

Financial Technology - HW1: Trading Valley Strategy Design

Student ID: 109062211

Name: 張惇媛

I. Explanation

- Design a strategy which makes the maximum profit and high sharp ratio.

利用機器學習設計分類器模型，預測每日股價的漲跌勢。交易數據集分為訓練集(資料期間－[2016-01-01 ~ 2021-12-31])與測試集(資料期間－[2022-01-01 ~ 2022-09-30])，為了使分類器表現最佳化，因此，取用單一投資標的－美股 COCA COLA CO (KO)，希望能在此檔股票中獲得高報酬，但是，單一投資組合存在著風險高的劣勢，因此加入「恐慌指數」作為風險管理的指標，以考量收益與風險兩者之間的平衡。分類器模型以五力指標中的趨勢(trend)作為分析指標，取前三天的趨勢指標(S0, S1, S2)與 VIX 指數(VIX)作為輸入值，以及股價資料的漲(True)或跌(False)作為輸出值(T)。(詳細說明於 Process 的部分)

- What's your strategy to decide the entry and the exit?

由於不考慮任何交易費用(fee & tax)且每次交易投入全部資金(all-in in one trade)，因此採用高頻交易(high-frequency trading)的方式－交易頻率為當日進出。透過機器學習的分類器模型(詳細說明於 Process 的部分)，每日判斷今日的股價可能為漲或跌，若判斷為漲勢，預期獲得正報酬，則於當日開盤價買進、收盤價賣出，以獲得今日價差的報酬，反之，預期獲得負報酬，則當日不做任何交易，以避免可能的虧損。

- How's your strategy's performance? Including profit, sharp ratio and asset change chart.

詳細說明於 Result 的部分

II. Process

1. 獲得數據

分類器的 input(輸入值)：

(1) 選擇五力指標中的趨勢指標

因趨勢指的是股價持續沿著一個方向行進，分數越高（低）代表個股處於多頭（空頭）趨勢，我認為可代表股價預期的漲跌，與我分類器的輸出值有高度正相關。另外選擇一個時間段的所有趨勢指標，是因為希望反映變化而不是單一時間點的數值，最後對於分類器維度過大容易造成 overfitting 的考量，以三天作為時間段，因此取 Growin 數據中前三日(含當日)的趨勢指標(S0(今天), S1(昨天), S2(前天))作為分類器的特徵(feature)。

(2) 選擇額外的指標

VIX 指數(VIX)是用來衡量市場波動率的指標，大致反映了大盤的預期走向，可做為管理系統風險的評估，因此加入分類器的特徵(feature)。

分類器的 output(輸出值)：

取 Yahoo Finance 數據中的開盤價與收盤價，將收盤價大於開盤價的標記為 True (股價漲)；收盤價小於或等於開盤價的標記為 False (股價跌)，作為分類器的標籤(class label)。

另外計算每日漲跌幅((收盤價-開盤價)/開盤價)，用於最後投資分析的數據。

2. 分類器模型

接著因趨勢指標採取等第劃分，等第間差距容易使數據產生中間模糊段，而且 SVC 能夠切出最大 margin，可以減少邊界中間模糊段誤差的特性，所以選用 Support Vector Classification(SVC)進行分類，其中參數設定 kernel 的選擇經過比較評估的三項分數，挑選最佳的 kernel - 'linear'，還有固定 random_state 的數值，以獲得可重現之數據，最後將訓練集的數據建立分類器。

3. 評估

對於分類器模型，以 precision_score(精準率), recall_score(召回率), f1_score 此三項分數，來評估模型的優劣，在訓練集與測試集的結果如下：

```
Train data:
precision_score: 0.573869
recall_score: 0.733933
f1_score: 0.644106
-----
Test data:
precision_score: 0.598291
recall_score: 0.666667
f1_score: 0.630631
```

4. 帶入回測

將分類器對於訓練集與測試集的預測帶入股票市場進行回測，結果於 Result 部分展現

III. Results

對於投資報酬的結果，以獲利、總報酬率、日平均報酬率(總報酬率的算術平均數)、年化報酬率(以日平均報酬率計算 365 日複利後的報酬率)以及 Sharpe Ratio 此五項與報酬相關的數據，來評估投資策略的優劣。其中，Sharpe Ratio 的計算如下

