

**National Tsing Hua University**  
**11220IEEM 513600**  
**Deep Learning and Industrial Applications**  
**Homework 4**

Name: 曾聖閔

Student ID: 112034564

**Due on 2024/05/02.**

**Note: DO NOT exceed 3 pages.**

1. (15 points) Experiment with different window sizes and steps. Train the model using **3** different combinations of window size and step. Evaluate the Mean Squared Error (MSE) for each configuration. Report the MSEs using a table and analyze the results. (Approximately 100 words.)

Window sizes	Steps	Mse
10	15	434.31823...
20	10	183.2007...
50	25	1559.273071

挑選合適的窗口大小，較大可以捕捉較多信息，但較小容易忽略。步長大小容易影響資料重疊程度。窗口大小和步長大小的選擇需要在信息捕獲能力、模型複雜性、計算成本和時間序列模式等方面進行權衡，並根據具體情況選擇最合適的值。選擇20window sizes以及一半的步長效果感覺較好。

2. (Approximately 200 words.)

- (i) (15 points) Include 'Volume' as an additional input feature in your model. Discuss the impact of incorporating 'Volume' on the model's performance.

以window sizes=20，steps=10為例，加入Volume後MSE變1374.8441162109375，可能有以下原因

1. 特徵重要性差異：Volume 特徵與其他特徵（如 Open、High、Low、Close）可能有很大的重要性差異。如果 Volume 的範圍遠大於其他特徵，則模型可能更關注 Volume 特徵，而忽略其他特徵，導致對股價的預測準確度降低。
2. 特徵相關性：Volume 特徵與股價之間可能存在較弱的相關性。如果 Volume 與股價的相關性較低，則將 Volume 特徵添加到模型中可能會引入噪音，從而降低模型的預測性能。
3. 模型過擬合：當添加額外的特徵時，模型的複雜度增加，容易導致過擬合。如果 Volume 特徵未能提供足夠的信息來改善模型的泛化能力，則模型可能會過度擬合訓練數據，導致在驗證集或測試集上的性能下降。

- (ii) (15 points) Explore and report on the best combination of input features that yields the best MSE. Briefly describe the reasons of your attempts and analyze the final, optimal input combination.

將特徵選擇從2個到5個為輸入，若要預測股價，開盤價、收盤價、High、Low都放進input得到的MSE最小，而Volume的數據資料太大，跟其他不太匹配，故只選擇四個特徵作為輸入。

3. (15 points) Analyze the performance of the model with and without normalized inputs in Lab 4. You can use experimental results or external references (which must be cited) to support your conclusions on whether normalization improves the model's performance. (Approximately 100 words.)

若要將Volume作為輸入特徵，在沒有進行normalize的情況下，MSE會到1374，因為Volume的值跟其他差太多，但若將Volume進行normalize後，將範圍縮小到0~1之間，得到的MSE是150.675。

4. (10 points) Why should the window size be less than the step size in Lab 4? Do you think this is correct? If you use external sources, please include references to support your response. (Approximately 50 words.)

如果窗口大小小於步長，意味著模型在每次移動時會丟失部分歷史數據，這可能會導致模型無法充分利用所有可用的歷史信息進行訓練，從而降低模型的性能和預測準確性。

相反，如果窗口大小大於步長，模型將能夠利用更長的歷史數據序列進行訓練，有助於更好地捕獲數據的時間依賴關係和模式，從而提高模型的泛化能力和預測性能。因此，通常情況下，窗口大小應該要大於步長，以確保模型能夠充分利用歷史數據進行學習和預測。

GPT

5. (15 points) Describe one method for data augmentation specifically applicable to time-series data. Cite references to support your findings. (Approximately 100 words.)

針對時間序列數據的一種數據擴增方法是時間偏移 (Time Shift)。這種方法通過將時間序列中的觀察值向前或向後偏移一定的時間步長，以生成新的樣本。例如，可以將觀察值向前或向後移動幾個時間步，從而創建一系列時間序列的變體，這有助於增加數據集的多樣性，從而提高模型的泛化能力。

GPT

[How to Develop LSTM Models for Time Series Forecasting](#)

6. Discuss how to handle window size during inference in different model architectures (approximately 150 words):

(i) (5 points) Convolution-based models

通常使用固定大小的卷積核進行特徵提取。在推論階段，可以將整個時間序列切割為固定大小的窗口，然後將每個窗口作為輸入餵入模型進行預測。這樣做可以確保模型在推論階段處理的輸入大小與訓練階段相同。

(ii) (5 points) Recurrent-based models

窗口大小通常由序列長度指定。在推論階段，可以將整個序列或固定大小的窗口作為輸入，並將其餵入模型進行預測。這些模型能夠處理不同大小的輸入序列。

(iii) (5 points) Transformer-based models

窗口大小可以通過調整遮罩來控制。在推論階段，可以將整個序列切割為固定大小的窗口，然後使用適當的遮罩來確保模型只關注當前窗口的信息。