

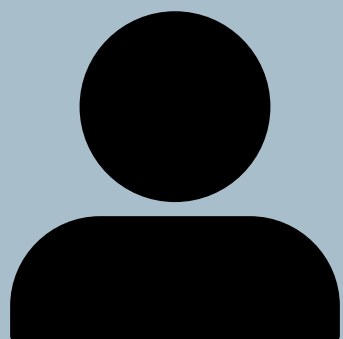
# 美股公司財報分析

以美股財報用字預測未來毛利率/ROA 之模型分析

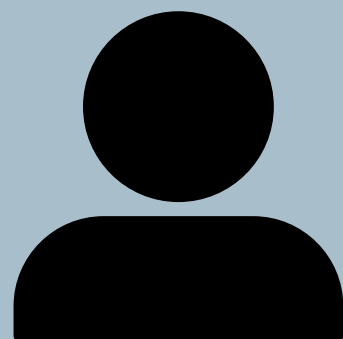
27 December, 2024



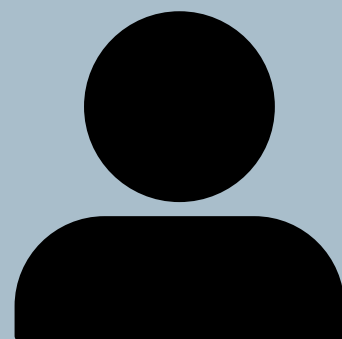
# 小組成員



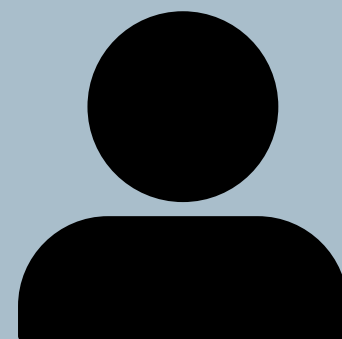
張念詠  
擷取文本  
整理數據  
Rocchio  
KNN



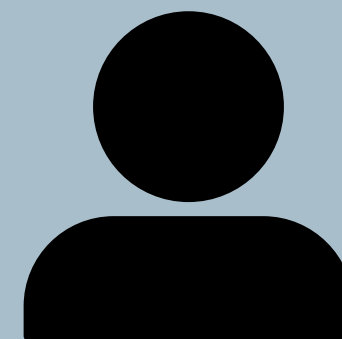
鄭益丞  
擷取文本  
SVM  
K-means



莫家希  
擷取文本  
前處理  
SVD  
LDA



蕭皓澤  
Naive Bayes  
Bernoulli NB  
Neural Network  
Merge Models



蔡明龍  
HAC  
RNN  
proposal  
Merge Models



# 研究動機

- 過去主流研究公司財務表現皆以量化的數字為主，如:毛利率、淨利...等
- SEC規定美股上市公司每年需提交10-K 年報，其中Item 7 MD&A 需揭露管理層營運分析
- MD&A 潛在資訊多於財務數字，包含公司未來營運策略、風險來源、產業趨勢
- 管理層擁有較多內部資訊，藉由MD&A了解其想法，分析財務表現與公司字彙使用關聯



# 以 *NVIDIA* 為例



---

During fiscal year 2022, we made substantial strides in **broadening** our supply base to **scale** our company and better serve customer demand. We expect to remain **supply-constrained** into the first half of fiscal year 2023, primarily in Gaming and Networking. We have placed non-cancellable inventory orders for certain supply in advance of our historical lead times, paid premiums and provided deposits to secure future supply and capacity and may need to continue to do so in the future.

---





# 研究目的



- 以美股標普500指數裡的342家公司的財報，判斷其下一年財務指標變化方向
- 探討多種分類（Classification）與分群（Clustering）模型在財報分析中的適用性與表現，並最佳化其預測能力，且期待得到能明確區分分群的關鍵字
- 希望能發展出一套穩定且準確的財報分析框架，能有效地運用於投資領域上