對於 Sharpe Ratio - 以日平均報酬率來計算，假設無風險資產的日平均報酬率為 0.01% (因作業設無風險年利率為 2%，則日平均利率=2%/365=0.005...%≐0.01%)
公式為 (日平均報酬率 - 無風險日平均利率) / (每日報酬標準差) x (252**0.5)
(x (252**0.5)是因為大約有 252 個交易日/年，將波動數值從每日調整成每年)

Reference: <https://rich01.com/what-sharpe-ratio/>

(註：

With classifier – 使用我的投資策略，於每日判斷是否進行交易；

Without classifier – 每日皆進行交易，可視為股價的變化。

*交易皆為當日開盤價買進、收盤價賣出)

訓練集 (2016-01-01~2021-12-31)

Train data -

With classifier:

獲利: \$185393

總報酬率(%): 185.393440

年化報酬率(%): 56.449856

日平均報酬率(%): 0.122696

Sharpe Ratio: 3.035256

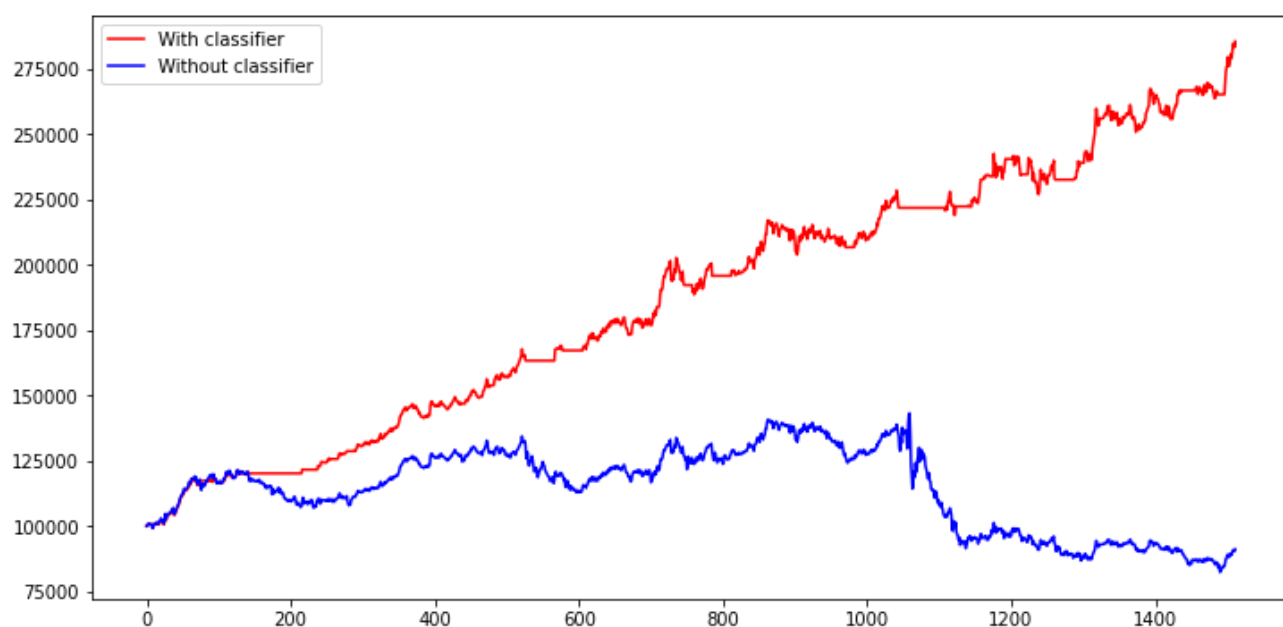
Without classifier:

總報酬率(%): -8.891332

年化報酬率(%): -2.124968

日平均報酬率(%): -0.005884

資產變動折線圖(初始資金皆為\$100000)



測試集 (2022-01-01~2022-09-30)

Test data -

With classifier:

