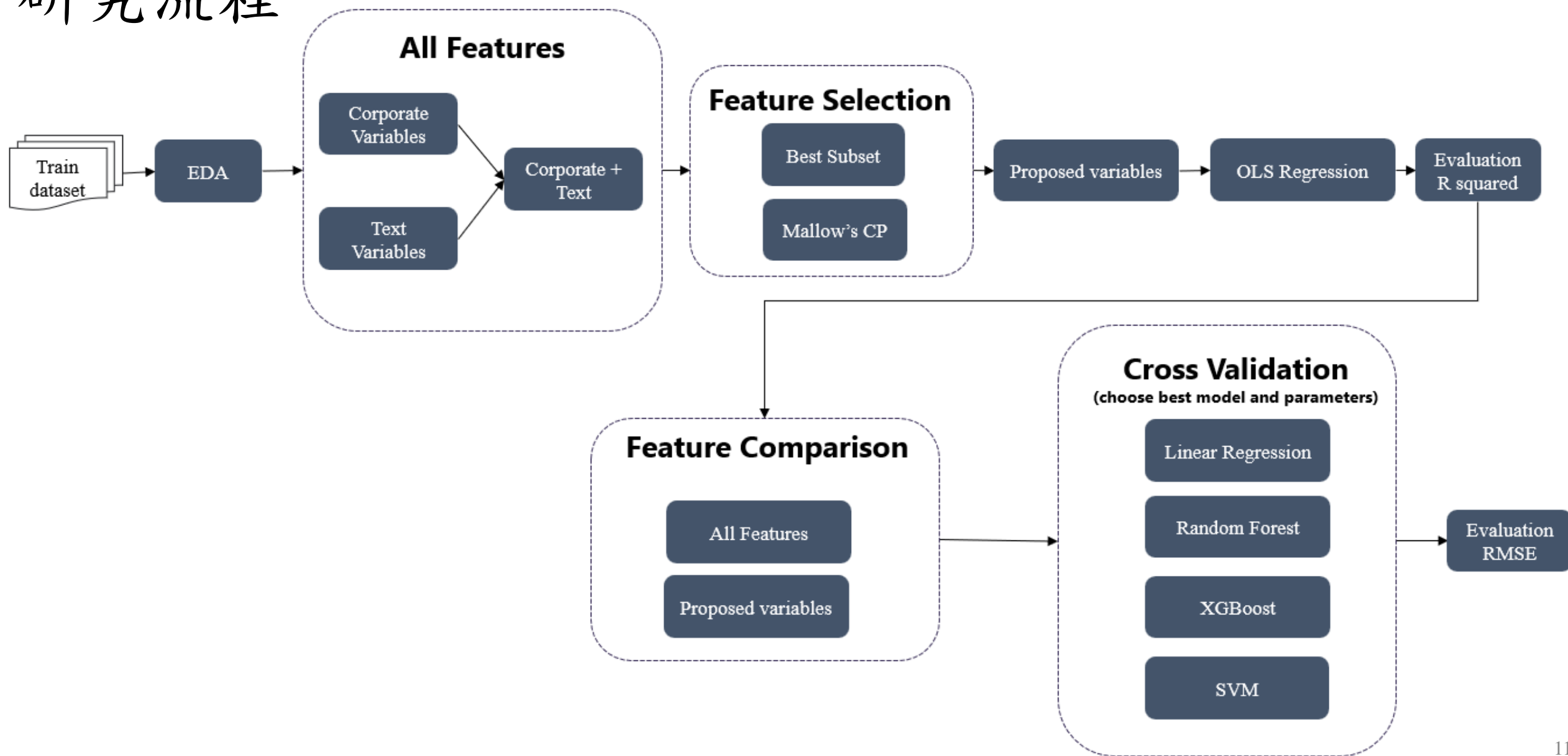
彰化銀行2022年淨收益有72%以上來自利息，因此彰化銀行特別關注氣候變遷議題對授信戶的影響。

2022年度轉型風險層面之情境分析，盤點彰化銀行授信對象屬於國內環保署納管之應盤查登錄溫室氣體排放量之排放源者，瞭解未來政策與法規對授信對象營運之影響，進而影響彰化銀行信用風險之預期損失金額；於實體風險層面，以彰化銀行全臺營運據點及所有授信戶之不動產擔保品為分析對象，觀察擔保品所在區域於2022、2025、2030、2040，以及2050年之氣候風險值(VaR)變化，以掌握極端氣候對資產減損影響之金額。

08 氣候策略

10

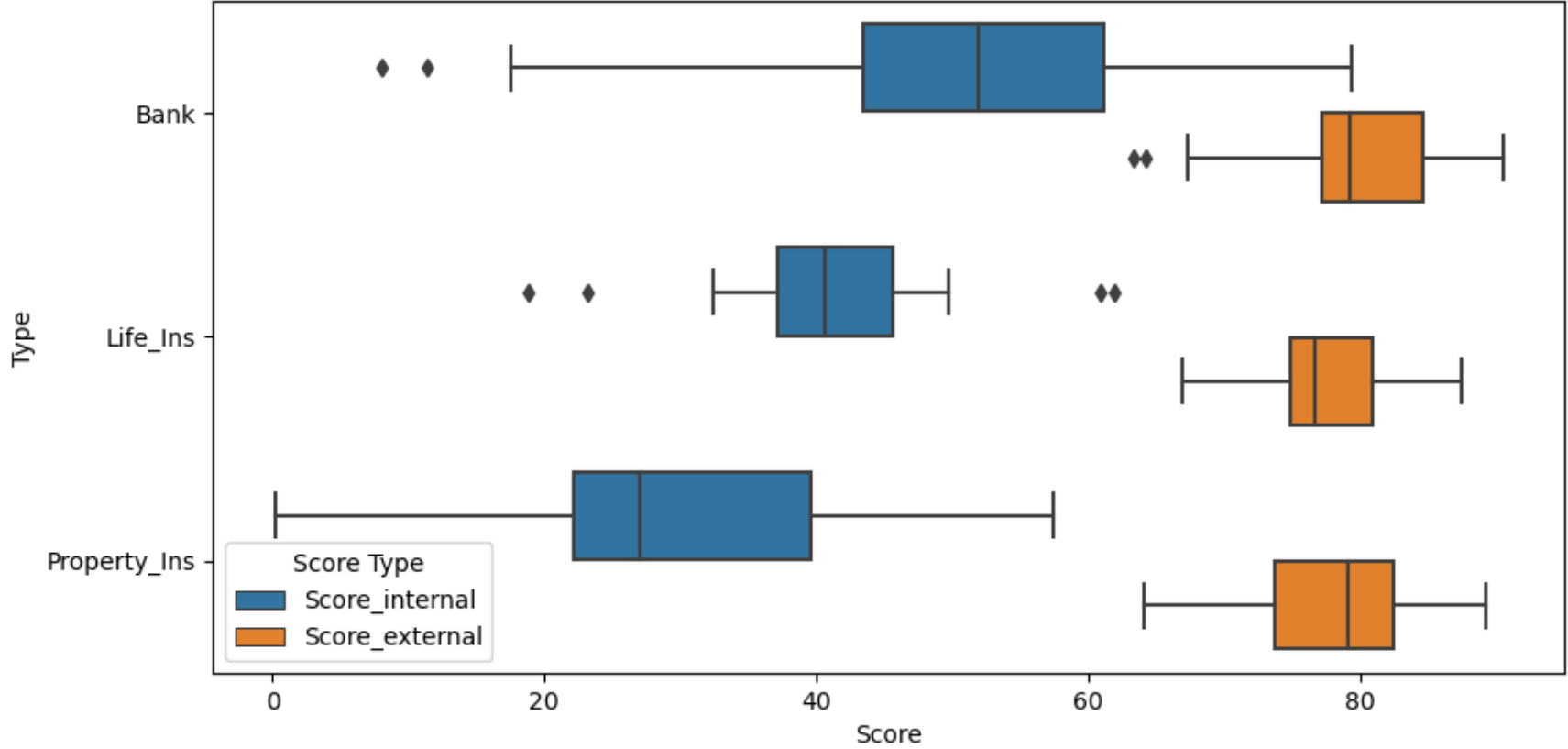
# 研究流程



# 變數設計

| 公司變數         |              | 文字變數                  |  |
|--------------|--------------|-----------------------|--|
| 銀行           | 壽險、產險        | 銀行、壽險、產險              |  |
| 1 銀行狀態       | 1 有金控母公司     | 1 總字數                 |  |
| 2 公股         | 2 金控母公司是公股   | 2 不同字數                |  |
| 3 資本額        | 3 資本額        | 3 隨機抽樣500字之平均TTR      |  |
| 4 TCFD_會計師確信 | 4 TCFD_會計師確信 | 4 隨機抽樣500字之平均Entropy  |  |
| 5 TCFD_BSI查核 | 5 TCFD_BSI查核 | 5 隨機抽樣500字之平均Simpson  |  |
| 6 溫室氣體驗證     | 6 溫室氣體驗證     | 6 總詞數                 |  |
| 7 有任第三方驗證    | 7 有任第三方驗證    | 7 不同詞數                |  |
|              |              | 8 隨機抽樣500詞之平均TTR      |  |
|              |              | 9 隨機抽樣500詞之平均Entropy  |  |
|              |              | 10 隨機抽樣500詞之平均Simpson |  |
|              |              | 11 詞向量                |  |

