

國立政治大學金融學系

碩士學位論文

機器學習下建構 ESG 股息波動因子投資組合

Constructing ESG Portfolio with Factor Investing for Dividend
Yield and Volatility by Machine Learning

指導教授：林士貴 博士

研究生：賴晨心 撰

中 華 民 國 一 一 〇 年 六 月

謝辭

隨著論文的完成，短暫的兩年碩士生涯也畫下句點，研究過程中曾想過無數次完稿後該如何致謝，真正到了此刻卻不知從何下筆，要感謝的人太多了。首先，由衷感謝指導教授林士貴老師悉心的指導，花了許多時間與我討論，處處都能感受到老師的用心與付出，真的非常謝謝老師。感謝口試委員廖四郎博士和黃俊仁博士對本篇論文的細心審閱，並給予寶貴的意見。再來特別感謝博班顯吉學長，對於我的研究提供許多建議，每一次討論都能深深感受到學長對於研究的熱忱。

此外，感謝在永豐金證券策略交易部實習的日子，不僅培養我做研究的精神，也提升我在程式撰寫上的能力，對於論文研究有相當大的助益。同時也要感謝金融所同學們，平時互相交流與學習，讓我在碩士生活更加歡樂與充實。

謝謝家人的支持，尤其在疫情嚴峻時期，讓我在家安心完成論文最後一哩路，並且陪伴我度過熬夜研究的時光，給予我無比的溫暖與關心。最後，我要感謝自己選擇就讀研究所，短暫而忙碌的兩年期間成長了許多，也完成很多過去從未想像的成就，如果沒有當初堅持念研究所的選擇，就不會有現在的自己，儘管仍有需多不足之處，還是帶著種種經歷與成長邁向人生下一個階段，繼續努力成為更好的自己。

賴晨心 謹誌於

國立政治大學 金融學系研究所

民國一一〇年六月

摘要

因應投資人需求，全球資產管理規模成長迅速，其中ESG (Environmental, Social and Governance, ESG)、股息與波動為長期投資人熱門選擇標的。本文使用美國證券市場於2003到2019年的資料，透過機器學習XGBoost與歷史因子投資法預測未來股息波動因子特性，並以粒子群最佳化 (Particle Swarm Optimization, PSO) 建構限制資產數量與權重的最佳化投資組合，本文探討議題與實證結果歸納為以下四點：(1) 比較歷史因子投資與機器學習兩種方法之預測能力，兩者皆具相當程度的預測能力，且機器學習預測能力較佳，其中機器學習之重要特徵變數為過去殖利率、波動度、本益比；(2) 分別針對歷史因子投資與機器學習預測法建構Markowitz投資組合，機器學習下之因子投資最接近正確股息波動投資組合表現；(3) 利用PSO配置限制資產數量的投資組合，能夠達到Markowitz全樣本投資組合之績效；(4) 比較全體與ESG資料集結合股息波動因子表現，ESG結合股息波動因子對於投資組合的績效表現有正向關係。

關鍵字：XGBoost、粒子群最佳化、因子投資、投資組合理論、ESG

Constructing ESG Portfolio with Factor Investing for Dividend Yield and Volatility by Machine Learning

Student : Chen-Hsin Lai

Advisor : Shih-Kuei Lin

Department of Money and Banking, National Chengchi University

Abstract

In response to the needs of investors, the scale of global asset management has grown rapidly. ESG, high dividends, and low volatility are popular choices for investors in long-term. In the study, data from U.S. securities market from 2003 to 2019 are used to predict the characteristics of future dividend and volatility factors through machine learning XGBoost model and historical factor investing method. Furthermore, PSO is used to construct optimized portfolio with limits of the number of assets, maximum and minimum weight. The empirical results and main topics are summarized into the following three points: (1) Compare the predictability of dividend and volatility between historical factor investing and machine learning methods, both have great predictive ability and ability of machine learning is better. The important characteristic variables of machine learning prediction are historical dividend, volatility, and price-to-earnings ratio. (2) The performance of portfolio with dividend yield and volatility by machine learning is closer to correct data than historical factor investing method. (3) Using PSO to construct portfolio with a limited number of assets can achieve the performance of Markowitz's full sample portfolio. (4) ESG combined with high dividend and low volatility has a positive relationship with portfolio performance.

Keywords: *XGBoost, Particle Swarm Optimization, Factor Investment, Portfolio Theory, ESG*

目次

第一章 緒論	1
第一節 研究動機與目的	1
第二節 研究架構	3
第二章 文獻探討	5
第一節 因子投資	5
第二節 機器學習方法	6
第三節 Markowitz 投資組合理論與粒子群最佳化	7
第四節 ESG 與投資組合報酬	8
第三章 研究方法	11
第一節 XGBoost 機器學習方法	11
第二節 Markowitz 投資組合理論與實務問題	15
第三節 粒子群最佳化	19
第四節 預測與績效評估指標	23
第四章 實證結果	26
第一節 資料描述與模型設定	26
第二節 歷史因子投資與機器學習預測績效	28
第三節 歷史因子投資與機器學習投資組合報酬績效	32
第四節 Markowitz 與粒子群最佳配置投資組合報酬績效	36
第五節 ESG 與股息波動投資組合報酬績效	38
第五章 結論與未來展望	47
第一節 研究結論	47
第二節 未來展望	49
參考文獻	50
附錄：機器學習模型之輸入變數	52

表次

表 1：混淆矩陣(Confusion Matrix)	24
表 2：歷史因子投資與機器學習預測比較.....	29
表 3：股息波動平均個股數量.....	32
表 4：股息波動 Markowitz 投資組合績效：2004Q1-2019Q4.....	33
表 5：不同配置方法之股息波動投資組合績效：2004Q1-2019Q4	36
表 6：不同資料集之股息波動投資組合績效：2004Q1-2019Q4	39
表 7：不同資料集之股息波動投資組合績效：2012Q3-2019Q4	43
表 8：ESG 股息波動投資組合之公司出現次數前十名.....	44
表 9：ESG 股息波動投資組合之公司加權比重前十名.....	45



圖次

圖 1：研究架構.....	4
圖 2：XGBoost 分類概念示意圖	12
圖 3：XGBoost 演算流程	12
圖 4：粒子在 Markowitz 投資組合理論移動之示意圖	20
圖 5：粒子群最佳化演算流程.....	23
圖 6：全體資料集下重要特徵變數前十名出現次數.....	31
圖 7：前 20%股息波動 Markowitz 投資組合之累積報酬：2004Q1-2019Q4.....	33
圖 8：前 25%股息波動 Markowitz 投資組合之累積報酬：2004Q1-2019Q4.....	34
圖 9：前 30%股息波動 Markowitz 投資組合之累積報酬：2004Q1-2019Q4.....	34
圖 10：前 35%股息波動 Markowitz 投資組合之累積報酬：2004Q1-2019Q4.....	35
圖 11：不同配置方法之股息波動投資組合累積報酬：2004Q1-2019Q4.....	37
圖 12：不同資料集之 10 檔股息波動投資組合累積報酬：2004Q1-2019Q4	39
圖 13：不同資料集之 20 檔股息波動投資組合累積報酬：2004Q1-2019Q4	40
圖 14：不同資料集之 30 檔股息波動投資組合累積報酬：2004Q1-2019Q4	40
圖 15：不同資料集之 40 檔股息波動投資組合累積報酬：2004Q1-2019Q4	41
圖 16：不同資料集之 50 檔股息波動投資組合累積報酬：2004Q1-2019Q4	41

第一章 緒論

因子投資在全球資產管理業中發展蓬勃，從經常被使用的低波動、品質、價值、動能與市值等財務相關因子，到近年流行的主題式投資，如 ESG (Environmental, Social and Governance, ESG)永續投資、5G 未來通訊與生技醫療等都是影響報酬率的關鍵因子。其中在股票市場中產生大波動時，像是受到 2008 年金融海嘯與 2020 年 COVID-19 疫情影響，投資人更加重視投資組合的穩定性與公司體質，高股息低波動與 ESG 因子的投資組合成為市場上長期熱門的投資標的。全球各地有多檔 ETF 追蹤「美國 S&P 500 高股息低波動指數」，包括台灣 2017 年成立的國泰標普低波高息基金 (00702)與美國 2012 年成立的 Invesco 標普 500 高股利低波動 ETF (Invesco S&P 500 High Dividend Low Volatility ETF, SPHD)；而據全球永續投資聯盟 (Global Sustainable Investment Alliance) 統計，全球 ESG 資金規模逐年成長，從 2014 年的 18.2 兆美元到 2020 年的 40.5 兆美元，顯示 ESG 永續投資為長期投資人熱門投資標的。

另外，臺灣金管會在 2020 年也提出四大金融機構管理目標，第一為強化金融體系韌性；第二為推動創新金融，發展金融科技；第三為發展永續金融，促進環境與永續平衡；第四為落實普惠金融，提供更多元化的金融商品；顯示金融業對於永續投資、金融科技與普惠金融投資的重視。

第一節 研究動機與目的

在結合永續金融、金融創新、以及普惠金融三大要素下，本研究欲針對近年投資熱絡的股息波動與 ESG 因子建構投資組合。選定影響報酬率的因子後，不

論是主動式基金或被動式 ETF 投資，在編制上不同選股方法與權重的配置皆會影響整體投資組合的表現，在選擇成分股的方法中，目前市面上股息波動特性的基金與 ETF 主要以過去歷史資料為篩選依據，例如以過去一年的財報資料包括平均殖利率與波動度來排名，選擇過去排名高的作為高息低波的個股，然而，如何確認此種歷史因子投資篩選方式能夠正確預測未來個股具有相同因子特性，過去排名高的高息低波個股並不一定代表未來同樣也是高息低波，因此針對預測股息波動的因子特性問題，本研究採用機器學習的預測方法與歷史因子投資預測方法做比較，期望能找到能更準確預測未來具有相同因子特性的方法。

另外，在權重配置的部分，現有資產配置、基金與 ETF 多半採用歷史資料排名的分數高低來選擇權重、市值加權或是加一些限制條件如同一產業不能超過特定百分比標準等，也有許多資產配置使用傳統 Markowitz 最佳化投資組合的建構方法，然而，Markowitz 最佳化投資組合的方法並未考量到權重平衡與持有資產數量等實務上的限制，如果加上實務上的基數限制，Markowitz 無法找出最佳公式解，形成實務上權重配置的困難，因此針對求解基數限制的 Markowitz 最佳化投資組合問題，本研究以粒子群最佳化演算法 (Particle Swarm Optimization, PSO) 求解並建構實務上可行的最佳化投資組合。

綜合以上動機歸納出本文之研究方向與目的為：(1) 歷史因子投資與機器學習法預測股息波動特性的能力，並與目標正確股息波動投資組合進行績效比較。(2) 觀察 Markowitz 與 PSO 兩種不同配置方法對投資組合的影響。(3) 觀察在機器學習與 PSO 下，ESG 結合股息波動因子長期對投資組合之影響。

第二節 研究架構

本文針對近年投資熱絡的股息波動與 ESG 因子建構美國股市的股票投資組合。先以機器學習方法 XGBoost (eXtreme Gradient Boosting) 預測出不同股息波動特性的股票標的，預測分成兩組：(1) 第一組為高息低波；(2) 第二組為非高息低波，高股息定義為股票殖利率前 30% 之個股，而低波動定義為股票之季年化波動度後 30% 之個股，高息低波則為兩者交集。除了以機器學習方法預測股息波動特性外，也採用歷史因子投資的方法，將過去為高息低波 30% 的個股，當成未來也具有高息低波特性，觀察同樣預測為高息低波個股下，機器學習或是歷史因子投資方法的預測能力較佳，以及何種預測方式透過 Markowitz 建構投資組合後，其績效表現能夠較接近正確股息波動的資料。

經機器學習與歷史因子投資法預測出高息低波個股並建構投資組合後，再進一步加入 PSO 演算法來建構實務可行的投資組合，解決 Markowitz 投資組合最佳化限制之問題，經 PSO 配置出最佳權重與資產數量限制後，比較 Markowitz 與 PSO 配置最佳投資組合之績效表現，觀察在限制資產數量的 PSO 方法是否能達到全部高息低波股票的 Markowitz 投資組合表現。最後，針對全體與 ESG 資料集，同時建構機器學習加上 PSO 配置的投資組合，可以觀察 ESG 因子對投資組合報酬與風險的影響程度，同時，也比較 ESG 加上股息波動因子經 PSO 配置後的長期績效表現與 S&P 500、市面上股息波動因子投資商品之差異。本文研究架

構如圖 1 所示。



圖1：研究架構

本文共分為五個章節，第一章緒論介紹研究動機、目的與架構；第二章文獻探討主要分為因子、預測方法、投資組合理論與演算法相關的四個小節；第三章說明本文會使用到的研究方法，包括機器學習方法 XGBoost、Markowitz 投資組合理論與實務問題、以及粒子群最佳化 PSO；第四章則為實證結果，涵蓋資料描述與模型設定、歷史因子投資與機器學習預測績效、Markowitz 與粒子群最佳配置投資組合報酬績效、ESG 與股息波動投資組合報酬績效；最後第五章為本文的結論與未來展望。

第二章 文獻探討

本章共四小節，第一節為因子投資相關文獻，主要探討股票殖利率、波動度對股票報酬的影響，為本文挑選高息低波因子的依據；第二節為機器學習方法 XGBoost 與其他方法比較，在本文也將驗證 XGBoost 預測股息波動特性之能力；第三節為 Markowitz 投資組合理論與粒子群最佳化 PSO，利用 PSO 能夠求解基數限制的 Markowitz 最佳化投資組合問題，實現實務上可行之投資組合權重與資產數量配置；第四節為 ESG 與投資組合報酬，探討 ESG 永續投資因子對投資組合報酬之影響程度，本文將透過實證觀察 ESG 因子對投資組合的影響。

