

多目標遺傳演算法優化 S&P 500 股票 預測模型的特徵選擇

研究生：郭東輝 財金碩一

1

Introduction

2

Data and Basic Characteristics

3

Method

4

Main Result

5

Conclusion

Introduction



摘要

- 本研究主要基於多目標遺傳演算法，用於選擇最佳的股票特徵指標並應用於美國股市預測(Han 等人 (2019))
- 同時最佳化股票波動度和報酬率的股票特徵，並將這些特徵應用於預測股票價格的波動性和報酬率
- 將適應度函數設定為機器學習模型均方誤差(mse)(Zhang 等人 (2019))
- 並篩選同時 排名波動度前 30%低和報酬率前 30%高股票作為投資組合依據，並每四個月重新預測 一次波動率與報酬率並重新建立投資組合(Guo 等人 (2019))



Data and Basic Characteristics

Data

- 標準普爾 500 指數資料來源：yahoo finance 股價資料庫
- 股票數量：250支股票(標準普爾 500隨機篩選)
- 資料期間：2019/01/01 至 2023/04/30 資料

資料類型	區間
多目標演算法訓練	2019/01/01 至 2019/10/30
建立投資組合	2019/10/30 至 2023/04/30

Data

- 每一支股票都計算 174 種股票特徵，建構250隻股票特徵資料庫

資料類型	資料類型
價量資料	統計函數
週期指標	波動性指標
數值變換	交易量指標
動量指標	重疊研究指標
圖表形態識別	價格變換



Method

多目標基因演算法

- 本研究主要透過多目標基因演算法，用於解決特徵選擇的最佳化問題。該問題的目標為從一組股票特徵中選擇最佳的子集
- 本研究假設族群大小為 20，代表隨機生成 20 個以二進制為基礎的向量，每個基因由二項分布獨立同分布生成，總共生成長度 174，由 0 和 1 組成的基因序列

舉例:(以長度10為例，實際長度為174)

單一股票特徵資料

column	0	1	2	3	4	5	6	7	8	9
基因	1	0	1	0	0	1	1	0	1	0
	考慮			捨棄			考慮			捨棄

適應度函數

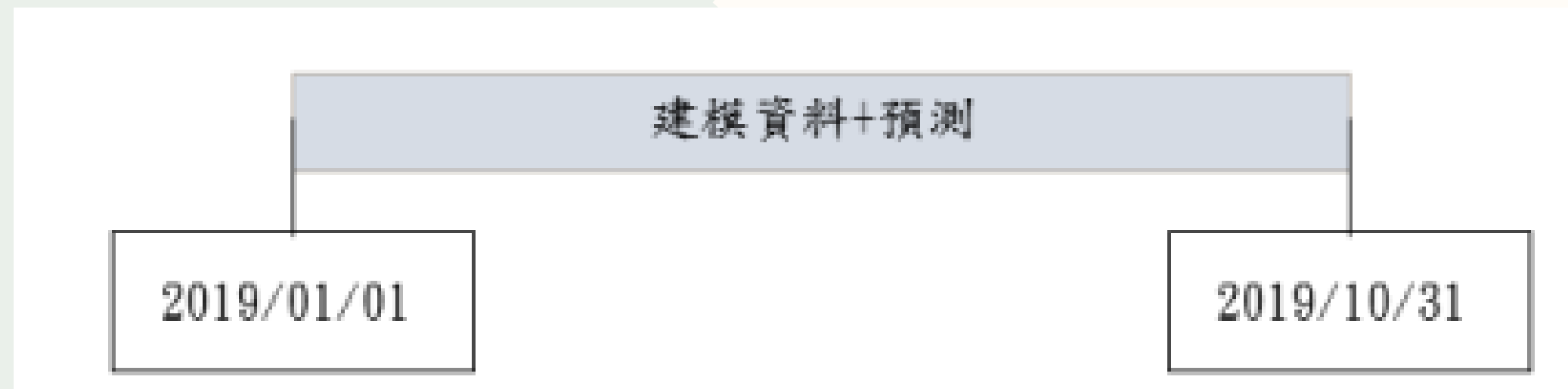
- 本研究將個體的適應度函數分為兩個部分

適應度1

機器學習模型在所有股票，共 250 支股票，對**波動度**預測訓練集上的均方誤差 (Mean Squared Error)

適應度2

機器學習模型在所有股票，共 250 支股票，對**報酬率**預測訓練集上的均方誤差 (Mean Squared Error)




投資組合建構

- 透過多目標演算法找到最佳的股票特徵後，透過這些特徵對每支股票進行波動與報酬預測，並透過篩選的方式，建立低波動高收益投資組合

預測波動度

預測報酬率

篩選	<30%	30%~70%	>70%
<30%			
30%~70%			
>70%			




Main Result

投資組合績效-未調整

篩選	報酬率	最大回落程度	Calmar ratio
投資組合	6.6%	3.71%	1.78
投資組合	10.54%	2.03%	5.2
投資組合	5.99%	5.37%	1.12
投資組合	5.12%	5.0%	1.02
投資組合	7.49%	11.91%	0.63
投資組合	-1.08%	18.35%	-0.06
投資組合	9.85%	8.66%	1.14

投資組合績效-經市場報酬調整

篩選	調整後報酬率	最大回落程度	Calmar ratio
 投資組合	0.07%	3.71%	1.78
 投資組合	4%	2.03%	5.2
投資組合	-3.63%	5.37%	1.12
 投資組合	7.74%	5.0%	1.02
 投資組合	3.66%	11.91%	0.63
 投資組合	7.78%	18.35%	-0.06
投資組合	0.88%	8.66%	1.14



Conclusion

本篇研究的主要實證結果

- 在大多數測試期間，投資組合在報酬率和風險方面表現優於標準普爾 500 指數
- 投資組合能夠有效地過濾低報酬和高風險的股票，就使得整體投資組合的報酬率更高，風險水平更低。
- 七個投資組合的走勢與標準普爾 500 指數之間存在高度相關性，這種相關性的存在使得本研究的股票選擇方法能夠更好地應對市場波動 風險



Thank you!