

經濟論文
中央研究院經濟研究所
47:3(2019),395-448

漫步於隨機森林— 輔以多數決學習的台股指數期貨交易策略

鄭仁杰

元大證券計量交易部

江彌修*

國立政治大學金融學系

關鍵詞: 多數機器決學習、隨機森林、交易策略、台灣加權股價指數期貨、
卡馬比率

JEL 分類代號: G11, G17

* 聯繫作者: 江彌修, 國立政治大學金融學系, 台北市 116 文山區指南路二段 64 號。電話: (02) 2939-3091 分機 81265; 傳真: (02) 2939-8004; E-mail: mhchiang@nccu.edu.tw。本文承蒙科技部給予經費補助 (MOST 107-2410-H-004-071), 在此感謝, 並同時感謝責任編輯委員及兩位匿名審查人的寶貴意見。

摘 要

應用隨機森林演算法來進行未來期貨價格漲跌的分類預測,本文以技術面與籌碼面指標作為模型訓練的特徵,進而建構輔以多數決學習的台股指數期貨交易策略。藉由參數的重要性衡量,我們辨識出爭議變數,並探究參數配置的屬性擾動之於演算法預測能力及策略績效的影響。利用2007年至2018年的台股指數期貨資料,本文以多重角度測試策略之績效與穩健性。實證結果顯示,在考量交易成本之下,本文所建構之多數決學習台股指數期貨交易策略,要能於其訓練區間及測試區間皆呈現穩定勝出大盤的績效,其隨機森林模型所共同具備的參數配置必須包含3-14日MA與RSI指標、遠月期貨交易量、現貨交易量、期貨外資未平倉量與買賣權未平倉比率。

1. 前言

近年來人工智慧 (artificial intelligence) 的研究與發展蔚為風潮, 其中, 機器學習 (machine learning) 作為一種辨認巨量多維度資料之間未知關聯性與規則的方法, 在硬體計算能力的快速躍進之下, 更加催化了許多過去所認為不可觸及的應用領域。有鑒於金融市場的龐大數據及其多元且複雜的內蘊資訊, 如何自海量數據中萃取出重要的價格規律及風險意涵, 並進行演算化的表徵學習 (representation learning) 及未來資訊預測, 以輔佐金融市場參與者在投資、交易與避險上的廣大需求, 實為一實務界與學術界所高度關注的共同議題。

不可否認的, 眾多學者在應用機器學習於金融市場上所作的諸多嘗試, 讓人們對於如何使用機器學習來輔佐交易的進行有了更多期待與想像。¹ 憑藉著與傳統金融建模 (financial modeling) 截然不同的演算思維, 機器學習模型並不受限於可胃納的資料型態, 因而得以同時汲取時間序列分析 (times series analysis) 和技術分析於具體化金融資產價格特徵的個別優勢。例如 Brooks et al. (2001) 的研究發現, FTSE 100 指數與其期貨價格之間存在著領先－落後的時間序列關係。他們利用此規律建構交易策略並驗證其套利的有效性。Easley and O'Hara (1992) 及 Blume et al. (1994) 以技術分析方法為根基, 他們提出交易量特徵－特別是未預期的交易量－具有預測股價指數的能力。這些主流文獻提醒著我們: 任何對價格和風險的預測, 都必須源自於其模型對資料特徵的具體掌握。²

在本文中, 我們視交易策略的本質為買賣訊號的生成, 而買賣訊號之生成, 則有賴於一個有能力的學習機制來辨識金融資產價格漲跌的潛在規律並進行預測。因為著重於漲跌的預測, 我們以機器學習中的分類問題 (classification) 來界定未來期貨指數價格漲跌的預測問題, 應用隨機森林來學習台灣

¹ Ren et al. (2006) 利用 1993 至 2002 年的 S&P 500 股票資料, 藉由建構決策樹機器學習演算法來提高選股的績效。Dutta et al. (2012) 運用羅吉斯迴歸找出能夠影響股票報酬的財務比率風險因子, 並以此建構預測股票報酬的模型。

² 國內文獻中, 巫和懋與許智翔 (2010) 的研究指出, 技術分析有助於提升投資者的預期效用, 之於波動率較大的資產, 技術分析所帶來的效益更為顯著。

股價指數期貨市場的資料特徵並進行未來價格漲跌的分類預測。基於此架構, 我們建構輔以多數決學習(ensemble learning)的台股指數期貨交易策略。在此構面上, 就針對台灣金融市場的獨特性來建構適應於其的人工智慧交易策略, 就我們的認知, 本文為文獻中的首度嘗試。

就資料特徵的學習而言, 僅僅藉由單一分類器(classifier)來表徵學習整體假說空間所無法避免的偏誤(bias), 在機器學習中是一個被普遍認知的問題。關於如何降低偏誤並有效地提昇模型的預測能力的探討, 促成了多數決學習模型理論的產生。現有的研究成果顯示, 多數決學習之於模型的預測能力確實能夠產生泛化的優越成效。例如 Ho et al. (1994) 的文章, 他們採用多元分類器的決策樹, 並以類別化的排序來增進了其多元決策樹的分類效果。Giacomel et al. (2015) 在另一方面, 採用了兩個全然不同架構的類神經網路來建構基於多數決學習的類神經網路模型的股票交易策略。他們以北美與巴西股票市場中選出個股作為對象, 實證的結果顯示, 所建構之多數決學習類神經網路交易策略對於個股的漲跌具有顯著的預測能力。Borovkova and Tsiamas (2018) 則以十二種個長短期記憶模型(long short-term memory, LSTM) 結合成為多數決長短期記憶模型。他們針對紐約證券交易所(NYSE)與那斯達克(NASDAQ)上市的 22 檔大型個股的五分鐘價格走勢進行預測。其預測結果顯示, 所提出之多數決長短期記憶模型大幅勝出於 LASSO (least absolute shrinkage and selection operator) 迴歸模型與脊(ridge)迴歸模型。

