



分析能源類股股價影響因子與預測漲跌 Explore Stock Price by Applying Fuzzy Clustering Method and Support Vector Machine

邱登裕¹
Deng-Yu Chiu

謝素真^{2*}
Su-Chen Hsieh

李玫郁³
Mei-Yu Li

蔡松偉⁴
Sung-Wei Tsai

摘要

股票是台灣投資人熱愛的投資工具之一，根據台灣證券交易所統計，110 年 12 月股票交易市場週轉率為 11.16 次，高於日本(8.88 次)與新加坡(3.17 次)等亞洲國家。由此可發現台灣投資人偏向短期投資型態，因此如何在找出影響股價因子並有準確的預測股價趨勢是投資人重要的課題。本研究目的在於找出股價影響因子組合，依據資料特徵值進行資料分群，探討各影響因子與股價變動關係。本研究資料為 1323 筆能源類別股資料，時間為 2014 年 1 月至 2015 年 9 月，公司樣本數 63 家，使用 Cart 找出影響股價因子組合，將訓練資料使用 Fuzzy C-means 依據特徵值進行分群，根據不同群聚資料建置不同持向量機預測模型後，將測試資料計算所屬群聚，導入各自所對應支持向量機預測器後並收集預測結果。支持向量機預測結果與實際輸出結果相比，本研究模型預測準確率達 71.28%，並根據預測結果進行模擬投資，平均月報酬率為 1.77%，年化報酬率為 21.24%%，優於其他方法。

關鍵詞: 決策樹、模糊分群法、支持向量機、股價趨勢預測。

Abstract

According to the statistics of securities market published by Taiwan Stock Exchange Corporation (TWSE), in Dec.2021 the volume turnover rate traded in TWSE was 11.16, which was much more than Hong Kong and Singapore. Investors in Taiwan tend to adopt short term investment strategy. How to discover factors influencing stock price and further create a model to predict stock prices is the important concern of investors. The study collected data from listed companies of energies related industry in TWSE from January 2014 to September 2015 which accounts to 63 companies and total number of data is 1,323. First of all, the classification and regression tree (CART) method was adopted to filter the most effective factors related to stock prices. After using a combination of factors to identify the impact of stock price, Fuzzy C-means clustering was used to cluster training data based on feature values. Then support vector machine (SVM) was applied to build different prediction models. Compare the results from SVM prediction with actual output, the model shows that 283 of 397 test record transactions are accurately predicted. The prediction

¹中華大學資管系教授 chiuden@chu.edu.tw

²中華大學財管系助理教授 schsieh@g.chu.edu.tw *通訊作者

³明新科大財務金融系助理教授 mylee@must.edu.tw

⁴宏基股份有限公司工程師 m10410030@chu.edu.tw

accuracy rate of this study is 71.28%, the average monthly return rate is 1.77 %, and annual rate of return is 21.24% the method by this study is much better than other methods.

Keywords: Decision Tree, Fuzzy C-means, Support Vector Machine, Stock Price

1. 緒論

根據台灣證券交易所統計，110 年 12 月股票交易市場週轉率為 11.16 次，高於日本(8.88 次)與新加坡(3.17 次)等亞洲國家，由此可知短期投資仍是台灣股市投資人的偏好形態。因此如何從市場資訊中找出值得投資的股票，以期在短期間能夠獲利對投資人來說是一個重要的課題。為了保護投資人的權益，證交所規定各上市櫃公司於每個月的 10 日之前需公告公司上個月的營業收入數據，營收與過去相比，成長率越高並且比同業還多代表公司經營能力越優異，營收增加對公司營運績效來說是一項正面消息，反之亦然；對短期投資的投資人來說每個月公佈的營收資料是投資機構人及投資人很重要的參考指標，也是影響各公司股價波動的一大因素。在影響因子眾多股票市場中，一般投資者往往無法解讀眾多資訊進而獲得投資參考的依據，如公司本身營運情形，國際原物料價格波動、總體經濟變數、法人買賣動向等等。但這些變數與指標並不是每個都擁有高度影響性，過多影響因子在趨勢預測上可能造成誤判並在投資者在解讀這些欄位上也有相當的難度，若是能將佔有絕對影響性因子欄位先行挑選出，或許可以高模型預測準率也可以提供投資者簡單明瞭的投資依據。

