

1. 我設定的三組 window size 和 step 組合與 MSE 如下表

組合	結果
window_size = 5, step = 2	Test MSE: 6.3304
window_size = 10, step = 15	Test MSE: 517.2670
window_size = 20, step = 5	Test MSE: 17.9673

window_size = 5, step = 2 組合取得最小的 MSE，表示此配置的預測效果最佳，推測是因為較小的 window_size 能捕捉到更精確的時序資訊，step 也適中，使模型能更好學習趨勢；window_size = 10, step = 15 組合的 MSE 較高，可能是因 step 較大，模型無法有效學到足夠的時序特徵，從而影響了預測準確性；window_size = 20, step = 5 組合的 MSE 介於兩者之間，雖然提供了更多的時序資訊，但 step 依然過大，導致預測效果稍遜。

2.

a. Test MSE: 1334.4651

加入 Volume 後，Test MSE 從 517.2670 上升至 1382.2622，預測表現下降。推測原因為 Volume 與價格波動關聯性不明顯，且數值範圍與其他特徵有較大差異，使模型學習無效資訊，甚至過擬合，故未改善結果。

b.

Input features 組合	MSE
'Open', 'Close'	Test MSE: 372.3008
'Open', 'Close', 'High'	Test MSE: 742.3602
'Open', 'Close', 'Low'	Test MSE: 458.7678
'High', 'Low'	Test MSE: 407.8056
'Open', 'High', 'Low'	Test MSE: 891.8744
'High', 'Low', 'Close'	Test MSE: 436.1296
'Open', 'High'	Test MSE: 735.5635
'Open', 'Low'	Test MSE: 595.5718

Input features 組合'Open', 'Close'的模型預測表現最佳，Test MSE 為 372.3008，顯示開盤價與收盤價對於預測收盤價趨勢有較高的關聯，且資訊足夠、雜訊較少。反觀加入其他特徵後，MSE 普遍上升，推測可能因 features 之間存在多重共線性，且部分 features 對預測收盤價的貢獻度有限，反而導致模型學習效果被分散。

3. 在未正則化狀態下，五次 Test MSE 如下，平均為 397.2256

Test MSE: 517.2670 Test MSE: 392.3911 Test MSE: 292.8851 Test MSE: 404.2083 Test MSE: 381.3763

在正則化狀態下，五次 Test MSE 如下，平均 Test MSE 為 328.4320

Test MSE: 292.5501 Test MSE: 426.9289 Test MSE: 276.5851 Test MSE: 349.7940 Test MSE: 296.3020

正則化後的模型平均 MSE 較低，且預測結果相對穩定，證明資料正則化有助於提升模型學習效率與預測準確性，此現象主要歸因於正則化能消除不同特徵尺度差異對模型學習造成的干擾，使模型能更均衡學習各特徵對價格預測的貢獻，從而提升整體的預測準確性。

4. Lab 4 中設定 window size 小於 step size 主要是為了降低樣本重疊率，避免資料過度重複，導致模型過擬合或學習效益不佳。然而此規則不適用所有情況，若資料變動頻繁，適當重疊能保留時間序列的連續性，幫助模型更好掌握時序特徵，因此實務上常見 window size 等於或大於 step size 的做法，反而更能穩定模型效能。

5. 我找的這篇研究提出一種結合二維轉換與生成對抗網路(GAN)的資料增強方法，解決時間序列資料量不足的問題。首先，利用 Gramian Angular Difference Field (GADF) 將一維時間序列轉換成二維圖像，保留其時序特徵與結構資訊；接著透過 CycleGAN 生成新的影像樣本，以達到增強資料的目的。此方法能在無需配對樣本的情況下產生高品質、多樣化的資料，實驗結果顯示能顯著提升分類模型的準確率與穩定性。

參考文獻：Chen, X., Jin, X., Zhang, H., Xiong, J., Deng, Y., & Zhang, X. (2025). Temporal 2D-cycle-generation framework for time series classification. *Applied Soft Computing*, 171, 112778.

<https://doi.org/10.1016/j.asoc.2025.112778>

6.

- a. Convolution-based models：CNN 擅長擷取局部特徵，因此推論時常使用固定長度的 window size，並根據 convolution kernel 大小和 layer number 確定 receptive field，然而其對長距離依賴的建模能力有限，因此若任務需考慮全域資訊，可能需擴大 window size 或搭配注意力機制來補充。
- b. Recurrent-based models：RNN 或 LSTM 根據時間序列逐步處理輸入，因此 window size 直接影響模型的記憶長度，推論時可根據訓練階段使用的 window size 調整，但過長的 window size 可能造成梯度消失或計算負擔，因此需要在效能和穩定性之間找到平衡。
- c. Transformer-based models：Transformer 架構透過自注意力機制能夠同時考慮整個輸入的 window size，非常適合處理變長序列，推論時 window size 可彈性調整，不受固定長度限制，而現代改良如 gated residual attention 能進一步提升效率與預測準確性，使其成為時序任務中處理 window size 的好選擇。