身為多數決學習的重要方法之一, 隨機森林(random forest)以多棵決策樹的最後表決結果作為其最終決策, 其抽樣的隨機性能使模型不致於過度配適, 且允許建模過程中模型錯誤率(out-of-bag error rate, 袋外錯誤率)的直接驗證。其出色的類別預測能力, 也令它開始被應用在金融資產的交易策略上。Kumar and Thenmozhi (2006) 透過線性區別分析(linear discriminant analysis, LDA)、羅吉斯迴歸(logistic regression)、類神經網路(neural network)、隨機森林以及向量支持機(support vector machine, SVM)等五種機器學習模型, 分別建構 S&P CNX Nifty 指數走勢預測模型, 並以樣本外預測命中率相互比較各模型之預測能力。其實證結果發現, 向量支持機與隨機森林的預測能力遠優於其餘三種演算法。從別的市場的角度出發, Sen and Chaudhuri (2017) 以印度國家證券交易所(National Stock Exchange, NSE)中的兩檔個股為實例, 比較三種分類模型與三種迴歸模型的預測能力優劣, 其中, 三種分類模型分

別為羅吉斯迴歸、隨機森林與向量支持機, 而最後的實證結果顯示, 隨機森林在所有分類模型中能夠達到最好的預測準確度。除此之外, 他們也呼應了 Kumar and Thenmozhi (2006) 的結論, 指出隨機森林與向量支持機確實具有很好的預測能力, 然而兩者的優劣排序, 會依不同的資料屬性、市場或商品而有所差異。

由上述文獻我們可以觀察得到, 應用機器學習於金融交易策略的研究多數著重於檢驗演算法的預測能力, 然而對於任何交易策略之建構所不可缺席的「風險承擔下的」績效衡量卻明顯地缺乏著墨, 遑論不同參數配置所造成的屬性擾動之於其策略獲利績效的影響。有鑒於此, 本文在檢驗模型的預測能力之外, 更著力於論證參數配置之於模型分類預測能力和策略獲利績效之間的權衡關係, 以彌補現有機器學習的金融市場應用性文獻於投資意涵上的廣泛缺憾。

就參數配置的適當性而言, 檢視配置參數之於模型的重要性是一重要的環節。在本文中, 我們藉由兩種量化配置參數重要性的方法來檢視本文所建構的多數決台股指數期貨交易策略。其一, 是從隨機森林分類預測結果的平均準確度減少 (mean decrease in accuracy, MDA) 來評定每個配置參數 (特徵節點) 的選取之於模型預測能力的貢獻程度。意即我們藉由打亂配置參數 (特徵節點) 的順序, 並且度量順序變動對模型的準確度的影響。其二, 則是從每個配置參數 (特徵節點) 的選取之於隨機森林分類預測結果的平均吉尼不純度減少 (mean decrease in Gini, MDG) 來度量。意即我們觀察以每個配置參數 (特徵節點) 對誤差 (以平均吉尼不純度來衡量) 的平均減少程度。在實證結果的呈現上, 本文以建構兩個交易策略來闡述參數配置之於模型分類預測能力及策略獲利績效之間的權衡關係: 交易策略 (一) 是以全部變數包含爭議變數 (attribute in dispute) 來建構。交易策略 (二) 則是以全部變數不包含爭議變數來建構。我們分別命名交易策略 (一) 與交易策略 (二) 為「Perturbed_R_F 策略」與「unPerturbed_R_F 策略」, 以區分在「受爭議變數擾動」和「未受爭議變數擾動」的隨機森林參數配置之下所衍生的期貨交易策略。在本文中, 爭議變數對於模型預測能力及交易策略績效的具體影響, 是藉由比較 Perturbed_R_F 與 unPerturbed_R_F 兩策略的混淆矩陣與績效衡量指標所呈現。

就策略「風險承擔下的」績效衡量上, 我們尤其關注其下方風險 (down-side risk), 因為在真實的交易世界裡, 令交易員恐慌的往往不是單次蒙受損

失,是在其蒙受損失的時間區塊過度持續。傾向於承受冗長損失區間的策略,在實務的應用上是可議的。本文主要選用的績效指標為卡馬比率(Calmar ratio),它的互相比較性是藉由策略採取相同的時間區塊為基準。基於2007年7月至2018年2月的台股指數期貨資料,本文的實證結果發現,在考量交易成本之下,Perturbed_R_F與unPerturbed_R_F兩策略於全部區間分別能達到32.20與29.39的卡馬比率,明顯勝於大盤的0.84卡馬比率。³即使在大盤更為強勢的測試區間,兩策略之卡馬比率也分別能達到勝過大盤的4.01與4.32。使用其他績效指標比較亦能得出相似結果。⁴故就績效的層面上,本文所建構的多數決隨機森林期貨交易策略,在受爭議變數擾動和未受爭議變數擾動的兩類模型配置情境下,皆有勝出大盤的績效的表現。