本研究藉由過去歷史資料探討影響未來股價變動規則，利用公司財務季報表當中的公司營收成長率、與去年同期比營收成長率、負債比率、股東權益報酬率、法人買超數量、與國際原油價格變動、公司股本大小，從上述欄位中找規則進而提高預測股價趨勢，並根據預測結果進行投資選擇，進而計算預測成功率時之投資月報酬率，希望提供機構投資者、投資人及後續對此議題之研究者之參考。

2. 文獻回顧

如何從投資中獲利是大眾所追求的目標，而股票投資相對其他投資工具對投資人而言，是最常用投資項目；因此如何準確的預測股價也是許多研究的主題。Fuzzy C-means 相較於 k-means 的分群，Fuzzy C-means 為不能明顯劃分族群的資料，提供了更彈性的分類方法。這個特性應用在醫學影像的判定上(Alruwaili, Siddiqi, & Javed, 2020; Zhao et al., 2020)。和醫學影像資料相同的是，響股價的因素十分複雜，且充滿雜訊，所以本研究利用 Fuzzy C-means 此特性來針對預測變數進行分群。如何準確的預測股價一直是眾多研究的重心(Schumaker, 2009)、(Kwan, 2014)；而人工智慧方法的應用，對於非線性、多雜訊的股價預測，提供了更多元的方法。(Göçken, Özçalıcı, Boru, & Dosdoğru, 2016)以類神經網路(Artificial Neural Networks, ANN)，(Sun, Guo, Reza Karimi, Ge, & Xiong, 2015)以模糊理論(Fuzzy)結合時間序列分，以現貨價格、收盤價成交量等變數，用來預測股價指數期貨。股價分時變化極大，(Long, Lu, & Cui, 2019)以建立多個過濾層的神經網路(neural network, NN)用股價每分鐘的變化為預測基礎，(Liu, Yeh, & Lee, 2012)則以模糊神經(neuro-fuzzy)模型，以日收盤價為預測基礎。(Bisoi, Dash, & Parida, 2019)以變分模態分解(Variational Mode Decomposition, VMD)模型，以每月股價、(Zhang & Lou, 2021)則以股價相關及簡單技術指標，用倒傳遞神經網路討論預測變數不同天數的預測效果。(Dash, Samal, Dash, & Rautray, 2019)以技術指標為預測變數，以理想解相似度偏好法

(Technique for Order of Preference by similarity to Ideal Solution, TOPSIS)預測股價的變動；長短期記憶(Long short-term memory, LSTM)方法也是常用的分析工具，(Li, Bu, Li, & Wu, 2020)以此法以財經網站上之訊息建立預測指數，以分析短期訊息對當時股價之影響程度，(Jing, Wu, & Wang, 2021)則除了訊息之文本內容外，佐以技術分指標，做為股價的預測。

(Chatzis, Siakoulis, Petropoulos, Stavroulakis, & Vlachogiannakis, 2018)將決策樹(decision tree)用於金融市場危機的預測，(Ince, 2014; Tsai, Lin, Yen, & Chen, 2011)以決策樹分群的優點結合人工智慧的方法，得到不錯的效果。而在眾多人工智慧方法中，支持向量機對於未來不確定值的預測是很好的方法(Wang & Pardalos, 2015); (Chen, Ong, Zheng, & Hsu, 2017)以 SVM 預測台北市的房價，(Xu, Jiang, & He, 2015)則用在風險價值(value at risk, VAR)的預測，研究發現(Xie, Luo, & Yu, 2010)以財務變數與經濟變數預測公司破產，SVM 的預測效果優於多變量區別分析方法(Multivariate Discriminant Analysis, MDA)，而(Das, Achary, & Padhy, 2016; Fenghua, Jihong, Zhifang, & Xu, 2014)研究發現在股價預測方面以 SVM 為基礎的預測方法有其優勢。基於以上分析，本研究結合 Fuzzy C-means 具有更彈性的分類方法的優點及利用決策樹可以做良好分類的特性，並結合 SVM 做股價的預測，並與其他方法做較，結果顯示此結合可以得到不錯的預測效果。