第一節 因子投資

股息與波動因子不僅在業界長期被作為影響股票報酬之重要因子，在學術界也是長久以來重要研究議題，相關文獻內容廣泛且龐大，Blume (1980) 針對美國股票市場以迴歸模型進行股票殖利率對股票報酬影響之探討，實證顯示高殖利率股票報酬表現較優異，股票殖利率與調整後報酬有顯著正向關係。Haugen & Baker (1991) 以美國市值前 1000 檔股票為基礎，建構每季最小變異數投資組合，發現低波動股票投資組合相較於市場加權投資組合 Wilshire 5000，有較高的報酬率與較低的標準差，印證市場加權法並非最有效率之投資組合。Gombola & Liu (1993) 探討股票殖利率對股票報酬影響程度在牛市與熊市的表現，使用 Compustat 資料庫中股價、股票殖利率、及公司規模資料進行迴歸分析，結果顯示在熊市下，殖利率和投資組合報酬率呈正向關係，在牛市下則為反向關係。Baskin (1989) 對美國股票資料進行迴歸分析後，發現股利政策的支付與配股率和股價波動度呈反向關係，即使在考慮公司規模、獲利波動性、負債比率、成長

性與產業後，仍支持股票殖利率高、股價波動度低的結果。

綜合股票殖利率與波動度因子相關文獻探討，可以歸納出四個主要論點，第一股票殖利率與報酬呈正向關係；第二股票波動度與報酬呈反向關係；第三股票殖利率與波動度呈反向關係；第四股票殖利率與報酬在牛市與熊市下表現不同，牛市下呈反向關係，熊市下呈正向關係。本文將於第四章實證結果中，觀察在不同景氣下，高殖利率與低波動度對股票投資組合報酬之影響程度。

第二節 機器學習方法

機器學習方法廣泛被運用於各領域，而近年來於財務領域亦不例外，Patel, Shah, Thakkar & Kotecha (2015) 利用四種機器學習模型預測個股股價與股票指數走勢，包括神經網路 (Artificial Neural Network, ANN)、支援向量機 (Support Vector Machine, SVM)、隨機森林 (Random Forest)、樸素貝葉斯 (Naive Bayesian) 四種方法，使用印度股市資料，實證結果顯示隨機森林預測表現優於其他三種機器學習模型。Zhongbin & Jinwu (2019) 使用六種機器學習模型建立股票選擇模型，包括主成分分析 (Principal Component Analysis)、羅吉斯迴歸 (Logistic Regression)、樸素貝葉斯、ANN、決策樹 (Decision Tree)、隨機森林、XGBoost 六種模型，資料為上海和深圳 300 指數之成分股，隨機森林、神經網路、以及 XGBoost 分類效果較佳，其中神經網路模型表現相對較不穩定且波動大，使用隨機森林與 XGBoost 所建構之投資組合績效優於平均。

Chen & Guestrin (2016) 提出 XGBoost 模型，為以樹模型為基礎的監督式機器學習模型，說明 XGBoost 在數學算法與程式化上，相較於傳統的梯度提升決

策樹 (Gradient Boosting Decision Tree, GBDT) 模型更加優化。

機器學習方法運用於財務領域之相關文獻探討可歸納出兩大觀點，第一在股票分類問題中，隨機森林與 XGBoost 預測效果與其投資組合報酬較佳；第二，XGBoost 為近年學者提出之更優化模型。因此，本文將以 XGBoost 機器學習模型預測高息低波之個股，觀察其預測能力與所建構投資組合之績效表現。

第三節 Markowitz 投資組合理論與粒子群最佳化

Markowitz (1959) 提出平均數-變異數投資組合模型 (Mean-Variance Model) 與效率前緣 (Efficient Frontier) 之概念，成為經常被使用的投資組合理論，平均數-變異數投資組合模型說明資產配置時，投資組合透過權重的調整，能得到不同的預期報酬與風險，而風險偏好或風險趨避取決於個別投資人的風險屬性。效率前緣則代表最有效率的投資組合，在效率前緣上任一點，預期報酬相同時，相對總風險最小；或在相同風險下，有最高的預期報酬。

投資人可藉由平均數-變異數投資組合模型調整資產權重配置，以更接近效率前緣曲線上的最佳投資組合，然而，Michaud (1989) 提出在 Markowitz 投資組合最佳化會忽略財務與現實上的限制，包括配置過多權重於單一資產上，且這些權重較大的資產具有預期高報酬及低波動特性，這些資產有可能使得最佳化過程產生較大的誤差，單一資產配置權重過度集中，亦會使投資組合的風險提高；另外，投資組合中資產持有數量、單一資產最大與最小權重也是實務上須考量的部分。Avramov & Zhou (2010) 指出在實務上無法做空時，Markowitz 最佳化投資組合更容易產生資產過度集中問題，使其權重配置不平衡，同時提出 Bayesian 方法

解決 Markowitz 投資組合最佳化的問題。

由於原本 Markowitz 平均數-變異數模型加上實務限制式後，無法有效求解最佳化的投資組合，Deng, Lin & Lo (2012) 提出粒子群最佳化 (Particle Swarm Optimization, PSO) 演算法可用於求解基數限制下的 Markowitz 平均數-變異數模型 (Cardinality Constrained Markowitz Portfolio Optimization problems, CCMPO problems)，並和其他三種演算法進行比較，包括遺傳演算法 (Genetic algorithm)、禁忌搜索演算法 (Tabu Search)、模擬退火演算法 (Simulated Annealing)，PSO 在準確率、分散性、穩定度、以及誤差上皆優於其他三種方法。Kennedy & Eberhart (1995) 提出粒子群最佳化演算法為，主要概念為模擬社會行為，在社群中個體行為會同時受到過去自身經驗與外在社會所影響，粒子會藉此調整移動速度與更新位置，以更接近最佳解。

經相關文獻探討後，可以發現雖然 Markowitz 投資組合理論被廣泛運用，但 Markowitz 投資組合理論仍存在實務限制，加入限制後難以求解最佳化投資組合，因此運用 PSO 能幫助求解基數限制下的 Markowitz 最佳化投資組合問題。

第四節 ESG 與投資組合報酬

ESG 包含環境、社會責任、公司治理三大概念，其中社會責任投資 (Socially Responsible Investment, SRI) 早期大部分為已開發國家如歐美地區發展較快，許多相關文獻研究長短期社會責任投資對投資組合的影響，近年來投資人除了注重企業財務面的表現外，環境、社會責任與公司治理也備受重視，陸續有文獻探討已開發與開發中國家之 ESG 排名對投資組合報酬、風險分散程度的影響。

Renneboog, Ter Horst & Zhang (2008) 探討歐洲、英國、美國、亞洲四大區之社會責任投資基金表現，實證結果顯示美國、英國、歐洲大陸和亞太國家的社會責任基金表現落後於國內基準，除了法國、日本和瑞典等國家外，社會責任基金的風險調整後報酬與傳統基金表現在統計上並無明顯差異，然而，可以透過社會責任的因子識別出表現不好的基金，顯示其有相當程度的基金選擇能力。Nofsinger & Varma (2014) 分析美國社會責任基金長期表現，發現在金融危機時，這些社會責任基金表現優於一般基金，而在非金融危機時則表現較差，顯示社會責任投資因子對投資組合報酬影響，與市場狀況存在一定的關係，對於想要保護下行風險的投資人，社會責任投資為較佳的選擇。

Verheyden, Eccles & Feiner (2016) 比較 46 個已開發與開發中國家股票中，ESG 排名對報酬、風險與資產分散程度的影響，以 Sharpe ratio 來看，ESG 排名前 10% 與 25% 的股票僅略優於未分類之股票，而以最大回撤 (Maximum Draw Down, MDD) 來看，高 ESG 排名最大回撤優於未分類股票，顯示 ESG 因子投資能獲得較高報酬與較低的下檔風險。

Ashwin Kumar et al. (2016) 對道瓊永續指數 (Dow Jones Sustainability Index, DJSI) 成分股與非成分股進行分析，道瓊永續指數藉由企業永續評鑑方法 (Corporate Sustainability Assessment, CSA) 選出評分最高前 10% 企業，意即 ESG 表現良好的公司，結果顯示 ESG 因子對報酬與風險的影響會根據不同行業而表現不同，對原物料、金融、能源和科技業影響較大，ESG 的實踐能夠幫助公司降低風險，且在大多數行業中，ESG 公司也獲得更高的報酬，尤其是能源、食品以及醫療保健行業更為明顯，此與傳統上低風險低報酬的觀念相反，作者指出原因在於不同行業的特性導致，例如能源產業中，多數公司未以友善環境與永續方式

經營，因此有做 ESG 的石油或天然氣公司會有較佳的表現，整體來說，在考慮 ESG 實踐所帶來的隱藏價值後，ESG 公司確實能提供較高的風險調整後報酬、Sharpe ratio 越高。

綜合 ESG 因子對投資組合報酬影響之相關文獻，可以歸納出四個主要探討方向：(1) 過去研究結果不完全一致，有些文獻結果顯示 ESG 之投資組合擁有較高報酬、另外也有一部分文獻認為無明顯差異。(2) 多數文獻認為 ESG 雖然未必能獲得絕對高報酬，然而，在考量風險後，能得到較低的波動度與下檔風險、較高的 Sharpe ratio。(3) ESG 因子對報酬的影響會隨市場景氣與產業類別而有所不同，當金融危機發生時，ESG 公司擁有較佳的抗跌特性；其在不同產業對報酬影響程度也不同。本文將透過不同股票資料集（全體與 ESG）長期投資組合表現之比較，觀察 ESG 因子對投資組合報酬與風險的影響程度。

第三章 研究方法

本章研究方法共分為三節，對應研究流程與架構，先以機器學習方法 XGBoost 預測個股股息波動特性，而歷史因子投資法單純以過去股息波動資料作為未來同樣具有相同特性的依據，因此本文預測股息波動特性之研究方法以介紹 XGBoost 機器學習模型為主。第二節將介紹 Markowitz 投資組合理論與其實務會遇到之問題。第三節則引進粒子群最佳化演算法，以求解基數限制下的 Markowitz 最佳化投資組合，實現最佳化的高息低波投資組合。

第一節 XGBoost 機器學習方法

一、XGBoost 概念介紹

XGBoost 模型由 Chen & Guestrin (2016) 提出，為監督式學習的一種，其目標函數包含預測誤差 (Loss) 與正則項 (Regularization)，XGBoost 的目標函數將於後面詳細說明。XGBoost 主要概念以樹模型為基礎，聚集許多分類和回歸樹 (Classification and Regression Tree, CART) 模型形成強分類器，回歸樹擁有與決策樹一樣的分支方式，在每個葉端 (Leaf) 會有一個對應的預測分數 (Prediction Score)，且每添加一棵樹就會學習一個新的函數，並擬合上次預測的殘差，最後根據每顆樹對應的分數加總，即為預測結果，本文將 XGBoost 用於股息波動特性之預測分類，主要概念示意圖如圖 2、演算流程如圖 3 所示。

輸入財報資料變數：殖利率、波動度...

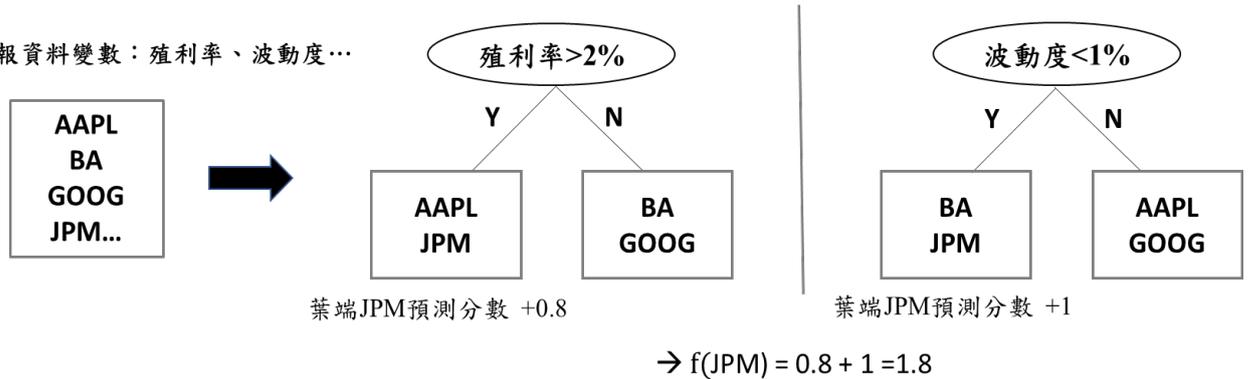


圖2：XGBoost分類概念示意圖

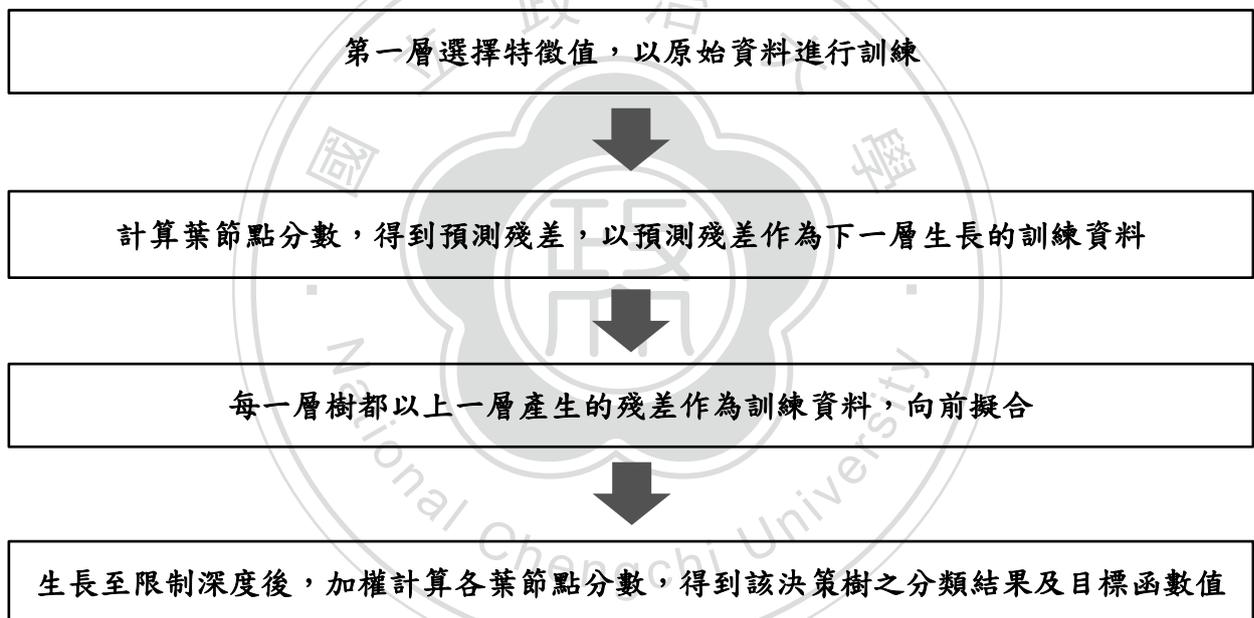


圖3：XGBoost演算流程

二、目標函數與學習方式

XGBoost 之目標函數如下：

$$\text{Obj} = \sum_i l(\hat{y}_i, y_i) + \sum_k \Omega(f_k) \text{ , where } \Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2 \quad (3.1)$$

