

Cloud Computing and Big Data Analytics-

Homework 2 Report

309555025 資電亥客-羅文笙

1. Algorithm/Implementation/Model description

主要是使用簡單的 LSTM network，藉由輸入 5 年內的 open/high/low/close/adj close/volume 加上美國納斯達克指數的 adj close，總共 7 個 features，來預測該股票的 adj close price。

執行方法的話是使用每個股票都去跑 10 次模型，然後訂定一個 threshold(根據前一天的表現做調整，為一個超參數)，只要漲(0)/跌(2)出現的次數超過這個 threshold，就預估當天為漲(0)/跌(2)，反之則預測為平穩(1)。

LSTM 模型的參數如下：

Epoch	Batch size	Feature number	Learning rate	Sequence length	Num of Layer
100	32	7	0.005	10	1

2. Reference

老師的講義

3. Suggestion

其實這次作業不太確定到底能不能做出結果，我這次自己想的演算法似乎太簡單了，可能會需要再去看一下其他進階的 paper 是怎麼做的，有時候面對當天股票變動大一點就只能對個 8~10 個，五天下來也總共才拿到 60 分左右。希望之後老師上課能暗示一下我們結果到底能做到什麼程度，或是這次作業的用意是什麼，不然我們會有點不知所措，成績方面可能也不太理想。

4. ReadMe

Jupyter notebook 的程式碼應該是直接按 run 就可以開始跑了，要記得檢查一下讀取資料的那個 FILE_PREFIX 變數，目前預設是在 Windows 系統下，如果助教要在 Linux/Mac 下跑的話可能會有錯。

跑完之後會看到 20 個 1x3 的陣列，在這三格裡面分別代表跑了 10 次模型之後得到的預測結果，假設 threshold 為 3，例如[1 9 0]，則代表這一支股票我會預測是平穩(1)，而[3 7 0]，則代表這一支股票我會預測是漲(0)。這個部分的話就是要用人去看，我沒有寫成程式。而 5/7 這一天我設的 threshold 是 5。