為了深入了解績效的實際樣貌,本研究更以年與月兩種時間框架來全盤檢視策略於2007年至2018年資料區間的損益。以年為週期來看,除去僅有兩個月資料的2018年,Perturbed_R_F和unPerturbed_R_F兩策略皆能呈現穩定的每年度獲利績效,特別是在逆勢環境中,兩策略的績效表現更遠勝於大盤。直覺上,如此熊市中的績效表現並不意外,它極可能直接受惠於我們採用下方風險的衡量指標來訓練模型及衡量策略績效所產生的助益。而就勝率而言,Perturbed_R_F和unPerturbed_R_F兩策略全部年度最差表現也分別能達到57.32%與54.78%,全部區間平均則有六成以上的美好表現。⁵透過上述實證結果可以得出,本研究建構之交易策略不僅在獲利能力上相當穩定,其勝出大盤績效表現自然提供了許多市場摩擦可以耗損的空間和因應投資人不同風險胃納的彈性。站在交易員角度來觀察,月損益的分析結果顯示,Perturbed_R_F和unPerturbed_R_F兩策略皆僅發生兩次連續三個月虧損的情況,表示其波段虧損情況的發生率不高,策略在現實市場中有一定可被執行的價值。

最後,本文相互比較Perturbed_R_F和unPerturbed_R_F兩策略於訓練區間

³ 本文在計算獲利時會扣除手續費與滑價(合稱費用)。其扣除方法為一口期貨契約開倉、平倉皆會扣除一次費用,而部位反轉(多單轉空單,抑或相反)則會扣除兩次費用,單次費用為三點期貨點數。詳細定義請見附錄2。

⁴ 我們採用的指標包括Omega、索丁諾比率(Sortino ratio)、潛在上漲比率(upside potential ratio)與赫斯特指數(Hurst exponent)。

⁵ 在本文中,相對於「錯誤率」以類別中有多少比例沒有被預測到為機器學習預測能力的衡量基準,「勝率」計算的是策略的每筆交易最後損益是否為正。

與測試區間的實證結果, 研究發現去除爭議變數的 `unPerturbed_R_F` 策略, 其模型預測能力確實有提昇, 然而其訓練區間卡馬比率卻因此而下降, 顯示其預測能力提昇的成本來自於其策略績效的妥協。另一方面, 兩策略於測試區間的卡馬比率卻呈現相反結果。進一步的實證結果顯示, 爭議變數的捨棄會提昇模型在預測下跌的能力, 然而同時降低模型的上漲預測能力。就訓練區間台股指數期貨長期資料所呈現的多頭市場特性而言, `Perturbed_R_F` 策略所採用的(包含爭議變數)參數配置, 相較於 `unPerturbed_R_F` 策略, 呈現出更適應於資料期間多頭市場特性的策略取向, 其績效因而更能勝出。測試區間的情況則更為特殊, 兩策略於短期強勁多頭市場中的決策做多比例皆大幅上升, 以致策略鑒別度降低。然而此時 `unPerturbed_R_F` 策略在測試區間所勝出的績效來自於它勝出的預測下跌能力。

本文後續之結構安排如下: 在第二節中我們提供隨機森林分類器與其底層演算法的重點式描述。第三節為研究方法, 我們介紹本研究之架構流程, 敘述本研究之訊號與特徵生成方法、模型衡量方法與績效指標等關鍵因子。第四節為本文之實證結果, 我們以台灣加權股價指數期貨為標的, 實現前述之步驟以建構基於多數決的隨機森林交易策略, 並對模型的預測能力和績效表現做詳盡的檢測。第五節為本文的結論及後續研究建議。

2. 多數決隨機森林決策樹

2.1 模型定義

假設 X 為樣本空間, Y 為二元類別標識集合, 則 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ 為整體訓練集。另外, 假設樣本空間劃分為 M 個區域, 分別為 R_1, R_2, \dots, R_M , 在每個區域都有一個對應的二元樣本類別, 分別為 c_1, c_2, \dots, c_M , 因此模型可以表示為:

$$f(x) = \sum_{m=1}^M c_m I(x \in R_m). \quad (1)$$

$$I(x \in R_m) = \begin{cases} 1, & \text{if } x \in R_m, \\ 0, & \text{else.} \end{cases} \quad (2)$$

其中, I 為一指標函數, x 為樣本空間中的樣本點, R_m 為根據樣本空間劃分出的區域, c_m 為該區域以不純度函數找出的預測類別。

隨機森林決策樹的建構過程中必須倚賴不純度函數 (impurity function) 來找尋最佳化的節點, 而其最佳化節點的判定準則即為找到一極小化切點 A :

$$\min : \text{IF}(A) = \phi[P(y = 1|A)]. \quad (3)$$

其中, IF 與 ϕ 代表不純度函數, A 代表某切點, y 為二元類別樣本點。基於上述的最佳化節點準則, 如此持續在各個子葉的節點不段產生分支, 最後使所有樣本類別標示都被完全分類, 此即完成一棵樹的建構。