# 探索資料特色 | 外部評審對各產業的看法較相近

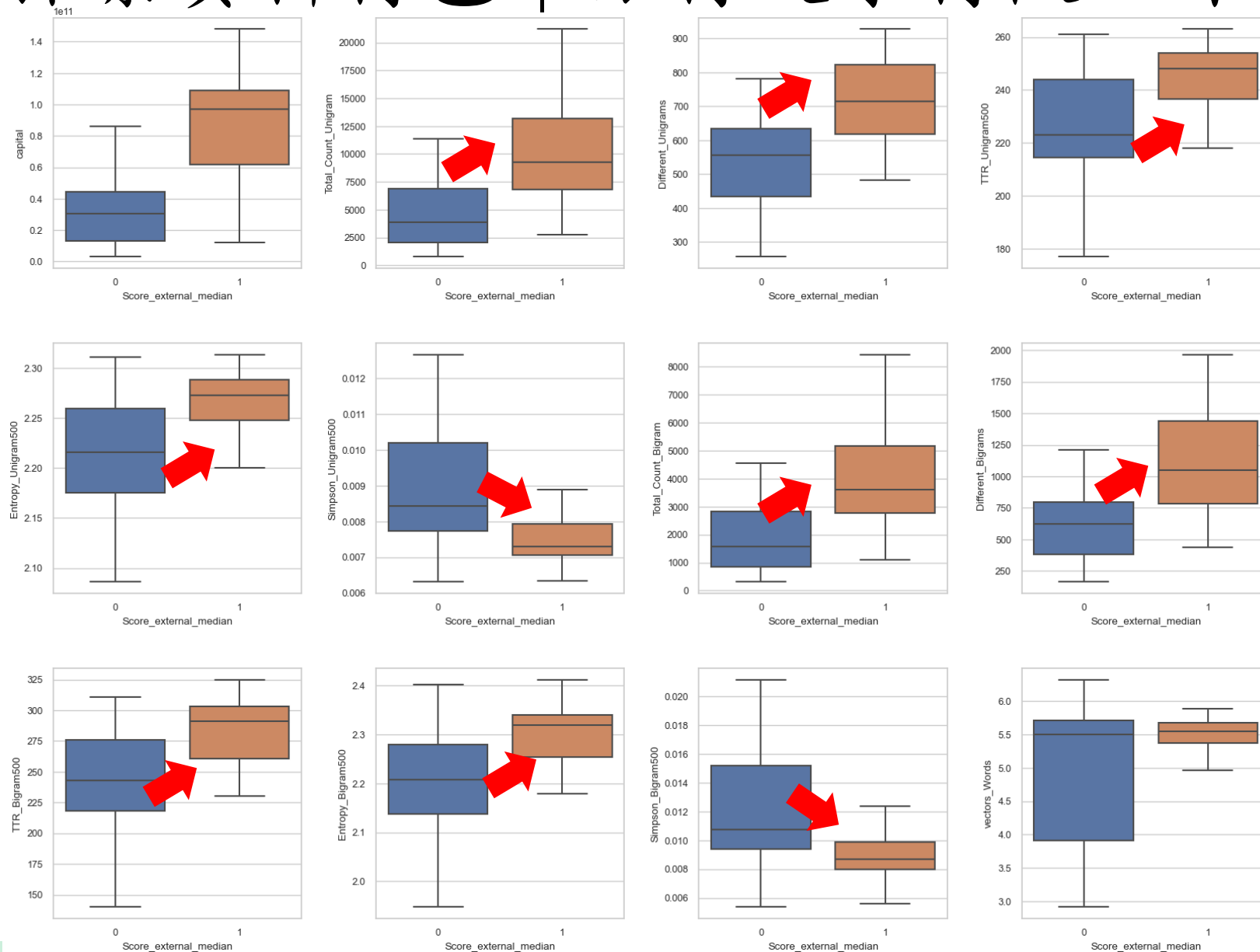


| 專家分數 | 內部   | 外部   |
|------|------|------|
| 銀行   | 0.25 | 0.08 |
| 壽險   | 0.24 | 0.07 |
| 產險   | 0.49 | 0.09 |

變異係數 =  $\frac{\sigma}{\mu}$



# 探索資料特色 | 銀行文字特徵VS外部分數



- 外部評審偏好各產業的報告書篇幅長、字詞種類多和豐富度高
- 內部評審銀行同上，但保險無明顯偏好

# 多變量迴歸模型建置 | 變數篩選

| 全樣本         | 模型類別  | 內部    |           | 外部    |           |
|-------------|-------|-------|-----------|-------|-----------|
|             |       | 變數量   | R squared | 變數量   | R squared |
| 銀<br>n = 38 | 公司    | 2     | 0.52      | 2     | 0.67      |
|             | 文字    | 4     | 0.67      | 4     | 0.65      |
|             | 公司+文字 | 9 (6) | 0.87      | 9 (5) | 0.90      |
| 壽<br>n = 21 | 公司    | 1     | 0.14      | 2     | 0.43      |
|             | 文字    | 1     | 0.18      | 1     | 0.17      |
|             | 公司+文字 | 6 (4) | 0.65      | 7 (4) | 0.77      |
| 產<br>n = 23 | 公司    | 1     | 0.34      | 1     | 0.36      |
|             | 文字    | 1     | 0.28      | 1     | 0.40      |
|             | 公司+文字 | 4 (3) | 0.66      | 4 (3) | 0.73      |

| 僅TCFD       | 模型類別  | 內部    |           | 外部    |           |
|-------------|-------|-------|-----------|-------|-----------|
|             |       | 變數量   | R squared | 變數量   | R squared |
| 銀<br>n = 31 | 公司    | 2     | 0.58      | 2     | 0.69      |
|             | 文字    | 4     | 0.71      | 4     | 0.70      |
|             | 公司+文字 | 9 (6) | 0.89      | 9 (5) | 0.90      |
| 壽<br>n = 15 | 公司    | 1     | 0.11      | 2     | 0.45      |
|             | 文字    | 1     | 0.05      | 1     | 0.11      |
|             | 公司+文字 | 6 (4) | 0.64      | 7 (4) | 0.76      |
| 產<br>n = 11 | 公司    | 1     | 0.53      | 1     | 0.42      |
|             | 文字    | 1     | 0.31      | 1     | 0.64      |
|             | 公司+文字 | 4 (3) | 0.67      | 4 (3) | 0.80      |

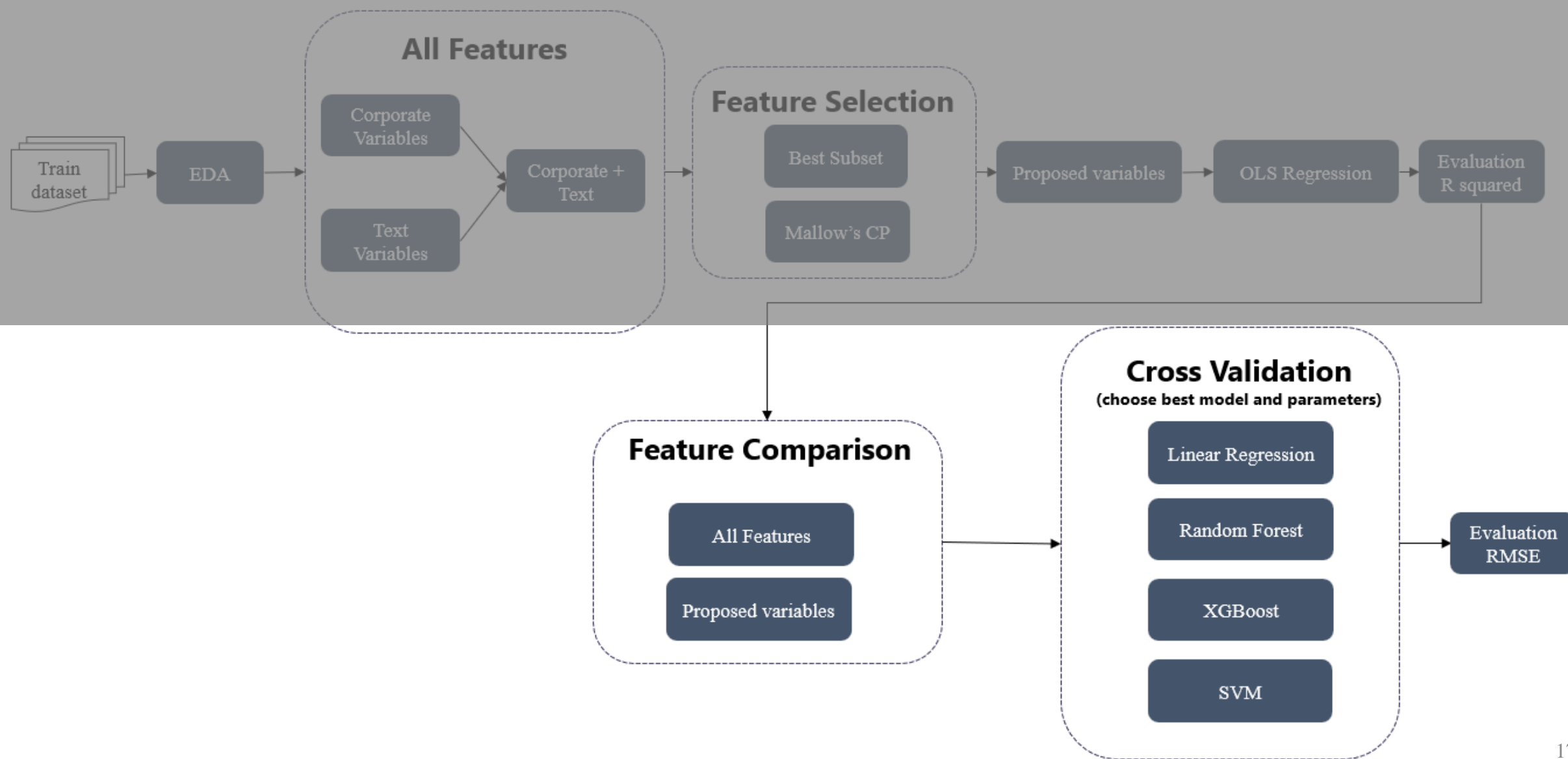
- 公司文字變數合併使用迴歸模型解釋力大幅提高，挑選出的變數過半是文字變數
- 公司+文字模型較可以捕捉外部評審的評分想法。
- 銀行、壽險、產險公司+文字模型外部R<sup>2</sup> 0.90、0.77、0.73，內外部皆銀行最佳
- TCFD、ESG皆可以用迴歸模型評分

# 多變量迴歸模型建置 | 滿足假設檢定

| 產業 | 模型類別  | 分數 | Goldfeld-Quandt test | Rainbow test | Kolmogorov-Smirnov test |
|----|-------|----|----------------------|--------------|-------------------------|
| 銀行 | 公司    | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |
|    | 文字    | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |
|    | 公司+文字 | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |
| 壽險 | 公司    | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |
|    | 文字    | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |
|    | 公司+文字 | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |
| 產險 | 公司    | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |
|    | 文字    | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |
|    | 公司+文字 | 內部 | V                    | V            | V                       |
|    |       | 外部 | V                    | V            | V                       |

- 常態性檢定Kolmogorov-Smirnov test、變異數同質性檢定Goldfeld-Quandt test、適合度檢定Rainbow test
- 迴歸分析之標準化殘差結果無離群值

# 研究流程



# 機器學習模型驗證 | 各產業Proposed RMSE

|          | 銀行 (9)      | 壽險 (7)       | 產險 (4)      |
|----------|-------------|--------------|-------------|
| 外部分數平均   | 80          | 77           | 78          |
| 驗證集的RMSE | 3.51~4.41   | 4.55~8.08    | 5.50~6.86   |
| 誤差比例     | 4.39%~5.51% | 5.91%~10.49% | 7.05%~8.79% |

