



# 財經新聞情緒因子多因子選股策略

Presenter: 吳翼名

Mentor: 張哲瑋、陳育理

## 目錄

策略構想與說明

情緒分析模型

情緒分析因子

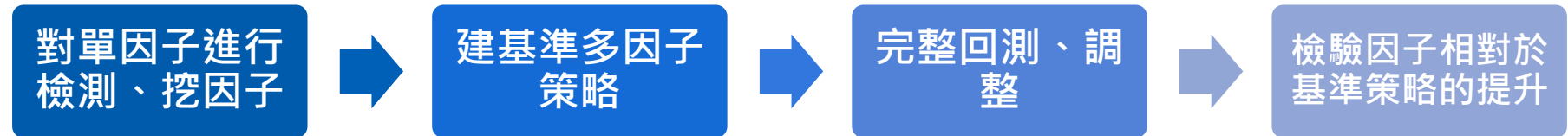
回測績效

策略優化

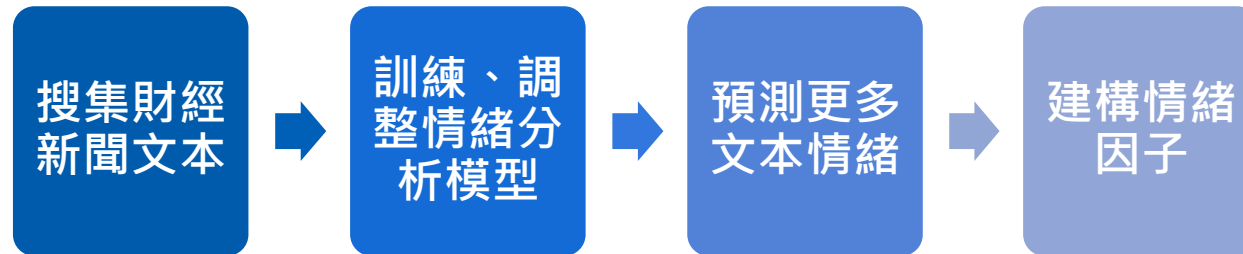


# 策略構想與說明

## 基準多因子策略



## 情緒分析模型



# 情緒分析模型

---

## 目標

- 根據新聞文本內容對於股價的影響，給予正面(1)、中立(0)、負面(-1)評價

## 訓練與Validation資料集

- 訓練新聞資料來源：TEJ金融大事記
- 人工Label情緒
- 新聞數：原始收集文本數2,049 → Down Sample至比例為(-1:0:1 = 1:3:2)

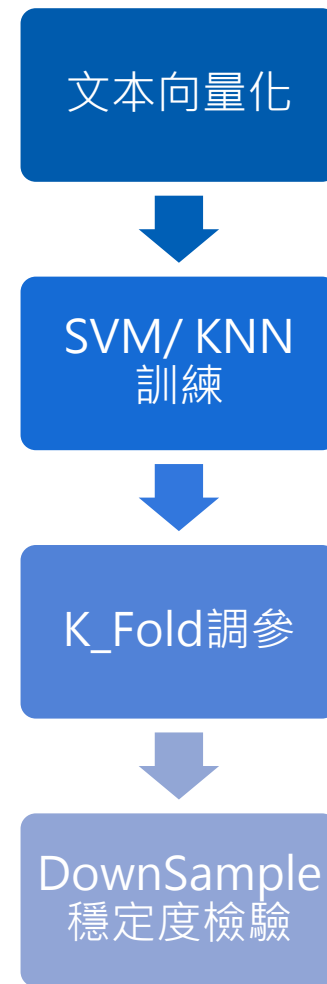
## Model選擇

- 預訓練的BERT (RoBERTa)進行embedding，取CLS Token
- 以RBF SVM, KNN進行Classification

# 情緒分析模型測試

Random Seed 0-99下的模型表現

	RBF SVM	KNN
Avg. F1_Macro	$0.7874 \pm 0.0214$	$0.7683 \pm 0.0215$
Avg. Precision	$0.7901 \pm 0.0239$	$0.7746 \pm 0.0220$
Avg. Recall	$0.7877 \pm 0.0225$	$0.7678 \pm 0.0232$
Avg. Accuracy	$0.7888 \pm 0.0206$	$0.7669 \pm 0.0205$
Best F1	0.8333	0.8145
Worst F1	0.7312	0.7128



