



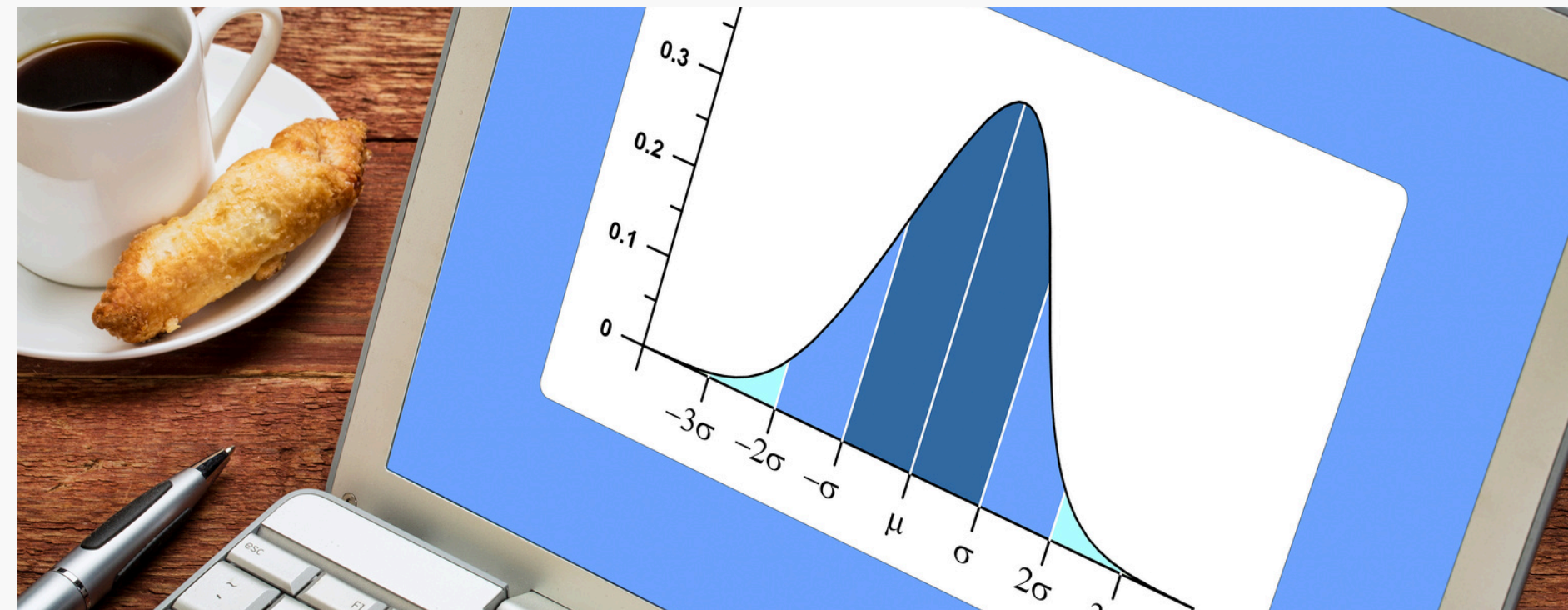
# 結合大離差理論與機器學習之 預測風險指標建構

110071025 計財大四 張智傑

09 June, 2025



# 目錄



01 研究動機與目的

03 整體效能評估

02 資料與風險指標建構

04 結論與延伸應用

# 研究動機與目的



- 在金融市場中，極端事件雖然發生機率低，但其對投資組合的淨值與策略報酬往往造成非對稱且劇烈的衝擊。由於此類事件多屬於**系統性風險**，難以透過傳統的投資組合多角化手段加以分散，因此如何有效應對成為投資管理中的一大挑戰
- 為了對抗此類風險，投資人常透過期貨與選擇權等衍生性金融商品，為原本的資產部位進行**避險操作**，降低面臨市場劇烈波動時的資本損失。然而，若**無法有效掌握極端事件發生的時機**，長期持有看跌選擇權或不斷放空期貨，將可能大幅侵蝕投資報酬，使避險策略本身成為績效拖累來源
- 因此，若能設計一套具備前瞻性、可預警極端風險的判斷機制，協助投資人在適當的時點啟動風險控制行動，將有助於提升避險效率，並保留多數時間的資產報酬潛力

# 摘要



本研究導入**大離差理論** (Large Deviation Principle, LDP) 之觀念，建構出反映報酬異常偏離程度之風險函數 (Rate Function)，再結合**市場統計特徵** (如 **Z-score**、**偏態**、**峰態**、**動能與VaR**) 設計Future Risk Score作為極端事件預測指標。