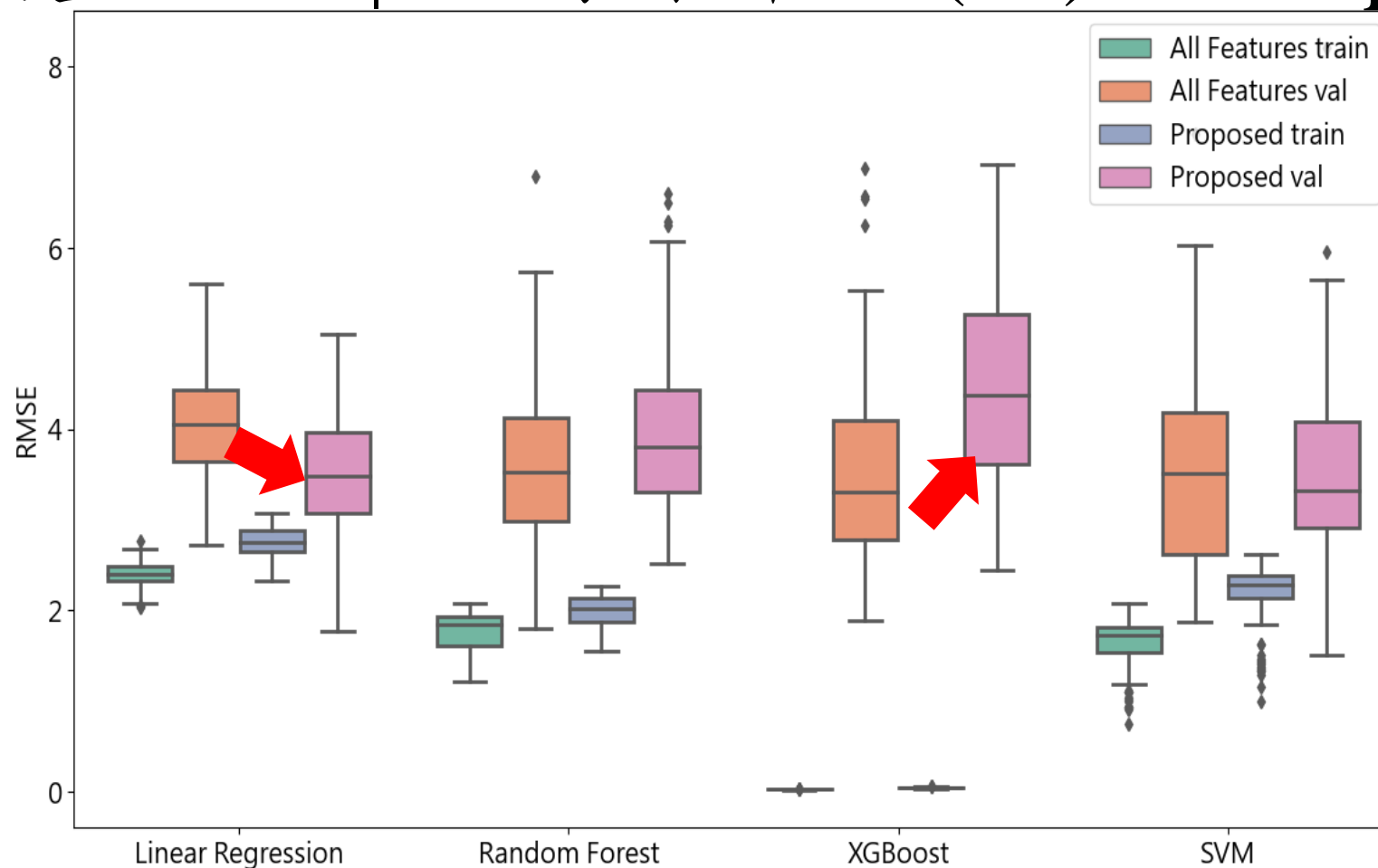
|          | 銀行 (9)       | 壽險 (6)        | 產險 (4)        |
|----------|--------------|---------------|---------------|
| 內部分數平均   | 51           | 41            | 30            |
| 驗證集的RMSE | 10.48~12.47  | 9.55~12.00    | 10.74~12.29   |
| 誤差比例     | 20.5%~24.45% | 23.29%~29.27% | 35.80%~40.97% |

$$RMSE = \sqrt{\sum_{i=1}^n \frac{(\hat{y}_i - y_i)^2}{n}}$$

- 銀行/壽險/產險全部特徵各有56/54/54個變數
- 以篩選變數建模，外部分數誤差比例<10.5% (RMSE/平均分數)
- 內部分數預測結果較差且波動度大

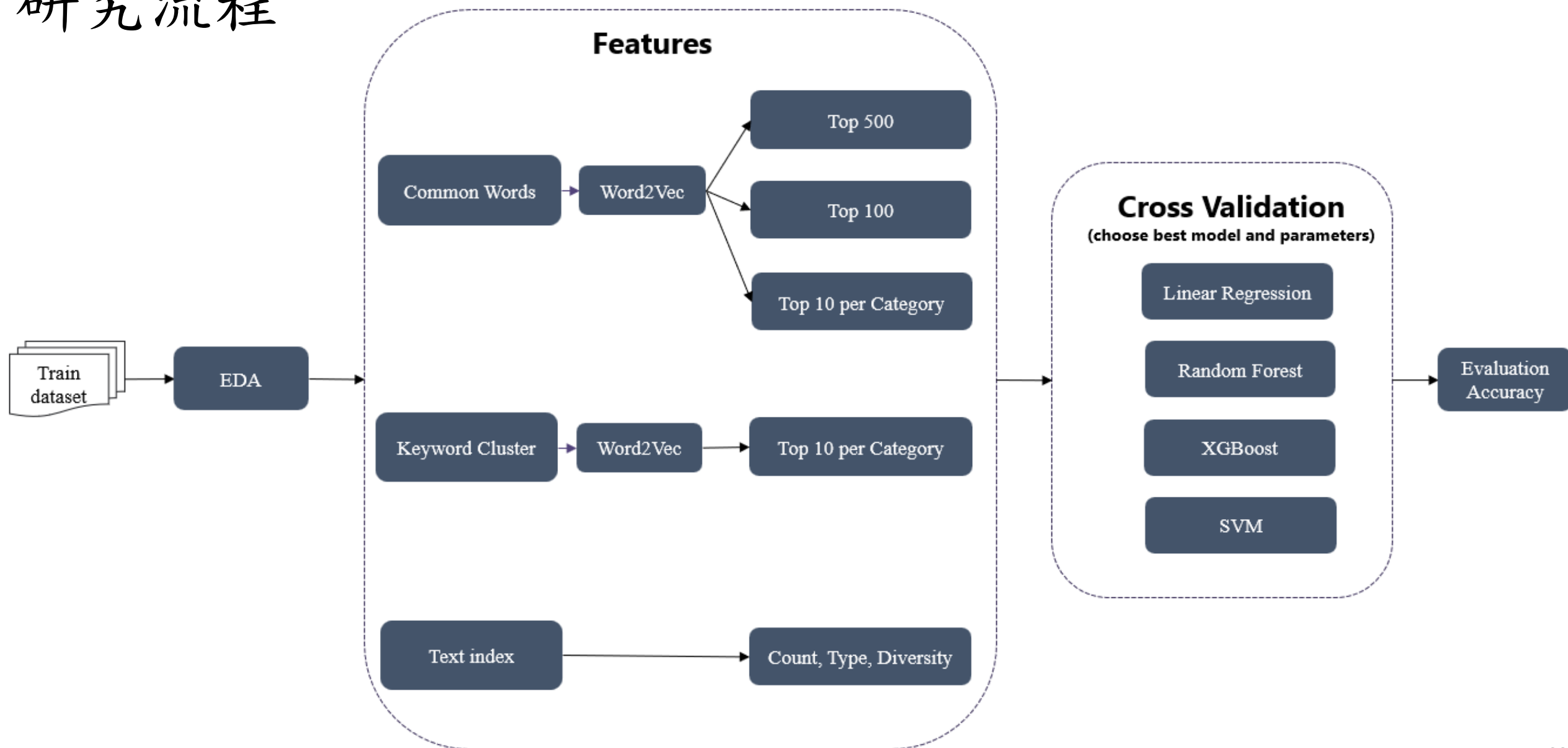


# 機器學習模型驗證 | 銀行外部All(56) vs Proposed(9)



- 驗證集下，Linear Regression 使用挑選的變數建模後RMSE顯著下降，其他機器學習模型結果RMSE提升但皆不顯著→**篩選的變數能使迴歸模型具穩健性**
- **篩選變數可讓統計模型優於機器學習模型**

# 研究流程



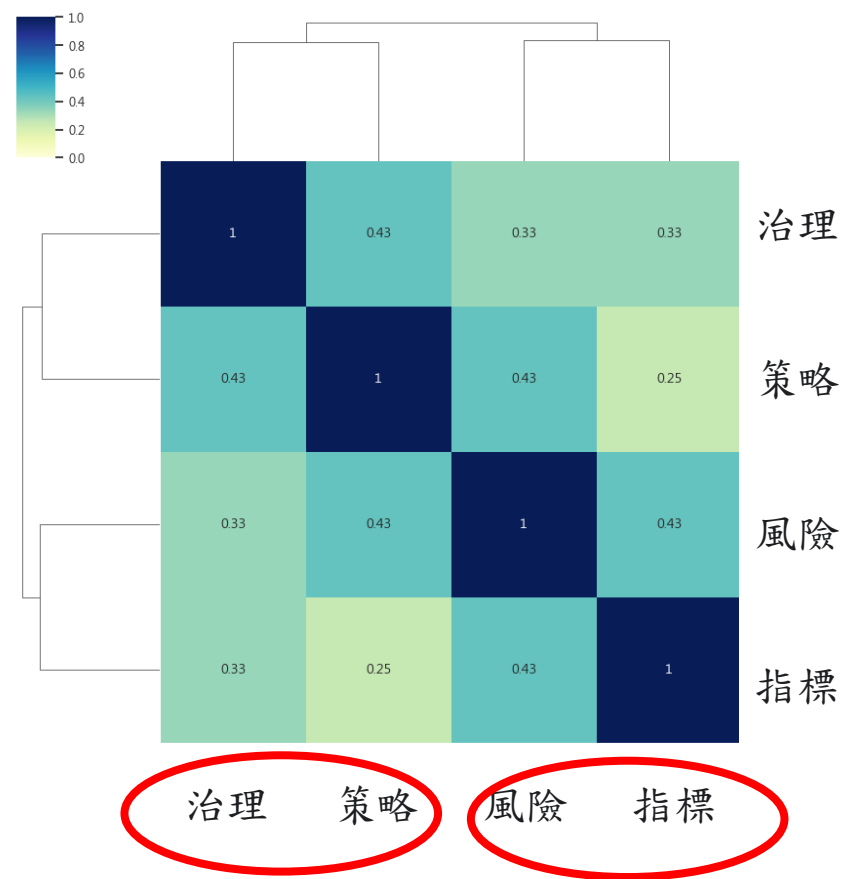
多樣性 | 論文摘要 > 銀行 > 壽險 > 產險

| 資料類別    | 子類別     | 字    |         |         | 詞    |         |         |
|---------|---------|------|---------|---------|------|---------|---------|
|         |         | TTR  | Entropy | Simpson | TTR  | Entropy | Simpson |
| 商管論文摘要  | 其他商業及管理 | 1154 | 6.238   | 0.004   | 4042 | 7.526   | 0.002   |
|         | 企業管理    | 1129 | 6.209   | 0.004   | 3865 | 7.426   | 0.002   |
|         | 風險管理    | 1022 | 6.122   | 0.004   | 3408 | 7.287   | 0.002   |
|         | 財務金融    | 1017 | 6.094   | 0.004   | 3589 | 7.275   | 0.002   |
|         | 會計      | 960  | 6.037   | 0.004   | 3276 | 7.134   | 0.003   |
|         | 醫管      | 1090 | 6.196   | 0.004   | 3561 | 7.318   | 0.002   |
|         | 行銷與流通   | 1107 | 6.188   | 0.004   | 3667 | 7.348   | 0.002   |
|         | 貿易      | 1059 | 6.167   | 0.004   | 3452 | 7.321   | 0.002   |
|         | 一般商業    | 1108 | 6.224   | 0.004   | 3753 | 7.483   | 0.002   |
|         | 財政      | 994  | 6.120   | 0.004   | 3268 | 7.287   | 0.002   |
| 銀行永續報告書 | 治理      | 648  | 5.489   | 0.008   | 1428 | 5.802   | 0.013   |
|         | 策略      | 865  | 5.947   | 0.004   | 2490 | 6.685   | 0.005   |
|         | 風險      | 808  | 5.830   | 0.006   | 2223 | 6.438   | 0.008   |
|         | 指標      | 834  | 5.945   | 0.004   | 2298 | 6.764   | 0.003   |
| 壽險永續報告書 | 治理      | 625  | 5.456   | 0.008   | 1194 | 5.690   | 0.013   |
|         | 策略      | 827  | 5.871   | 0.005   | 2226 | 6.539   | 0.006   |
|         | 風險      | 775  | 5.783   | 0.006   | 1973 | 6.351   | 0.009   |
|         | 指標      | 725  | 5.832   | 0.005   | 1507 | 6.436   | 0.004   |
| 產險永續報告書 | 治理      | 546  | 5.403   | 0.009   | 788  | 5.539   | 0.014   |
|         | 策略      | 769  | 5.773   | 0.006   | 1692 | 6.266   | 0.008   |
|         | 風險      | 744  | 5.695   | 0.008   | 1617 | 6.157   | 0.012   |
|         | 指標      | 665  | 5.725   | 0.005   | 1120 | 6.263   | 0.004   |