其中 $\sum_i l(\hat{y}_i, y_i)$ 代表損失函數 (Loss Function) 即預測誤差，計算經分類後葉節點所得權重分數 \hat{y} 與實際 y 之預測殘差， $\hat{y}_i = \sum_{k=1}^k f_k(x_i)$ 為 k 棵樹對資料 x_i 的總預測； $\sum_k \Omega(f_k)$ 代表正則函數，藉由 $\Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2$ 限制單一決策樹生長層樹與深度，讓模型更穩定，其中 T 表示葉子節點的個數、 γ 能控制葉子節點的個數、 w 表示葉子節點的分數、 λ 可以控制葉子節點的分數讓他不會過大，並防止模型過度擬合。透過尋找損失函數與正則函數的平衡，找出模型適合資料集的設定。

在 XGBoost 模型中，使用於回歸樹的學習方式為 Gradient Boosting，也就是先前主要概念提及，每次預測時都會加入一個新的學習函數，且保留前面的學習函數，學習方式如下：

$$\begin{aligned} \hat{y}_i^{(0)} &= 0 \\ \hat{y}_i^{(1)} &= f_1(x_i) = \hat{y}_i^{(0)} + f_1(x_i) \\ \hat{y}_i^{(2)} &= f_1(x_i) + f_2(x_i) = \hat{y}_i^{(1)} + f_2(x_i) \\ &\dots \\ \hat{y}_i^{(t)} &= \sum_{k=1}^t f_k(x_i) = \hat{y}_i^{(t-1)} + f_t(x_i) \end{aligned} \quad (3.2)$$

其中 $\hat{y}_i^{(t)}$ 表示第 t 次訓練的總預測，即為前 t 次加入的所有函數值總和， $\hat{y}_i^{(t-1)}$ 為保留前面的學習函數， $f_t(x_i)$ 則為新加入的學習函數，而此新加入的函數取決於對目標函數做優化，過程如下：

$$\begin{aligned} \text{Obj}^t &= \sum_{i=1}^n l(y_i, \hat{y}_i^{(t)}) + \sum_{i=1}^{t-1} \Omega(f_i) + \Omega(f_t) \quad (3.3) \\ &= \sum_{i=1}^n l(y_i, \hat{y}_i^{(t-1)} + f_t(x_i)) + \Omega(f_t) + \text{constant} \end{aligned}$$

經多項式的泰勒展開後得到：

$$\text{Obj}^t \cong \sum_{i=1}^n [l(y_i, \hat{y}_i^{(t-1)}) + g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t) + \text{constant} \quad (3.4)$$

$$g_i = \frac{\partial l(y_i, \hat{y}_i^{(t-1)})}{\partial \hat{y}_i^{(t-1)}}, \quad h_i = \frac{\partial^2 l(y_i, \hat{y}_i^{(t-1)})}{\partial (\hat{y}_i^{(t-1)})^2} \quad (3.5)$$

其中 g_i 、 h_i 分別為誤差函數的一次與二次微分， $l(y_i, \hat{y}_i^{(t-1)})$ 為常數，因此將常數省略後可以將目標函數簡化如下：

$$\text{Obj}^t \cong \sum_{i=1}^n [g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i)] + \Omega(f_t), \quad (3.6)$$

$$\text{where } \Omega(f) = \gamma T + \frac{1}{2} \lambda \|w\|^2$$

將葉端分數定義為權重 w ，每次分類所得的分數為 $f_t(x_i) = w_q(x)$ ， q 代表分類到第幾個葉端，因此， n 筆資料進入回歸樹後會被分類到其中一個葉端權重上，也就是 n 筆資料分配至 T 個集合中， $I_j = \{i | q(x_i) = j\}, j = 1, \dots, T$ ，代入下列目標函數中：

$$\text{Obj}^t \cong \sum_{i=1}^n [g_i w_q(x_i) + \frac{1}{2} h_i w_j^2] + \gamma T + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2 \quad (3.7)$$

$$= \sum_{j=1}^T \left[\left(\sum_{i \in I_j} g_i \right) w_j + \frac{1}{2} \left(\sum_{i \in I_j} h_i + \lambda \right) w_j^2 \right] + \gamma T$$

將 $G_j = \sum_{i \in I_j} g_i$, $H_j = \sum_{i \in I_j} h_i$ 代入可得目標函數之最佳權重和分數，目標函數的值越低，代表模型優化效果越佳：

$$Obj^t = \sum_{j=1}^T \left[G_j w_j + \frac{1}{2} (H_j + \lambda) w_j^2 \right] + \gamma T \quad (3.8)$$

$$w_j^* = -\frac{G_j}{H_j + \lambda} \quad (3.9)$$

$$Obj_{optimal} = -\frac{1}{2} \sum_{j=1}^T \frac{G_j^2}{H_j + \lambda} + \gamma T \quad (3.10)$$

在 Boosting 學習中的分支方式為，將葉端每個資料皆嘗試做分支，透過計算每次分支得到的目標分數增益 (Gain) 找出最佳分支方法，目標分數增益公式如下，中括弧內第一項代表左分支的分數、第二項為右分支分數、第三項為不分支的分數， γ 則作為增加分支複雜度之代價：

$$Gain = \frac{1}{2} \left[\frac{G_L^2}{H_L + \lambda} + \frac{G_R^2}{H_R + \lambda} - \frac{(G_L + G_R)^2}{H_L + H_R + \lambda} \right] - \gamma \quad (3.11)$$

本文將運用 XGBoost 模型來預測股息波動特性的個股，主要觀察準確率 (Accuracy)、精確率 (Precision)、召回率 (Recall)、F1-score 以判斷模型之預測能力，詳細模型之實驗設計與研究流程將於第四章實證結果中呈現。

第二節 Markowitz 投資組合理論與實務問題

一、Markowitz 投資組合理論

Markowitz (1959) 提出投資組合理論，透過平均數-變異數模型及效率前緣的概念，提供投資人建構報酬與風險最佳化投資組合之方法，平均-變異數模型之目標函數為：

$$\min \lambda \left[\sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} \right] - (1 - \lambda) \left[\sum_{i=1}^N w_i u_i \right] \quad (3.12)$$

$$\text{subject to } \sum_{i=1}^N w_i = 1, 0 \leq w_i \leq 1 \quad (3.13)$$

其中 N 代表不同的資產數量； w_i 為投資組合中資產 i 的權重，所有資產權重總和為 1 且單一資產權重必須介於 0 到 1 之間； σ_{ij} 為資產 i 和資產 j 的共變異數； u_i 為資產 i 的預期報酬； λ 為風險趨避參數，用以平衡風險與報酬的參數，且介於 0 與 1 之間， λ 越大代表投資人越風險趨避， $\lambda = 1$ 時，投資人為完全風險趨避， $\lambda = 0$ 時，投資人為完全風險偏好，風險趨避程度取決於個別投資人的風險趨避程度。

在 Markowitz 目標函數中，以極小化目標函數來最佳化投資組合，第(3.12)式之第一項 $\sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}$ 為降低投組總風險、第二項 $\sum_{i=1}^N w_i u_i$ 為提高投組總報酬，因此投資人決定各自風險趨避程度 λ 後，即可透過此目標函數求解出最佳投資組合。而不同的 λ 會產生不同的目標函數，在不同目標函數之下，也會產生不一樣的報酬與風險組合，效率前緣即為這些組合連線而成，在效率前緣上任一點，都是給定預期報酬相同下，相對總風險最小；或在給定相同風險下，有最高的預期報酬，形成最佳投資組合。

二、Markowitz 最佳化投資組合之實務問題

Markowitz 平均數-變異數模型之最佳化投資組合概念被廣泛使用於資產配置上，然而，Michaud (1989) 提出在 Markowitz 投資組合最佳化理論未考量實務上的限制，例如投資組合權重配置不平衡，過度集中於單一資產，或是未限制投資組合資產數量，這些皆為實務上會遇到的問題，參考 Fernandez & Gomez (2007) 公式之下，實務上基數限制的 Markowitz 平均數-變異數模型為：

$$\min \lambda \left[\sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij} \right] - (1 - \lambda) \left[\sum_{i=1}^N w_i u_i \right] \quad (3.14)$$

$$\text{subject to } \sum_{i=1}^N w_i = 1, 0 < w_i < 1 \quad (3.15)$$

$$\sum_{i=1}^N z_i = K \quad (3.16)$$

$$\varepsilon_i z_i \leq w_i \leq \delta_i z_i, i = 1, \dots, N \quad (3.17)$$

$$z_i \in \{0,1\}, i = 1, \dots, N \quad (3.18)$$

其中新增的符號為，K 代表投資組合的總資產數量；二元變數 $z_i \in \{0,1\}$ 為 1 代表資產 i 在投資組合中，反之為 0； ε_i 為投資組合最小持有權重， δ_i 為投資組合最大持有權重。而除了原本平均數-變異數模型目標函數與限制外，新增了(3.16)式代表總資產數量 K 的限制式，第(3.17)式則表示單一資產加入最小及最大權重的限制，加入此式可以避免投資組合權重過度集中於個別資產。

加入實務限制式後的 Markowitz 平均數-變異數模型，其最佳化投資組合問

題並不存在有效公式解，因此 Deng et al. (2012) 提出粒子群最佳化演算法，用於求解基數限制下的 Markowitz 平均數-變異數模型，本文在使用機器學習 XGBoost 模型預測出高息低波個股後，需藉由粒子群最佳化求解 Markowitz 最佳化投資組合，在考量實務資金與交易成本下，限制投資組合數量於 10、20、30、40 與 50 檔個股 (K)，由粒子群最佳化挑選出最佳投資組合。

在 Markowitz 平均數-變異數模型之目標函數中，由於風險趨避參數 λ 取決於個別投資人，自行設定 λ 的值可能不夠一般化，因此將極小化目標函數改為極大化 Sharpe ratio，更新之目標函數與限制式如下：

$$\max \text{ Sharpe ratio} = \frac{\sum_{i=1}^N w_i (u_i - r_f)}{\sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}} \quad (3.19)$$

$$\text{subject to} \quad \sum_{i=1}^N w_i = 1, \quad 0 < w_i < 1 \quad (3.20)$$

$$\sum_{i=1}^N z_i = K \quad (3.21)$$

$$\varepsilon_i z_i \leq w_i \leq \delta_i z_i, \quad i = 1, \dots, N \quad (3.22)$$

$$z_i \in \{0,1\}, \quad i = 1, \dots, N \quad (3.23)$$

原本基數限制下 Markowitz 最佳化投資組合之目標函數不同之處在於第 (3.19) 式，雖然第 (3.19) 式極大化 Sharpe ratio 目標函式與第 (3.14) 式極小化原目標函數公式不同，但其實極大化 Sharpe ratio 只是原本目標函數的一個特例，觀察第 (3.19) 式 $\frac{\sum_{i=1}^N w_i (u_i - r_f)}{\sum_{i=1}^N \sum_{j=1}^N w_i w_j \sigma_{ij}}$ 可以發現增加了 r_f 符號，代表無風險利率，極大化 Sharpe ratio 為無風險利率與效率前緣曲線的切點，意即其同為 Markowitz 最佳化投資組合的其中一個點。

更新後的目標函數代表本文所建構之投資組合目標為，追求極大化承擔相同風險之下的預期報酬，除了避免受參數 λ 影響外，Sharpe ratio 也是現今投資人著重關注的指標，尤其在高息低波與 ESG 因子探討中，報酬與風險的平衡更加重要，因此本文將以粒子群演算法求解第(3.19)-(3.23)目標函數與限制式，得到實務限制條件下，且更加一般化的最佳投資組合。

第三節 粒子群最佳化

一、粒子群最佳化介紹

Kennedy & Eberhart (1995) 提出粒子群最佳化主要概念來自模擬鳥群覓食的社會行為，在社會群體中，個體行為會受到過去的自身經驗影響，也會受到外在社會行為影響，粒子會不斷調整移動速度並更新位置，快速在問題空間中利用這個移動與更新規則找到最佳解，更新規則如下：

$$w_{i,j}^{n+1} = w_{i,j}^n + v_{i,j}^{n+1} \quad (3.24)$$

$$v_{i,j}^{n+1} = x * v_{i,j}^n + c_1 * r_1 * (p_{i,j} - w_{i,j}^n) + c_2 * r_2 * (p_g - w_{i,j}^n) \quad (3.25)$$

