

姓名:李鎮帆 學號:112034703

1.

	Window size	step	MSE
Case1	10	15	123.3362
Case2	10	5	9.3961
Case3	10	3	1.6125

2. (i) 在未加上 Volume 時的 MSE 為 123.3362 在加上後，MSE 變為 1135.4211，明顯大出很多。表示加入後可能讓模型過度擬合，以至於讓預測的表現不好。
- (ii)在不加入 Volume 一變數的情況下，step size 又小於 window size 時(step size =3>window size=10)時，MSE 為 1.6125，使用兩者情況的原因為避免 overfitting 的同時又能讓更多讓模型更能捕捉到資料的變化。
3. 在 normalized 之前的 MSE 為 123.3362，而 Normalize 後 model 的 MSE 為 6.4360，明顯減少了很多，表示其在標準化後有了更好的數據範圍(較統一)，有助於模型訓練時的收斂。且以下這篇也同樣說到標準化的效果。
<https://medium.com/%E5%B1%95%E9%96%8B%E6%95%B8%E6%93%9A%E4%B8%A9%E7%94%9F/python%E6%A9%9F%E5%99%A8%E5%AD%B8%E7%BF%92-108-%E8%B3%87%E6%96%99%E6%A8%99%E6%BA%96%E5%8C%96%E8%88%87%E9%80%B2%E9%9A%8E%E7%9A%84%E6%A2%AF%E5%BA%A6%E9%81%9E%E6%B8%9B-5fa02e578da0>
4. 由第一題的表格中可以發現，當 window size 維持不變時，step 若從大於 window size 的情況下轉變為小於 window size 的值時 MSE 有明顯的下降許多，這表示可能在 step 較小的情況下，模型更能捕捉到資料的隨之變化。
5. Dynamic Time Warping Barycentric Average (DTWBA)，是一種基於 Dynamic Time Warping (DTW)的 augmentation，方法為衡量時間序列中的相似度 (similarity)，並將其同步「sync」。其他還有像 Gaussian noise，其背後是假設在時間序列中添加少量 noise 可能不會顯著改變其本身，但會增加資料集中此類雜訊樣本的數量。
<https://towardsdatascience.com/time-series-augmentations-16237134b29b>
6. CNN: window size 通常由 convolutional kernel size 決定。較大的 kernels 可以捕捉更廣泛前後數據，但可能導致過擬合。因此，需要通過交叉驗證或其他技術來選擇合適的 window size。

RNN: window size 由 time step 和序列長度(sequence length.)決定。較長的 time step 長可以增加模型的記憶能力，但可能導致資訊丟失或過度平滑。因此，需要平衡 time step 和序列長度，來保持適當的前後資訊。

Transformer: window size 由 self-attention mechanism 中的 attention mechanism 決定。這些模型通常使用 self-attention mechanism 來捕捉不同位置之間的依賴關係。較大的 window size 可以捕捉更長範圍的依賴關係，但可能導致計算成本增加。因此，需要權衡 window size 和計算成本。