# 財務指標判斷



## 毛利率成長率 (GROSS MARGIN)

衡量企業盈利能力的指標，越高則意味著企業獲利越好或控制成本越好。

以 $\pm 1\%$ 為基準：

小於 $-1\%$ 為表現不佳的公司

位於 $-1\% \sim 1\%$ 則為表現平常的公司

大於 $1\%$ 為表現良好的公司



## 資產報酬率成長率 (ROA, RETURN ON ASSET)

衡量一間公司營運績效、賺錢的效率，可以看出公司資產使用效率、運用資產創造利潤的能力。

以 $\pm 1\%$ 為基準：

小於 $-1\%$ 為表現不佳的公司

位於 $-1\% \sim 1\%$ 則為表現平常的公司

大於 $1\%$ 為表現良好的公司



# 研究方法

---

擷取文本

前處理

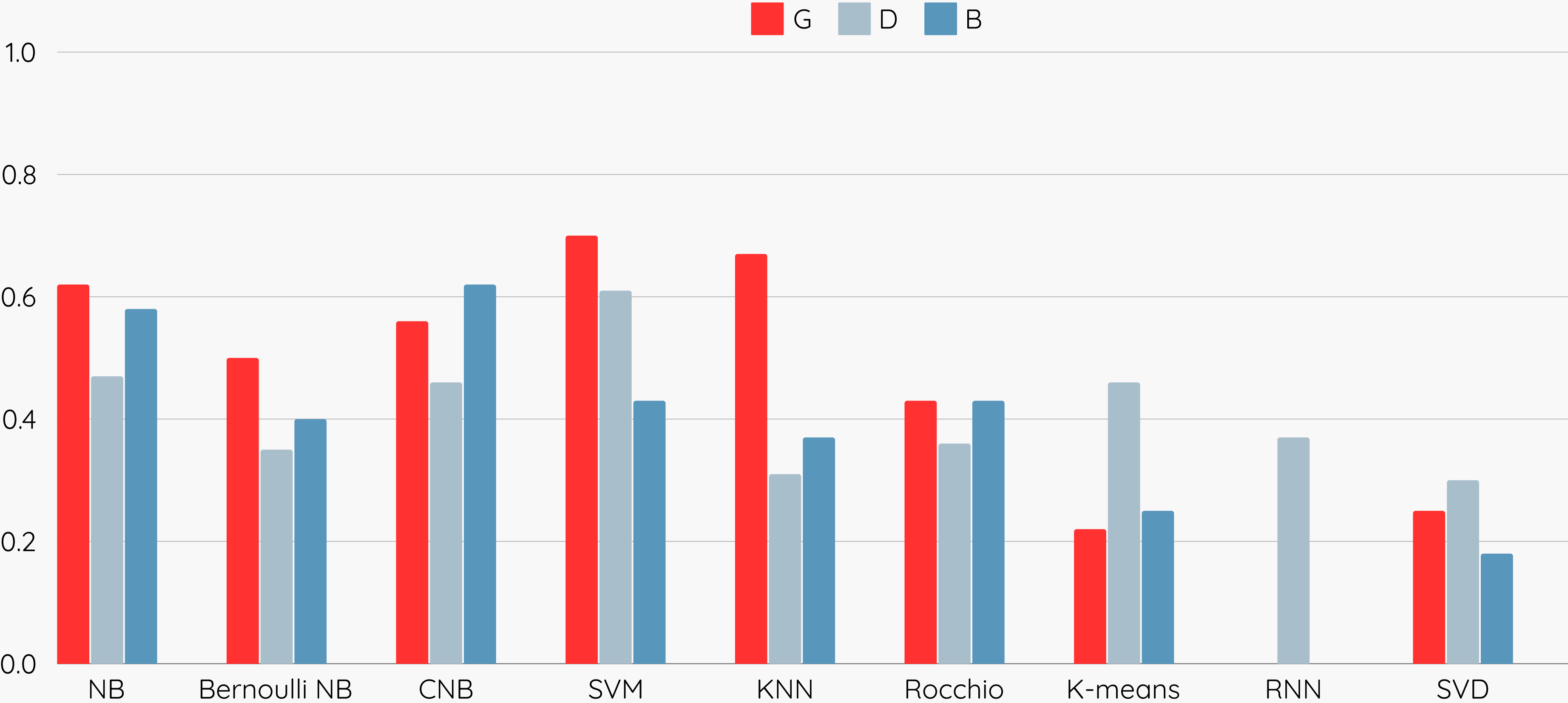
Naive Bayes  
Bernolli NB  
SVM  
Rocchio  
KNN  
K-means  
HAC  
SVD  
LDA  
Neural Network  
RNN