股價常用的分析方法有基本面分析及技術分析，這也成了研究中常用的分析變數。(Yu, Chen, & Zhang, 2014)以財報的財務比率為基礎，(Sadorsky, 2021; Zhang & Lou, 2021)以技術指標，(Hossain, Baten, & Mukta, 2021; Illa, Parvathala, & Sharma, 2021; Tripathy, 2021)則以股價做為預測變數，而近年由於資料探勘技術的廣為應用，(Duarte, Montenegro González, & Cruz, 2021; Hao, Kung, Chang, & Ou, 2021; Li et al., 2020; Suppawong et al., 2021)更利用公開的消息，來預測短期股價的波動及走勢。(Prasad & Seetharaman, 2021)則結合股價、總體經濟變數及網路資訊的做為股價預測的基準。雖然加入多元且大量變數可以更準確的預測股價走勢，然而對一般不具財經背景的投資人而言，其參考價值有限。有鑑於台灣股票市場的高週轉率，表示投資人持有股票期間偏向短期投資，而主管機關要求上市櫃公司每月 10 日前需公佈上個月的營收，這是很值得投資人參考的指標。本研究不同於以前的研究，將營收的變數納入預測的變數中，並加入基本面及股票籌碼面的相關變數，結果顯示，營收的資料對短期投資而言，確實是重要影響變數，本研究的結合多種人工智慧方優於其他的預測方法。

3. 研究流程與研究方法

本研究提出一個短期投資建議模型，藉由此模型來判斷股市價格走勢，藉此提供投資者在投資決策上有用的投資建議。首先用決策樹挑選出較具有影響性因子的欄位，經由決策樹篩選並不會有人為主觀的因素參雜，挑選出欄位可以較客觀。將影響力高低以樹狀圖方式呈現，再以 Fuzzy C-means 結合 Support Vector Machine 發展出多重 SVM 預測模型。此模型優勢在於先將資料特性較接近的資料歸類於同一群，再依據不同群聚資料發展出各自適合的 SVM 預測器。其目的在於經由 Fuzzy C-means 分群後將特性相同資料歸類至同一群聚中，避免單一 SVM 預測分類器進行預測時因不同資料特性或離群值造成預測上的誤判與誤差值的產生。

本研究共運用三個人工智慧方法，首先將能源類股月資料收集完全後，利用 Cart 決策樹得出最佳欄位變數組合，其後將影響性因子不高欄位去除，再將資料分

割為測試資料(70%)與訓練資料(30%)。將訓練資料進行模糊分群，分群數目設定 3 至 7 群。以試誤法與觀察資料分佈來判別是否為最佳分群數目，再透過計算不同群聚中心點計算出每筆訓練資料歸屬程度與各所屬群聚。當 Fuzzy C-Means 達到最佳分群數目並收斂後，依據不同群聚特性資料建置各群聚所屬 SVM 分類預測器，藉此建立多重 SVM 預測模型。例如:Fuzzy C-Means 最佳分群數為四個群聚，便會建立四個各自所對應的 SVM 預測模型。根據 SVM 模型預測結果進行投資，當預測模型預測 30 日後股價為上漲時，便買進此檔股票；反之當預測 30 日後股價走勢為跌，便不買進此張股票，最後計算此模擬投資之報酬率，收集每筆測試資料預測結果並計算最終報酬率，並與對照組進行比較。茲將研究流程與研究方法分別詳述如下：

3.1 研究流程

本研究建構一綜合應用的人智慧模型。茲將本研究之流程羅列如下：

1. 從事資料蒐集:資料來源為台灣經濟新報資料庫 (TEJ)，研究資料包含:石油價格變動率、去年同期月營收成長率、月營收成長率、負債比率、資本額、股東權益報酬率、法人買賣超張數以及月營收公布後三十日股票價格漲跌比率。
2. 得出最佳欄位變數組合:利用 Cart 決策樹從步驟 1 當中挑選出最具影響因子欄位組合後，影響性因子不高欄位去除後並將資料分割成測試資料(70%)與訓練資料(30%)。
3. 以 Fuzzy C-means 分群:將訓練資料進行模糊分群，分群數目設定 3 至 7 群。以試誤法與觀察資料分佈來判別是否為最佳分群數目，再透過計算不同群聚中心點計算出每筆訓練資料歸屬程度與各所屬群聚。
4. 訓練多重 SVM 分類預測器:當 Fuzzy C-Means 達到最佳分群數並收斂後，依據不同群聚特性資料建置各群聚所屬 SVM 分類預測器藉此建立多重 SVM 預測模型。例如:Fuzzy C-Means 最佳分群數為四個群聚，便會建立四個各自所對應的 SVM 預測模型。
5. 計算測試資料中心點與歸屬程度:計算每筆測試資料與各群聚中心點距離，藉由之間距離來判斷測試資料所屬群聚，每筆測試資料皆以最短距離為此群聚所歸屬。
6. 分析測試資料預測與結果:計算出每筆測試資料歸屬程度與所屬群聚後，依據不同群聚資料導入各自所屬的 SVM 預測模型進行預測，將不同 SVM 模型預測結果統整後與實際輸出結果比較並計算預測準確率。
7. 進行模擬投資:根據 SVM 模型預測結果進行投資，設當預測模型預測 30 日後股價為漲便買進此檔股票，反之當預測 30 日後股價走勢為跌便不買進此張股票，最後計算此模擬投資之報酬率，收集每筆測試資料預測結果並計算最終報酬率。
8. 對照實驗組結果:為證明此研究提出混和預測模型架構優於其他預測模型，因此與單一羅吉斯回歸預測模型、單一決策樹預測模型與單一類神經網路進行比較，證實本研究架構預測準確率上優於對照組結果。