# 情緒因子建構

## 初始因子值生成

- 取SVM, KNN預測平均  
(預測結果相關係數0.75)

```
Rbf SVM預測值分布factor_1
0      250587
-1     106568
1      105145
Name: count, dtype: int64
KNN預測值分布factor_2
1      179588
-1     143344
0      139368
Name: count, dtype: int64
因子值分布factor
0.0     134556
-1.0    103336
1.0     95775
0.5     89288
-0.5    39345
Name: count, dtype: int64
```

## 因子值加工

- 考慮新聞影響力非單日
- 衡量橫截面相對情緒比較

$$\tilde{X}_{t,i} = \begin{cases} X_{t,i}, & \text{若 } X_{t,i} \text{ 有值} \\ 0, & \text{若 } X_{t,i} \text{ 為 NaN} \end{cases}$$

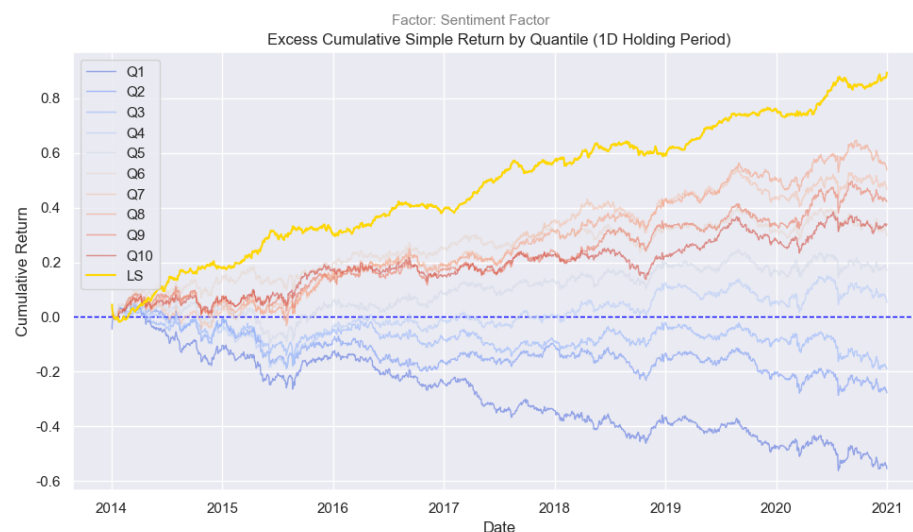
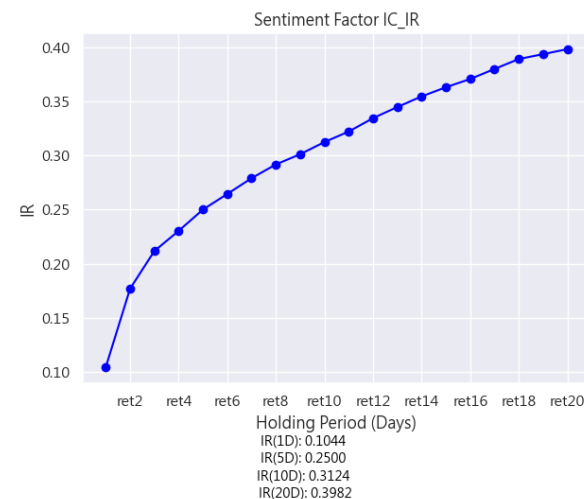
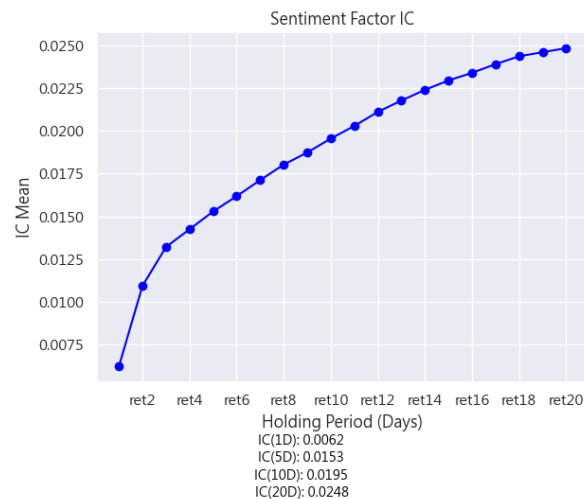
$$F_{t,i} = \text{cs\_rank}\left(\alpha \tilde{X}_{t,i} + (1 - \alpha) S_{t-1,i}\right), \quad \alpha = 0.85$$

