National Tsing Hua University

11220IEEM 513600

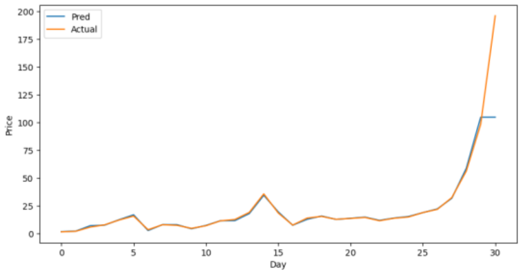
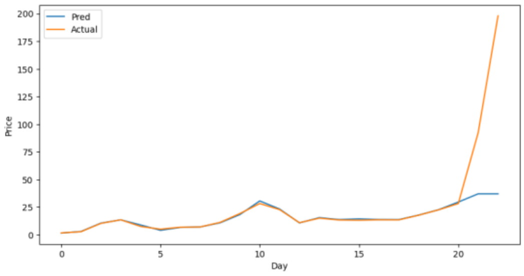
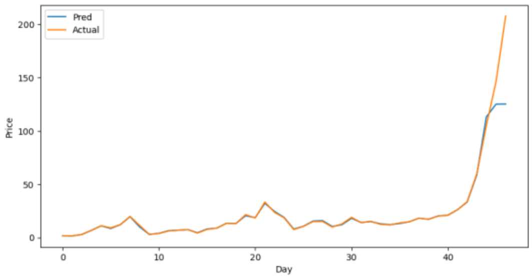
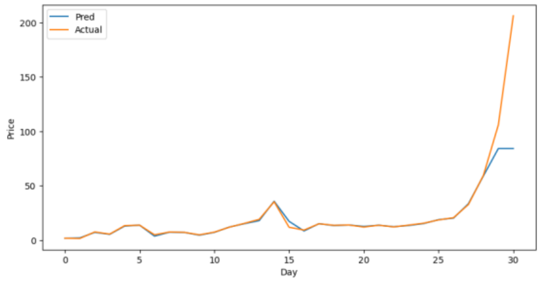
Deep Learning and Industrial Applications

Homework 4

Name: 陳奕均 Student ID: 112034604

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.)

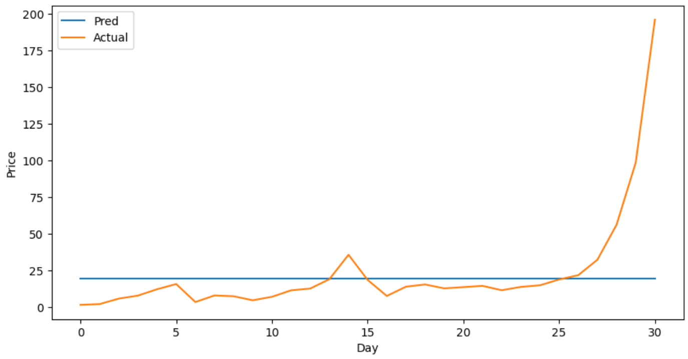
	Window Size=10, Step=15(Orig.)	Window Size=15, Step=20
MSE	 MSE: 339.0136	 MSE: 1138.1818
	Window Size=15, Step=10	Window Size=20, Step=15
MSE	 MSE: 155.7757	 MSE: 510.7410

從表格中可以看到window size=15, step=10的組合有最小的MSE，而window size=15, step=20的組合則產生最大的MSE，但無論window size < step還是step > window size的組合都沒有dominate另一組，因此應該透過實驗找到符合問題數據的最佳參數。

2. (Approximately 200 words.)