不同不純度函數決定不同決策樹演算法, 常見之不純度函數 (當 Y 為二元) 包括: 貝氏錯誤 (Bayes error): $\phi(p) = \min(p, 1 - p)$; 熵 (entropy): $\phi(p) = -p \log_b p - (1 - p) \log_b (1 - p)$; 吉尼不純度 (Gini impurity): $\phi(p) = p(1 - p)$ 。其中, ϕ 代表不純度函數, p 代表某一類別的比例。

在此, 我們可進一步將決策樹表達如下:

$$h(x, \Theta_k). \quad (4)$$

其中, \mathbf{x} 為 X 中採用拔靴法 (bootstrap aggregating/bagging) 抽出之樣本集合, 令其抽取樣本個數為 n 且 $n < N$; k 為隨機森林中的第 k 棵樹, 令 K 為建構隨機森林設定的決策樹數量; Θ_k 為建構第 k 棵樹所抽取出的特徵, $\Theta_1 \cdots \Theta_k$ 互為 i.i.d. (independent identically distributed), 抽取個數為 M 且 $M < G$ 。

因此, 多數決隨機森林模型可呈現為:

$$\text{Ensemble}(h(\mathbf{x}, \Theta_1) \cdots h(\mathbf{x}, \Theta_k)). \quad (5)$$

其中, Ensemble 代表隨機森林的多數決分類器 (函數), 其依照每棵決策樹分類結果採取多數決 (majority vote), 以多數樣本的投票決定最後該樣本點的分類。

2.2 分類器決策樹

隨機森林演算法被發表於 21 世紀，但其所根據的決策樹演算法在上個世紀就已被提出。做為一種機器學習模型，隨機森林的名稱來自於 Ho (1995) 提出之隨機決策森林 (random decision forest)，並於 2001 年時由 Leo Breiman 與 Adele Cutler 共同發表，用於解決預測的分類與數值問題 (Breiman, 2001)。

決策樹的概念被應用在諸多領域，其關鍵在於特徵節點的決定，而其，如前一節模型定義中我們所陳述，演算法中的評斷標準即為不純度函數。不純度函數是計算資料雜亂程度的方法，如同迴歸分析使用殘差平方和計算迴歸模型對於資料配適不佳的程度，決策樹同樣有不純度函數可以判斷以何者作為節點能提昇模型的配適度，但隨著處理問題的相異，就會有不同的不純度函數與之對應。過往研究中，若用以處理數值問題則以變異縮減 (variance reduction) 為主；若如同本研究欲用以處理分類問題，則以熵、信息增益 (information gain) 與吉尼不純度為較為常用的方法。

其中，熵為資料亂度的衡量方法，而信息增益是資料亂度的降低程度。熵在資訊理論又被稱為資訊熵或是夏農熵 (Shannon's entropy)，其概念來自於訊息雜亂度的衡量 (Shannon, 1948)。當資訊來源雜亂度越低，其熵值就會越小。反之，則其熵值越大，其計算公式如下：

$$H(X) = \sum_{i=1}^n P(x_i) I(x_i) = - \sum_{i=1}^n P(x_i) \log_b P(x_i). \quad (6)$$

其中， X 為一隨機變數， x_i 為第 i 個樣本， I 代表資訊含量， b 為對數底數。

信息增益則代表某節點之下熵值降低的程度，其公式如下：

$$IG(X) = H(X) - H(X|a). \quad (7)$$

其中， X 代表隨機變數， a 代表節點。

在另一方面，吉尼不純度是一種衡量在同一個集合之下，按照其元素機率分布隨機選擇其中的元素後，元素被錯誤標記的程度，其定義如下：

$$\text{Gini}(p) = 1 - \sum_{i=1}^J p_i^2. \quad (8)$$

其中, p 代表某類別佔全部類別的比例。

在處理不同資料型態的類別化問題上, 相對於熵比須倚賴先入為主的函數形式所連帶的模型風險, 吉尼不純度極有彈性的形式賦予其明顯的優勢, 如 Breiman et al. (1984) 在其 CART (classification and regression tree) 分類與迴歸決策樹演算法中所呈現。做為一種結合二元資料類別的決策樹 (binary tree), CART 演算法必須拆解每個變數為不同的二元組合, 再依不同組合計算個別的吉尼不純度。本文所建構的交易策略本質為類別化的決策樹, 基於 Breiman et al. (1984) 的經驗, 在取捨其特徵節點的參數配置上我們因而採用吉尼不純度。

2.3 多數決隨機森林 (分類器) 決策樹

單一決策樹有可能預測能力不足、容易過度配適, 而集成 (ensemble) 多棵決策樹的預測則有機會使模型更加完備。隨機森林會重複抽取部分資料與變數, 用以建構決策樹, 使其底層的多棵決策樹具有隨機性, 最後再用多數決決定最終預測結果。因而集成與多數決一詞, 往往在文獻中往往可以互相替代使用。

資料抽取的部分, 隨機森林使用拔靴法來完成, 每一棵決策樹抽取 n 個資料, 這筆資料稱為袋外資料 (bagging data)。另一方面, 每棵樹要抽取 m 個特徵, 並決定模型內決策樹總數量 ntree , 其參數限制如下:

$$n \leq N, \quad m \leq M, \quad 0 < \text{ntree} < \infty.$$

其中, N 為訓練資料的資料總數量, M 為訓練資料變數總數量, ntree 為隨機森林內決策樹總數量。

另外, 隨機森林中任意兩棵樹的相關性提高或是每棵樹的分類能力降低, 都會使模型整體錯誤率增加, 而 m 設定較低會使每棵樹的分類能力降低但樹之間的相關性也降低, 較高則會使每棵樹的分類能力提昇但樹之間的相