其中 $w_{i,j}^n$ 代表 n 次迭代下 j 維度下粒子 i 的位置 (此處的 $w_{i,j}^n$ 與 Markowitz 目標函數與限制式中的 w_i 代表意義不同)； $v_{i,j}^n$ 代表 n 次迭代下 j 維度下粒子 i 的速度； $p_{i,j}$ 代表在 j 維度的粒子 i 從最初時點開始，到過距離目標最近的位置； p_g 代表所有粒子從最初時點開始，到過距離目標最近的位置； r_1 與 r_2 是一個介於 0 到 1 之間的隨機變數； c_1 與 c_2 分別決定個別粒子與全體粒子貢獻程度的正係數，此係數越高，表示粒子移動速度較快，但經過的位置會較少； x 為前一期速度的保

留程度，Shi & Eberhart (1998) 用以促進粒子群最佳化的收斂， $x \geq 1$ 時，粒子移動速度會增加至最大速度，使其無法收斂， $x \leq 1$ 時，粒子移動速度會逐漸減少到 0，達成收斂，而最大速度計算方式為對粒子位置的上下界差距，取特定比率計算。

粒子群最佳化初始為一群隨機粒子，在資產配置的問題中，每一個粒子代表一個包含不同資產與權重的投資組合，粒子的位置 $w_{i,j}^n$ 代表 n 次迭代下投資組合 i ，投資組合內 j 個資產分別的權重則以維度 j 表示，透過 n 次迭代後，不斷移動並調整速度，第(3.24)式代表第 n 次迭代下 j 維度下粒子 i 的位置 $w_{i,j}^n$ 會根據速度 $v_{i,j}^{n+1}$ 更新，達到 $w_{i,j}^{n+1}$ 位置，第(3.25)式為速度 $v_{i,j}^{n+1}$ 的更新規則。每個粒子的移動都代表著每個投資組合中資產權重的改變，最後達到目的地，運用於 Markowitz 平均數-變異數模型上，目的地即為效率前緣曲線，本文將目標函數設定為極大化 Sharpe ratio，目的地為無風險利率與效率前緣曲線之切點，不同資產權重組合的粒子，會根據自身和周遭粒子群的位置來更新位置，粒子所在位置對應投資組合的報酬與風險，粒子在 Markowitz 投資組合理論中移動示意圖如圖 4。

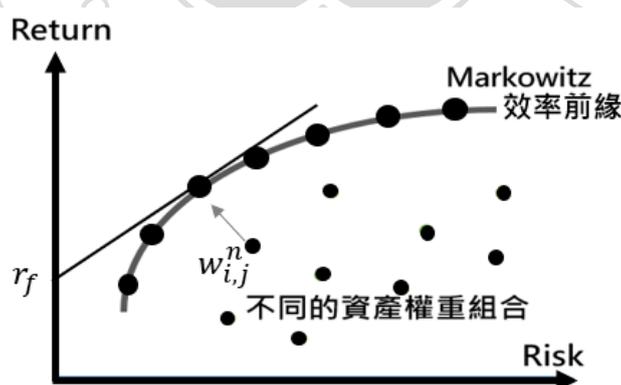


圖4：粒子在Markowitz投資組合理論移動之示意圖

二、粒子群最佳化求解基數限制下的 Markowitz 最佳化投資組合

如本章第二節所述，Markowitz 最佳化投資組合加入實務限制條件後，其目標函數與限制式不存在有效公式解，Deng et al. (2012) 以粒子群最佳化演算法求解此問題，因此本文使用相同方法，在機器學習 XGBoost 預測出高息低波公司後，接著再以粒子群最佳化求解基數限制下的 Markowitz 最佳化投資組合，有效得到限制持有資產數量、單一資產權重的投資組合。為了讓代表投資組合的粒子移動到目標效率前緣曲線上，粒子群最佳化過程中有三個主要條件與參數的設定，第一為初始化粒子群的設定。在初始化粒子群的設定中，會先建立 m 個隨機產生粒子的粒子群，代表 m 個投資組合內的資產會先給定隨機權重。第二必須滿足限制式資產數量 K 、最小 ε_i 與最大 δ_i 權重設定，所做的調整。粒子移動到目的地的過程中，每一次的迭代都要使總持有資產數量 $K^n = K$ ， K 為給定的總資產數量，如果 $K^n < K$ ，模型會隨機選新的資產加進投資組合內，並先給予最小權重，如果 $K^n > K$ ，模型會從原本投資組合內刪除權重最小之資產，持續調整直到 $K^n = K$ 。投資組合內資產的最小 ε_i 與最大 δ_i 則須滿足第(3.21)式 $\varepsilon_i z_i \leq w_i \leq \delta_i z_i$ ， $i = 1, \dots, N$ ，且投資組合內每個資產的最小最大權重必須介於 0 和 1 之間 $0 \leq \varepsilon_i z_i \leq w_i \leq \delta_i z_i \leq 1$ ，在每一個迭代過程中的資產權重 a_i 須滿足此限制，如果 $a_i < \varepsilon_i$ ，最小權重 ε_i 會取代 a_i ，如果 $a_i > \delta_i$ ，則資產的權重會根據以下式子重新計算：

$$w_i = \varepsilon_i + \frac{a_i}{\sum_{j \in Q, a_j > \varepsilon_j} a_j} * (1 - \sum_{j \in Q} \varepsilon_j) \quad (3.26)$$

其中 Q 為投資組合內包含 K 個資產的集合，第(3.26)式能夠減少重要性較低的資產之權重。

最後一個條件設定為參數的調整，包括前一次速度對更新速度的影響程度 x 、分別決定個別粒子與全體粒子貢獻程度的正係數 c_1 與 c_2 。 x 表示前一次速度的慣

性，這個慣性影響程度越大，粒子越會在空間裡大範圍的移動，反之，慣性影響程度越小，粒子會在空間裡往鄰近的位置移動；而在求解最佳化投資組合過程中，權重在初期應有較大幅度的變動，也就是使其速度慣性影響程度 x 越大，在越靠近最佳化目標時效率前緣曲線時，權重變動的幅度應縮小，也就是使其速度慣性影響程度 x 越小、找尋鄰近位置的最佳解，速度慣性影響程度 x 的調整依據 Shi & Eberhart (2007)指出其需隨迭代次數增加線性減少，速度慣性影響程度 x 與 n 迭代次數的關係為：

$$w(n) = [w(0) - w(N_n)] * \frac{N_n - n}{N_n} + w(N_n) \quad (3.27)$$

其中 $w(n)$ 為迭代次數 n 的權重； $w(0)$ 為初始權重值，參考 Deng et al. (2012)，通常設為 0.9； $w(N_n)$ 為最終權重，通常設為 0.4； N_n 則為最大迭代次數。

另外，個別粒子與全體粒子貢獻程度的正係數 c_1 與 c_2 兩個值通常會設為相等，讓自身與社群影響移動程度不會差異過大，而也有文獻提供更佳 c_1 與 c_2 的公式，如 Ratnaweera, Halgamuge & Watson (2004) 將 c_1 與 c_2 的值設為隨迭代次數而線性改變，如以下公式：

$$c_1(n) = (c_{1,min} - c_{1,max}) * \frac{n}{N_n} + c_{1,max} \quad (3.28)$$

$$c_2(n) = (c_{2,max} - c_{2,min}) * \frac{n}{N_n} + c_{2,min} \quad (3.29)$$

其中假設 $c_{1,max} = c_{2,max} = 2.5$ 、 $c_{1,min} = c_{2,min} = 0.5$ ， $c_1(n)$ 會隨迭代次數增加而減少， $c_2(n)$ 則相反，代表在粒子移動初期，會更重視自身的影響程度，而在後期越接近目標函數時，則會較重視群體的影響程度。

最後，介紹完粒子群最佳化的核心概念、移動與更新位置方式、以及最佳化過程中三個條件與參數的設定後，可用圖 5 呈現粒子群最佳化的演算流程。

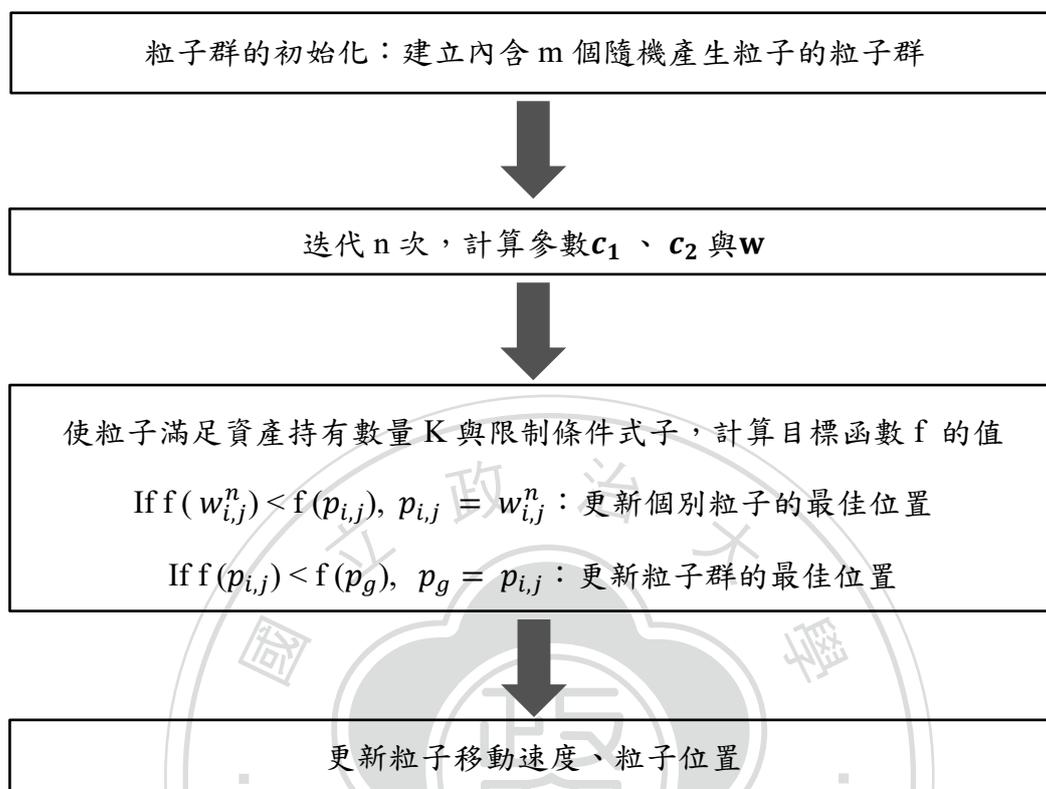


圖 5：粒子群最佳化演算流程

第四節 預測與績效評估指標

一、預測能力評估指標

本文實證使用兩種方法來預測股息波動特性，分為高息低波與非高息低波兩類，分類模型中較常見的評估預測能力指標為 Accuracy、Precision、Recall 與 F1-score，這四項指標計算來自混淆矩陣 (Confusion Matrix)，並由四個元素組成混淆矩陣，如表 1 所示。

表1：混淆矩陣(Confusion Matrix)

Confusion Matrix		True	
		Positive	Negative
Predict	Positive	TP (True Positive)	FP (False Positive)
	Negative	FN (False Negative)	TN (True Negative)

其中 True 代表實際分類，Predict 代表預測分類，TP (True Positive) 為正確預測真實為正的樣本，運用於本研究中即為正確預測真實為高息低波個股；TN (True Negative) 為正確預測真實為負的樣本，也就是正確預測非高息低波個股；FP (False Positive) 為錯誤預測為正，實際上為負樣本，代表實際上為非高息低波個股，卻被誤分類為高息低波，此為型二誤差 (Type II Error)；FN (False Negative) 為錯誤預測為負，實際上為正的樣本，代表實際上為高息低波，卻誤被分類為非高息低波個股，此為型一誤差 (Type I Error)。Accuracy、Precision、Recall 與 F1-score 四項常見的評估分類能力指標由混淆矩陣建構而成：

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (3.30)$$

$$Precision = \frac{TP}{TP + FP} \quad (3.31)$$

$$Recall = \frac{TP}{TP + FN} \quad (3.32)$$

$$F1 - score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \quad (3.33)$$

Accuracy 為所有樣本皆預測正確的比率，在分類樣本不平衡下，易發生全部分為多數的類別、高 Accuracy 的情況發生，但更加關注的是少數分類是否正確，因此較不會使用此指標；Precision 代表預測為真下，有多少個股實際上為高息低波；Recall 代表真實為高息低波個股下，有多少個股預測正確；F1-score 則為

Precision 與 Recall 的調和平均數，綜合評斷 Precision 與 Recall。在本實驗設計中，使用歷史因子投資與機器學習方法預測股息波動特性，期望預測為高息低波的個股，確實為高息低波，因此較關注 Precision 的表現，並輔以綜合指標 F1-score 來評估預測與分類能力。

二、投資組合績效評估指標

經過機器學習與歷史因子預測出高息低波的股票群後，使用粒子群最佳化演算法建構高息低波投資組合，由於高息低波的因子投資主要為追求穩定獲利、並降低風險，因此評估投資組合配置績效除了報酬率與標準差(波動度)外，將以綜合報酬與風險的幾個指標來評估績效：

$$\text{Sharpe ratio} = \frac{R_p - R_f}{\sigma_p} \quad (3.34)$$

$$\text{Sortino ratio} = \frac{R_p - R_f}{\sigma_{p-\text{negative}}} \quad (3.35)$$

第(3.34)式中， R_p 為投資組合整體報酬， R_f 為無風險利率，本文使用美國 10 年期公債殖利率， σ_p 為投資組合整體風險，Sharpe ratio 衡量每承擔一單位的風險，會獲得多少超額報酬，其為投資人普遍衡量報酬與波動度的指標之一。第(3.35)式中， $\sigma_{p-\text{negative}}$ 為投資組合的下方風險，Sortino ratio 代表每承擔一單位下方風險，能夠獲得多少超額報酬，Sortino ratio 更關注於跌幅的波動。在高息低波投資組合中，將著重觀察投資組合長期報酬與波動，預期高息低波因子投資組合擁有較低的波動，將於下一部份進行投資組合配置績效之實證分析。