Precision

混淆矩陣

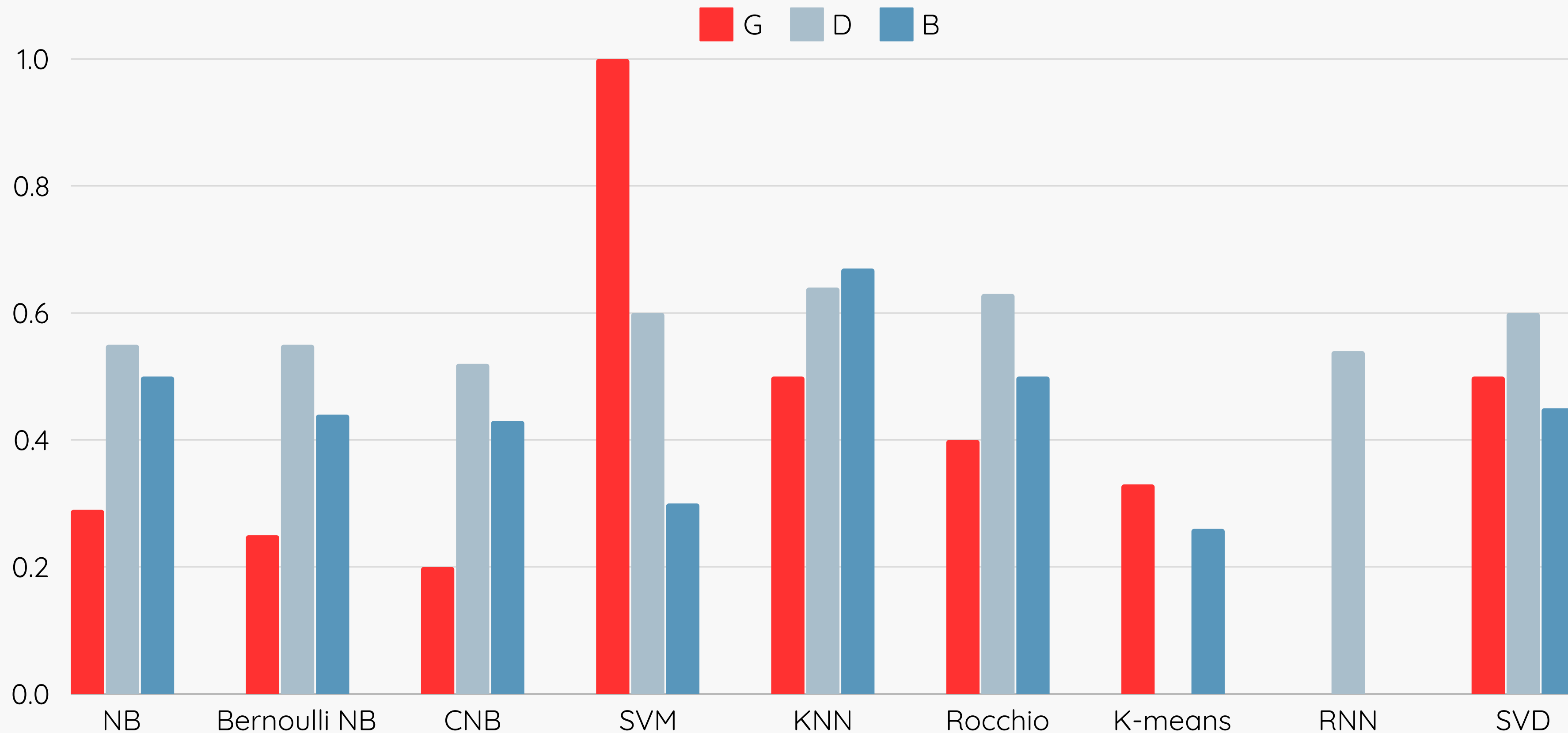
混合模型

# 研究結果 *Gross Margin Precision*





# 研究結果 *ROA Precision*



# 研究結果 *Confusion Matrix*

Predict \ Real	G	D	B
G	9	6	2
D	4	26	7