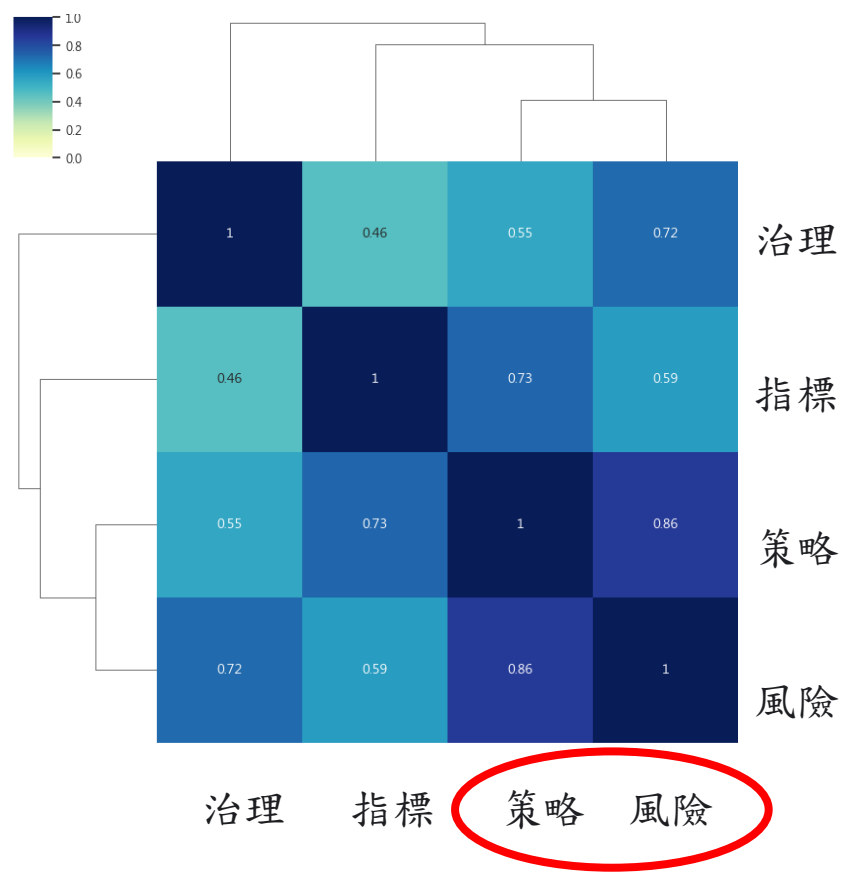
法遵文件多樣性低

# 常見詞彙探索 | 銀行分群 | 相似度指標&詞彙篩選

Jaccard Index (每類十大)



Yue Index (全部百大)



- 策略、風險撰寫風格相似，治理最不同

# 常見詞彙探索 | 驗證

銀行

| 詞彙&分群 | 相似度指標              | 詞彙篩選 | 詞彙數 | 集群一   | 集群二   | 集群三   | 集群四 | Accuracy          |
|-------|--------------------|------|-----|-------|-------|-------|-----|-------------------|
| 1     | Cosine/Yue         | 每類十大 | 21  | 治理    | 策略    | 風險    | 指標  | <div></div> 0.972 |
| 2     | Cosine/Yue/Jacarrd | 百大   | 100 | 治理    | 策略    | 風險    | 指標  | <div></div> 0.750 |
| 3     | Cosine/Yue/Jacarrd | 五百大  | 500 | 治理    | 策略    | 風險    | 指標  | <div></div> 0.774 |
| 4     | Cosine/Yue         | 每類十大 | 21  | 治理    | 指標    | 策略、風險 | --  | <div></div> 0.973 |
| 5     | Cosine/Yue         | 百大   | 100 | 治理    | 指標    | 策略、風險 | --  | <div></div> 0.885 |
| 6     | Cosine/Yue         | 五百大  | 500 | 治理    | 指標    | 策略、風險 | --  | <div></div> 0.897 |
| 7     | Jacarrd            | 每類十大 | 21  | 治理、策略 | 風險、指標 | --    | --  | <div></div> 0.999 |

- 7種可能詞彙與分群組合
- 每類十大常見詞彙語意已經可以有相當好的分類效果



# 常見詞彙探索 | 驗證

壽險

| 詞彙&分群 | 相似度指標              | 詞彙篩選 | 詞彙數 | 集群一 | 集群二 | 集群三   | 集群四 | Accuracy          |
|-------|--------------------|------|-----|-----|-----|-------|-----|-------------------|
| 1     | Jacarrd/Yue/Cosine | 每類十大 | 21  | 治理  | 策略  | 風險    | 指標  | <div></div> 0.783 |
| 2     | Jacarrd/Yue/Cosine | 百大   | 100 | 治理  | 策略  | 風險    | 指標  | <div></div> 0.505 |
| 3     | Jacarrd/Yue/Cosine | 五百大  | 500 | 治理  | 策略  | 風險    | 指標  | <div></div> 0.503 |
| 4     | Jacarrd/Yue/Cosine | 每類十大 | 21  | 治理  | 指標  | 策略、風險 | --  | <div></div> 0.821 |
| 5     | Jacarrd/Yue/Cosine | 百大   | 100 | 治理  | 指標  | 策略、風險 | --  | <div></div> 0.687 |
| 6     | Jacarrd/Yue/Cosine | 五百大  | 500 | 治理  | 指標  | 策略、風險 | --  | <div></div> 0.653 |

產險

| 詞彙&分群 | 相似度指標              | 詞彙篩選 | 詞彙數 | 集群一 | 集群二 | 集群三   | 集群四 | Accuracy          |
|-------|--------------------|------|-----|-----|-----|-------|-----|-------------------|
| 1     | Jacarrd/Yue/Cosine | 每類十大 | 22  | 治理  | 策略  | 風險    | 指標  | <div></div> 0.632 |
| 2     | Jacarrd/Yue/Cosine | 百大   | 100 | 治理  | 策略  | 風險    | 指標  | <div></div> 0.445 |
| 3     | Jacarrd/Yue/Cosine | 五百大  | 500 | 治理  | 策略  | 風險    | 指標  | <div></div> 0.375 |
| 4     | Jacarrd/Yue/Cosine | 每類十大 | 22  | 治理  | 指標  | 策略、風險 | --  | <div></div> 0.727 |
| 5     | Jacarrd/Yue/Cosine | 百大   | 100 | 治理  | 指標  | 策略、風險 | --  | <div></div> 0.500 |
| 6     | Jacarrd/Yue/Cosine | 五百大  | 500 | 治理  | 指標  | 策略、風險 | --  | <div></div> 0.500 |

- 銀行、壽險有作出要素的差異化，但產險內容較重疊

# 關鍵詞叢 | 銀行十大二連詞

13 { 4 {

| 排序 | 治理  |     | 策略   |     | 風險   |     | 指標   |     |
|----|-----|-----|------|-----|------|-----|------|-----|
|    | 先行詞 | 關聯詞 | 先行詞  | 關聯詞 | 先行詞  | 關聯詞 | 先行詞  | 關聯詞 |
| 1  | 氣候  | 風險  | 氣候   | 風險  | 氣候   | 風險  | 氣候   | 風險  |
| 2  | 風險  | 管理  | 實體   | 風險  | 管理   | 風險  | 風險   | 氣候  |
| 3  | 管理  | 風險  | 風險   | 氣候  | 風險   | 管理  | 實體   | 風險  |
| 4  | 氣候  | 管理  | 轉型   | 風險  | 風險   | 氣候  | 指標   | 目標  |
| 5  | 風險  | 氣候  | 情境   | 風險  | 相關   | 風險  | 高碳排  | 產業  |
| 6  | 管理  | 氣候  | 分析   | 風險  | 本行   | 風險  | 氣候   | 目標  |
| 7  | 氣候  | 相關  | 情境   | 分析  | 氣候   | 管理  | 相關   | 風險  |
| 8  | 相關  | 風險  | 氣候變遷 | 風險  | 評估   | 風險  | 本行   | 氣候  |
| 9  | 本行  | 風險  | 相關   | 風險  | 氣候變遷 | 風險  | 溫室氣體 | 排放  |
| 10 | 相關  | 管理  | 風險   | 影響  | 實體   | 風險  | 氣候   | 相關  |

灰色：「氣候」、「風險」、「管理」、「相關」  
紅色：只出現在單一要素的特色二連詞

# 關鍵詞叢 | 保險比銀行報告書風格單調保守

壽險十大二連詞

| 排序 | 治理  |     | 策略   |     | 風險   |     | 指標   |     |
|----|-----|-----|------|-----|------|-----|------|-----|
|    | 先行詞 | 關聯詞 | 先行詞  | 關聯詞 | 先行詞  | 關聯詞 | 先行詞  | 關聯詞 |
| 1  | 氣候  | 風險  | 氣候   | 風險  | 氣候   | 風險  | 氣候   | 風險  |
| 2  | 風險  | 管理  | 相關   | 風險  | 管理   | 風險  | 氣候變遷 | 風險  |
| 3  | 管理  | 風險  | 風險   | 氣候  | 相關   | 風險  | 相關   | 風險  |
| 4  | 氣候  | 管理  | 氣候變遷 | 風險  | 風險   | 管理  | 風險   | 氣候  |
| 5  | 相關  | 風險  | 實體   | 風險  | 風險   | 氣候  | 公司   | 風險  |
| 6  | 氣候  | 相關  | 氣候   | 相關  | 公司   | 風險  | 風險   | 管理  |
| 7  | 風險  | 氣候  | 情境   | 風險  | 氣候   | 管理  | 管理   | 風險  |
| 8  | 管理  | 氣候  | 影響   | 風險  | 氣候變遷 | 風險  | 氣候   | 管理  |
| 9  | 公司  | 風險  | 轉型   | 風險  | 氣候   | 相關  | 實體   | 風險  |
| 10 | 相關  | 管理  | 分析   | 風險  | 風險   | 相關  | 氣候   | 相關  |