本研究主要以**邏輯回歸模型**為基礎，利用 2015 至 2019 年之歷史資料訓練，2020 至 2025 年資料驗證模型效能，成功建立可預測未來市場**是否出現超過 -3% 報酬變動**的極端風險事件的分數機制。經 AUC 分析與混淆矩陣評估，結果顯示該風險指標對於重大市場波動具有一定前瞻性，並可作為日後選擇權策略 (如 Covered Call) 調整機制之核心判斷依據。

# 投資組合之設計

---



- 資料期間：2015.01.01至 2025.05.31
- 資料頻率：日資料（每日投資組合報酬率）
- 投資組合：以 0050/0051 混合策略為主，主要想複製大盤指數，0050配置70%、0051配置30%
- 分析目標：預測未來是否出現極端市場事件，作為風險控制依據
- 模型訓練與驗證分割：
- 訓練集：2015-2019 年（歷史資料訓練）
- 測試集：2020-2025 年（模擬實戰預測）



# 極端事件定義

---



- 定義依據：根據投資組合每日報酬率
- 事件標準：當日報酬率跌幅超過 -5%，即視為一次極端市場事件（避免樣本數過少）  
選擇只關注「下跌」事件，原因如下：
- 通常極端事件都是代表市場短期內快速重挫，好比4月份的川普帶來之不確定性
- 大漲時即使錯過上行獲利，仍可保有原始資產報酬，不構成實質風險，只是少賺





# 風險指標建構

## • Rate Function 指標設計

### 前提假設

- 假設每日報酬  $r_t$  近似服從常態分布：

$$r_t \sim \mathcal{N}(\mu_t, \sigma_t^2)$$

- 參考大型偏差理論 (Large Deviation Principle, LDP)，  
我們關心隨機變數偏離其期望時，  
機率衰減速度所對應的 **rate function**  $I(x)$ 。

### 一般形式——Cramer Rate Function

$$I(x) = \sup_{\theta \in \mathbb{R}} \left\{ \theta x - \log \mathbb{E}[e^{\theta X}] \right\},$$

其中

$$\mathbb{E}[e^{\theta X}]$$

為隨機變數  $X$  的 mgf (moment generating function)。

### 套用至常態分布

對  $X \sim \mathcal{N}(\mu, \sigma^2)$  而言，mgf 為

$$\mathbb{E}[e^{\theta X}] = \exp\left(\theta\mu + \frac{1}{2}\theta^2\sigma^2\right).$$

將之帶入Cramer Rate Function，取  $\sup_{\theta}$  後可得

$$I(x) = \frac{(x - \mu)^2}{2\sigma^2}.$$

### 本研究之 Rate Function 指標

因此，對每日報酬  $r_t$ ，令

$$\mu_t = \text{rolling\_mean}_{60}(r_{t-1}, \dots, r_{t-60}), \quad \sigma_t = \text{rolling\_std}_{60}(r_{t-1}, \dots, r_{t-60}),$$

則 大離差風險指標 定義為

$$\text{rate\_func}_t = \frac{(r_t - \mu_t)^2}{2\sigma_t^2}.$$

rate function數值越大，代表當日報酬對歷史分布而言越罕見，  
亦即潛在極端事件風險越高。

# 風險指標建構



$$\mu_t = \frac{1}{60} \sum_{i=1}^{60} r_{t-i}, \quad \text{Var}_t = \frac{1}{60} \sum_{i=1}^{60} (r_{t-i} - \mu_t)^2, \quad \sigma_t = \sqrt{\text{Var}_t}.$$

- Z-score

標準化的偏離量，保留「向上或向下」方向資訊，利於捕捉跌勢前兆。

- Skewness（偏態）

描述報酬分布是否向下或向上傾斜，負偏態顯示下尾肥厚，預示市場更可能出現大幅下跌。

- Kurtosis（峰度）

衡量分布尾部肥厚與尖銳程度，峰度越高，代表極端波動（跳躍）發生機率上升

- Abs Momentum（絕對動能）

觀察近期趨勢力度，動能過大時，市場更易出現劇烈回檔或急速反轉，對避險決策具前瞻性。

- VaR<sub>5</sub>（歷史 5% VaR）

意義：過去 60 日內最差 5% 報酬的臨界值。量化近期尾部損失水位，給出系統性風險「底線」參考。

$$\text{Skew}_t = \frac{1}{60} \sum_{i=1}^{60} \left( \frac{r_{t-i} - \mu_t}{\sqrt{\text{Var}_t}} \right)^3$$

$$\text{Kurt}_t = \frac{1}{60} \sum_{i=1}^{60} \left( \frac{r_{t-i} - \mu_t}{\sqrt{\text{Var}_t}} \right)^4 - 3$$

$$\text{AbsMom}_t = \left| \frac{1}{60} \sum_{i=1}^{60} r_{t-i} \right|$$

$$\text{VaR}_t^{(\alpha)} = - \inf \left\{ x \in \mathbb{R} : F_{t-1}(x) \geq \alpha \right\}$$