B	1	8	6

## Ex. NB's ROA predict

- the precision : 0.64 , 0.65 , 0.40

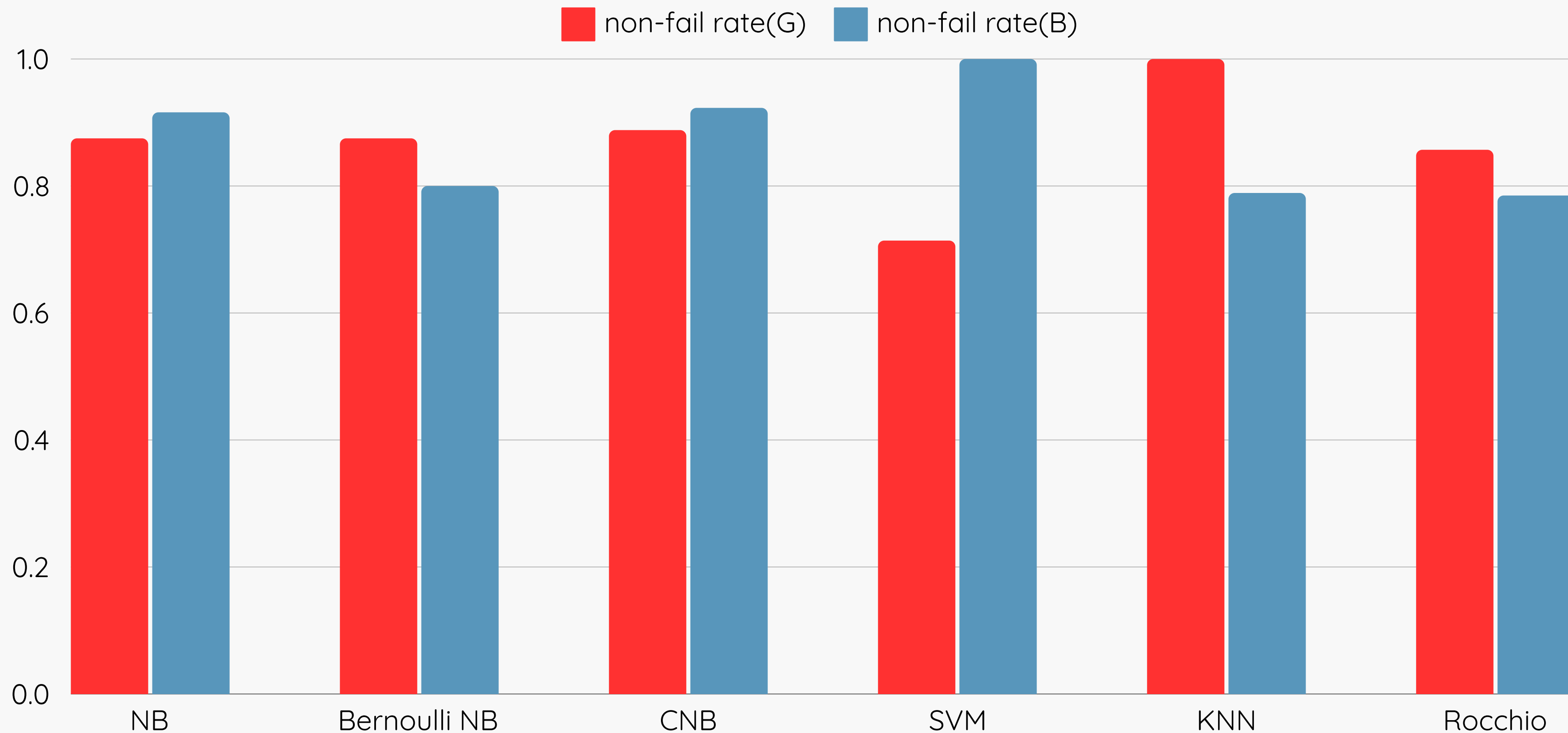
不失敗率(G) =  $\text{realG} + \text{realD} / \text{predictG}$

- is ~93% !