3.2 研究方法

1. 決策樹:利用決策樹找出影響因子，首先定義資料輸入與輸出欄位；輸入欄位為石油價格變動率(Oil Price)、去年同期營收成長率(Revenue(last))、營收成長率(Revenue(Now))、資產負債比(Debt ratio)、公司股本(Capital stock)、股東權益報酬(Return on equity)、法人買賣數量(Legal)，輸出欄位為營收公布日三十日後股

- 價漲跌，從上述欄位中找出輸入與輸出欄位中較具影響性因子。
2. 模糊分群:用模糊分群法來針對股票證券資料進行分群，藉由資料分群散佈與試誤法來判斷是否為最佳分群數，以輸出類別作為依據，用以觀察被分群後的資料其每一群資料的輸出結果是否相同。利用 Fuzzy C-means 每次分群結果不同特性，探討不同分群數資料分布情形，初始設定的分群數為 3，但由於資料數量的關係，最大分群數設定為 7，而其是否達到最大分群數將以支持向量機的預測成功率作為評斷。
 3. 支持向量機:支持向量機(Support Vector Machine)透過高維度切割方式進行資料分類，但如果資料當中有產生雜訊資料或離群值存在，皆會產生誤差值，因此本研究先以 Fuzzy C-means 將資料特性接近資料分類為同一群，在各建立不同群聚所屬分類預測模型，目的在於避免不同資料特性造成誤差值產生進而提高預測準確率。經由 Fuzzy C-means 演算法分群後，依據不同群聚建立各所屬 SVM 分類預測器，建立多個 SVM 分類預測器。建置完成多個 SVM 分類預測器後，計算每筆測試資料與各群聚中心點距離且計算各自歸屬程度，依據所屬不同群聚導入各自所對應 SVM 分類預測進行預測。

4. 研究結果

本研究資料來源為台灣經濟新報(TEJ)資料庫，研究樣本為台灣證券交易所及櫃台買賣中心掛牌交易之公司，產業類別為汽車工業、航運、油電燃氣與太陽能類股。區間為 2014 年 1 月 1 日至 2015 年 9 月 30 日。資料皆以月為單位，資料期間共 21 個月，總計料樣本數為 1,323 筆。以總樣本數 70% 資料為訓練樣本，以總樣本數 30% 為測試樣本。資料欄位包含:月營收成長率(Revenue)、去年同期月營收成長率(Last Revenue)、負債比(Debt Ratio)、石油價格變動率(Oil Price)、公司股本大小(Capital Stock)、股東權益報酬(Return on equity)、法人買賣超數量(Legal); 輸出欄位為月營收公布後三十日股票價格漲跌

首先以 Cart 決策樹對研究資料欄位進行篩選，經由決策樹篩選出具有較高影響性欄位，因而可以減少訓練模型時過多不要的雜訊與離群值造成模型預測上的誤判與產生誤差值。Cart 篩選欄位後所得輸入欄位與輸出欄位以階層樹狀圖如圖 1 所示:

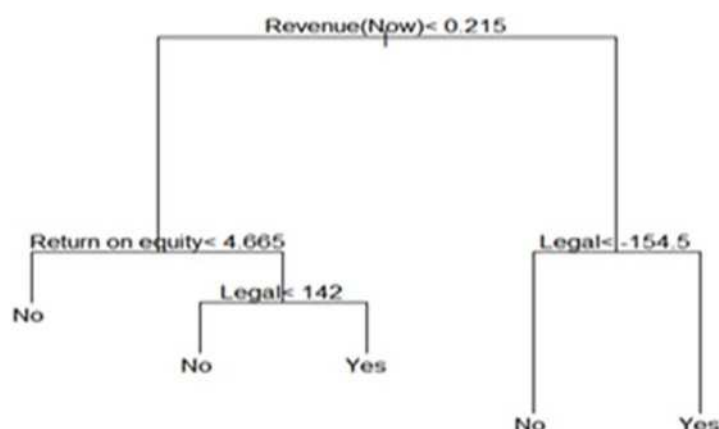


圖 1. 決策樹分析結果

實驗資料經由 Cart 欄位篩選去除影響性較低欄位後，利用 Fuzzy C-means 模糊分群將資料特性較接近歸類至同一群，減少在後段訓練訓練 SVM 預測模型時因離

