

HW4 113034557 楊興元

1.

Window size	step	MSE
10	5	17.6615
20	5	17.6387
30	10	191.2341

從實驗結果來看，當 window size 為 10 與 20 且 step 為 5 時，模型的 MSE 表現相近，預測效果穩定。而當 window size 提高到 30 且 step 為 10 時，MSE 明顯上升，推測可能是樣本過少導致模型學習不足，或長序列反而造成過擬合與預測不穩。適中長度的輸入窗口有助於捕捉有效特徵。

2.

(i)在固定 window size = 20、step = 5 的情況下，我將 Volume 納入原本的輸入特徵（Open, High, Low, Close），作為第五個輸入維度。原始模型的驗證 MSE 為 17.6387，而加入 Volume 後 MSE 明顯上升至 1095.7299。這代表 Volume 這個特徵對模型反而產生了負面影響，可能是因為其數值範圍遠大於其他價格類特徵（平均達千萬等級），若未經適當標準化或正規化，會導致模型在訓練時重心偏移，導致學習困難。因此，在處理股價類資料時，加入量價特徵前應先進行縮放處理。

(ii) 我比較了以下幾組特徵組合：1. Close 單獨 2. Open, High, Low, Close 3.

Open, High, Low, Close + Volume

其中，僅使用 Open, High, Low, Close 四個特徵時的 MSE 最佳（17.6387），比單獨用 Close 更穩定，且比加上未標準化的 Volume 效果好。這顯示股價本身的價格變動特徵對模型預測力貢獻較大，且不適當的特徵可能會干擾學習效果。未來若要加入 Volume，應先進行 Min-Max 或 Z-score 標準化處理，再重新評估其貢獻。

3. 參考: <https://arxiv.org/abs/1502.03167>

在 Lab4 的實驗中，當我沒對 Volume 特徵進行正規化時，模型的 MSE 從 17.6387 急劇上升至 1095.7299，顯示未經處理的數值落差會嚴重干擾模型學習。

正規化可將不同尺度的特徵拉到相近範圍，使模型更容易收斂、提升穩定性。

根據 Ioffe 和 Szegedy (2015) 的研究，Batch Normalization 技術能加速深度神經網路的訓練，提升模型的泛化能力。因此，正規化通常是深度學習中不可或缺的前處理步驟。

4.參考： <https://ieeexplore.ieee.org/document/279181>

在時間序列建模中，window size 通常應大於或等於 step，以確保資料樣本間有重疊、提升學習效果。若 $\text{step} > \text{window size}$ ，可能導致樣本過於稀疏。網路上的研究都提到建議重疊序列能提升模型泛化能力。

5. 參考： <https://arxiv.org/abs/2004.08780>

一種適用於時間序列資料的資料增強方法是「引導式時間扭曲（Guided Time Warping）」。此方法利用動態時間扭曲（Dynamic Time Warping, DTW）和 shapeDTW 的元素對齊特性，將樣本模式的特徵扭曲以匹配參考模式的時間步驟，從而生成新的訓練樣本。此外，該方法引入了一個判別式教師（Discriminative Teacher）作為引導參考，以進一步提高增強資料的品質。在 2020 年的研究中，Iwana 和 Uchida 將此方法應用於 85 個 UCR 時間序列資料集，並在深度卷積神經網路（CNN）和循環神經網路（RNN）上進行評估，結果顯示該方法能有效提升模型的泛化能力。

6.

(i) Convolution-based models

參考：

<https://amirsh15.github.io/Prediction%20Time%20Series%20with%20CNN.pdf>

在處理時間訓練的卷積模型中，輸入的 window size 決定了每個樣本的感受野。推論時需要提供與訓練時相同長度的序列作為輸入，確保卷積核能掃描完整特徵結構。若使用 causal convolution，則能逐步輸出未來預測。

(ii) Recurrent-based models

參考： <https://clgiles.ist.psu.edu/papers/MLJ-finance.pdf>

RNN/LSTM 模型需要完整的時間序列長度做序列記憶，window size 表示模型記憶的時間範圍。推論時應保持與訓練相同的輸入長度。若使用滑動窗口方式，可逐步進行未來點預測。

(iii) Transformer-based models

參考：<https://arxiv.org/abs/2310.06625>

Transformer 架構基於自注意力機制，能同時考慮輸入中的所有時間點。
window size 控制的是輸入的序列上下文範圍。模型在推論時需提供足夠的上下文長度，且需搭配 **position encoding** 確保順序資訊完整。