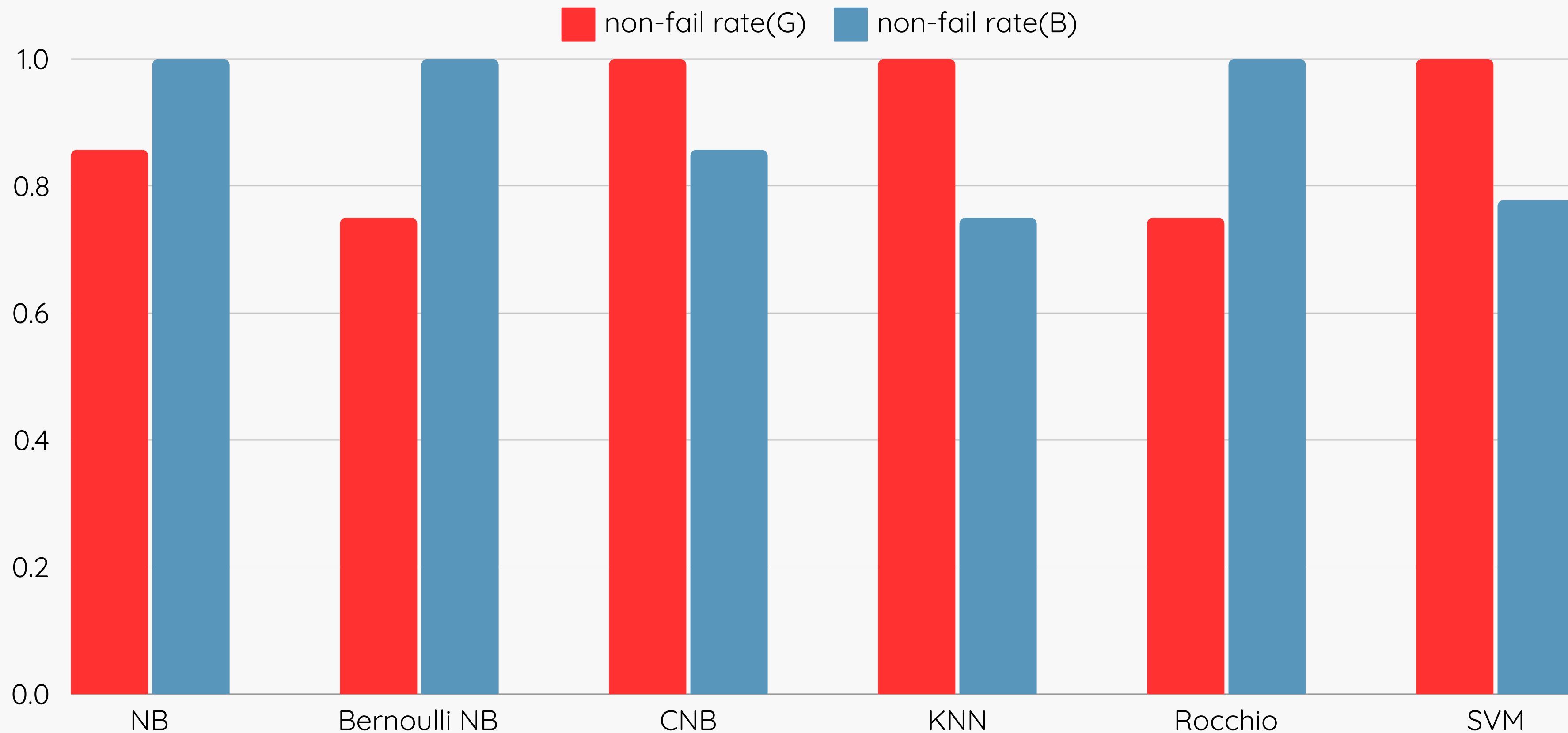
不失敗率(B) =  $\text{realB} + \text{realD} / \text{predictB}$

- is ~87% !