關性也隨之提昇, 此即指出相關性與分類能力之間存在抵換關係, 而 m 也成為模型中的關鍵參數。

3. 研究方法

在本節中, 我們首先介紹多數決隨機森林交易策略之建構, 以訊號與特徵之生成方法為主, 後續的小節進而建立模型衡量與策略執行準則, 最後則講述本研究使用之關鍵績效指標, 並具體提供本研究之交易策略的關鍵步驟說明。

3.1 多數決隨機森林交易策略之建構

3.1.1 交易訊號生成

建構模型之前, 必須找到訓練模型所需目標, 此即為訊號生成。訊號生成方式為影響模型預測效果的關鍵因素, 因此, 本文的訊號生成方式會借鏡於過去文獻。本研究的訊號生成定義將沿用 Khaidem et al. (2016), 其定義如下:

$$\text{Signal}_i = \text{sign}(\text{close}_{i+d} - \text{close}_i). \quad (9)$$

其中, d 代表天數, 是本文訊號生成過程中唯一的參數, 因此訊號生成的定義可以解釋為某日收盤價格與其經過 d 天後的收盤價格漲跌。

由於近月期貨合約每個月都會結算, 在結算日當天的收盤價格與隔天的開盤價格都會有段跳空, 而那段跳空對於績效來說不存在, 因此本文中生成訊號與計算績效的資料, 都會使用扣除跳空後的修正後近月期貨價格。另一方面, 本文中的訊號生成參數與持有天數可以不同, 因為研究中建構模型的特徵包含的資訊量有許多都超過一天, 也就是交易策略看的是波段漲跌, 而單看隔天漲跌可能會使模型訓練過程遭受雜訊干擾, 故本研究選用參數 $d = 2$ 。

3.1.2 特徵生成

如我們在前言中所指出, 任何對價格和風險的預測, 都必須源自於其模型對

資料特徵的具體掌握。要建構適應於台股指數期貨市場特徵的交易策略, 必須先能夠具體化其市場特性。此階段, 在機器學習中或稱之為表徵學習。本研究選用的市場特徵分為兩大類型, 其一為技術面指標, 另一為籌碼面指標。技術面指標包含許多市場價量與動能資訊, 能在交易中產生良好的買賣點; 籌碼面指標能量化不同類型投資人對市場的想法與行動, 讓策略根據市場情緒與籌碼水位修正。我們在以下篇幅中提供關於特徵更詳細的定義、方法以及所承襲的文獻。

3.1.2.1 技術面指標

(1) 相對強弱指數 (relative strength index, RSI)

我們定義 RSI 公式如下:

$$RSI(n) = 100 - \frac{100}{1 + RS(n)}. \quad (10)$$

$$RS(n) = \frac{SMA(U, n)}{SMA(D, n)}. \quad (11)$$

$$U = \max(\text{Close}_{now} - \text{Close}_{previous}, 0). \quad (12)$$

$$D = \max(\text{Close}_{previous} - \text{Close}_{now}, 0). \quad (13)$$

其中, Close_{now} 代表今日價格, $\text{Close}_{previous}$ 代表前一日價格, n 代表天數, SMA 公式如移動平均段落所示。

隨機森林底層演算法是決策樹, 而決策樹本身具有降維能力, 因此使用較多特徵較不會影響模型預測能力, 且能讓模型獲得多重動能資訊。本文採用 3 至 14 日共 12 種 RSI 來學習市場有效的動能資訊。

(2) 移動平均 (moving average, MA)

如 Lo et al. (2000) 所指出, 平滑能使觀測誤差降低, 因此 MA 指標經常被應用為建模特徵。為了適應更複雜的市場環境, 許多 MA 指標的變化型態被提出並被應用為策略特徵, 如 Kim (2003) 便將數種 MA 指標變化型應用於向量支持機股價指數預測模型。MA 在傳統技術分析中廣為應用, 指標擁有降低雜訊、突顯趨勢的特性, 本文也因而將 MA 指標納為市場(資料)表徵學習中不可缺失的特徵。MA 指標有許多不同的定義可以採用, 本文採用的是所

謂簡單移動平均 (simple moving average, SMA), 其型式定義如下:

$$\text{SMA}(\text{Close}_t, n) = \frac{\text{Close}_t + \text{Close}_{t-1} + \cdots + \text{Close}_{t-(n-1)}}{n}. \quad (14)$$

其中, Close_t 代表第 t 日價格, n 代表天數。

然而若單純將價格的 SMA 放入隨機森林, 會導致模型被不同時期不同標準的市場價格影響。因此本研究對 MA 指標提出修正調整, 使其能保留 SMA 移動平均的良好特性, 又可以消除市場價格環境的影響, 其式如下:

$$\text{MA}^{\text{Adj.}}(\text{Close}_t, n) = \text{Close}_t - \frac{\text{Close}_t + \text{Close}_{t-1} + \cdots + \text{Close}_{t-(n-1)}}{n}. \quad (15)$$