\*若單日有複數篇新聞則疊加初始值

# 情緒因子建構

## 情緒因子檢定

主要資訊	
IC Mean	0.0153
IC Std.	0.0611
IC_IR	0.25



# 回測假設

---

- 商品：台股上市櫃不含金融、KY股
- 時間頻率：5日換倉
- 樣本內區間：2014/01/01 ~ 2020/12/31
- 樣本外區間：2021/01/01 ~ 2025/06/30
- 交易成本：買進0.1425%、賣出(0.1425% + 0.3%)



# 基準策略邏輯

---

## 建立單因子資料庫

- 對各類價量、財報、籌碼因子測試RankIC, IR, Sharpe, CAGR等
- 以TPE演算法組合因子並測試，擴充因子庫

## 挑選因子建立多因子策略

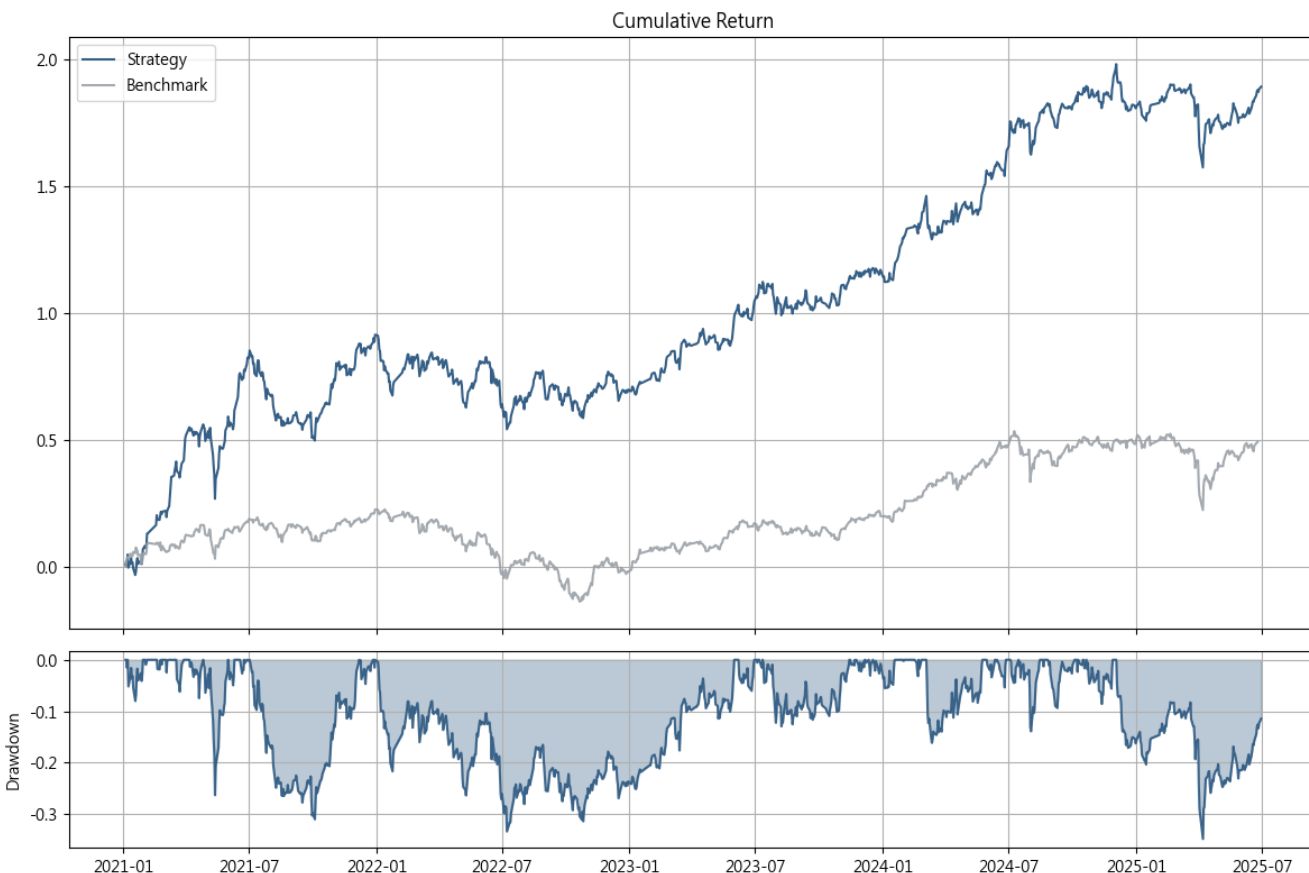
- 篩選IC, Sharpe值最高的40個因子
- XGBoost預測建立多因子策略 (樣本內拆8:2訓練、測試再全數訓練)  
→ 設定early stopping避免overfit