# 研究結果 *Gross Margin non-fail rate*



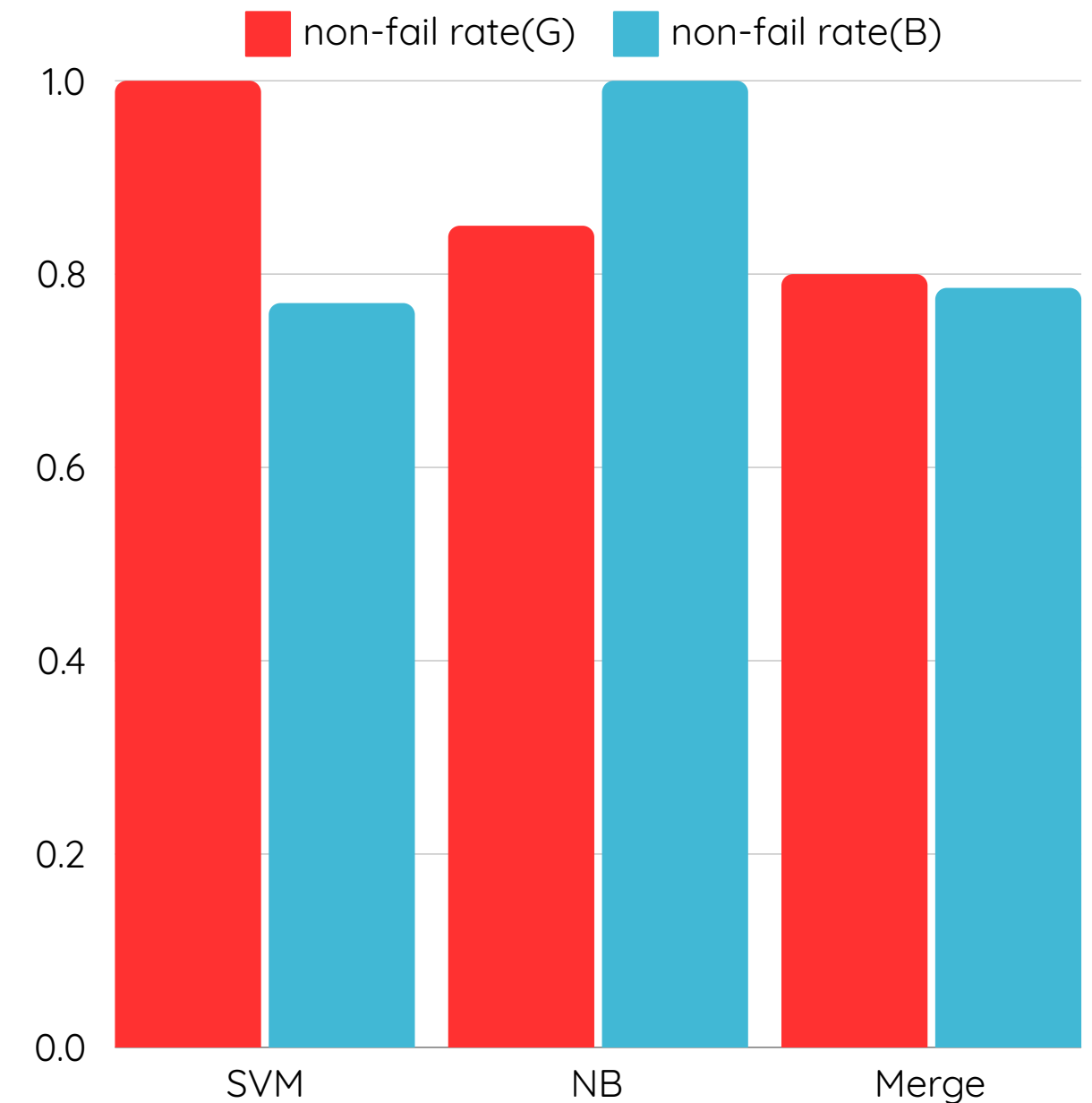