群值或極端值產生預測規則上的誤判。預測模型預測準確率在模糊分群為 3 時結果最佳。表 1 為各群聚中心點:

表 1. FCM 分群中心點表

名稱	群聚	月營收長率	股東權益報酬	法人買賣超數量
股票資料	I	0.3530383	0.4513611	149993.505
	II	2.1537097	1.3695304	-4212.127
	III	1.0752616	4.3844825	1021.772

在 Fuzzy C-means 分群過程中，當公司歷史資料月營收成長率接近 1.07% 且股東權益報酬靠近 4.38% 和法人買賣超數量接近 1,022 張，這資料在計算與各群聚中心點距離時與群聚三的距離會最短且與群聚三的歸屬程度為最大值。測試資料適當分群後，依照不同群聚導入個群聚所屬的預測模型，並與實際輸出結果進行比較進而探討預測準確率與模擬投資報酬率。表 2 為預測結果與實際結果比較，0 為月營收公布後三十日後股價為下跌，反之 1 為為月營收公布後三十日後股價為上漲。

表 2. 多重 SVM 預測結果與實際結果比較

predict	output
0	1
1	0
1	1
0	1
0	0
1	1
0	1

模型預測結果為下跌且實際結果為下跌共有 213 筆，預測結果為下跌但實際結果為上漲共有 75 筆，預測結果為上漲且實際結果為上漲共有 70 筆，預測結果為上漲但實際結果為下跌共有 39 筆，397 筆測試資料中，模型預測結果與實際結果相符資料筆數有 283 筆，整體模型預測成功率達 71.24%。圖 2 為多重 SVM 預測結果與對照組實驗結果。

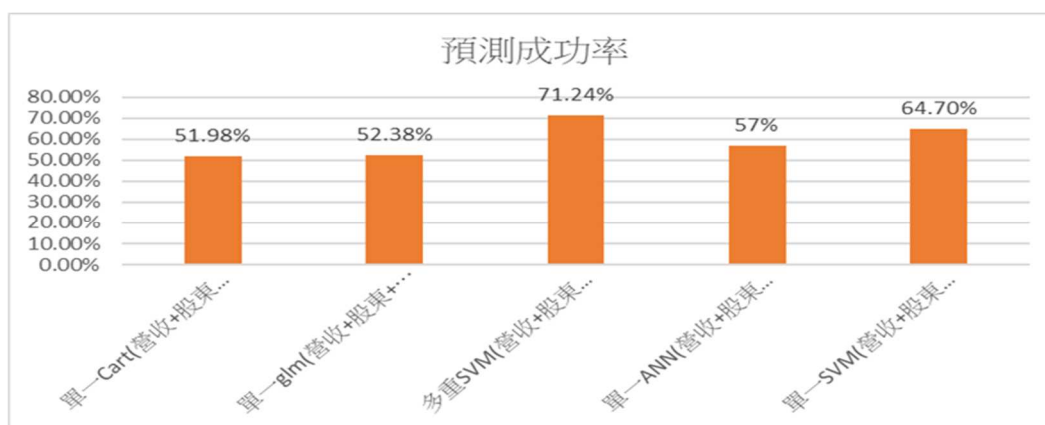


圖 2. 多重 SVM 預測結果與對照組實驗結果

由 Cart 篩選較具影響性欄位後，直接建置 SVM 預測模型與經由模糊分群加入多重概念分類預測器進行比較，實驗證實經由模糊分群後建置多重 SVM 預測模

型，在預測成功率上優於單一 SVM 預測模型。

本研究利用公司營收成長率、與營收成長率公布前三十天法人買賣超數量與上季股東權益報酬等歷史資料預測營收成率公布後三十天後股價漲跌，藉由模型預測結果進行投資決策並從中獲利，當模型預測結果為 1 時，表示該家公司股票三十天後會上漲反之 0 為下跌，預測結果顯示為上漲時，此時就買進股票並持有三十天後賣出後計算平均報酬率，圖 3 可以看出本研究所提預測模型在投資報酬率上皆優於其他單一人工智慧方法。