獲利: \$27351

總報酬率(%): 27.351226

年化報酬率(%): 70.001187

日平均報酬率(%): 0.145485

Sharpe Ratio: 2.676116

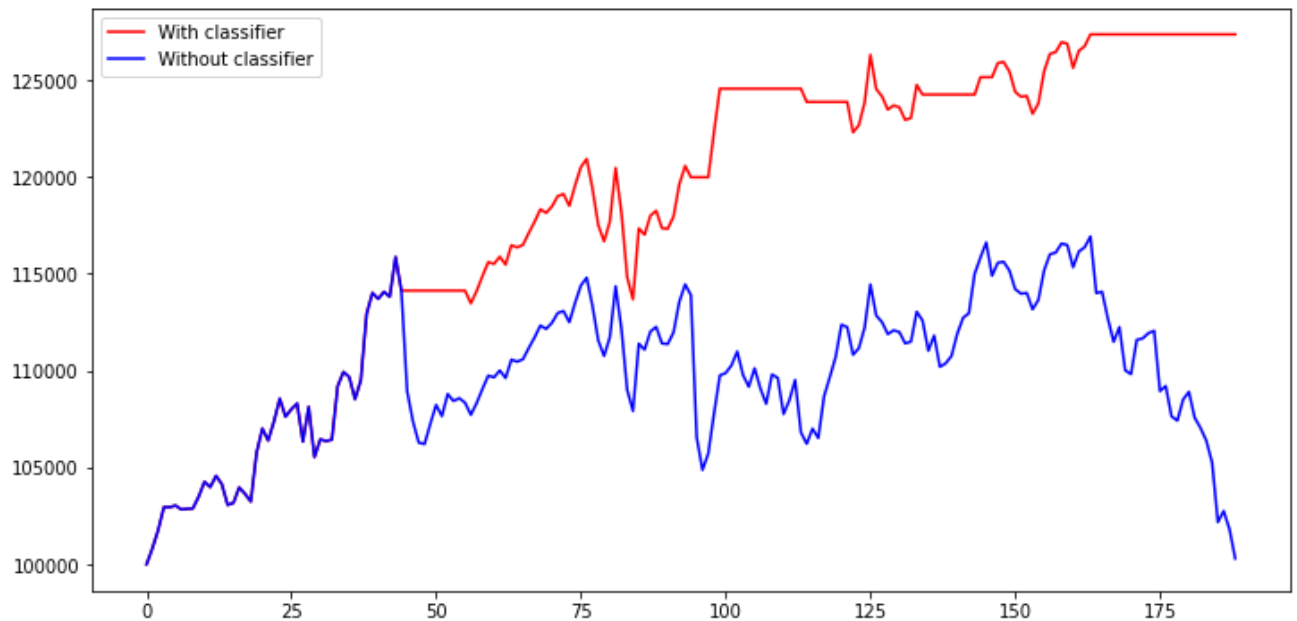
Without classifier:

總報酬率(%): 0.293593

年化報酬率(%): 0.571631

日平均報酬率(%): 0.001562

資產變動折線圖(初始資金皆為\$100000)



IV. Discussion

1. 由於本次作業不考慮任何交易費用，因此可以選擇每日操作買賣，但在實務上每筆交易應加入交易費用進行考量，嚴格地定義所謂的正報酬必須要扣掉額外的交易費用，因此分類器的標籤。另外若長期皆預測為漲勢，可以持有一段時間再賣出，降低交易次數，以降低其中的交易費用。
2. 由資產變動折線圖可以看出在使用的策略當長期預測為跌勢時，在此期間不會進行任何交易，因此資產會維持在同一水平線上，對於投資策略來說，可以於此時間段選擇將資產放置在使用的策略預測為漲勢的其他檔投資標的進行交易，以獲得更多的報酬。
3. 由於本次作業不強制考慮資產配置，但單一標的的投資組合無法分散非系統風險，會影響到 Sharpe Ratio 的表現，因此實務上應考慮自身風險能夠承擔的程度，進行適度的資產配置，以降低投資組合存在的風險。
4. 另外，由於我的作法採取隨機選擇一檔股票，但是並未考量到 COCA COLA CO 屬於防禦性消費，波動率相較同產業與整體市場相對較低，較適合長期持有，所以實務上選擇的投資標的應考慮其產業特性、股票類型等等來做出組合比例。

Reference:

防禦性投資 <https://rich01.com/what-is-defensive-investment/>

投資風險 <https://rich01.com/systematic-risk-and-unsystematic-risk-invest/>