# 模型設計與回歸分析結果



- 採用Logistic Regression  
(二元分類，輸出未來極端事件機率)
- 設定5個指標都是使用過去60日之報酬資料去計算，設定不斷rolling操作
- 訓練集：  
2015.01.01~2019.12.31
- 測試集：  
2020.01.01~2025.05.31

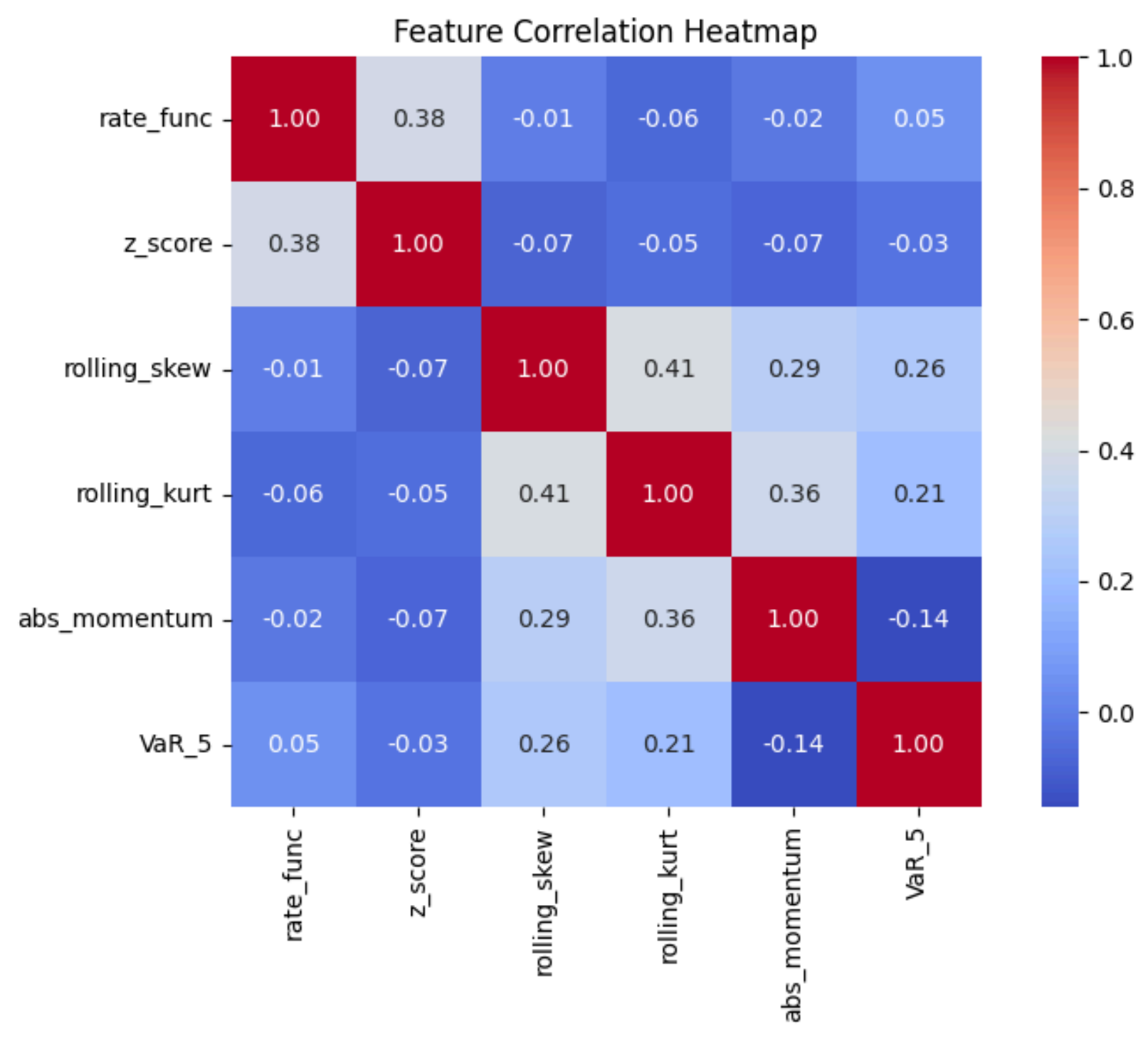
```
Optimization terminated successfully.
Current function value: 0.162947
Iterations 10

Logit Regression Results
=====
Dep. Variable:      extreme_event    No. Observations:      2468
Model:              Logit           Df Residuals:          2462
Method:              MLE            Df Model:              5
Date:               Mon, 09 Jun 2025  Pseudo R-squ.:        0.2560
Time:               06:07:07          Log-Likelihood:       -402.15
converged:           True             LL-Null:              540.49
Covariance Type:    nonrobust         LLR p-value:          1.029e-57
=====
```

	coef	std err	z	P> z	[0.025	0.975]
rate_func	0.2164	0.031	6.922	0.000	0.155	0.278
z_score	-1.9835	0.131	-15.126	0.000	-2.240	-1.726
rolling_skew	-0.1617	0.049	-3.315	0.001	-0.257	-0.066
rolling_kurt	-0.0392	0.009	-4.179	0.000	-0.058	-0.021
abs_momentum	-203.6571	68.610	-2.968	0.003	-338.129	-69.185
VaR_5	90.5498	6.521	13.886	0.000	77.769	103.331

```
=====
```