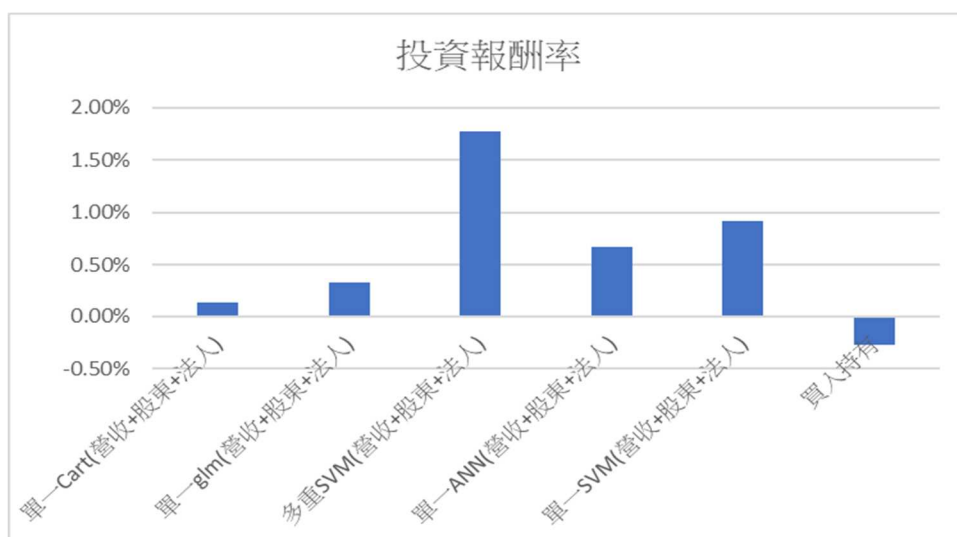


圖 3. 多重 SVM 與對照組投資報酬率

雖然在股價趨勢上漲預測準確率只有 48.27%，但在股價趨勢下跌預測準確率高達 84.52%，從中可以看出此模型不僅在預測上漲趨勢有不錯的效果，在下跌部分高預測成功率更能提供投資者避險的參考依據。

為提高本研究架構參考價值，本研究將究資料進行滾動式預測，以當季資料當作訓練資料下月資料為測試資料，舉列來說 2014 年 1、2、3 月為訓練資料 4 月為測試資料來測試本研究價各是否也具有高預測成功率，各季滾視窗預測如圖 4 所示。

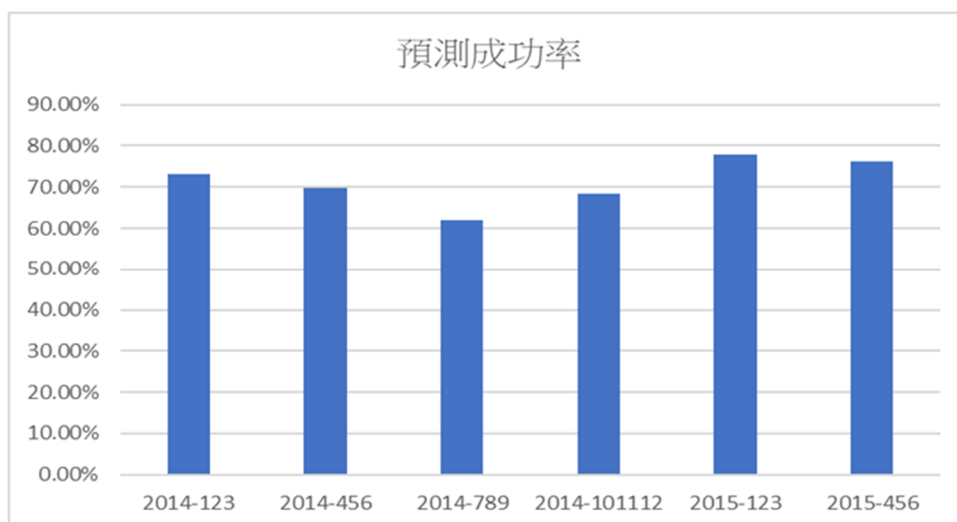


圖 4. 滾動式預測成功率

由上述實驗結果可知，經由決策樹挑選具有識別度的欄位組合有助於減少 Fuzzy C-means 的誤判，並提升模型的預測成功率，與不進行欄位篩選直接進行建置模型比較，經由 Cart 欄位篩選後擁有較好的預測成功率。而將資料以 Fuzzy C-means 分群後，建置各群聚所屬的預測模型，藉由多重 SVM 預測模型針對資料進行預測，結果顯示，多重 SVM 預測成功率高於單一 SVM。經由決策樹挑選具有識別度的欄位組合有助於 Fuzzy C-means 的誤判，並提升模型的預測成功率，與不進行欄位篩選直接進行建置模型比較，經由 Cart 欄位篩選後擁有較好的預測成功率。本研究結果顯示，與單一預測模型相比，適當加入不同人工智慧方法有助提升模型預測準確率。