第四章 實證結果

本文樣本資料為 2003 年到 2019 年美國上市公司資料，以機器學習方法實務建構 ESG 高息低波投資組合，實驗設計如研究架構所述，以 XGBoost 機器學習方法與歷史因子投資法預測高息低波個股，比較兩種方法之預測能力，並利用 Markowitz 配置方法每季建構投資組合，觀察預測能力高低與目標股息波動投資組合績效的關係；接著以 PSO 求解 Markowitz 基數限制下的最佳化投資組合，分析是否能夠以限制資產數量的 PSO 來建構與 Markowitz 相近的投資組合；最後，加入 ESG 資料集來比較 ESG 結合股息波動對長期投資組合報酬之影響。配合研究架構與實驗設計，本章節共分為五小節，第一節為資料描述與模型設定；第二節為歷史因子投資與機器學習預測能力之比較；第三節為歷史因子投資與機器學習績效比較；第四節 Markowitz 與粒子群最佳配置績效比較；第五節為 ESG 與股息波動投資組合績效比較。

第一節 資料描述與模型設定

一、資料介紹

本文實證資料使用美國上市公司財報資料與 ESG 分數資料，財報資料之來源為 Wharton Research Data Services (WRDS) 中的 Compustat 季資料，其中股價資料來源為 Center for Research in Security Prices (CRSP) 之日資料，財報資料包含所有 Compustat 可得的財報比率，再加上波動度，共 75 個變數，涵蓋面向有價值 (Valuation)、獲利 (Profitability)、資本 (Capitalization)、財務健全 (Financial Soundness)、償債能力 (Solvency)、流動性 (Liquidity)、效率 (Efficiency)、本益

比、波動度、平均波動度、平均殖利率、前一季分類與其他財務相比率，如附錄所示。ESG 分數資料來源為 Eikon with Datastream - Thomson Reuters 的 ESG Scores 年資料，包含三種評分方式，原始 ESG 分數、ESG 爭議分數以及 ESG 綜合分數。資料期間為 2003 年第一季到 2019 年第四季，其中，ESG 資料用來將個股分成不同資料集，以前一年有無 ESG 分數為下一年四個季度個股分在 ESG 資料集與否標準，例如 2018 年有 ESG 分數的個股，2019 年第一季到第四季就會被歸類在 ESG 資料集中，ESG 資料集平均 800 家公司，而全體資料集則是不分前一年有無 ESG 資料之全部個股，為美國紐約及那斯達克交易所所有上市公司，平均大約 3566 家公司。

二、模型設定

本文 XGBoost 模型實證部分，主要使用過去四季個股財報資訊，來預測未來一季的股息(殖利率) 與波動(年化標準差) 特性，依機器學習模型預測出之目標函數數量，高息低波之標準設定為殖利率前 30%、波動度後 30%之個股，目標變數設定如下式：

$$\begin{aligned} &\text{If Dividend Yield} \geq \text{Dividend Yield}_{30\%} \ \& \ \text{Volatility} \leq \text{Volatility}_{70\%} \quad (4.1) \\ &\text{then } Y = 1, \text{ else } Y = 0 \end{aligned}$$

XGBoost 模型分為訓練集與測試集，其中訓練集包含 30%驗證集，訓練方式採滾動式 (Rolling) 訓練，以四季為一訓練周期，共滾動 63 次，並對訓練集後一季進行測試。訓練期間為 2003 第一季到 2019 第三季，測試期間為 2003 第二季到 2019 第四季。XGBoost 參數設定上，採用 gbtree 模型 (XGBoost 參數設定程

式函數：booster)、樹的最大深度 (XGBoost 模型參數設定程式函數：max_depth) 設為 10、葉節點最小樣本數 (XGBoost 模型參數設定程式函數：min_child_weight) 為 3、樹模型之權重 (XGBoost 模型參數設定程式函數：eta) 為 0.2。而歷史因子投資法以過去四季平均殖利率前 30%與波動度後 30%交集之個股，作為未來同樣為高息低波個股之預測。

粒子群最佳化實證部分，根據機器學習預測出每季高息低波個股後，再由 PSO 建構 10、20、30、40 與 50 檔個股的最佳化投資組合，以過去一季報酬與波動度來計算最佳化 Sharpe ratio。其參數設定為迭代次數 (Iterations) 100 次；粒子個數 (Swarmsize) 500 個，代表初始給定 500 種權重組合；投資組合資產總數 (K) 為 10、20、30、40 與 50 檔，每季都只能持有該資產數量，如果原本預測出來高息低波個股群之總數未滿設定數量，投資組合資產總數參數 K 會自動設為預測出來高息低波個股群之總數；單一資產最大權重 (Maxweight) 設為 1；單一資產最小權重 (Minweight) 對應不同資產數量限制設為 0.01、0.005、0.0033、0.0025 與 0.002，以滿足所有最小權重佔比為 10%。

本研究實證過程中，皆假設不考慮滑價及手續費；另外，文中「正確答案」定義為下一季真正屬於高息低波 30%之個股，作為歷史因子投資、機器學習預測方法之比較基準，觀察正確選到未來為高息低波個股所建構之投資組合，其績效表現與兩種預測方法的差異。

第二節 歷史因子投資與機器學習預測績效

一、預測能力比較

在不同資料集之下，歷史因子投資與機器學習分別預測高息低波、非高息低波個股，並由前述四種指標 Accuracy、Precision、Recall 與 F1-score 評斷兩種方法之預測能力，並分為不同股息波動門檻之預測能力，包括 20、25、30、35%，預測結果如表 2，此為 2004 年第二季到 2019 年第四季 (測試集)，共 63 季預測之平均數值。

表2：歷史因子投資與機器學習預測比較

評估指標	預測方法	20%	25%	30%	35%
Accuracy	歷史因子投資	0.9548	0.9407	0.9288	0.8974
	機器學習	0.9559	0.9425	0.9316	0.9065
Precision	歷史因子投資	0.8623	0.8698	0.8791	0.8746
	機器學習	0.8670	0.8726	0.8794	0.8766
Recall	歷史因子投資	0.8485	0.8552	0.8627	0.8476
	機器學習	0.8530	0.8658	0.8789	0.8792
F1-score	歷史因子投資	0.8541	0.8614	0.8701	0.8544
	機器學習	0.8580	0.8674	0.8775	0.8735

在表 2 中，Accuracy 看來歷史因子投資與機器學習兩種預測方式之 Accuracy 皆達 0.89 以上，股息波動門檻則以前 20% 高息低波分類之預測能力較佳，然而，本實驗之分類為前 30% 高股息與低波動個股，易產生分類樣本不平衡的問題，因此觀察其他三種指標表現。以 Precision 來看，不管何種股息波動門檻或預測方法皆高於 0.86，代表使用歷史因子投資與機器學習方法所預測出下一季為高息低波的個股，實際上真的為高息低波，具有相當程度的預測能力，且在前 30% 股息波動門檻下有較佳的預測能力。而以 Recall 來評估，平均高於 0.84 以上，且機器學習預測下，前 35% 股息波動門檻有較佳的預測能力，歷史因子投資則為前 30%。最後，以綜合 Precision 與 Recall 的 F1-score 指標來看，兩種預測方法與不同股息波動門檻之下，皆高於 0.85，其中又以 30% 門檻為最佳。綜合 20、25、

30 與 35%股息波動之預測績效，發現前 30%之高息低波預測能力普遍高於其他門檻，原因在於以 35%門檻分類下，高股息門檻在 2008 年金融海嘯後為 0%，由於過少公司發放股息，導致股息門檻無分類與預測意義，因此本實證研究繼續以前 30%作為後續建構股息波動投資組合之門檻。

本研究進行預測高息低波個股的目的是在於，期望能選出真正為高殖利率與低波動度的股票，也就是預測為高息低波中真正為高息低波的比例，因此主要以 Precision 為評斷預測能力之指標。以 Precision 評估機器學習與歷史因子方法的預測能力，結果顯示機器學習模型平均 63 季的 Precision 皆優於歷史因子投資法，但兩者差異並不大，且皆高於 0.86 以上，顯示不論是以過去平均股息波動資料做為未來具有相同特性的方法，還是以財報資料放進機器學習 XGBoost 模型中訓練，都擁有良好的預測效果；除了 Precision 外，四種評估指標皆顯示機器學習預測能力較佳。歷史因子方法利用過去平均一年股息波動為前 30%高息低波的个股，作為預測未來也是高息低波，由於股票殖利率與波動度的特性通常存在慣性，高息低波的公司下一季很有機會繼續維持為高息低波，也呼應這些公司確實為低波動，股息波動特性不易有過大的變動。而在機器學習預測法之下，運用大量財報資訊作為預測股息波動的學習標準，結果同樣為預測效果良好，且每個股息波動門檻下，皆以機器學習預測表現都比歷史因子更佳，顯示在因子投資中加入機器學習能夠有效幫助投資人篩選出更接近該因子特性的股票。

二、機器學習之重要特徵變數

如前所述，機器學習預測法為使用財報資訊來學習股息波動特性並進行預測與分類，且預測能力優於歷史因子投資法，本文將進一步探討大量財報資訊中的重要特徵變數，找出對於預測股息波動特性貢獻較大的變數，以滾動 63 季中，重要特徵變數前十名出現次數，如圖 6 所示。

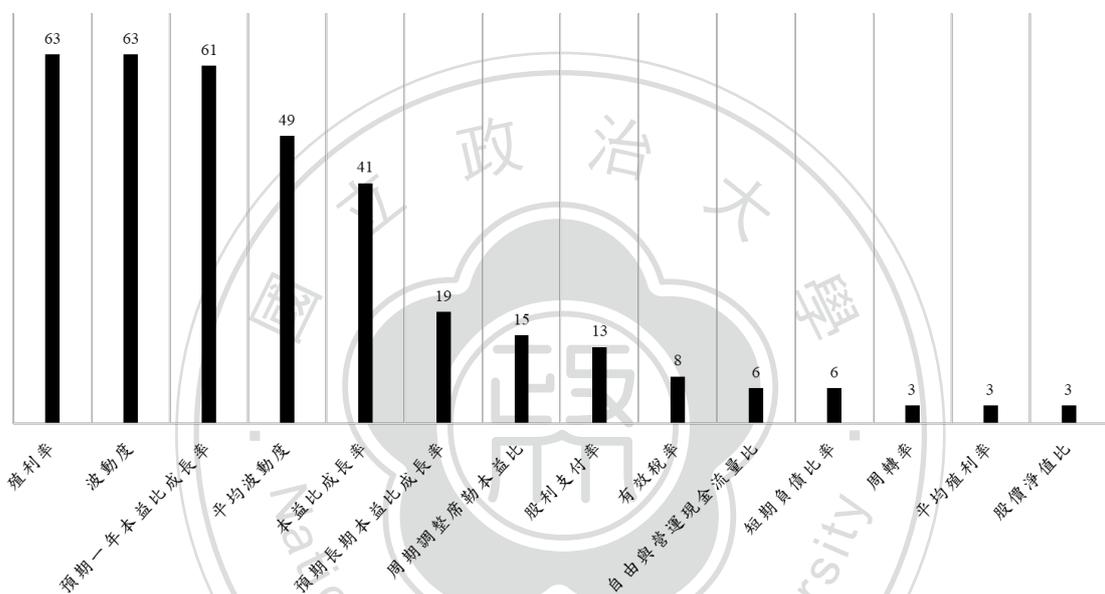


圖6：全體資料集下重要特徵變數前十名出現次數

觀察全體資料集之下，採機器學習預測股息波動特性，其重要特徵變數前十名在 63 次預測中的出現次數，出現最多次的前五個重要特徵變數為：(1) 股票殖利率；(2) 波動度；(3) 預期未來一年的本益比成長率；(4) 過去四季平均波動度；(5) 過去本益比成長率。顯示過去殖利率與波動度對於預測未來的股息波動特性極為重要，此結果也印證歷史因子投資法有良好的預測能力。另外，本益比為股價與每股盈餘相除，不同產業會有不同的合理範圍，又股息波動特性也會受到產

業的影響，例如金融業整體普遍為高股息與低波動，故本益比可能間接因為產業的關係而為重要影響變數。其餘重要影響變數包含預期未來長期本益比成長率、周期調整席勒本益比、股利支付率、有效稅率、自由現金流量與營運現金流量比、短期負債比率、周轉率、過去四季平均殖利率、股價淨值比，除了股利與本益比外，主要為短期流動性相關之比率對於預測股息波動特性有影響。

第三節 歷史因子投資與機器學習投資組合報酬績效

經歷史因子投資與機器學習預測績效比較後發現，機器學習在不同股息波動門檻下，不論以何種預測績效評估指標來看，其預測能力皆優於歷史因子投資法。另外，觀察股息波動之敘述統計，如表 3 為前 20、25、30 與 35%股息波動平均個股數量，前 30%真實高息低波個股數量為 605 檔，歷史因子投資預測高息低波個股數量為 573 檔，機器學習預測高息低波個股數量為 609 檔，顯示預測能力較佳的機器學習法，其股息波動平均個股數量也較接近正確答案。

表3：股息波動平均個股數量

預測方法	20%	25%	30%	35%
正確答案	309	448	605	860
歷史因子投資	297	428	573	732
機器學習	300	444	609	880

接著，本研究進一步以 Markowitz 方法建構不同預測方法與正確答案的投資組合，觀察預測高息低波個股能力越高，其投資組合績效是否也越接近真實高息低波個股所建構的投資組合。在不同預測方法下，預測前 20、25、30 與 35%股息波動股票，並以 Markowitz 建構投資組合，其投資組合報酬績效如表 4 與累積報酬圖 7、8、9 與 10。