產險十大二連詞

| 排序 | 治理  |     | 策略  |     | 風險   |     | 指標   |     |
|----|-----|-----|-----|-----|------|-----|------|-----|
|    | 先行詞 | 關聯詞 | 先行詞 | 關聯詞 | 先行詞  | 關聯詞 | 先行詞  | 關聯詞 |
| 1  | 氣候  | 風險  | 氣候  | 風險  | 氣候   | 風險  | 氣候   | 風險  |
| 2  | 氣候  | 相關  | 相關  | 風險  | 管理   | 風險  | 相關   | 風險  |
| 3  | 風險  | 管理  | 風險  | 氣候  | 風險   | 管理  | 風險   | 管理  |
| 4  | 相關  | 風險  | 氣候  | 相關  | 相關   | 風險  | 氣候變遷 | 風險  |
| 5  | 氣候  | 管理  | 評估  | 風險  | 公司   | 風險  | 氣候   | 相關  |
| 6  | 管理  | 風險  | 管理  | 風險  | 風險   | 氣候  | 公司   | 風險  |
| 7  | 風險  | 氣候  | 風險  | 管理  | 氣候   | 管理  | 風險   | 氣候  |
| 8  | 相關  | 管理  | 相關  | 氣候  | 氣候   | 相關  | 管理   | 風險  |
| 9  | 管理  | 氣候  | 影響  | 風險  | 保險   | 風險  | 氣候   | 管理  |
| 10 | 風險  | 相關  | 公司  | 風險  | 氣候變遷 | 風險  | 風險   | 相關  |

# 模型分類結果

| 模型<br>編號 | 銀行特徵數  | 特徵   |                 |         | Accuracy |
|----------|--------|------|-----------------|---------|----------|
|          |        | 字詞指標 | 常見詞的詞向量         | 關聯詞的詞向量 |          |
| SVM1     | 10     | 字詞指標 |                 |         | 0.403    |
| SVM2     | 500+10 | 字詞指標 | 全部五百大           |         | 0.764    |
| SVM3     | 500+10 | 字詞指標 | 全部五百大、全部百大      |         | 0.801    |
| SVM4     | 500+10 | 字詞指標 | 全部五百大、每類十大      |         | 0.833    |
| SVM5     | 500+10 | 字詞指標 | 全部五百大、全部百大、每類十大 |         | 0.870    |
| SVM6     | 100+10 | 字詞指標 | 全部百大            |         | 0.747    |
| SVM7     | 100+10 | 字詞指標 | 全部百大、每類前十大      |         | 0.862    |
| SVM8     | 21+10  | 字詞指標 | 每類十大            |         | 0.969    |
| SVM9     | 24+10  | 字詞指標 | 每類十大            | 每類十大二連詞 | 0.975    |

| 模型<br>編號 | 壽險特徵數  | 特徵   |                 |         | Accuracy |
|----------|--------|------|-----------------|---------|----------|
|          |        | 字詞指標 | 常見詞的詞向量         | 關聯詞的詞向量 |          |
| SVM1     |        | 字詞指標 |                 |         | 0.422    |
| SVM2     | 500+10 | 字詞指標 | 全部五百大           |         | 0.541    |
| SVM3     | 500+10 | 字詞指標 | 全部五百大、全部百大      |         | 0.635    |
| SVM4     |        | 字詞指標 | 全部五百大、每類十大      |         | 0.661    |
| SVM5     |        | 字詞指標 | 全部五百大、全部百大、每類十大 |         | 0.665    |
| SVM6     |        | 字詞指標 | 全部百大            |         | 0.533    |
| SVM7     |        | 字詞指標 | 全部百大、每類前十大      |         | 0.715    |
| SVM8     | 21+10  | 字詞指標 | 每類十大            |         | 0.822    |
| SVM9     | 22+10  | 字詞指標 | 每類十大            | 每類十大二連詞 | 0.865    |

# 模型分類結果

| 模型<br>編號 | 產險特徵數  | 特徵   |                 |         | Accuracy |
|----------|--------|------|-----------------|---------|----------|
|          |        | 字詞指標 | 常見詞的詞向量         | 關聯詞的詞向量 |          |
| SVM1     | 10     | 字詞指標 |                 |         | 0.327    |
| SVM2     | 500+10 | 字詞指標 | 全部五百大           |         | 0.264    |
| SVM3     | 500+10 | 字詞指標 | 全部五百大、全部百大      |         | 0.419    |
| SVM4     | 500+10 | 字詞指標 | 全部五百大、每類十大      |         | 0.502    |
| SVM5     | 500+10 | 字詞指標 | 全部五百大、全部百大、每類十大 |         | 0.527    |
| SVM6     | 100+10 | 字詞指標 | 全部百大            |         | 0.267    |
| SVM7     | 100+10 | 字詞指標 | 全部百大、每類前十大      |         | 0.494    |
| SVM8     | 22+10  | 字詞指標 | 每類十大            |         | 0.519    |
| SVM9     | 22+10  | 字詞指標 | 每類十大            | 每類十大二連詞 | 0.533    |

- 銀行、壽險的分類表現佳，表示各要素有扮演好自己的角色，但是產險要素間寫作差異較模糊不清，揭露資訊時有重複



## 結論與建議 | 評分

### 提供永續報告書文字分析方法及建構量化評分模型

- 銀行、壽險及產險的公司+文字模型在全樣本外部分數解釋力0.90、0.77、0.73
- 文字變數佔挑選總變數的一半以上
- Linear Regression 使用挑選的變數建模後大幅優化RMSE→篩選的變數能使迴歸模型更具穩健性
- 對機器學習模型使用全部變數通常效果較佳，而篩選變數能讓統計模型優於機器學習模型

## 結論與建議 | 分類

區隔四大要素的寫作風格，檢視要素是否分工明確揭露資訊

- 與單一學類的台灣碩博士論文摘要比較，各產業永續報告書用詞遣句較單一保守
- 二連詞反覆在「氣候」、「風險」、「管理」、「相關」詞彙打轉
- 四要素的差異化來說，最終銀行、壽險、產險可以達Accuracy 0.975/0.865/ 0.632→銀行、壽險各要素可以勝任自身角色，但產險要素間揭露資訊重疊
- 為提升氣候財務揭露的多樣性及差異化，建議政府參考國際經驗，制定更具體的揭露指引，並且強化揭露事項的獨特性

# 限制與展望

## 共同限制

只使用2022年的TCFD或ESG報告書→蒐集更多年資料  
交叉驗證未使用外部樣本測試

## 評分模型限制

- 未將常見詞、關鍵詞叢列為解釋變數
- 未各別考量四個核心要素的人工分數→各要素多變量迴歸模型
- 報告書形式上符合模型指標，但內容缺乏實質意義→BERT語境分析

## 寫作風格分析限制

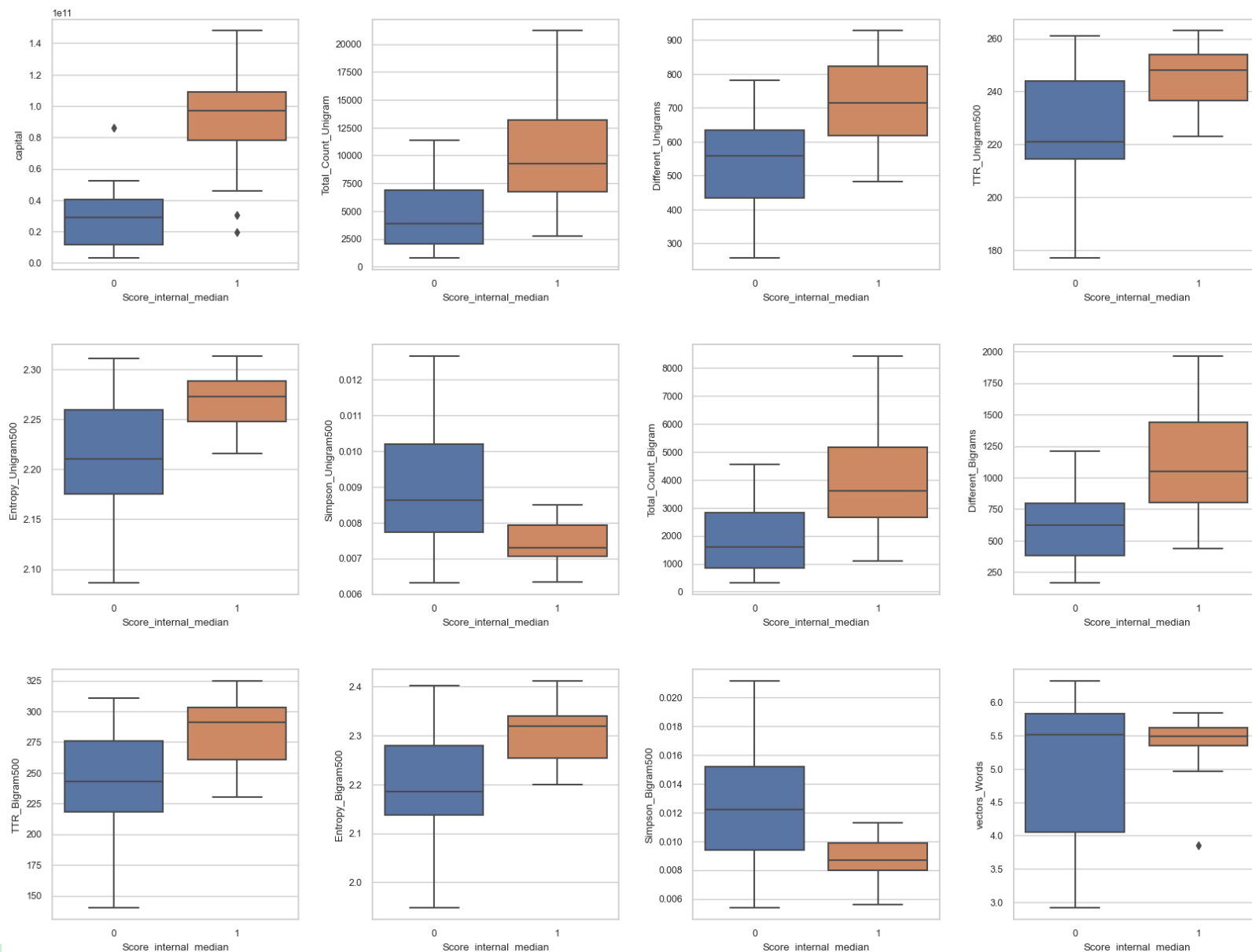
- 要素間分工明確也未必代表內容符合揭露事項→主題建模
- 未實現自動識別四要素的關鍵詞彙