本研究以能源類股為研究對象，然而台灣證券交易所及櫃台買賣中心掛牌交易的股票有水泥、鋼鐵、紡織、金融、電腦及週邊設備、電子通路業等等共 28 種類股，雖然同是掛牌公司的股票，但各種類股有其特殊性，本研究的限制為侷限能源類股，未來可將本研究樣本擴大至其他產業，並比較各產業之差異，可作為投資人的投資參考。

參考文獻

1. Alruwaili, M., Siddiqi, M. H., & Javed, M. A. (2020). A robust clustering algorithm using spatial fuzzy C-means for brain MR images. *Egyptian Informatics Journal*, 21(1), 51-66. doi: 10.1016/j.eij.2019.10.005
2. Bisoi, R., Dash, P. K., & Parida, A. K. (2019). Hybrid Variational Mode Decomposition and evolutionary robust kernel extreme learning machine for stock price and movement prediction on daily basis. *Applied Soft Computing*, 74, 652-678. doi: 10.1016/j.asoc.2018.11.008
3. Chatzis, S. P., Siakoulis, V., Petropoulos, A., Stavroulakis, E., & Vlachogiannakis, N. (2018). Forecasting stock market crisis events using deep and statistical machine learning techniques. *Expert Systems with Applications*, 112, 353-371. doi: 10.1016/j.eswa.2018.06.032
4. Chen, J.-H., Ong, C. F., Zheng, L., & Hsu, S.-C. (2017). Forecasting Spatial Dynamics of the Housing Market Using Support Vector Machine. *International Journal of Strategic Property Management*, 21(3), 273-283. doi: 10.3846/1648715X.2016.1259190
5. Das, S. P., Achary, N. S., & Padhy, S. (2016). Novel hybrid SVM-TLBO forecasting model incorporating dimensionality reduction techniques. *Applied Intelligence*, 45(4), 1148-1165. doi: 10.1007/s10489-016-0801-3
6. Dash, R., Samal, S., Dash, R., & Rautray, R. (2019). An integrated TOPSIS crow search based classifier ensemble: In application to stock index price movement prediction. *Applied Soft Computing*, 85. doi: 10.1016/j.asoc.2019.105784
7. Duarte, J. J., Montenegro González, S., & Cruz, J. C., Jr. (2021). Predicting Stock Price Falls Using News Data: Evidence from the Brazilian Market. *Computational Economics*, 57(1), 311-340. doi: 10.1007/s10614-020-10060-y
8. Fenghua, W. E. N., Jihong, X., Zhifang, H. E., & Xu, G. (2014). Stock Price Prediction based on SSA and SVM. *Procedia Computer Science*, 31, 625-631. doi: 10.1016/j.procs.2014.05.309
9. Göçken, M., Özçalıcı, M., Boru, A., & Dosdoğru, A. T. (2016). Integrating metaheuristics and Artificial Neural Networks for improved stock price prediction. *Expert Systems with Applications*, 44, 320-331. doi: 10.1016/j.eswa.2015.09.029
10. Hao, P.-Y., Kung, C.-F., Chang, C.-Y., & Ou, J.-B. (2021). Predicting stock price