其中, Close_t 代表第 t 日價格, n 代表天數。

在模型訓練上, 我們用 3 到 14 日共 12 種移動平均的 MA 特徵來學習台股指數期貨市場的長中短期趨勢。

(3) 交易量

交易量與價格走勢的關係一直是眾所皆知, 也因此交易量經常被納為交易策略重要因子之一。Blume et al. (1994) 透過建立一個以「交易員得到的資訊品質不同」為假設的均衡模型, 提出並證明交易量能夠帶來價格資料無法表現的資訊量品質的資訊。交易量是推動價格的關鍵, 也因此交易量的大小同時也會代表每日多空的力道, 當模型加入市場的交易量特徵, 便有機會辨認出更具品質的價格趨勢。

本研究共加入三種有關交易量的特徵, 其分別是現貨當日交易量、近月期貨當日交易量以及遠月期貨當日交易量。在模型中不僅放入近月期貨交易量, 更加入數種不同的相關特徵是因為, 本研究認為近遠月期貨以及現貨的交易量關係, 有可能反映出不同到期時間的契約與其標的現貨的價格資訊品質, 能夠使模型更精準地預測未來走勢。

3.1.2.2 籌碼面指標

(1) 期貨外資未平倉量

依照目前的國際情勢, 台灣仍位於從發展中國家過渡到已發展國家的階段, 也因此許多人認為, 當外國企業或金融機構決定將其龐大的金流注入台

灣金融市場, 這股熱錢能讓市場產生難以忽視的變化。諸多文獻都顯示出外資對台灣市場的影響, 例如 Richards (2005) 便發現國外資金流入對亞洲發展中市場有正向影響, 且其影響比過去文獻認為的更大, 而實證結果中, 台灣亦有出現相同現象。

在執行上, 由於本研究定義的期貨外資未平倉量在盤後才有資料, 為了符合現實條件, 本研究會將期貨外資未平倉量資料再往前平移一天, 代表模型會以昨天的期貨外資未平倉量預測明日價格走勢。

(2) 買賣權未平倉量比率

真實的市場情緒無法被測量, 故通常會以代理變數, 像是 VIX、買賣權比率等, 讓市場情緒能成為一個「看得到」的值。Simon and Wiggins (2001) 就將 VIX、ARMS Index 以及 Put-Call Ratio 視為市場情緒的衡量指標, 並運用這些指標預測未來一段期間的 S&P 500 報酬率。在他們提出的實證結果中, 發現這些指標確實具有統計及經濟意義上的顯著性, 且皆具有逆勢性質, 也就是當市場極為恐慌時, 經常就是很好的買點。

買賣權未平倉比率的重要性在眾多文獻中被充分地驗證。其不僅能從橫斷面角度判斷市場當下的多空意見, 又可從時間序列角度去看投資人的情緒水位, 因此本研究將其納為模型特徵之一。與外資期貨未平倉量一樣, 為了符合現實條件, 本研究會將買賣權未平倉比率資料再往前平移一天, 代表模型會以昨天的資料預測明日價格走勢。

3.2 隨機森林模型預測能力衡量

3.2.1 混淆矩陣 (confusion matrix)

隨機森林用以衡量其預測能力的方法稱之為袋外錯誤率 (簡稱為 OOB error rate), 名稱中的 Bag 是 Bootstrap Aggregating 的簡稱, 因為該方法會使用拔靴法得來的資料計算錯誤率。計算的第一步會先運用拔靴法抽取的資料建構第 K 棵樹, 而沒有被抽到的資料就叫做第 K 棵樹的袋外樣本 (OOB sample), 因此, 每個訓練資料都會有各自沒有參與建構的樹, 也就是每個資料點都有可能作為某幾棵樹的袋外樣本。最後, 將袋外樣本放入決策樹並做出多數決, 最後誤分資料與全部資料的比率即是該隨機森林模型的袋外錯誤率。袋外錯誤率是隨機森林的一大優點, 因為它可以在生成模型的過程中直接對誤差

表 1 混淆矩陣

	真實為陽性	真實為陰性	
預測為陽性	真陽性 (True Positive, TP)	偽陽性 (False Positive, FP)	精確率
預測為陰性	偽陰性 (False Negative, FN)	真陰性 (True Negative, TN)	陰性預測率
	召回率	特異率	
	盛行率	準確率	F ₁ 值

註: 真陽性代表真實為陽性且被預測正確的類別, 偽陽性代表真實為陰性但被預測為陽性的類別; 偽陰性代表真實為陽性但被預測為陰性的類別, 真陰性代表真實為陰性且被預測正確的類別。在本文中, 應用在台股指數期貨價格的漲跌預測上, 陽性代表的是上漲, 而陰性則代表下跌。

產生不偏估計, 而無需做非常耗時的交叉驗證 (cross validation) 來估計模型的錯誤率。

除了袋外錯誤率以外, 由袋外資料生成的混淆矩陣還可產生更多利於衡量模型好壞的比率, 如表 1 所示。

其中, 精確率 (precision): 又被稱為真實預測比率 (positive predicted value, PPV), 代表預測為陽性的結果中有多少比率是真實為陽性, 我們定義其公式如下:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (16)$$

陰性預測率 (negative predicted value, NPV) 則代表預測為陰性的結果中有多少比率是真實為陰性, 其公式如下:

$$\text{NPV} = \frac{\text{TN}}{\text{TN} + \text{FN}} \quad (17)$$

召回率 (recall): 又被稱為敏感度 (sensitivity) 或是真實陽性比率 (true positive rate, TPR), 代表所有真實為陽性的資料中多少比率被預測正確, 其公式如下:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (18)$$