# 特徵之相關性觀察



VIF 結果：

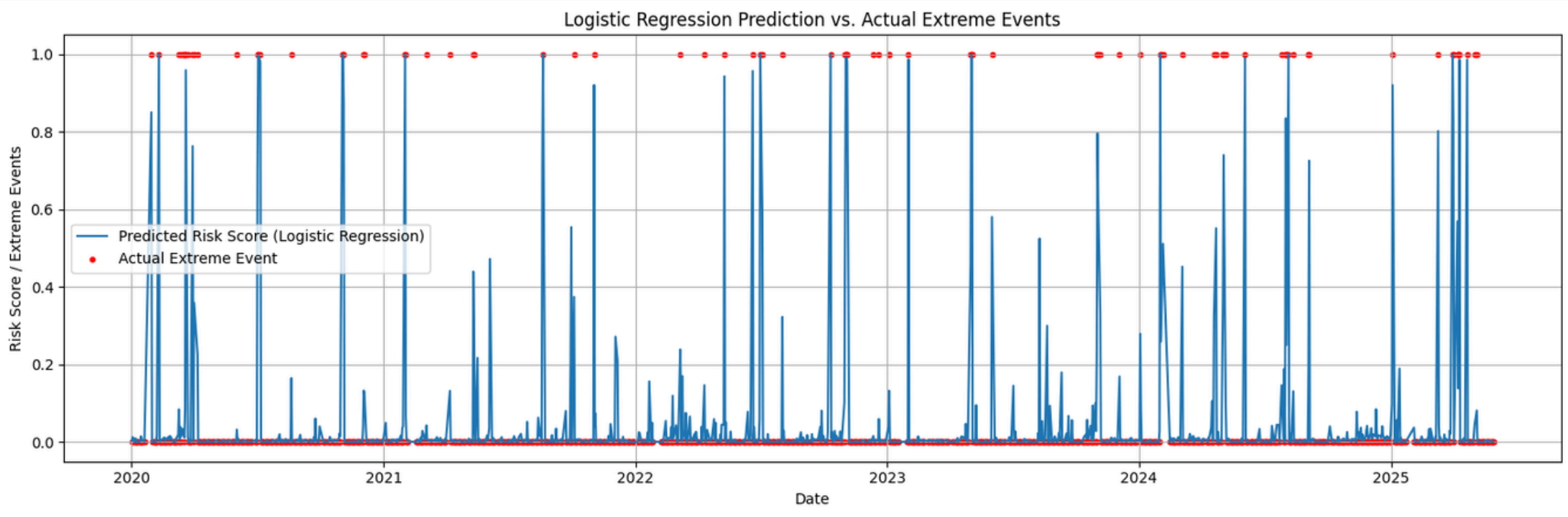
	Feature	VIF
0	const	5.980186
1	rate_func	1.181814
2	z_score	1.181708
3	rolling_skew	1.321346
4	rolling_kurt	1.372431
5	abs_momentum	1.296589
6	VaR_5	1.193991

# 整體預測能力



	precision	recall	f1-score	support
0	0.974441	0.996732	0.985460	1224.000000
1	0.931034	0.627907	0.750000	86.000000
accuracy	0.972519	0.972519	0.972519	0.972519
macro avg	0.952738	0.812320	0.867730	1310.000000
weighted avg	0.971591	0.972519	0.970003	1310.000000

	預測:正常日	預測:極端日
實際:正常日	1220	4
實際:極端日	32	54



# 結論與延伸發展

---



透過大離差等5個指標，結合機器學習模型的檢測可發現，對於極端事件的預測能力已經超過5成，可協助規避掉一半以上的下行風險。

未來會再去著手的是將其加入到Covered Call策略中，觀察是否表現可以勝過原本的 Covered Call 策略以及大盤

- 指標判別市場可能出現下行時，投資組合+ Buy Put
- 指標判別市場處於穩定or 小漲小跌，投資組合 + Short Call
- 指標判別存在大漲可能性時，投資組合 + Buy Call

2. 將大離差模型推展至Heston或其他模型去做指標，讓其可以更好詮釋股價之跳點與厚尾分布

3. 極端之設計優化





*Thank you*