表4：股息波動Markowitz投資組合績效：2004Q1-2019Q4

股息波動門檻	預測方法	年化報酬%	年化標準差%	Sharpe ratio	Sortino ratio
20%	正確答案	8.15%	12.49%	0.8966	1.1180
	歷史因子投資	8.37%	13.37%	0.8732	1.0225
	機器學習	8.15%	12.92%	0.8671	0.9751
25%	正確答案	8.70%	13.23%	0.9374	1.0854
	歷史因子投資	8.46%	14.12%	0.8425	0.9749
	機器學習	8.38%	13.27%	0.8827	0.9946
30%	正確答案	8.77%	13.69%	0.9176	1.1383
	歷史因子投資	6.84%	14.87%	0.5621	0.6452
	機器學習	8.05%	14.22%	0.7733	0.8985
35%	正確答案	9.99%	14.32%	1.0694	1.2262
	歷史因子投資	7.38%	15.23%	0.6253	0.6859
	機器學習	8.56%	15.63%	0.7759	0.8869
SPY		7.78%	14.08%	0.7386	0.8202

註：SPY 為 SPDR 標普 500 指數 ETF (SPDR S&P 500 ETF Trust, SPY)



圖7：前20%股息波動Markowitz投資組合之累積報酬：2004Q1-2019Q4

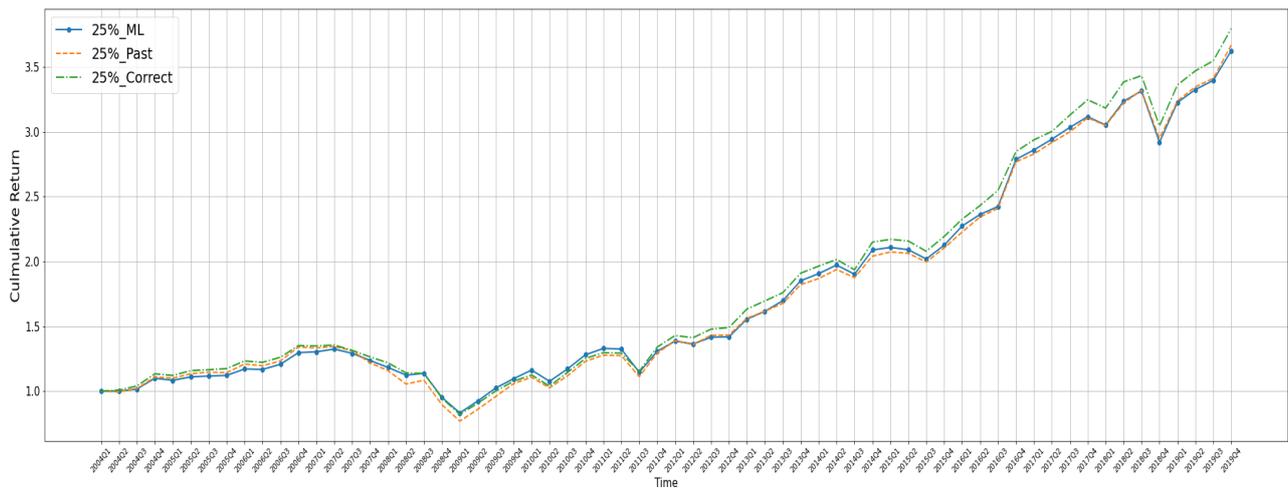


圖8：前25%股息波動Markowitz投資組合之累積報酬：2004Q1-2019Q4

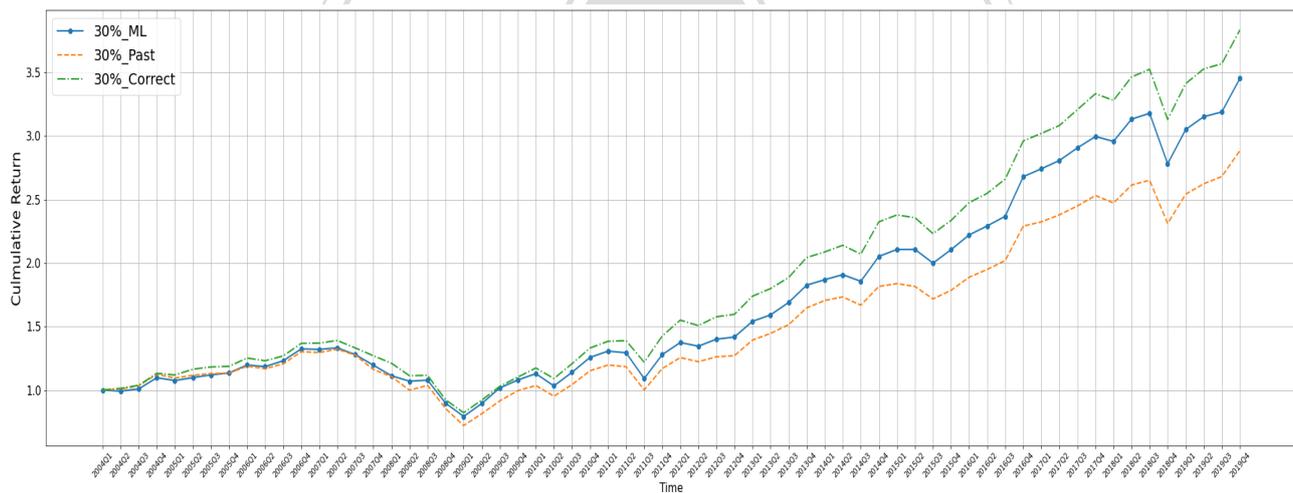


圖9：前30%股息波動Markowitz投資組合之累積報酬：2004Q1-2019Q4

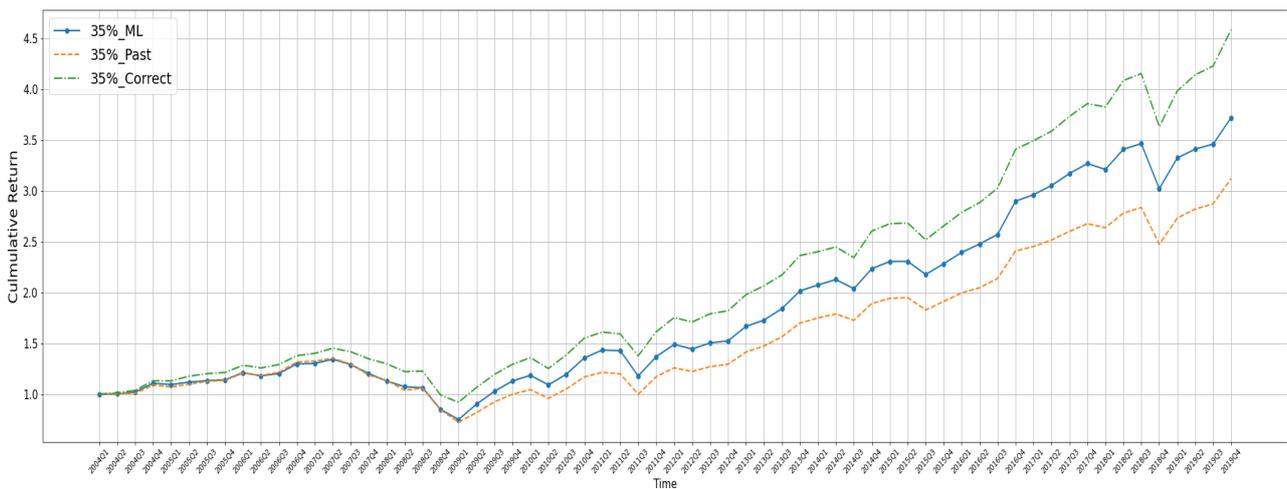


圖10：前35%股息波動Markowitz投資組合之累積報酬：2004Q1-2019Q4

觀察表 4 及圖 7、8、9 與 10 不同股息波動門檻下之 Markowitz 投資組合績效，不論是以 20、25、30 與 35%股息波動門檻，皆為機器學習預測法所建構之 Markowitz 投資組合報酬績效，較歷史因子投資法接近正確答案，也就是真實股息波動投資組合。顯示預測能力較佳的機器學習，其 Markowitz 投資組合報酬績效也會與正確答案較相近，且股息波動門檻越寬鬆，兩者差異越明顯。

另外，不論是以前 20、25、30 或 35%股息波動個股所建構之 Markowitz 正確答案投資組合，績效表現皆優於 SPDR 標普 500 指數 ETF (SPDR S&P 500 ETF Trust, SPY)，以綜合報酬與風險的 Sharpe ratio 與 Sortino ratio 評估指標來看，正確前 20%股息波動投資組合 Sharpe ratio 為 0.8966、Sortino ratio 為 1.1180，正確前 25%股息波動投資組合 Sharpe ratio 為 0.9374、Sortino ratio 為 1.0854，正確前 30%股息波動投資組合 Sharpe ratio 為 0.9176、Sortino ratio 為 1.1383，正確前 35%股息波動投資組合 Sharpe ratio 為 1.0694、Sortino ratio 為 1.2262，顯示較高息與低波的股票投資組合，績效表現確實優於標普 500，此結果也呼應 Blume (1980) 實證顯示股票殖利率與調整後報酬有顯著正向關係，以及 Haugen & Baker (1991)

發現低波動股票投資組合表現優於市場加權投資組合 Wilshire 5000。然而，現實中無法預知未來真實為高息低波的個股，因此才需要以機器學習或歷史因子投資法預測股息波動特性，期望建構出更接近正確的高息低波投資組合，本實證研究發現機器學習預測能力優於歷史因子投資法，且投資組合報酬績效確實也較貼近正確答案，故使用機器學習方法繼續進行實證分析，並以預測能力最高且較有分類意義的前 30%作為門檻值。

第四節 Markowitz 與粒子群最佳配置投資組合報酬績效

本小節進一步加入粒子群最佳配置方法，在機器學習預測前 30%高息低波個股下，比較 Markowitz 與 PSO 兩種不同配置方法所建構之投資組合報酬績效，在實務考量下，PSO 建構限制資產數量 10、20、30、40 與 50 檔的投資組合，期望以限制資產數量的粒子群最佳配置即可達到 Markowitz 全樣本投資組合表現，報酬績效比較如表 5、累積報酬比較如圖 11。

表5：不同配置方法之股息波動投資組合績效：2004Q1-2019Q4

配置方法	個股數量	年化報酬%	年化標準差%	Sharpe ratio	Sortino ratio
Markowitz	609	8.05%	14.22%	0.7733	0.8985
PSO	10	8.86%	18.39%	0.7958	1.1747
	20	5.66%	14.56%	0.3992	0.5194
	30	4.93%	16.04%	0.2654	0.3521
	40	5.01%	14.33%	0.3093	0.4082
	50	3.91%	15.23%	0.1396	0.1803

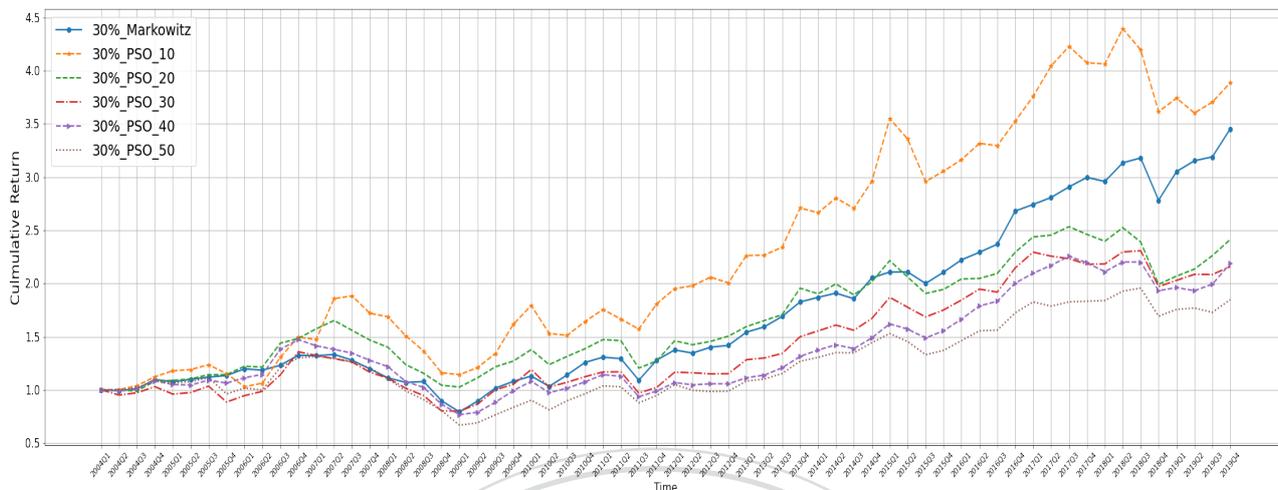


圖11：不同配置方法之股息波動投資組合累積報酬：2004Q1-2019Q4

實證結果發現以 PSO 建構 10 檔股票投資組合之報酬績效即可達到 Markowitz 全樣本 609 檔投資組合表現，PSO 之 10 檔高息低波投資組合年化報酬為 8.86%，為最接近 Markowitz 方法的 8.05%，而年化標準差則為 18.39% 高於 Markowitz 全樣本的 14.22%，Sharpe ratio 為 0.7958、Sortino ratio 為 1.1747，同樣高於 Markowitz 全樣本的 0.7733、0.8985，此結果顯示以限制資產數量為 10 檔之下，便可達到全樣本的投資組合績效，甚至較 Markowitz 全樣本投資組合表現優異，原因可能為 PSO 配置方法不只加入了資產數量的限制 (K)，同時也加入「最小權重限制」。在 Markowitz 配置方法下有權重不平衡的問題，Michaud (1989) 指出在 Markowitz 投資組合最佳化會忽略財務與現實上的限制，單一資產配置權重過度集中，使投資組合的風險提高；Avramov & Zhou (2010) 也指出在實務上無法做空時，Markowitz 最佳化投資組合更容易產生資產過度集中問題，導致權重配置不平衡。由於 Markowitz 投資組合可能存在單一資產權重過度集中的問題，使其績效表現並不一定會優於資產數量較少的 PSO 投資組合。因此，利用粒子

