

National Tsing Hua University

11220IEEM 513600

Deep Learning and Industrial Applications Homework 4

Name: 陳建宏

Student ID: 111003807

Due on 2024/05/02.

Note: DO NOT exceed 3 pages.

1.

	windows	strep	MSE
10 日均線	10	30	Train loss: 572.4265, Val loss: 195.7969, Best Val loss: 195.7969
30 日均線	30	30	Train loss: 222.2696, Val loss: 0.5754, Best Val loss: 0.4804
季線	90	30	Train loss: 670.8779, Val loss: 608.4763, Best Val loss: 608.4763

表格展示了使用不同窗口大小對時間序列數據進行分析時的模型訓練結果，其中「窗口」指的是在計算移動平均或其他統計時考慮的連續數據點的數量。「步長」則是指在進行滾動計算時移動的數據點數。每一行代表一種窗口大小的設定，並展示了相關的訓練損失、驗證損失和最佳驗證損失。具體解釋如下：

10 日均線

訓練損失: **572.4265** - 表示在訓練數據上模型的均方誤差。

驗證損失: **195.7969** - 表示在驗證數據上模型的均方誤差。

最佳驗證損失: **195.7969** - 表示在所有驗證周期中觀測到的最低驗證損失。

30 日均線

訓練損失: **222.2696**

驗證損失: **0.5754**

最佳驗證損失: **0.4804** - 這是在 **30** 日窗口設置下觀測到的最佳表現，表明此設定對模型的學習和泛化能力極為有利。

季線（90 天）

訓練損失: **670.8779**

驗證損失: **608.4763**

最佳驗證損失: **608.4763** - 在 **90** 天窗口下，模型的表現未見改善，可能因為太長的窗口尺寸使得模型難以捕捉到更短期的市場變動。

總結來看，不同窗口大小的設定對模型的性能有顯著影響。**30** 日均線設定在本例中顯示出了最佳的驗證性能，可能因為這個窗口大小既能捕捉到足夠的數據趨勢，又不至於過度平滑化數據，從而遺漏重要的變化。

2.

	windows	strep	MSE
10 日均線	10	1	Train loss: 950.6421, Val loss: 893.3427, Best Val loss: 893.3422
30 日均線	30	1	Train loss: 964.2217, Val loss: 864.7733, Best Val loss: 864.6610
季線	90	1	Train loss: 941.2268, Val loss: 1030.6542, Best Val loss: 1030.6542

分析

從提供的數據來看，**30** 日均線的設定在驗證損失上表現最優，最佳驗證損失達到了 **864.6610**，這是所有配置中最低的。這表明在 **30** 天的窗口大小下，模型能夠更好地泛化，捕捉到股價數據中的趨勢和模式，同時避免了過度擬合或對短期波動的敏感性。

嘗試的理由:選擇不同的窗口大小主要是為了評估模型對於長短期數據變化的敏感性。較小的窗口（如 10 日）可能使模型對短期波動更敏感，而較大的窗口（如 90 日）可能會捕捉到更長期的趨勢，但可能忽略了一些重要的短期信號。

最佳特徵組合:考慮到 30 日均線在驗證數據上的表現最佳，結合成交量（Volume）作為額外的特徵可能提供了關於市場活躍度和潛在價格變動壓力的有價值信息，這有助於模型更準確地預測股價變動。因此，將「開盤價」、「最高價」、「最低價」、「收盤價」以及「成交量」作為輸入特徵，配合 30 日的窗口大小，可能是最佳的特徵組合。

結論:綜上所述，30 日窗口大小，結合成交量作為輸入特徵，提供了最優的模型性能。這種組合可能最好地平衡了捕捉足夠的市場動態與避免過多的噪音，從而優化了模型的預測準確性。在實際應用中，可以進一步實驗不同特徵的組合和調整窗口大小來細化模型的性能。

3. 在實驗室 4 中，比較模型在使用標準化與未標準化輸入的表現顯示了特徵縮放對於神經網絡訓練的重要性，特別是對於金融市場這樣的時間序列數據。標準化通常可以通過確保每個特徵對學習過程的平等貢獻來提高模型性能，避免由於尺度較大的特徵而主導。實驗結果常常顯示，使用標準化輸入的訓練有更好的收斂速度和更穩定的訓練過程

（Goodfellow, I., et al. (2016) Deep Learning. MIT Press, Cambridge, MA.）。在金融時間序列預測中，標準化有助於模型更有效地捕捉微妙的趨勢和波動性，從而提高了精確度和在不同市場條件下的泛化能力

4. 在實驗室 4 中，窗口大小應該小於步長的說法是不正確的。通常情況下，窗口大小應大於或等於步長，這樣才能保證數據的連續性和覆蓋性，避免遺漏重要信息。步長太大可能會導致數據樣本之間的重疊不足，影響模型學習效果。

5. 一種適用於時間序列數據的數據增強方法是使用滑動窗口技術。這種技術通過在整個數據集上滑動一個固定大小的窗口來創建新的訓練樣本，從而增加數據的多樣性和量。窗口可以按照一定的步長滑動，每次滑動產生的序列稍有不同，這有助於模型學習更多的時間依賴性和泛化能力。這種方法尤其適用於股價預測、氣象預報等領域，可以有效提高模型對未來事件的預測準確性。

Arthur Le Guennec, Simon Malinowski, Romain Tavenard. Data Augmentation for Time Series Classification using Convolutional Neural Networks. ECML/PKDD Workshop on Advanced Analytics and Learning on Temporal Data, Sep 2016, Riva Del Garda, Italy.

6.在不同模型架構中，處理推理階段的窗口大小有其特定方法：

(i) Convolution-based models

在基於卷積的模型中，窗口大小通常對應於卷積核的接受域，決定了模型一次可以處理的輸入數據範圍。在推理時，窗口大小應保持一致，以確保模型能夠捕捉到訓練階段學到的空間特徵。這對於確保模型輸出的一致性和準確性至關重要。

(ii) Recurrent-based models

對於循環神經網絡（如 LSTM 或 GRU），窗口大小決定了每次前向傳播時輸入序列的長度。在推理時，保持窗口大小與訓練時相同是重要的，這樣模型才能正確地利用學到的時間依賴性。然而，也可以通過調整窗口大小來適應不同長度的輸入序列，但這可能需要額外的技術，如序列填充或截斷。

(iii) Transformer-based models 模型通常對序列長度具有靈活性，因為它們通過自注意機制全局處理輸入。在推理時，可以靈活調整窗口大小而不損害模型性能，因為 Transformer 能夠處理不同長度的輸入。這使得變壓器特別適合於需要處理長期依賴關係的應用，如機器翻譯或文本生成。