

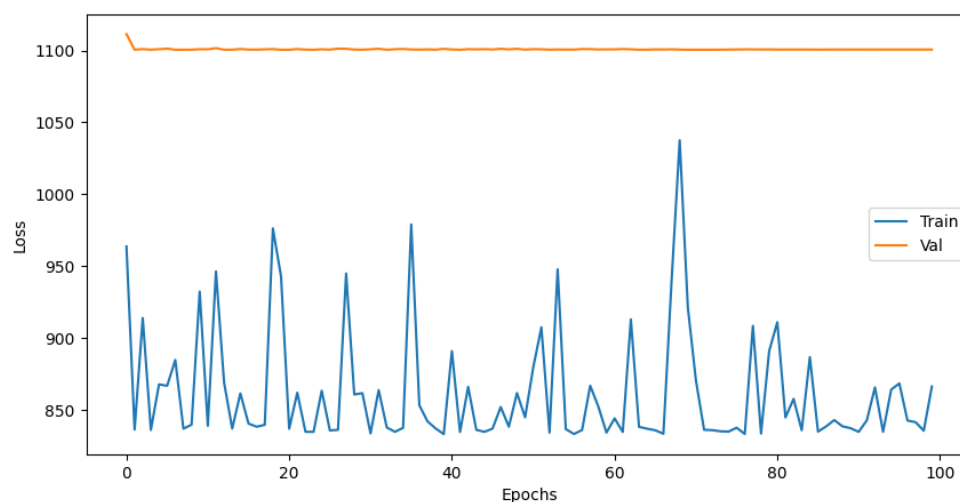
1. 針對不同的 Window Size 和 Step 的組合，將每種組合實驗 3 次的 MSE 取平均後記錄結果，如下表所示：

	Window Size	Step	MSE
組合 1	5	5	14.3622
組合 2	5	10	105.2178
組合 3	10	15	285.4951

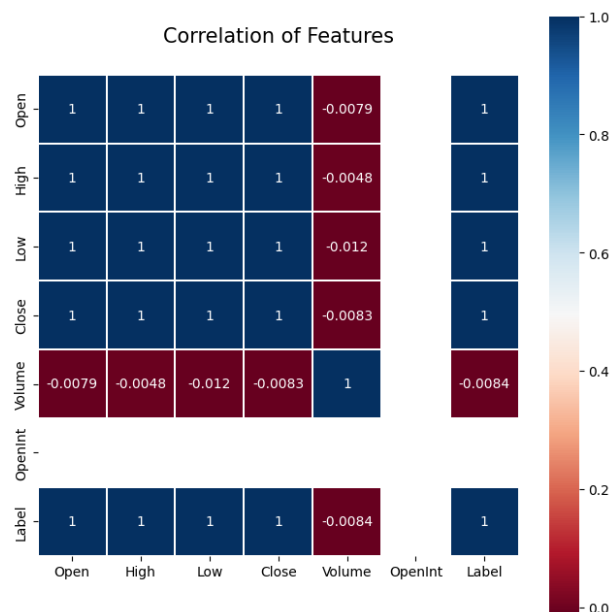
根據實驗結果，在 Window Size=5，Step=5 時模型會有較佳的訓練表現。推測是因為股價資訊並不需要參考太久以前的資料，加上資料只有紀錄周一至周五，所以用 5 天作為切分大小似乎比較適合。又因為 Step 較小可以產生較多組訓練資料，所以模型可能因此有較佳表現。

2.

- (i) 加入'Volume'這個特徵後，訓練結果如下，可以發現模型對訓練資料表現尚可，但對驗證資料表現不佳，似乎有過擬合的傾向。
(Window Size=5、Step=5)



對每個特徵進行相關性分析，可以觀察到'Volume'與'Label'的相關性極低，因此導致加入'Volume'這個特徵後訓練表現反而下降的情況。



- (ii) 由於在相關性分析中，除了 Volume 以外的其餘特徵相關性皆為 1。因此對不同的特徵組合進行實驗，同樣將 3 次實驗結果取平均，而表現最佳的是採用 'Open'、'High'、'Low'、'Close' 這四項特徵的組合。

(Window Size=5、Step=5)

Open	High	Low	Close	Volume	MSE
V	V	V	V	V	1093.564
V	V	V	V		14.3622
V	V	V			15.1761
V	V		V		16.6305
V		V	V		15.0305
	V	V	V		18.0592

3. 針對不同的 Normalization 方法，將每種方法實驗 3 次的 MSE 取平均後記錄結果，如下表所示：

(Window Size=5、Step=5、Feature=Open, High, Low, Close)

	平均 MSE
Without Normalized	14.3622
MinMax-Normalization	0.00001626776
Standard-Normalization	0.00164324891

在使用 MinMax-Normalization 時模型有最好的預測表現，推測是因為 MinMax Normalization 不會改變數據的分佈形狀，所以比較適合本次分析。

4. 由於 lab4 是要預測給定時間隔天的股價，而股價變化非常快速，且波動幅度可能很小，因此需要更精細的時間解析度來捕捉這些事件的變化，所以較小的 window size 可以讓模型更加靈活地捕捉這些瞬間性的波動，並更及時地做出反應。
5. 針對時間序列的資料增強方法有下列幾項，都可以幫助模型更好地理解不同時間尺度下的模式和趨勢，從而提高其泛化性能：
 - Jittering / Gaussian noise：直接對數據加上雜訊。
 - Magnitude Warping：通過對時間軸上的數據進行拉伸或壓縮，改變每個樣本的尺度。
 - Window Warping：先切分不同大小的 window size，再將其縮放至相同大小進行訓練。

參考網址：<https://towardsdatascience.com/time-series-augmentations-16237134b29b>

6.
 - (i) Convolution-based models：通常使用固定大小的卷積核來處理輸入序列。可以使用滑動窗口的方式來處理序列，即將一個 window size 大小的子序列作為輸入，並在整個序列上進行卷積操作，從而獲得預測結果。這樣可以適應不同長度的輸入序列。
 - (ii) Recurrent-based models：通常使用固定大小的序列長度來處理輸入序列。需要確定每個時間步的輸入大小，以便與模型相匹配。通常會使用固定大小的 window size 來處理序列，並將其作為模型的輸入。
 - (iii) Transformer-based models：由於其注意力機制可以處理不同長度的輸入序列，因此並不需要固定大小的 window size。相反，可以將整個序列作為模型的輸入，並使用適當的注意力機制來對不同部分進行加權，從而獲得預測結果。