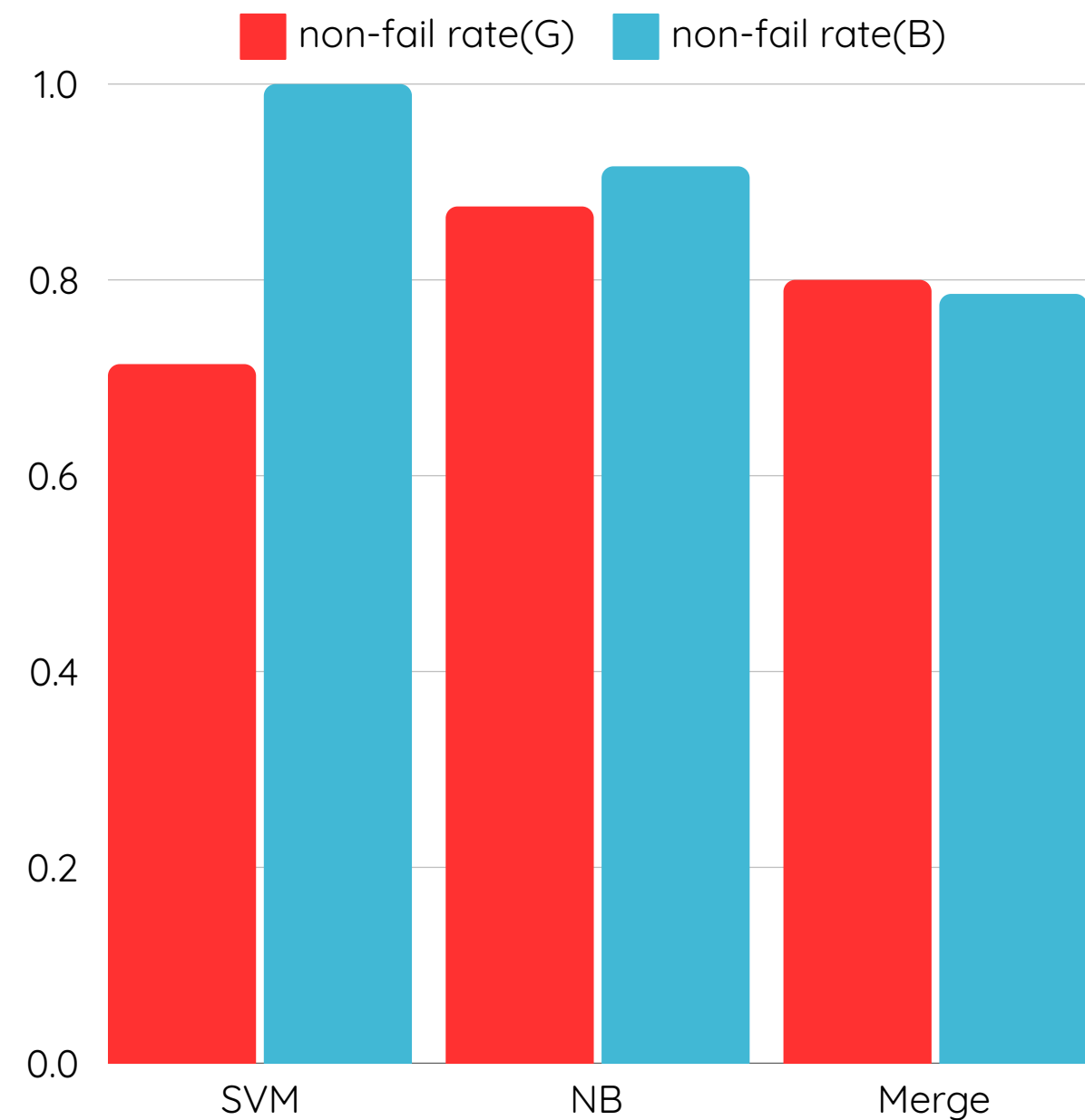
# 研究結果 *ROA non-fail rate*