群最佳配置之高息低波投資組合，為有效達到 Markowitz 全樣本投資組合表現，顯示 PSO 為實務上有效取代 Markowitz 的配置方法。

第五節 ESG 與股息波動投資組合報酬績效

本小節以前面實證結果為基礎，再加入 ESG 資料集建構高息低波投資組合，主要針對全體與 ESG 資料集，使用機器學習方法預測前 30% 高息低波個股，並以粒子群最佳配置高息低波投資組合，分析全體資料集與 ESG 資料集之股息波動投資組合報酬績效表現差異，最後再與 SPY 以及市面上高息低波 ETF 的 SPHD 進行比較，並進一步分析 ESG 結合高息低波投資組合內之公司組成。本小節分為兩部分，第一為 ESG 與股息波動因子；第二為 ESG 與股息波動投資組合公司；分別比較 ESG 因子效果與分析其組成公司。

一、ESG 與股息波動因子

首先，分別針對全體與 ESG 資料集，以機器學習預測前 30% 股息波動個股，並用 PSO 建構限制資產數量為 10、20、30、40 與 50 檔股票的投資組合，全體與 ESG 資料集下，機器學習結合粒子群最佳配置之投資組合報酬績效表現如表 6，以及累積報酬圖 12、13、14、15 與 16，分別代表不同資料集之 10、20、30、40 與 50 檔股票投資組合累積報酬，報酬計算期間為 2004 年第一季至 2019 年第四季。

表6：不同資料集之股息波動投資組合績效：2004Q1-2019Q4

資料集	個股數量	年化報酬%	年化標準差%	Sharpe ratio	Sortino ratio
全體	10	8.86%	18.39%	0.7958	1.1747
	20	5.66%	14.56%	0.3992	0.5194
	30	4.93%	16.04%	0.2654	0.3521
	40	5.01%	14.33%	0.3093	0.4082
	50	3.91%	15.23%	0.1396	0.1803
ESG	10	8.98%	15.46%	0.8441	1.3508
	20	11.05%	14.70%	1.2054	1.5953
	30	9.34%	13.42%	1.0305	1.4260
	40	8.91%	12.20%	1.0549	1.3454
	50	9.07%	12.61%	1.0486	1.3878
SPY	500	7.78%	14.08%	0.7386	0.8202



圖12：不同資料集之10檔股息波動投資組合累積報酬：2004Q1-2019Q4

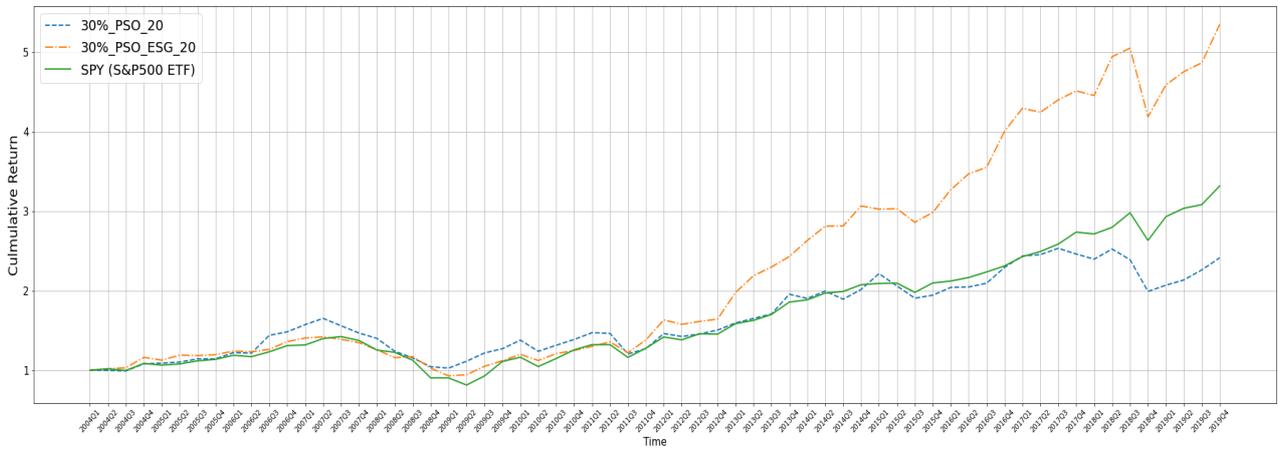


圖13：不同資料集之20檔股息波動投資組合累積報酬：2004Q1-2019Q4

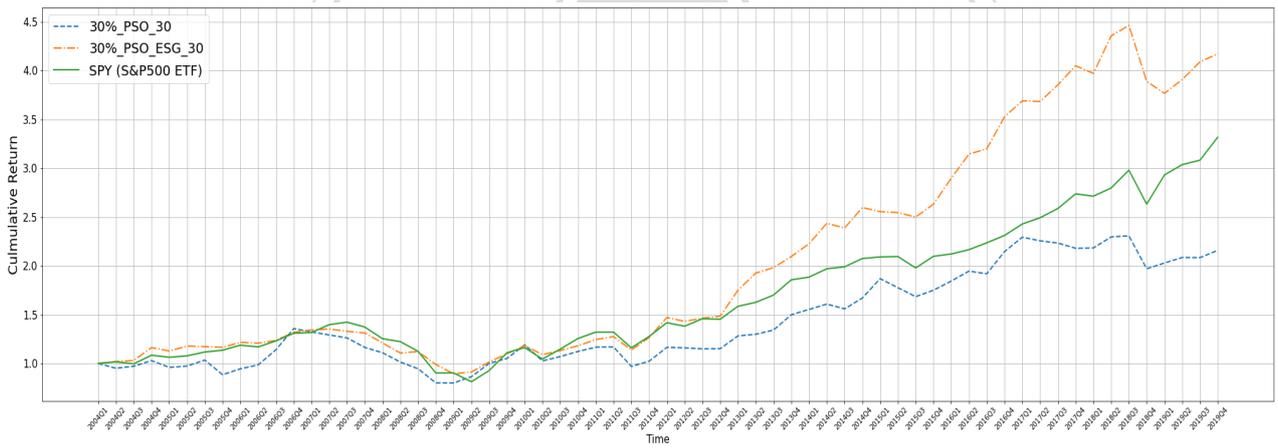


圖14：不同資料集之30檔股息波動投資組合累積報酬：2004Q1-2019Q4

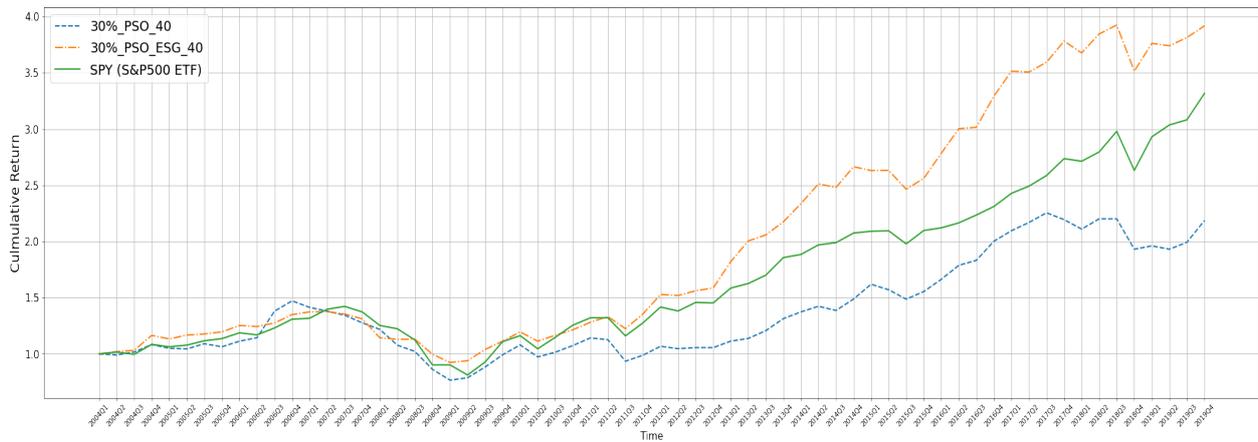


圖15：不同資料集之40檔股息波動投資組合累積報酬：2004Q1-2019Q4



圖16：不同資料集之50檔股息波動投資組合累積報酬：2004Q1-2019Q4

觀察表 6 不同資料集之投資組合績效發現，不論 10、20、30、40 或 50 檔股票投資組合，ESG 資料集皆優於全體資料集，在前一小節實證結果顯示全體資料集下限制資產數量為 10 檔的投資組合表現最佳，針對 ESG 資料集同樣建構 10 檔投資組合，其年化報酬為 8.98%、年化標準差為 15.46%、Sharpe ratio 為 0.8441、Sortino ratio 為 1.3508，而全體資料集之 10 檔投資組合年化報酬為 8.86%、年化

標準差為 18.39%、Sharpe ratio 為 0.7958、Sortino ratio 為 1.1747，ESG 投資組合整體績效表現較全體資料集高息低波投資組合佳。另外，以 20、30、40 與 50 檔資產數量來看，也擁有同樣的結果，顯示從有做 ESG 公司的資料集中建構 30% 高息低波投資組合，明顯比用全體公司資料集來建構高息低波投資組合表現優異，加入 ESG 因子所建構之股息波動投資組合，能有效提升報酬、降低風險、提高 Sharpe ratio 與 Sortino ratio，且優於 SPY 表現。另外，觀察累積報酬圖 12、13、14、15 與 16，在圖 12 中全體與 ESG 資料集 10 檔投資組合累積報酬圖差異不大，而其他 20、30、40 與 50 檔投資組合的累積報酬兩種資料集之差異較大，ESG 資料集之累積報酬明顯較佳，其中全體資料集下 10 檔為最佳資產數量，ESG 資料集下最佳資產數量為 20 檔，代表投資組合最佳資產數量在不同資料集下不一定會相同。

接著本文加入市面上高息低波 ETF 的 SPHD 來比較，SPHD 建構方式為，從 S&P500 中的各產業挑選過去平均 12 個月殖利率排名前 75 名個股，且每個產業不超過 10 檔，再根據過去 252 天股價走勢來選出波動最小的 50 檔股票，並按照殖利率為權重配置標準，以及加入單一產業不能超過 25%、單一股票不能超過 3% 的限制，由於 SPHD 為 2012 年 10 月發行，因此表 7 將本文建構的投資組合績效計算期間與 SPHD 切齊，為 2012 年第三季開始至 2019 年第四季，以利分析這些投資組合與 SPHD 之間的差異。

表7：不同資料集之股息波動投資組合績效：2012Q3-2019Q4

資料集	個股數量	年化報酬%	年化標準差%	Sharpe ratio	Sortino ratio
全體	10	9.17%	15.20%	0.8840	1.3355
	20	7.22%	13.20%	0.6944	0.8450
	30	9.08%	12.12%	1.0903	1.3001
	40	10.56%	10.78%	1.5330	1.9387
	50	9.08%	11.23%	1.1782	1.1621
ESG	10	13.78%	16.40%	1.4619	2.1538
	20	17.99%	14.86%	2.2943	1.9932
	30	15.54%	12.51%	2.2490	2.8774
	40	13.53%	11.13%	2.0975	2.6645
	50	14.15%	11.24%	2.2071	2.8355
SPY	500	12.01%	9.50%	2.0870	1.6262
SPHD	50	10.99%	9.05%	1.9321	3.8960

由表 7 比較 ESG 結合股息波動投資組合與 SPHD 績效，ESG 結合股息波動投資組合普遍擁有較高的年化報酬率，但其波動度在 2012 年第三季以後相對較大，增加資產數量才能降低 ESG 投資組合的風險，以 Sharpe ratio 來說，20、30、40 與 50 檔 ESG 高息低波投資組合表現優於 SPHD，Sortino ratio 則為 SPHD 較佳，整體來說，在 2012 年第三季至 2019 年第四季期間，ESG 結合股息波動投資組合確實有較全體資料集、SPY 與 SPHD 高的報酬，且整體績效優於不分類的全體資料集與 SPY，惟在這段期間的下方風險較 SPHD 大。

二、ESG 與股息波動投資組合之組成分析

接著，本文將進一步探討 ESG 結合高息低波投資組合績效表現較佳之影響因素，以績效表現最佳的 20 檔為例，如表 8 與 9 所示。

表8：ESG股息波動投資組合之公司出現次數前十名

公司	產業	次數
麥當勞 (MCD)	餐飲	18
威瑞森通訊 (VZ)	電信	17
高樂氏 (CLX)	家用/清潔用品	16
寶僑 (PG)	消費品	16
大眾服務企業集團 (PEG)	能源	16
輝瑞 (PFE)	生物製藥	14
第一能源 (FE)	能源	13
菲利普莫里斯 (PM)	煙草、食品	13
奧馳亞集團 (MO)	煙草	13
杜克能源 (DUK)	能源	13
安特吉 (ETR)	能源	13
南方 (SO)	能源	13

在 63 季中平均被選到次數最多的前十名個股，依序為麥當勞 (McDonalds, MCD)、威瑞森通訊 (Verizon Communications, VZ)、高樂氏 (Clorox Company, CLX)、寶僑 (Procter & Gamble, PG)、大眾服務企業集團 (Public Service Enterprise Group Inc., PEG)、輝瑞 (Pfizer, PFE)、第一能源 (FirstEnergy, FE)、菲利普莫里斯 (Philip Morris International Inc., PM)、奧馳亞集團 (Altria Group, MO)、杜克能源 (Duke Energy, DUK)、安特吉 (Entergy, ETR)、南方 (Southern, SO)，其中 FE、PM、MO、DUK、ETR 與 SO 並列第十名，其中產業涵蓋能源、菸草、食品、家用清潔用品、消費品、食品與電信產業，能源佔比較高，主要為電力與天然氣相關業務。

