NTU FinTech Fall 2019 Final Project Report NTU CSIE R08922024 胡安鳳

1. 前言: 使用LSTM, 或同作業二採技術指標即可?

答案: 以CP值而論

很明顯本期末專題是時序性資料預測,一定想到用回饋式神經網路RNN或長短期記憶網路LSTM訓練。神經網路固然強大,卻可能存在一個致命缺陷: 過擬合 (Overfitting): 在initail submission很好的分數,有可能在最後十四天的評測中,大大落敗 (而14天也較短,無法看出以多年資料作訓練集從長計議的優勢,例如模型以半年至一年的scale有賺頭,但在未來十四天的評測,因時間長度不夠,短線而言可能無法給出好結果)。或許只要layer by layer fine-tuning,搭配Drop, pruning即可避免。但會耗不少時間,效果亦未知。

期末繁忙下,以類似作業二方式實作,會是較高投資報酬率 (指所付出的時間心力和所獲得的成果)的方法。

2. 何種技術指標? 如何優化?



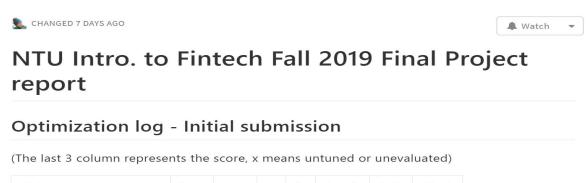
由上台指期走勢圖得知,指數雖沒有如亞馬遜公司在近幾年大幅竄升,卻是穩穩成長(納入股市型態學考量)。選擇最基本技術

指標MA、RSI即可。第二次作業,我用 RSI作指標,搭配短期 RSU數值限制在某範圍配合各支股票fine tuning,而獲得了高達 5.5萬分的高分成績位居全班第八。

```
rsi_s = float((up + 1) / (up + down + 1))
if rsi_s > rsi_l or (rsi_s > a and rsi_s < au):
    action = 1
elif rsi_s < rsi_l or (rsi_s < b and rsi_s > bl):
    action = -1
else:
    action = 0
```

MA不採用乃因測試數次後效果皆不如 RSI (MA單純採用**移動平均**, 易受到**極端值**影響;然RSI是**量化的結果**, 可清楚看出股票 "**潜在勢能**")。

本期末專題捨棄短期RSI額外設上下界 (上圖的 rsi_s > a and rsi_s < au), 以免過擬合, 純粹採用 rsi_s > rsi_l 作為買的指標, 再暴力搜尋s、l的參數(天數)即可。 又在HackMD作搜尋結果log, 分別記錄以分鐘和以天為單位兩個csv檔參數搜尋結果. 避免重複找參數區間而顯冗贅。



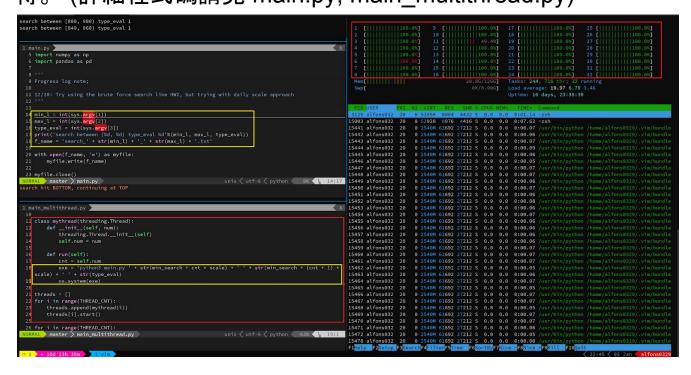
date	1	s	a	b	local	init	final
1221 daily [0, 100]	49	47	Х	х	0.040	508	399
1222 daily [100, 200]	126	124	Х	Х	0.042	535	426
1222 daily [200, 300]	253	240			0.038		
1226 daily [300, 400]	324	305			0.038	515	406
1222 daily [1000, 1200]							
1222 minly [0, 100]	68	63	Х	Х	0.036	473	365
1222 minly [200, 300]	234	225	X	Х	0.040	505	396
1226 minly [300, 800]	441	431	X	х	0.045	557	448
1228 minly [800, 1000]	916	903	Х	х	0.038		

最後,不論initial 或 final submission,我都獲得了**第一名**,可見該參數模型具有一定的股價評判能力 (final submission 到

正式開始評判前,已有超過 2019/12/16的資料,**完全不存在助教給的csv**中,可視為testing set概念,而得以客觀評價)程式碼**既短、快速又高分。**



3. 如何加速參數搜尋? 以達最少時間最大效果 暴搜參數相當曠日廢時,用python multithread,各自搭配 os.system('python main.py a b'), 搜尋 [a, b)區間傳 入 argv,既不用煩惱critical section問題,又能自動化,一舉數 得。(詳細程式碼請見 main.py, main_multithread.py)



- 4. 評估(如果有 1/10的結果再補寫這裡)
- 5. 未來展望

以上方法跟知名金融科技公司模型預測比可能小巫見大巫,但短期交易(炒短線)也算能派上用場。

不久將來,若以程式作自動交易、股價追蹤,RSI MA KD三個最基本指標,會是除LSTM等方法最需要納入模型中,甚至某些情況以上三個指標的表現可能不亞於神經網路。

但完全靠程式發大財, 我認為不可能, 影響股價的因素太多了: 預期心理、天然災害、政治戰爭情勢、甚至有權有勢的川普、 普丁、Elon Musk 一篇推文, 都牽一髮而動全身。 這時加入人 工修正就相當重要, 以人工為主、程式為輔, 會是最適切的組 合。

6. 心得

雖然目前仍未公布總體排名,但不論好壞,直接下市場廝殺是投資學相關科目最好的課程指導老師。

結果好則繼續精進、差強人意也無須氣餒,至少自己**學到相當 多金融科技的知識,這些是金錢與分數無法衡量的無價瑰寶!**

人生中本就有許多不確定因素,三分天註定,七分靠打拼,在能力所及之內,做好自己應盡本份,我想這就是無愧於心了。