# 研究結果 *Merge Model*

Merge : SVM , NB

- 動態調整權重+門檻值設定：算出最具信心的類別





# 研究結論

## 改進*Precision*

- 引入non-fail rate
  - 部分模型表現預測佳
  - 可以增加商業操作的成功率

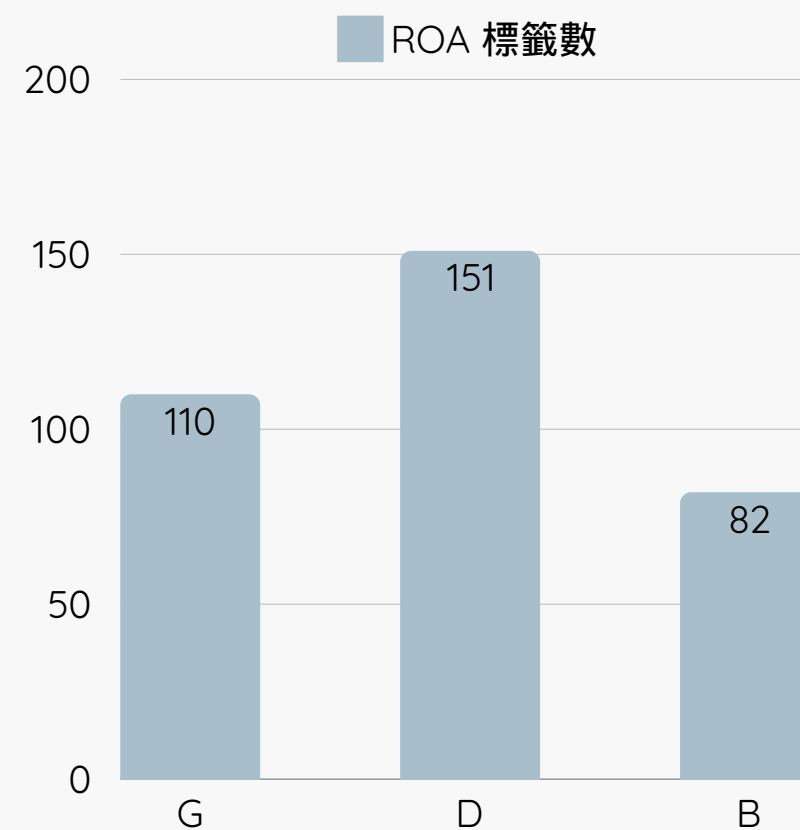
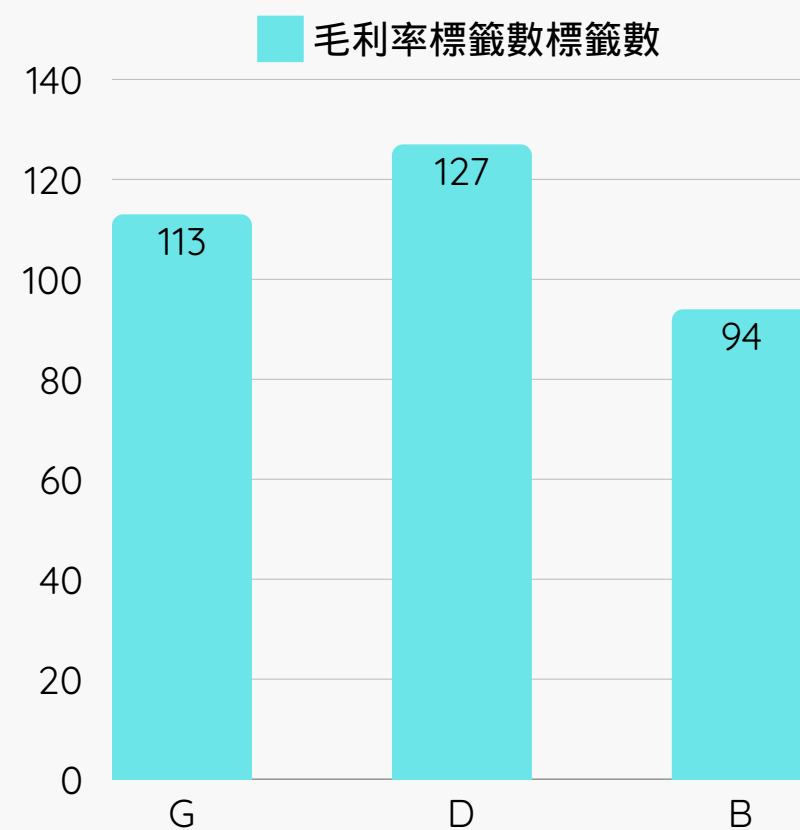
## 模型整合

- 整合方法
  - 組合表現好的模型
  - 權重動態調整
  - 門檻值(置信水平)
  - 試著進一步提升Precision、non-fail rate

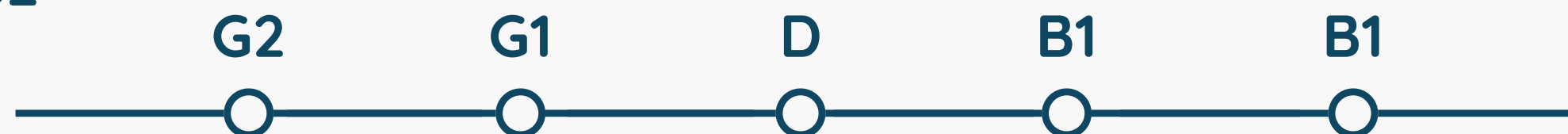
# 未來展望

-商業應用: 預估公司成長、衰退情形

-資料收集



-更細緻的分類: 可以分為 G2 , G1 , D , B1 , B2



-更多模型: 引入傳統機器學習的模型 (決策樹、隨機森林)



*Thank you*