# - 附錄 -

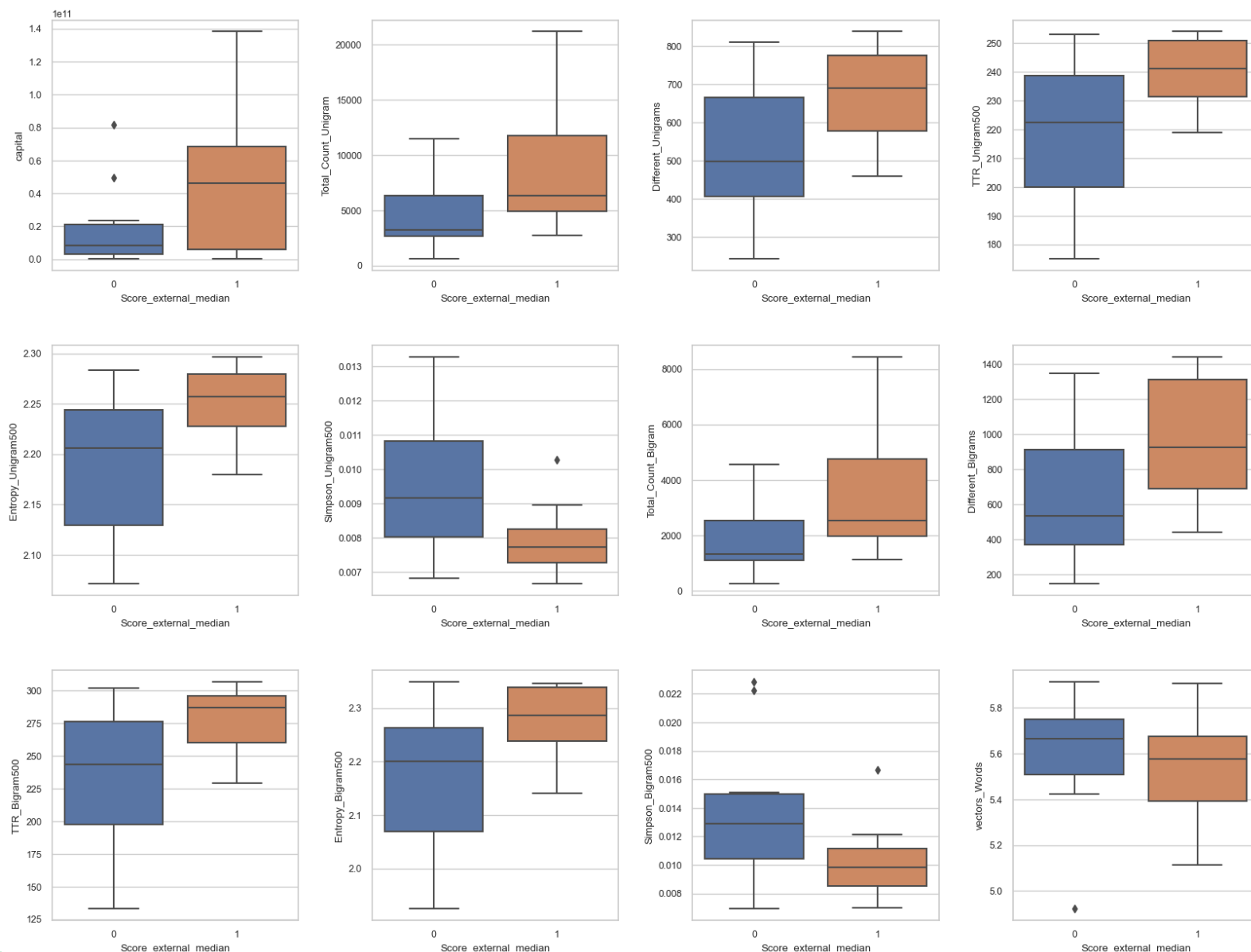
# 探索資料特色 | 銀行文字特徵VS內部分數

- 外部評審偏好各產業的報告書篇幅長、字詞種類多和豐富度高
- 內部評審銀行同上，但保險無明顯偏好



# 探索資料特色 | 壽險文字特徵VS外部分數

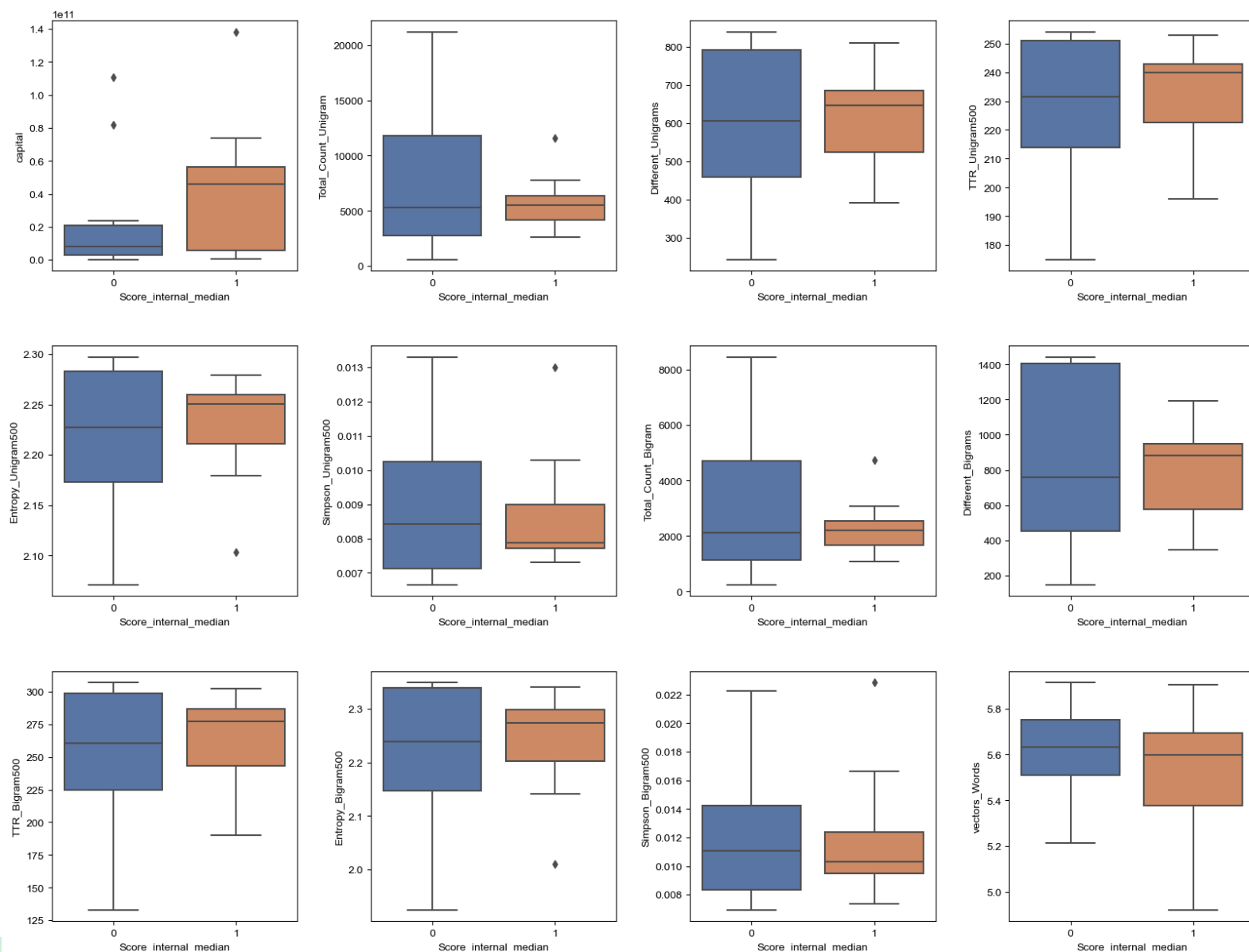
- 外部評審偏好各產業的報告書篇幅長、字詞種類多和豐富度高
- 內部評審銀行同上，但保險無明顯偏好





