

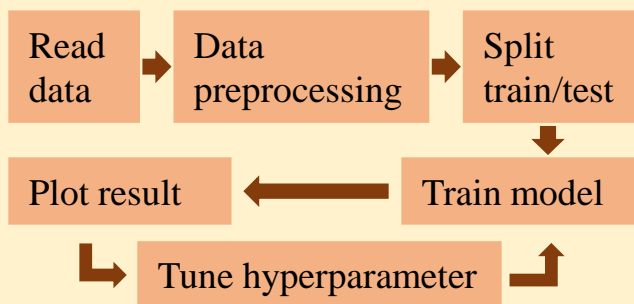
# Financial Index Trend Prediction Based on Machine Learning 利用機器學習預測金融指數走勢

組別：B132 組員：吳騏佑 黃昱翔 指導教授：翁詠祿

## Introduction

我們以能夠準確預測股價的漲跌為目標，利用金融歷史數據訓練機器學習模型。我們使用 Standard Bar 的方式整理股票資訊，並匯入 LSTM(長短期記憶)模型，依據過去數筆股票資訊，預測未來下一筆股價。

## Method



### 1. Standard bar

原始的金融數據是以時間為採樣頻率，為了獲得股票漲跌的規律，我們可以採用不同的採樣方式來分析，比如說每隔一定的交易量、交易金額或交易次數採樣一次。

### 2. LSTM (Long Short Term Memory)

LSTM 是目前 RNN(Recurrent Neural Network) 中最常使用的模型，用於解決資料有時間順序的問題，透過將隱藏層將結果儲存在記憶體裡，當下次輸入資料時，會同時考慮上一次存在記憶體的值進行計算。

### 3. 效能評估

預測股票的目的是掌握交易時機，要先能在股票漲時搶先預測股票會漲，再來討論漲多少才有意義。我們評估模型好壞的標準為模型預測漲跌的準確率，待準確率夠高後再來探討漲跌幅的預測差距。

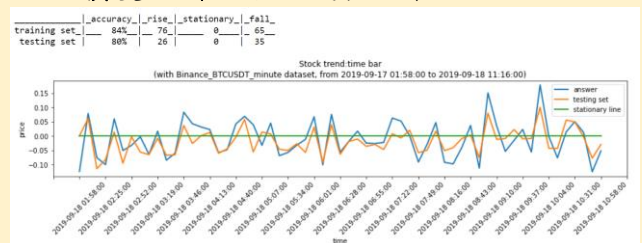
## Result

以下的股價都有經過正規化縮放到 0 和 1 之間。以輸入 20 筆 bar 預測下一筆 bar 為範例：

### 1. 股價預測圖如下



### 2. 股價變化圖如下，代表股票的漲跌幅



### 3. 模型準確率如右

我們使用的資料為取自 Binance.com 的金融歷史數據。不同採樣方式與不同數據集的準確率如右，輸入筆數代表使用了幾筆過去的資料來做預測。

採樣方式 與採樣頻率	數據集 輸入筆數	Training set	Testing set
Time bar (每 10 分鐘)	5	85%	88%
	10	77%	73%
	20	84%	80%
	30	59%	47%
	40	65%	64%
Volume bar (每 70 股)	5	70%	65%
	10	74%	71%
	20	70%	69%
	30	55%	57%
	40	61%	57%
Dollar bar (每 130 萬元)	5	82%	87%
	10	62%	50%
	20	81%	87%
	30	72%	68%
	40	74%	69%
Tick bar (每 100 筆交易)	5	53%	55%
	10	69%	72%
	20	62%	60%
	30	57%	57%
	40	66%	70%

## Conclusion

我們使用以分鐘為採樣頻率的金融歷史數據，極高的採樣頻率使得股價在小區間內的漲跌受股票消息面的影響極低，更加貼合我們用股票技術面預測股價的目標。在使用 standard bar 以適當的採樣方式與採樣頻率重新整理股票資訊後，模型的準確率都能夠達到 60% 以上甚至是 70%，已經算是可靠的模型了。

## Reference

[1] Lopez de Prado, Marcos, advances-in-financial-machine-learning, Wiley 2018