## 根據合成因子選股

- 轉倉日挑選因子值最高的10支個股等權做多

# 回測績效

## 基準策略 - 樣本外

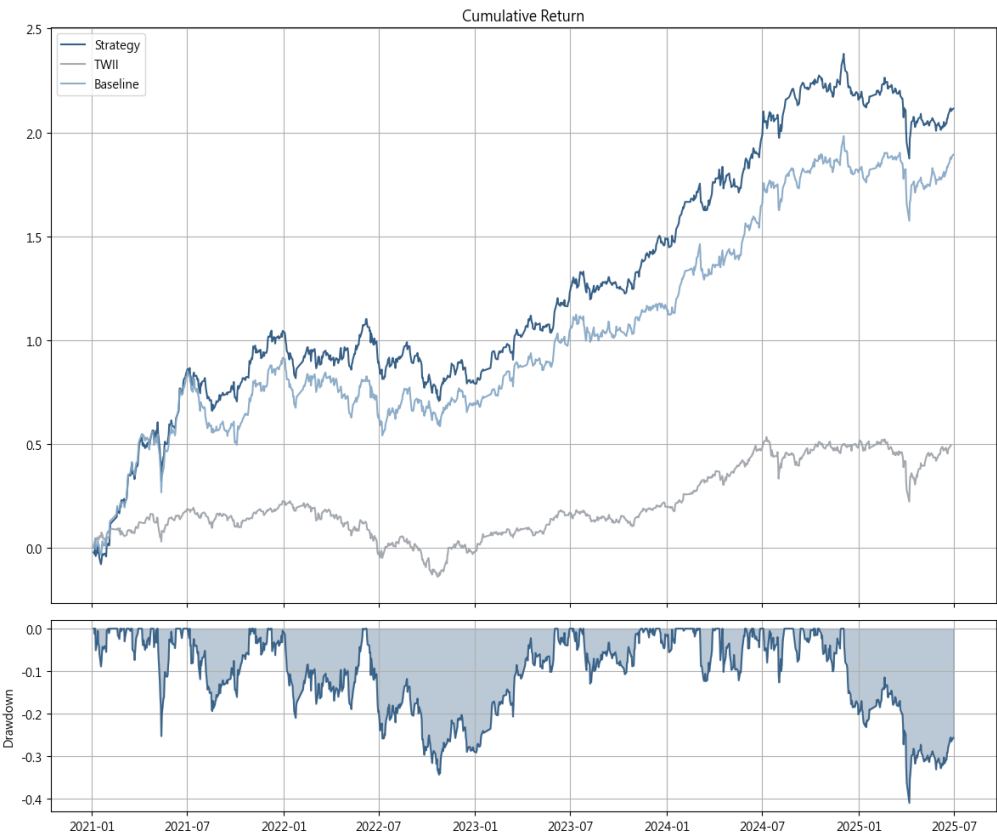


### 樣本外績效

累積報酬(%)	189.37
年化報酬(%)	46.7
年化波動(%)	33.25
MDD(%)	-34.84
年化Sharpe	1.32
年化Calmar	1.34
平均周轉率(%)	134.4