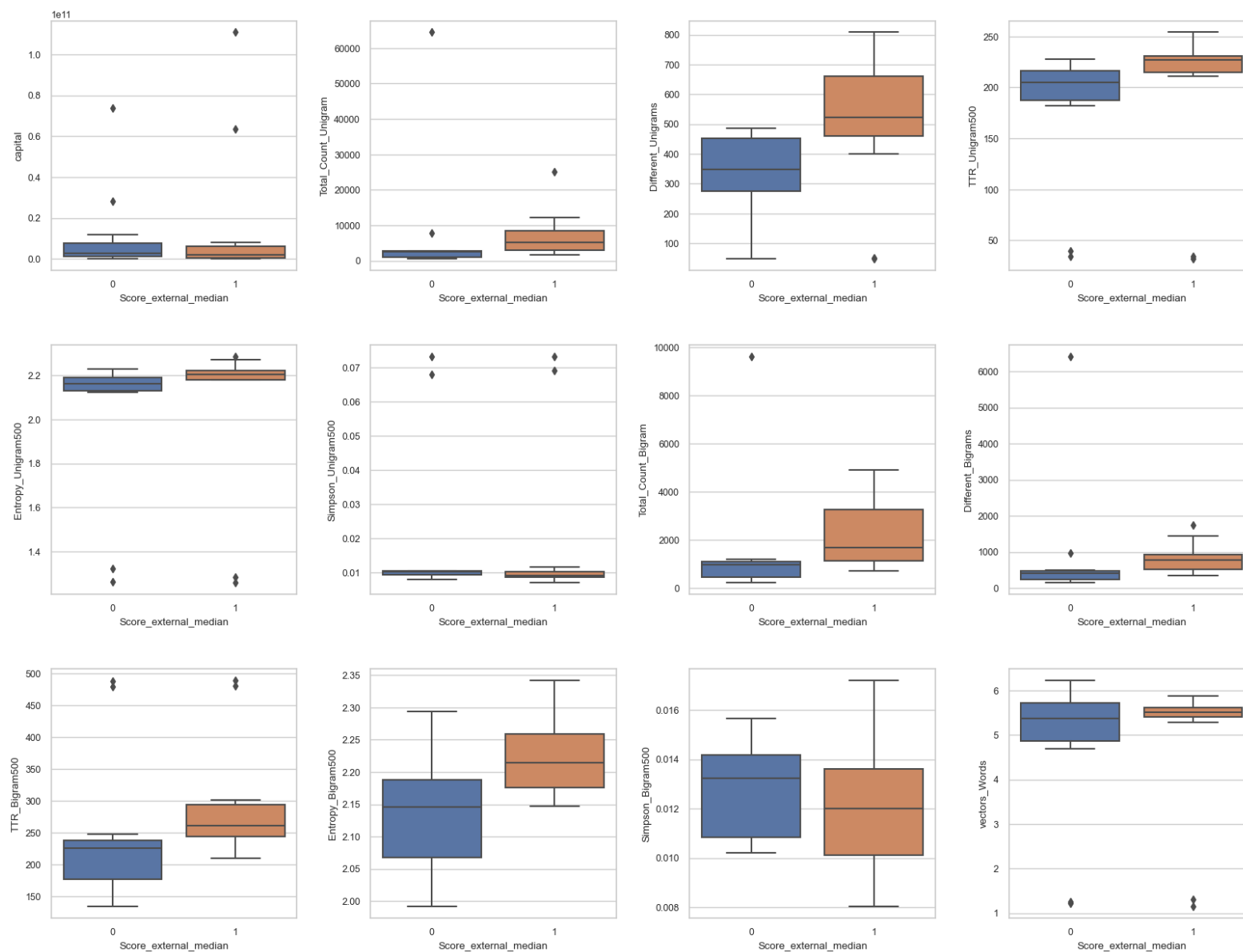
# 探索資料特色 | 壽險文字特徵VS內部分數

- 外部評審偏好各產業的報告書篇幅長、字詞種類多和豐富度高
- 內部評審銀行同上，但保險無明顯偏好



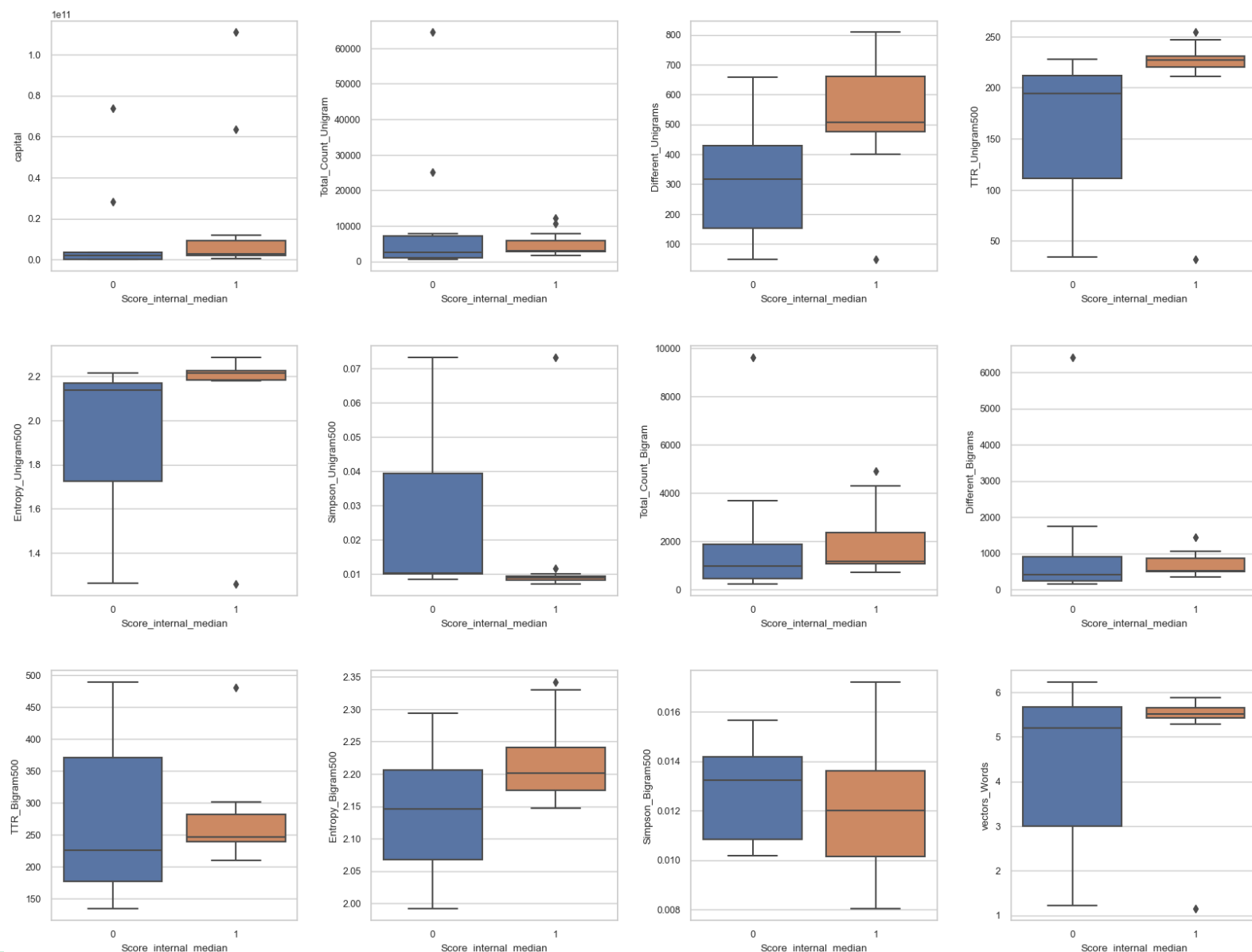
# 探索資料特色 | 產險文字特徵VS外部分數

- 外部評審偏好各產業的報告書篇幅長、字詞種類多和豐富度高
- 內部評審銀行同上，但保險無明顯偏好

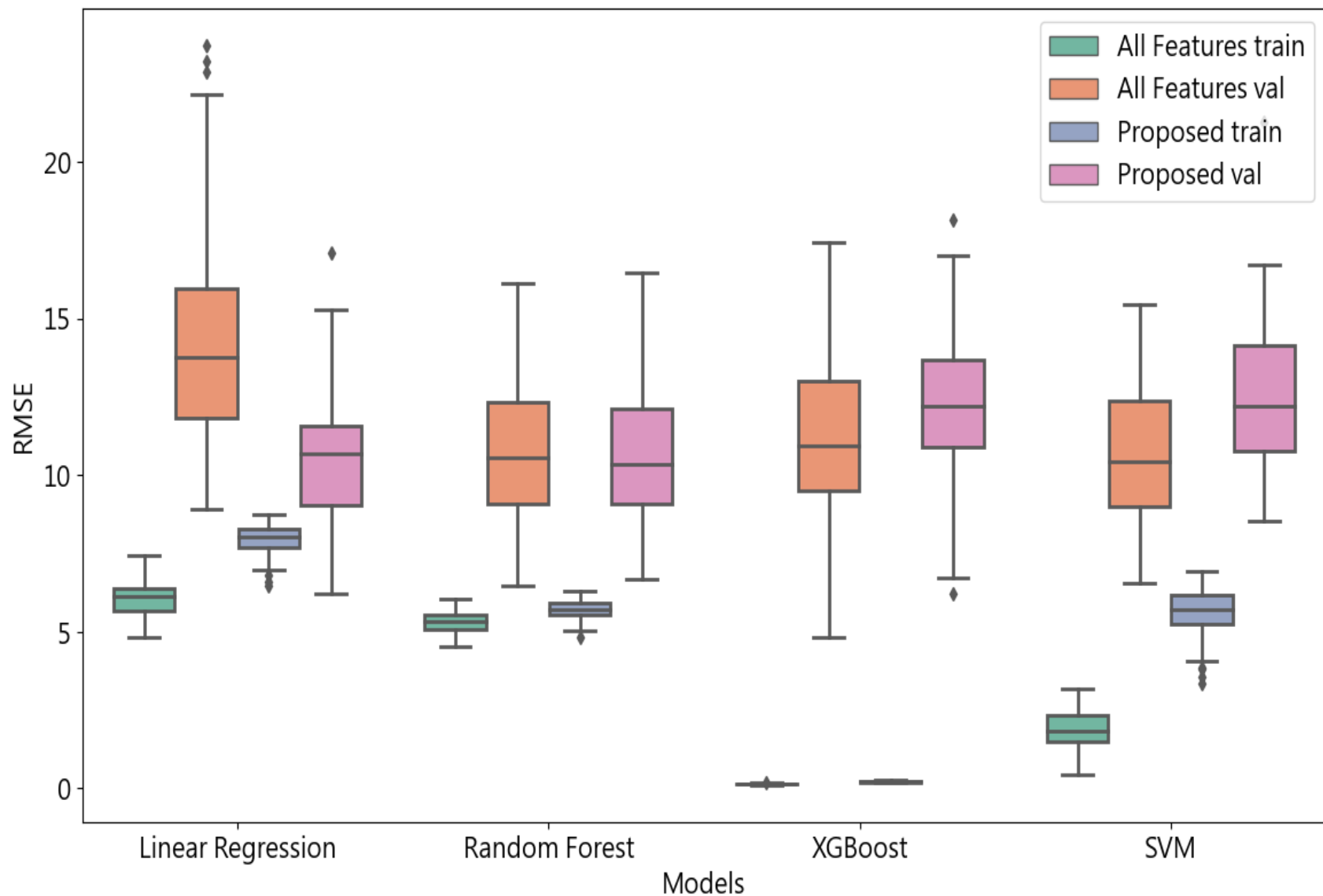


# 探索資料特色 | 產險文字特徵VS內部分數

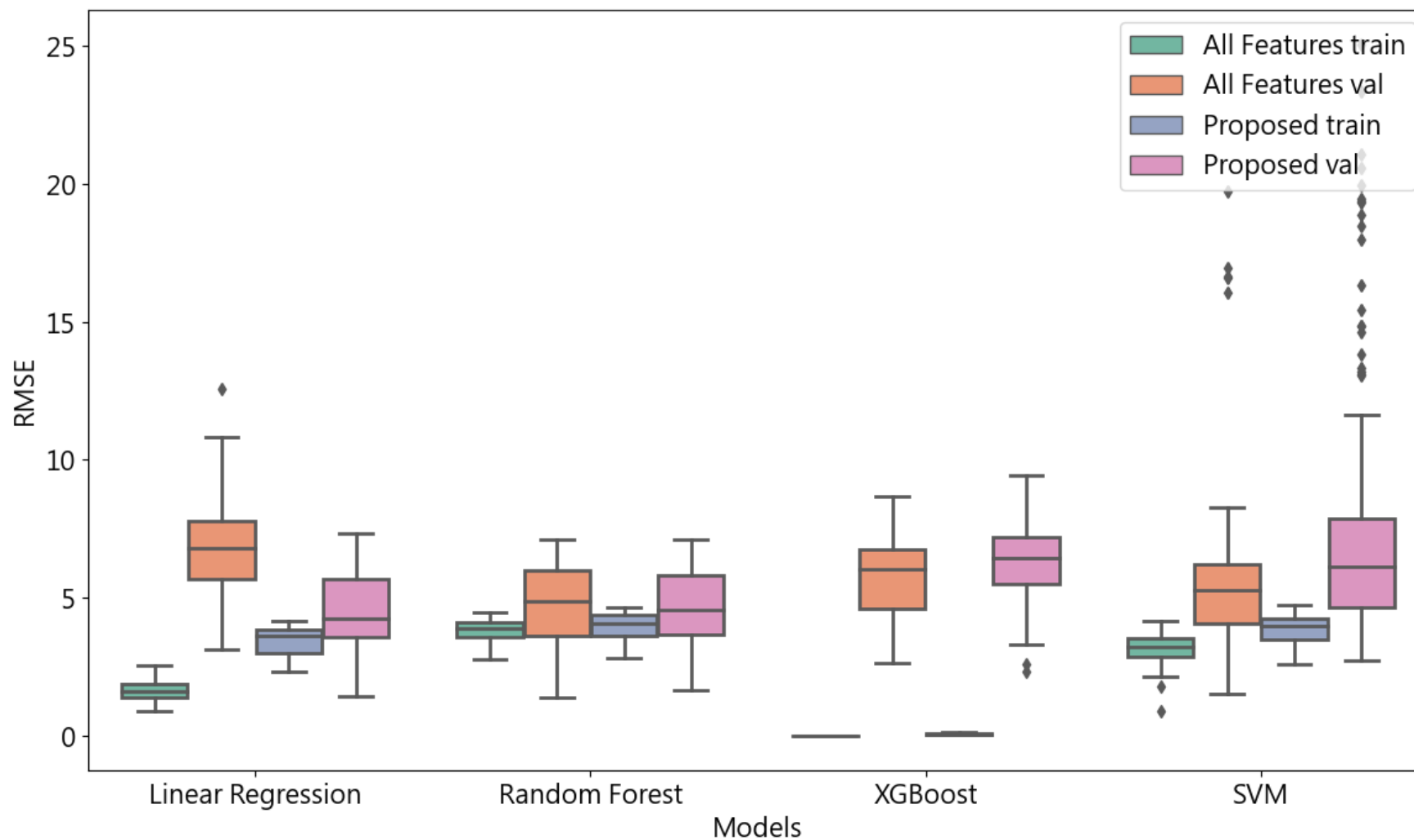
- 外部評審偏好各產業的報告書篇幅長、字詞種類多和豐富度高
- 內部評審銀行同上，但保險無明顯偏好