表9：ESG股息波動投資組合之公司加權比重前十名

公司	產業	總權重
必治妥施貴寶 (BMY)	生物製藥	1.67
高樂氏 (CLX)	家用/清潔用品	1.08
第一能源 (FE)	能源	1.07
艾迪生國際 (EIX)	能源	10.2
麥當勞 (MCD)	餐飲	0.99
好時 (HSY)	食品	0.98
默克藥廠 (MRK)	生物製藥	0.97
菲利普莫里斯 (PM)	煙草、食品	0.96
CMS 能源 (CMS)	能源	0.91
KAR Auction Services (KAR)	汽車	0.87

另外，在 63 季中被選到投資組合總權重最大的前十檔個股如表 9 所示，依序為必治妥施貴寶 (BMY)、高樂氏 (Clorox Company, CLX)、第一能源 (FirstEnergy, FE)、艾迪生國際 (Edison International, EIX)、麥當勞 (McDonalds, MCD)、好時 (Hershey, HSY)、默克藥廠 (Merck, MRK)、菲利普莫里斯 (Philip Morris International Inc., PM)、CMS 能源 (CMS Energy, CMS)、KAR Auction Services (KAR)，其中產業涵蓋能源、生物製藥、菸草、食品、家用清潔用品、餐飲與汽車。

無論是以平均選到個股次數或是總權重前十名之公司，這些公司大多都在標普 500、羅素 1000、羅素 3000、道瓊工業指數、與精選高股利指數 (Dow Jones U.S. Select Dividend Index) 中，其中 ESG 高息低波公司主要包括能源、生物製藥、煙草、家用清潔用品及食品等，尤其在能源產業更為明顯，包括天然氣與電力公司。這也印證了 Ashwin Kumar et al. (2016) 提出 ESG 因子對報酬與風險的

影響會根據不同行業而表現不同，ESG 的實踐能夠幫助公司降低風險，尤其是能源、食品以及醫療保健行業更為明顯，能源產業中，大多數公司並未以永續、友善環境及環保的形式經營，因此有做 ESG 的能源公司會有較佳的表現；本研究也發現除了能源、生物製藥及食品業外，煙草業中有做 ESG 的公司也相對有較佳的表現，在考慮 ESG 實踐所帶來的隱藏價值後，確實能提供較高的報酬、Sharpe ratio 與 Sortino ratio。



第五章 結論與未來展望

第一節 研究結論

一、 歷史因子投資與機器學習預測結果

本文使用兩種不同方法來預測股息波動特性，其一為以歷史資料作為未來有相同股息波動特性的歷史因子投資法，另一種方法為以機器學習 XGBoost 模型訓練財報資料，對未來做股息波動特性的預測。實證結果顯示，兩種預測方法皆有相當程度的預測能力，且機器學習之 Precision 優於歷史因子投資法。在機器學習模型中，重要特徵變數為三大變數，包括過去殖利率與波動度、本益比，顯示股息與波動因子易有慣性，延續上一季的股息波動特性，而本益比則可能受產業所影響而間接成為重要特徵變數，影響股息波動特性的預測。

在面對高息低波分類的問題上，歷史因子投資的篩選與預測方式確實能有效預測未來具有相同因子特性，原因在於高息低波的因子特性容易有慣性，也因此目前市面上因子投資的基金或 ETF 主要仍以歷史因子投資的方式篩選個股，然而，機器學習也同樣具有高預測能力且優於歷史因子投資預測法，顯示引入機器學習能夠提供投資人更正確篩選出真實股息波動因子特性。

二、 歷史因子投資與機器學習投資組合績效結果

預測結果顯示機器學習擁有較良好的預測股息波動能力，並以 Markowitz 配置方法建構不同預測方法下的股息波動投資組合，實證結果發現預測能力較佳的

機器學習法，所建構之 Markowitz 投資組合報酬績效也越接近正確股息波動投資組合績效，再度驗證當投資人欲建構高息低波投資組合時，機器學習為較有效且預測更正確的方法。

三、 Markowitz 與粒子群最佳配置投資組合績效結果

因應 Markowitz 配置方法有實務上資產數量及配置權重不平衡的問題，高息低波個股經由 PSO 求解基數限制下的 Markowitz 最佳化問題，得到限制資產數量的投資組合，並與 Markowitz 全樣本投資組合進行比較，實證顯示以限制資產數量的 PSO 所建構之最佳化投資組合，其績效能夠達到 Markowitz 全樣本投資組合，實務考量下，PSO 確實為有效取代 Markowitz 的投資組合配置方法。

四、 ESG 與股息波動投資組合績效結果

比較全體與 ESG 資料集下，以機器學習預測高息低波股票、再由 PSO 建構最佳化投資組合，實證結果為 ESG 與高息低波對投資組合的績效表現有正向影響關係，顯示有做 ESG 的公司為更加穩定獲利的公司，進一步分析投資組合平均篩選與配置最多的公司發現，大多為有做 ESG 的能源產業，顯示 ESG 對投資組合報酬的影響根據產業而不同，且 ESG 的實踐能有效提升投資組合績效表現。

綜合以上結論，機器學習下建構 ESG 股息波動因子投資組合能夠提供一般投資人建構實務上可行且更穩定獲利的投資組合，同時也提供金融業者作為基金或 ETF 商品發行的參考，以實踐永續投資、金融科技與普惠金融。

第二節 未來展望

一、考量其他因子的影響

ESG 與股息波動因子皆會受到產業所影響，例如有做 ESG 之能源公司表現較好、金融業股價表現普遍呈現低波動，雖然特定產業的 ESG 高息低波投資組合較佳，但仍不宜配置太多股票於同一產業中，在建構投資組合時如能平均分散個股於不同產業中，可避免整體產業下跌風險。另外，ESG 與股息波動因子也會在不同時空背景下受到動能因子所影響，因此未來可以分析動能因子對投資組合報酬績效的影響。

二、預測其他因子特性

由實證結果發現，由於股息波動特性通常存在慣性，過去為高息低波個股，通常下一季也較容易繼續為高息低波個股，因此，以歷史資料為依據的歷史因子投資法仍有相當程度的預測能力，未來可嘗試以機器學習法預測其他非線性特徵的因子，也許會有更顯著的預測能力差異。

參考文獻

1. Ashwin Kumar, N. C., Smith, C., Badis, L., Wang, N., Ambrosy, P., & Tavares, R. (2016). ESG factors and risk-adjusted performance: a new quantitative model. *Journal of Sustainable Finance & Investment*, 6(4), 292-300.
2. Avramov, D., & Zhou, G. (2010). Bayesian portfolio analysis. *Annual Review of Financial Economics*, 2(1), 25-47.
3. Baskin, J. (1989). Dividend policy and the volatility of common stocks. *Journal of Portfolio Management*, 15(3), 19.
4. Blume, M. E. (1980). Stock returns and dividend yields: Some more evidence. *The Review of Economics and Statistics*, 567-577.
5. Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).
6. Deng, G. F., Lin, W. T., & Lo, C. C. (2012). Markowitz-based portfolio selection with cardinality constraints using improved particle swarm optimization. *Expert Systems with Applications*, 39(4), 4558-4566.
7. Fernández, A., & Gómez, S. (2007). Portfolio selection using neural networks. *Computers & Operations Research*, 34(4), 1177-1191.
8. Gombola, M. J., & Liu, F. Y. L. (1993). Dividend yields and stock returns: Evidence of time variation between bull and bear markets. *Financial Review*, 28(3), 303-327.
9. Haugen, R. A., & Baker, N. L. (1991). The efficient market inefficiency of capitalization-weighted stock portfolios. *The Journal of Portfolio Management*, 17(3), 35-40.
10. Kennedy, J., & Eberhart, R. (1995, November). Particle swarm optimization. In *Proceedings of ICNN'95-International Conference on Neural Networks* (Vol. 4, pp. 1942-1948). IEEE.
11. Markowitz, H. (1959). *Portfolio Selection: Efficient Diversification of Investments*

(Vol. 16). New York: John Wiley.

12. Michaud, R. O. (1989). The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal*, 45(1), 31-42.
13. Nofsinger, J., & Varma, A. (2014). Socially responsible funds and market crises. *Journal of Banking & Finance*, 48, 180-193.
14. Patel, J., Shah, S., Thakkar, P., & Kotecha, K. (2015). Predicting stock and stock price index movement using trend deterministic data preparation and machine learning techniques. *Expert Systems with Applications*, 42(1), 259-268.
15. Ratnaweera, A., Halgamuge, S. K., & Watson, H. C. (2004). Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. *IEEE Transactions on Evolutionary Computation*, 8(3), 240-255.
16. Renneboog, L., Ter Horst, J., & Zhang, C. (2008). The price of ethics and stakeholder governance: The performance of socially responsible mutual funds. *Journal of Corporate Finance*, 14(3), 302-322.
17. Shi, Y., & Eberhart, R. (1998, May). A modified particle swarm optimizer. In *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence (Cat. No. 98TH8360)* (pp. 69-73). IEEE.
18. Verheyden, T., Eccles, R. G., & Feiner, A. (2016). ESG for all? The impact of ESG screening on return, risk, and diversification. *Journal of Applied Corporate Finance*, 28(2), 47-55.
19. Zhongbin, Z., & Jinwu, F. (2019, December). Empirical research about quantitative stock picking based on machine learning. In *2019 International Conference on Economic Management and Cultural Industry (ICEMCI 2019)* (pp. 138-141). Atlantis Press.

附錄：機器學習模型之輸入變數

- Valuation Ratios (13)
 1. Dividend Payout Ratio (dpr)
 2. Trailing P/E to Growth (PEG) ratio (PEG_trailing)
 3. Book/Market (bm)
 4. Shillers Cyclically Adjusted P/E Ratio (capei)
 5. Dividend Yield (divyield)
 6. Enterprise Value Multiple (evm)
 7. Price/Cash flow (pcf)
 8. P/E (Diluted, Excl. EI) (pe_exi)
 9. P/E (Diluted, Incl. EI) (pe_inc)
 10. Price/Operating Earnings (Basic, Excl. EI) (pe_op_basic)
 11. Price/Operating Earnings (Diluted, Excl. EI) (pe_op_dil)
 12. Price/Sales (ps)
 13. Price/Book (ptb)
- Profitability Ratios (15)
 14. Effective Tax Rate (efftax)
 15. Gross Profit/Total Assets (GProf)
 16. After-tax Return on Average Common Equity (aftret_eq)
 17. After-tax Return on Total Stockholders Equity (aftret_equity)
 18. After-tax Return on Invested Capital (aftret_invcapx)
 19. Gross Profit Margin (gpm)
 20. Net Profit Margin (npm)

21. Operating Profit Margin After Depreciation (opmad)
22. Operating Profit Margin Before Depreciation (opmbd)
23. Pre-tax Return on Total Earning Assets (pretret_earnat)
24. Pre-tax return on Net Operating Assets (pretret_noa)
25. Pre-tax Profit Margin (ptpm)
26. Return on Assets (roa)
27. Return on Capital Employed (roce)
28. Return on Equity (roe)
- Capitalization Ratios (4)
 29. Capitalization Ratio (capital_ratio)
 30. Common Equity/Invested Capital (equity_invcap)
 31. Long-term Debt/Invested Capital (debt_invcap)
 32. Total Debt/Invested Capital (totdebt_invcap)
- Financial Soundness (16)
 33. Inventory/Current Assets (inv_t_act)
 34. Receivables/Current Assets (rec_t_act)
 35. Free Cash Flow/Operating Cash Flow (fcf_ocf)
 36. Operating CF/Current Liabilities (ocf_lct)
 37. Cash Flow/Total Debt (cash_debt)
 38. Cash Balance/Total Liabilities (cash_lt)
 39. Cash Flow Margin (cfm)
 40. Short-Term Debt/Total Debt (short_debt)
 41. Profit Before Depreciation/Current Liabilities (profit_lct)
 42. Current Liabilities/Total Liabilities (curr_debt)

43. Total Debt/EBITDA (debt_ebitda)
 44. Long-term Debt/Book Equity (dltt_be)
 45. Interest/Average Long-term Debt (int_debt)
 46. Interest/Average Total Debt (int_totdebt)
 47. Long-term Debt/Total Liabilities (lt_debt)
 48. Total Liabilities/Total Tangible Assets (lt_ppent)
- Solvency Ratios (6)
 49. Total Debt/Equity (de_ratio)
 50. Total Debt/Total Assets (Total debt as a fraction of total assets) (debt_assets)
 51. Total Debt/Total Assets (Total liabilities as a fraction of total assets) (debt_at)
 52. Total Debt/Capital (debt_capital)
 53. After-tax Interest Coverage (intcov)
 54. Interest Coverage Ratio (intcov_ratio)
 - Liquidity Ratios (4)
 55. Cash Conversion Cycle (Days) (cash_conversion)
 56. Cash Ratio (cash_ratio)
 57. Current Ratio (curr_ratio)
 58. Quick Ratio (Acid Test) (quick_ratio)
 - Efficiency Ratios (7)
 59. Asset Turnover (at_turn)
 60. Inventory Turnover (inv_turn)
 61. Payables Turnover (pay_turn)
 62. Receivables Turnover (rect_turn)
 63. Sales/Stockholders Equity (sale_equity)

- 64. Sales/Invested Capital (sale_invcap)
- 65. Sales/Working Capital (sale_nwc)
- Other Ratios (4)
 - 66. Accruals/Average Assets (Accrual)
 - 67. Research and Development/Sales (RD_SALE)
 - 68. Advertising Expenses/Sales (adv_sale)
 - 69. Labor Expenses/Sales (staff_sale)
- Valuation Ratios (Require IBES subscription)(2)
 - 70. Forward P/E to 1-year Growth (PEG) ratio (PEG_1yrforward)
 - 71. Forward P/E to Long-term Growth (PEG) ratio (PEG_ltgforward)
- Others (4)
 - 72. Volatility (Vol)
 - 73. Average Volatility (Ave_vol)
 - 74. Average Dividend Yield (Ave_div)
 - 75. Last Classification (Last_Y)