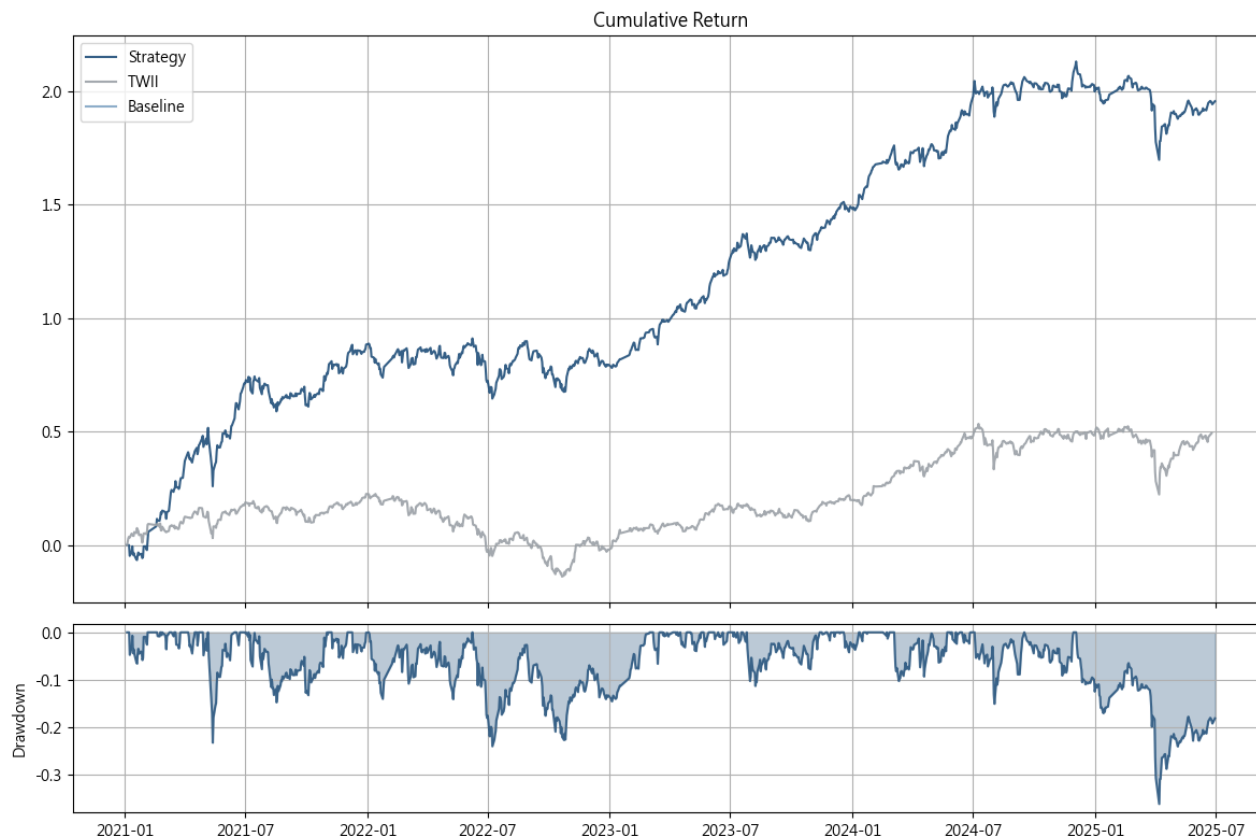
# 回測績效

## 加入情緒因子 - 樣本外



	情緒因子	基準策略	大盤
累積報酬(%)	211.50	189.37	52
年化報酬(%)	54.36	46.7	10.43
年化波動(%)	33.37	33.25	19.7
MDD(%)	-40.95	-34.84	-31.63
年化Sharpe	1.47	1.32	0.6
年化Calmar	1.33	1.34	0.33
平均周轉率(%)	141.6	134.4	N/A