特異度 (specificity): 又被稱為真實陰性比率 (true negative rate, TNR), 代表所有真實為陰性的資料裡, 有多少比率被預測正確, 其公式如下:

$$\text{Specificity} = \frac{\text{TN}}{\text{TN} + \text{FP}}. \quad (19)$$

盛行率 (prevalence): 代表所有資料中真實為陽性的比率, 其公式如下:

$$\text{Prevalence} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}. \quad (20)$$

準確率 (accuracy): 代表所有資料中有多少比率預測正確, 其公式如下:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}. \quad (21)$$

F_1 值 (F_1 score): F_1 值原始定義為精確率與召回率的調和平均 (harmonic mean), 其為陽性類別資料預測的綜合評量標準, 在本研究中, 我們不僅衡量針對陽性 (上漲) 資料的預測能力, 也必須衡量針對陰性 (下跌) 資料的預測能力。我們因此沿用 F_1 值的概念, 建立之於陽性 (上漲) 資料預測和陰性 (下跌) 資料預測的 F_1 值綜合評量標準, 我們各別稱之為 F_1P 值與 F_1N 值, 其公式具體如下:

$$F_1P = \frac{2}{\frac{1}{\text{Precision}} + \frac{1}{\text{Recall}}}. \quad (22)$$

$$F_1N = \frac{2}{\frac{1}{\text{NPV}} + \frac{1}{\text{Specificity}}}. \quad (23)$$

3.3 策略績效衡量與變數重要性衡量

3.3.1 績效衡量指標

有鑒於現有機器學習的金融市場應用性文獻於投資意涵上廣泛缺憾，本研究論證參數配置、分類預測能力和策略的獲利績效之間的權衡關係。因為交易策略的漲跌預測機制和其投資績效是息息相關的，策略的績效衡量因而在本研究中是不可缺失的。

夏普比率 (Sharpe ratio) 財務的各種績效衡量指標中最常被提及的。它以價格的變異數詮釋風險，和假設風險在牛市和熊市中的動能雷同，在策略績效的評量上，因為真實的價格報酬並非常態分配，其風險估計因而存在偏誤。在本文中，期貨為本研究之標的，而期貨採用保證金機制，無法與現貨商品一樣有一固定初始投入本金，故不適用以其商品價格的變化來求計算出其報酬率，因此本文會使用卡馬比率來取代夏普比率，作為訓練模型與呈現績效衡量的主要指標。卡馬比率與夏普比率概念相似，都是在每單位的風險承擔下來衡量一策略獲利的績效，不同的是，卡馬比率針對的是投資人所最嫌惡的下方風險，因而其風險的估算是以最大可能跌幅 (maximum drawdown) 來取代變異數。其具體公式如下：

$$\text{Calmar}_i = \frac{\text{Profit}_{i,T}}{-\text{MDD}_{i,T}}. \quad (24)$$

$$\text{MDD}_{i,t} = \min(\text{DD}_{i,t}, \text{MDD}_{i,(t-1)}). \quad (25)$$

$$\text{MDD}_{i,0} = 0. \quad (26)$$

$$\text{DD}_{i,t} = \text{Profit}_{i,t} - \max(\text{Profit}_{i,1}, \text{Profit}_{i,2}, \dots, \text{Profit}_{i,t}). \quad (27)$$

其中， $0 < t < T$ ， t 代表天數， T 代表最後一日，Profit 為累積報酬，max 為極大值運算符號，min 為極小值運算符號。

為了更進一步驗證策略績效的穩健性，本文也採用其他績效指標來反覆驗證策略績效。我們所採用的指標包含了：Omega、索丁諾比率、潛在上漲比率與赫斯特指數。

3.3.2 變數重要性衡量

參數配置將直接影響隨機森林的分類預測能力和交易策略的績效。之於模型配置的變數重要性衡量,因而是我們建模中一個非常重要的環節。經由辨認出爭議變數,並進而建構 Perturbed_R_F 和 unPerturbed_R_F 兩個策略,以區分受爭議變數擾動和未受爭議變數擾動的模型配置,本文才得以藉由各自的混淆矩陣與績效衡量指標所呈現的具體意義來探討爭議變數對於模型預測能力及交易策略績效的影響。

隨機森林在判斷變數重要性上,有兩種常見的方式:其一為平均準確度減少(MDA),另一為平均吉尼不純度減少(MDG)。意涵上,MDA 是以隨機排序某特徵來了解該特徵對模型準確度的影響程度,MDG 則是以每個特徵能降低每棵樹多少吉尼不純度為衡量標準。因為 MDA 是直接衡量每個變數對模型的貢獻程度,而 MDG 則僅考慮每個節點去衡量變數重要性,因此得出的結果可能偏向區域解,較容易帶出有偏誤的結論,因而一般認為 MDA 會比 MDG 有更好的衡量效果,以下我們敘述其求算步驟:

Mean Decrease in Accuracy (MDA)

1. 計算建構隨機森林的每棵決策樹的準確度,這裡稱為原始準確度。
2. 將每棵決策樹的某特徵的值隨機排序,再重新建構每棵決策樹,並計算其準確度,將其稱為隨機特徵準確度。
3. 計算隨機森林中每棵決策樹的隨機特徵準確度與原始準確度的差,再計算其平均並標準化,即為 MDA。