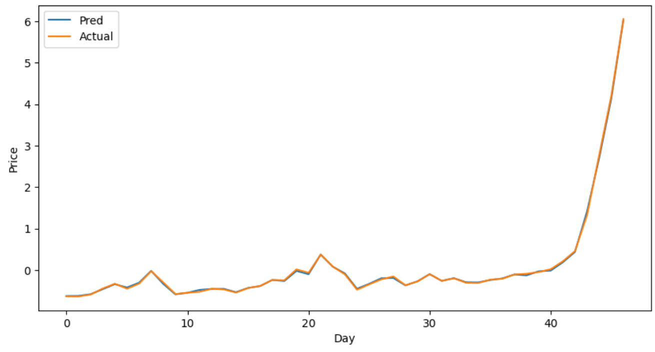
(i) (15 points) Include 'Volume' as an additional input feature in your model. Discuss the impact of incorporating 'Volume' on the model's performance.

加入Volume後，預測的趨勢圖如下表，MSE則為1345.7766，因為Volume的值遠大於其他 feature，加入後使得模型學習不到細微的變化，預測出的價格呈現無波動的水平線，MSE也大幅增加。



(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.

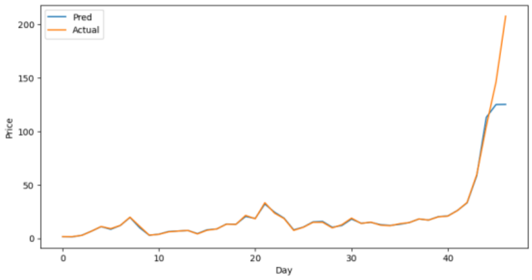
在上一題發現(window size, step) = (15, 10)的組合下有最小的MSE，因此沿用這一組參數然後加入normalization，並且不加入Volume，發現MSE能大幅降低至0.0147，在此基礎上繼續回去調動(window size, step) = (20, 10)，發現MSE能更進一步降到0.000724，因此認為最佳組合為：(window size, step) = (20, 10)、features = [Open, High, Low, Close]、normalization。

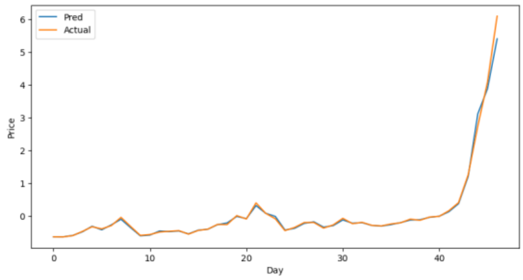


MSE: 0.000724

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.)

在將模型的參數組合設為(window size, step) = (15, 10)下，未使用normalization的MSE為155.7757，而加入normalization後MSE大幅降低至0.0147。normalization能減少離群值，也就是過大或過小值對模型預測結果的影響，也就能夠令模型預測地更準確。

<table>
<tr><th></th><th>Without Normalization</th><th>With Normalization</th></tr>
<tr><td>MSE</td><td>

MSE: 155.7757</td><td>

MSE: 0.0147</td></tr>
</table>

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.)

   一般會設定窗口大小大於步長，模型才能在每個時間步中觀察到不同的歷史信息，有助於更好地理解時間序列的動態特性和趨勢。

   ChatGPT

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

   現實中許多時間序列應用的標記數據可能是有限的，為提高模型的訓練數據規模和質量，可以利用Gaussian Noise或Generative machine learning model，前者是在時間序列中添加少量雜訊來增加樣本， 後者是以歷史數據去訓練機器學習模型，令它生成新的合成樣本。

   https://towardsdatascience.com/time-series-augmentations-16237134b29b

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

   (i) (5 points) Convolution-based models

   對於基於卷積的模型，在推論過程中處理窗口大小十，通常會對輸入的時間序列進行裁剪或填充操作，以確保輸入的大小與模型期望的窗口大小相匹配。

   (ii) (5 points) Recurrent-based models

   對於基於循環的模型，通常會使用滑動窗口方法來處理窗口大小，確保模型在推斷時能夠處理不同大小的輸入序列，並產生相應大小的輸出。

   (iii) (5 points) Transformer-based models

   對於基於Transformer的模型，因其自注意力機制能處理任意長度的序列，可以直接將整個輸入序列作為模型的輸入，而不需要使用窗口大小。