1. 當預測sine wave的時候，通常Dense裡有包含RNN的model會表現得比較好，因為RNN能夠捕捉一些時間相關的訊息，而Dense無法保留數據訊息，sine wave又是週期性數據，因此使用RNN通常能表現得更好。
2. 如果第一層RNN是(1, 16)，那我設計的RNN model會配置為 (16, 32)，第一個參數16是input，也就是第一層的output，第二個參數32會讓他比原本還大一點，以捕捉更多時序相關的feature，然後不敢選太大，因為要避免數字太大overfitting以及等太久。
3. RNN尺寸選大一點的話，如剛剛所說，可以提高他跟時序的關係，然後還可以學習更多更複雜的features，但是太大的也有可能會發生overfitting的狀況，或是train很久的情形，也有可能讓梯度直接消失，無法得到正解。