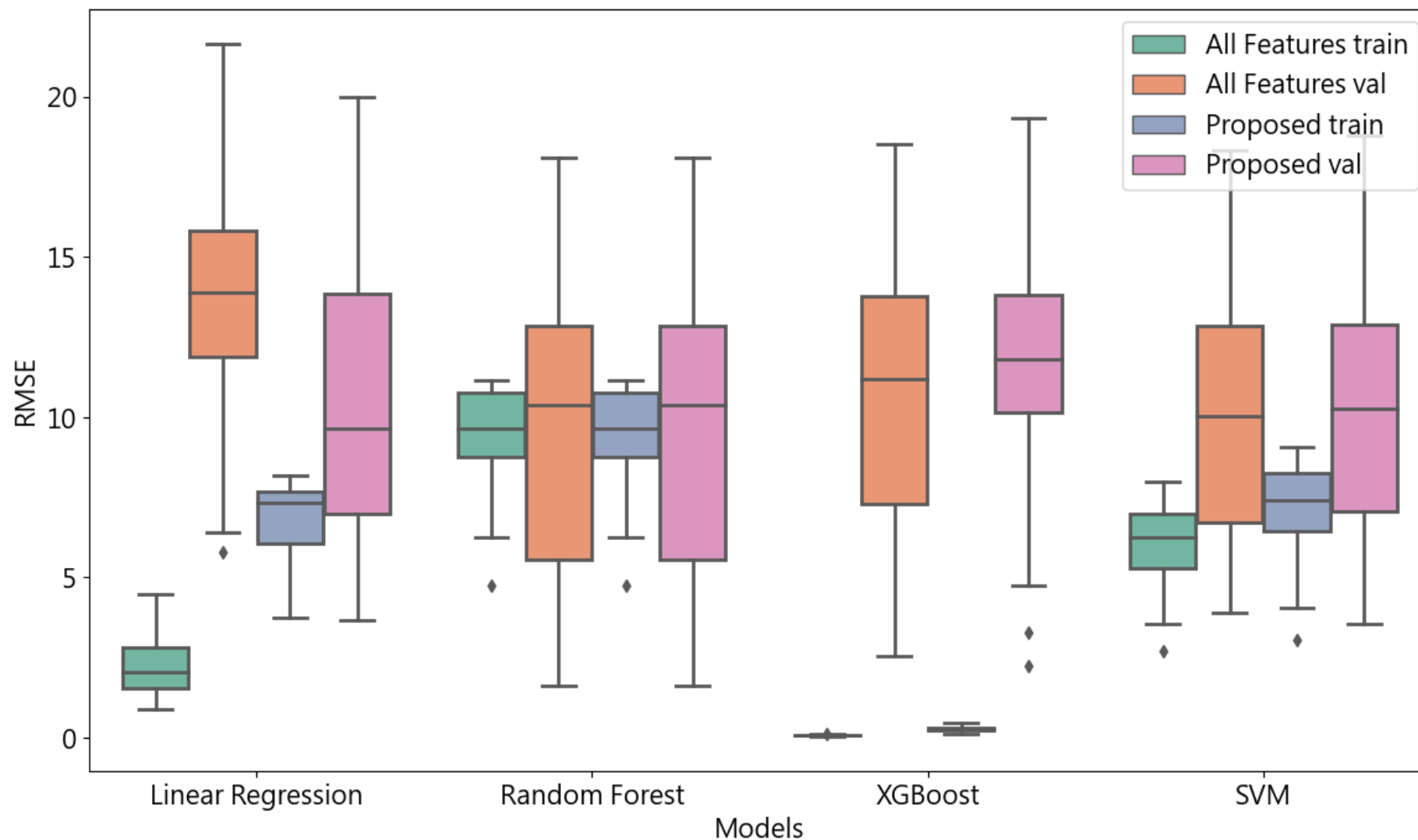
# 機器學習模型驗證 | 銀行內部All(56) vs Proposed(9)



# 機器學習模型驗證 | 壽險外部All(54) vs Proposed(7)

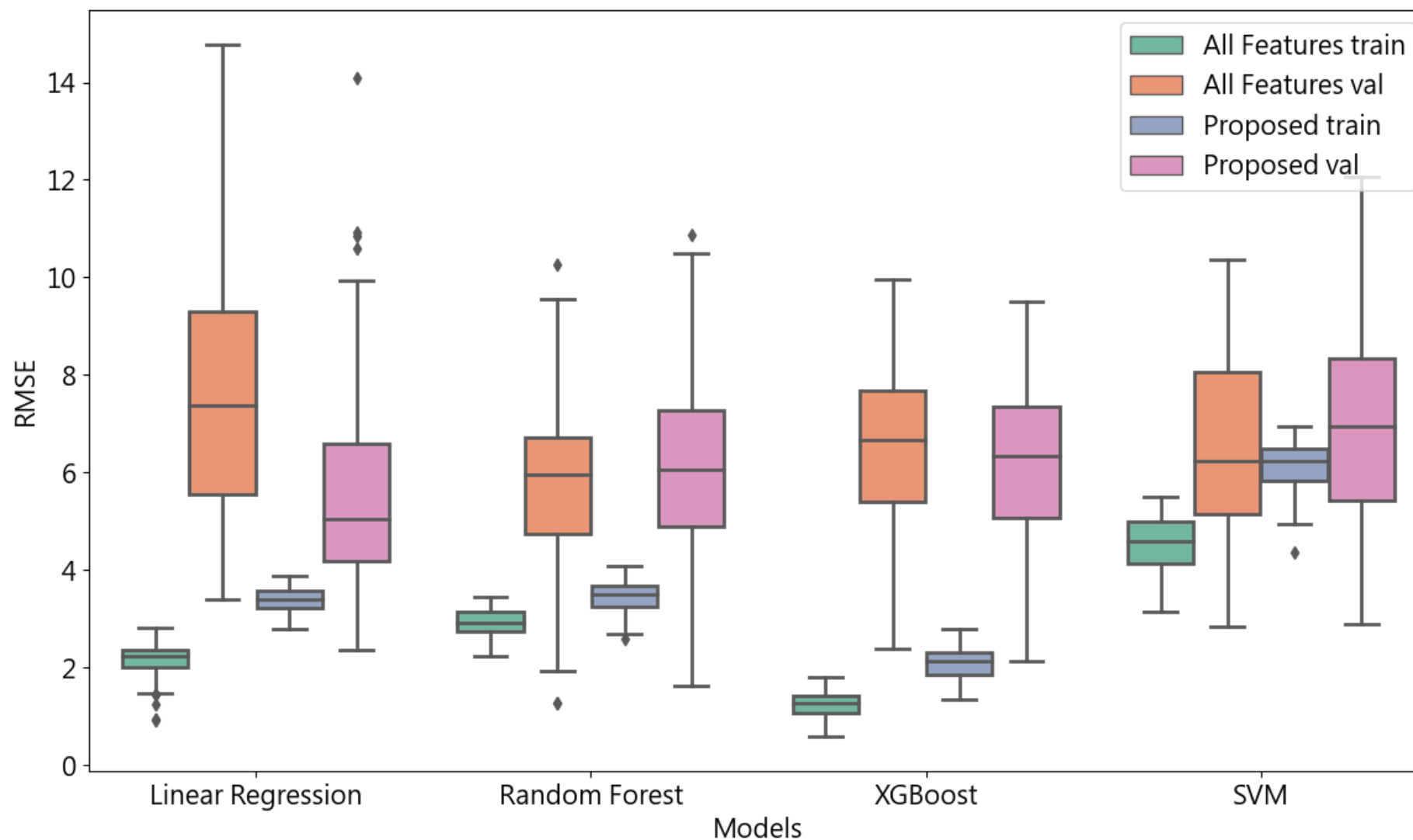


# 機器學習模型驗證 | 壽險內部All(54) vs Proposed(6)

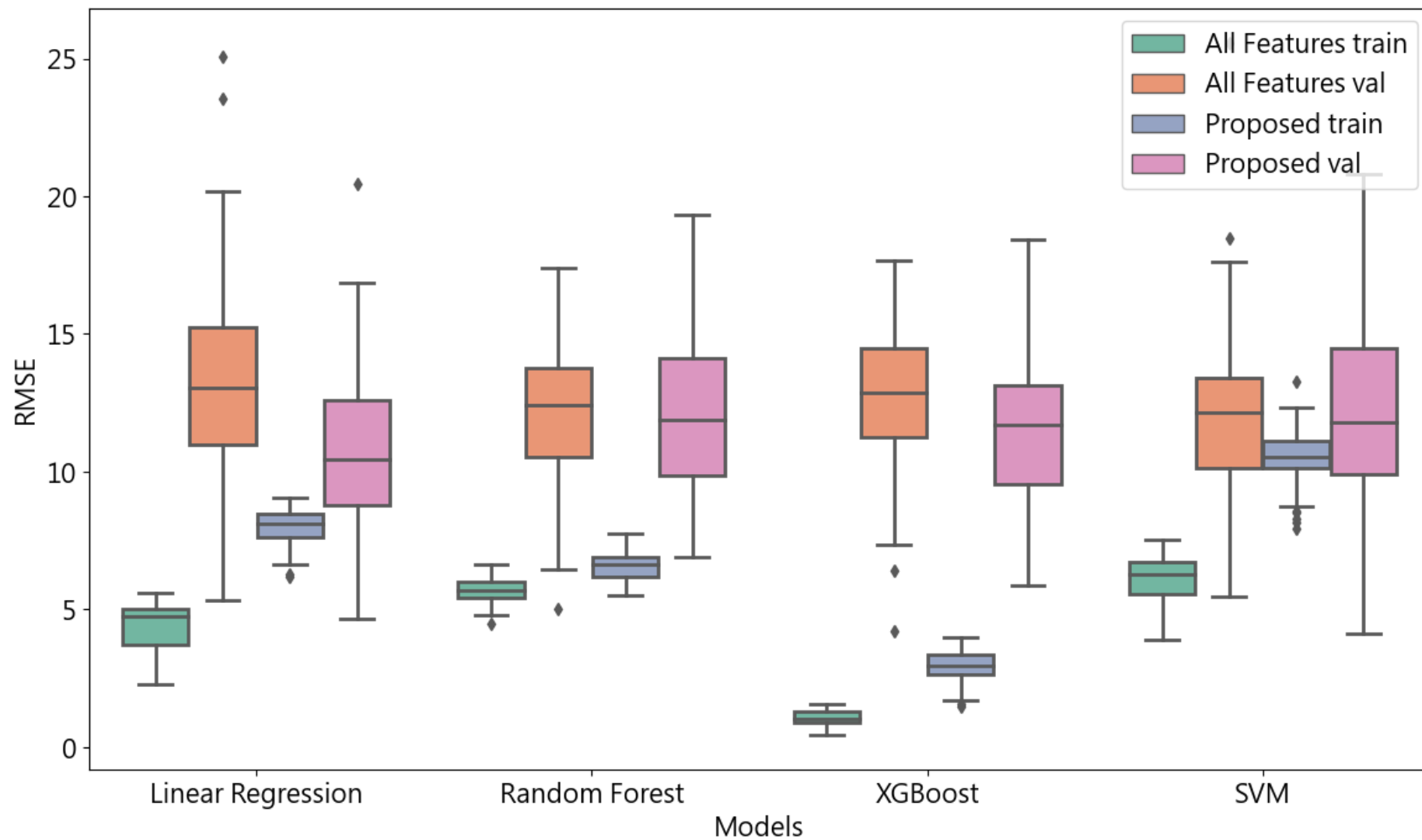




# 機器學習模型驗證 | 產險外部A11(54) vs Proposed(4)



# 機器學習模型驗證 | 產險內部All(54) vs Proposed(4)





- THANK YOU -