## 五個資金通道 – 樣本外



### 樣本外績效

累積報酬(%) 195.51

年化報酬(%) 50.89

年化波動(%) 28.73

MDD(%) -36.27

年化Sharpe 1.58

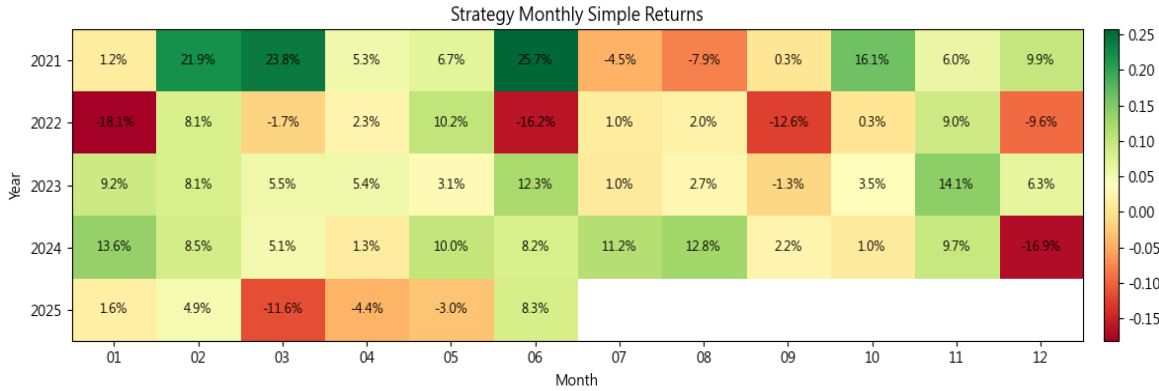
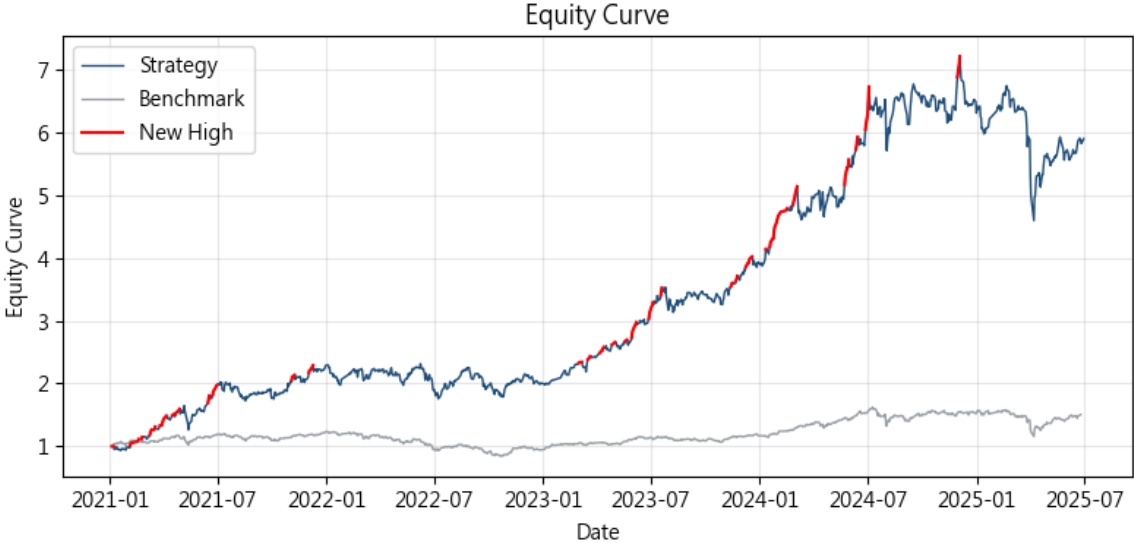
年化Calmar 1.4

平均周轉率(%) 28.67

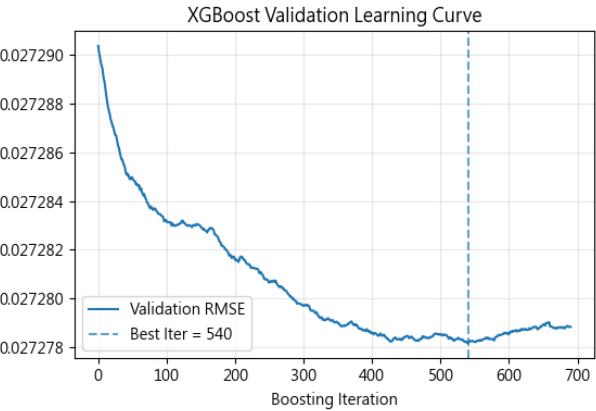
\*模擬五個資金通道，每日調整20%部位，避免選擇操作日有所偏差

# 策略優化

## 五個資金通道 – 其他回測圖表



情緒因子於XGBoost數據		
項目	數值	排名
Gain	195.51	21
Weight	825.0	4
Feature Importance	0.025	21



# 未來策略優化方向構想

---

## 情緒分析模型優化

- 提高訓練文本數量
- 數字資訊處理的優化 → 例如(2% vs. 200%)
- 考慮非線性模型、LLM → 更完整解讀語意

## 多因子策略優化

- 降低手續費磨損 → 周轉率過高
- 因子失效問題 → 滾動式計算IC篩選因子
- 對大型崩盤事件的避險 → 搭配做空或其他商品對沖風險



**感謝聆聽，敬請指教**