- trends based on financial news articles and using a novel twin support vector machine with fuzzy hyperplane. *Applied Soft Computing*, 98, 106806. doi: 10.1016/j.asoc.2020.106806
11. Hossain, M. S., Baten, M. A., & Mukta, F. B. (2021). Forecasting Volatility of Selected Banks of Dhaka Stock Exchange (DSE), Bangladesh with GARCH (p, q) Type Models. *Journal of Economic Cooperation & Development*, 42(1), 1-24.
 12. Illa, P. K., Parvathala, B., & Sharma, A. K. (2021). Stock price prediction methodology using random forest algorithm and support vector machine. *Materials Today: Proceedings*. doi: 10.1016/j.matpr.2021.10.460
 13. Ince, H. (2014). Short term stock selection with case-based reasoning technique. *Applied Soft Computing*, 22, 205-212. doi: 10.1016/j.asoc.2014.05.017
 14. Jing, N., Wu, Z., & Wang, H. (2021). A hybrid model integrating deep learning with investor sentiment analysis for stock price prediction. *Expert Systems with Applications*, 178. doi: 10.1016/j.eswa.2021.115019
 15. Kwan, Y. K., & Dong, J. (2014). Stock Price Dynamics of China: What Do the Asset Markets Tell Us About the Chinese Utility Function? *Emerging Markets Finance & Trade*, 50, 77-108. doi: 10.2753/REE1540-496X5003S305
 16. Li, Y., Bu, H., Li, J., & Wu, J. (2020). The role of text-extracted investor sentiment in Chinese stock price prediction with the enhancement of deep learning. *International Journal of Forecasting*, 36(4), 1541-1562. doi: 10.1016/j.ijforecast.2020.05.001
 17. Liu, C.-F., Yeh, C.-Y., & Lee, S.-J. (2012). Application of type-2 neuro-fuzzy modeling in stock price prediction. *Applied Soft Computing*, 12(4), 1348-1358. doi: 10.1016/j.asoc.2011.11.028
 18. Long, W., Lu, Z., & Cui, L. (2019). Deep learning-based feature engineering for stock price movement prediction. *Knowledge-Based Systems*, 164, 163-173. doi: 10.1016/j.knosys.2018.10.034
 19. Prasad, A., & Seetharaman, A. (2021). Importance of Machine Learning in Making Investment Decision in Stock Market. *Vikalpa: The Journal for Decision Makers*, 46(4), 209-222. doi: 10.1177/02560909211059992
 20. Sadorsky, P. (2021). A Random Forests Approach to Predicting Clean Energy Stock Prices. *Journal of Risk and Financial Management*, 14(2), 48. doi: 10.3390/jrfm14020048
 21. Schumaker, R. P., & Hsinchun, C. (2009). Textual Analysis of Stock Market Prediction Using Breaking Financial News: The AZFinText System. *ACM Transactions on Information Systems*, 27(2), 12:11-12:19. doi: 10.1145/1462198.1462204
 22. Sun, B., Guo, H., Reza Karimi, H., Ge, Y., & Xiong, S. (2015). Prediction of stock index futures prices based on fuzzy sets and multivariate fuzzy time series. *Neurocomputing*, 151, 1528-1536. doi: 10.1016/j.neucom.2014.09.018
 23. Suppawong, T., Wettayakorn, P., Ponpat, P., Traivijitkhun, S., Lim, S., Thanapon, N., & Tipajin, T. (2021). DAViS: a unified solution for data collection, analyzation, and visualization in real-time stock market prediction. *Financial Innovation*, 7(1), 7-56.
 24. Tripathy, N. (2021). How Investors Leveraging Gain In Stock Market Investment: A Predictive Analysis. *Academy of Marketing Studies Journal*, 25(5), 1-11.
 25. Tsai, C.-F., Lin, Y.-C., Yen, D. C., & Chen, Y.-M. (2011). Predicting stock returns by classifier ensembles. *Applied Soft Computing*, 11(2), 2452-2459. doi: 10.1016/j.asoc.2010.10.001
 26. Wang, X., & Pardalos, P. M. (2015). A Survey of Support Vector Machines with Uncertainties. *Annals of Data Science*, 1(3-4), 293-309. doi: 10.1007/s40745-014-0022-8

27. Xie, C., Luo, C., & Yu, X. (2010). Financial distress prediction based on SVM and MDA methods: the case of Chinese listed companies. *Quality & Quantity*, 45(3), 671-686. doi: 10.1007/s11135-010-9376-y
28. Xu, Q., Jiang, C., & He, Y. (2015). An exponentially weighted quantile regression via SVM with application to estimating multiperiod VaR. *Statistical Methods & Applications*, 25(2), 285-320. doi: 10.1007/s10260-015-0332-9
29. Yu, H., Chen, R., & Zhang, G. (2014). A SVM Stock Selection Model within PCA. *Procedia Computer Science*, 31, 406-412. doi: 10.1016/j.procs.2014.05.284
30. Zhang, D., & Lou, S. (2021). The application research of neural network and BP algorithm in stock price pattern classification and prediction. *Future Generation Computer Systems*, 115, 872-879. doi: 10.1016/j.future.2020.10.009
31. Zhao, Z., Zhao, J., Song, K., Hussain, A., Du, Q., Dong, Y., . . . Yang, X. (2020). Joint DBN and Fuzzy C-Means unsupervised deep clustering for lung cancer patient stratification. *Engineering Applications of Artificial Intelligence*, 91, 103571. doi: 10.1016/j.engappai.2020.103571