Mean Decrease in Gini (MDG)

1. 計算每棵決策樹的某特徵能降低其不純度的程度,稱為 DG。
2. 計算隨機森林中每棵決策樹 DG 的平均,即為 MDG。

4. 實證結果

本文的交易策略執行步驟如下:

1. 蒐集預測未來漲跌所需的今日(或前日)特徵資料,將其代入建構完成的隨機森林交易策略,依照預測結果判斷未來的漲跌並做出決策,共有決策做

多（隔日或數日後上漲）與決策做空（隔日或數日後下跌）兩種可能。

2. 本研究建構之交易策略操作規則如下：

- (1) 無持倉部位時，決策做多則進場一口多單；反之則進場一口空單。
- (2) 持倉多部位時，決策做多則不做任何部位變動；反之則將部位轉為一口空單。
- (3) 持倉空部位時，決策做多則將部位轉為一口多單；反之則不做任何部位變動。

3. 每月近遠月期貨交割時，會固定換倉繼續持有部位。

若持有期間超過一天，則策略訊號出現後必須持有至少該期間的長度，滿足該期間長度後才可轉換持倉部位，本研究建構之交易策略持有天數為一天。

本節呈現本研究的實證結果。我們首先提供資料描述與敘述統計，並詳述建構交易策略的過程與參數配置，並衡量各特徵的重要性。透過篩選出可能干擾建模過程的特徵，我們辨認出爭議變數，並進而建構 `Perturbed_R_F` 和 `unPerturbed_R_F` 兩個策略。我們會從混淆矩陣與績效指標來檢視兩策略的差異，並在最後分割為樣本內外來驗證模型預測能力好壞，再對其做詳細的績效衡量。

4.1 資料描述與敘述統計

本次研究標的為台灣期貨交易所發行的台灣證券交易所股價指數近月期貨，而採用的資料期間為 2007/07/03 到 2018/02/21，以 2007 年 7 月為資料期間起始點的原因為，建構交易策略的重要特徵－外資期貨未平倉量之最早可取得時間為 2007/07/02，但需延遲一天方可使用於交易策略，故以 2007/07/03 為起始日。這段期間共有 2,634 筆日資料，其中共有一個預測目標－台股期貨漲跌標籤，與 29 個特徵，分別為 3-14 日 MA 與 RSI 指標、近遠月期貨交易量、現貨交易量、期貨外資未平倉量與買賣權未平倉比率，資料來源為台灣期貨交易所與台灣證券交易所。

表 2 為修正前台股期貨日報酬率的敘述統計，透過本表可發現這筆資料有非常態及厚尾的特性，因此極端事件發生頻率會較常態高出許多；另外，其

表 2 修正前台股期貨價格報酬率敘述統計

	平均數	中位數	最小值	最大值	標準差	偏態	峰態
報酬率	0.02%	0.00%	-8.40%	7.14%	1.38%	-0.25	8.41

註: 此表資料為修正前台股期貨價格報酬率, 期間為 2007/07/03 到 2018/02/21, 可以發現資料有非常態且厚尾特性, 另外, 雖然平均數與中位數皆近於零, 但屬左偏分布。

平均數與中位數雖皆近於零, 但為左偏分布, 代表日報酬率密度較集中於正數。

另外, 每個月期貨交割日收盤價與其隔日開盤價有段跳空, 因此減去每個月跳空的修正後台股期貨價格才是這段時間真正的價格走勢, 而此累積報酬率接近 90%, 且此區間還包含了 2008 年的全球性金融海嘯。我們用以訓練模型的資料區間為 2007/07/03 到 2015/12/31, 共 2,113 筆日資料, 在此稱為「訓練區間」; 2016/01/04 到 2018/02/21 在此稱為「測試區間」, 共 521 筆資料。

4.2 模型參數配置

本研究的漲跌訊號是由 d 天後的漲跌生成, 因此單看隔日漲跌會使過多短期雜訊被帶入。有鑒於此, 本文的訊號由兩日後的修正後價格漲跌生成, 以過濾掉短期價格不穩定的特性。

表 3 和表 4 為本文使用之訊號於不同區間的漲跌日數、漲跌比例、卡馬比率與其他參考績效指標。如研究方法中所述, 卡馬比率為本文呈現績效的衡量指標, 其值愈大代表累積總報酬所負擔的風險愈小, 或是每單位風險的下方風險承擔能獲得更高的累積總報酬。Omega、索丁諾比率以及潛在上漲比率的概念皆類似, 都是獲利指標與風險指標的比率, 故其解釋方法與卡馬比率相同; 赫斯特指數是用以驗證時間序列資料是否具有持續性的指標, 當其值大於 0.5 代表該時間序列具有持續性, 小於 0.5 則具有反持續性, 等於 0.5 則顯現該時間序列資料游走方式為隨機漫步。在本文中, 我們保持對於其具有持續性與否之於投資績效的客觀立場。意即, 我們將以赫斯特指數客觀呈現不同頻率之下(以一個交易日、一個月、一年等為單位)策略損益的時間序列資料於各頻率段落之間的相關性。為符合真實情況, 本文提及之策略的績效指標皆已扣除手續費。

表 3 資料漲跌日數與比例

	上漲日數	下跌日數	上漲比例(%)	下跌比例(%)
全部區間	1,437	1,197	54.56	45.44
訓練區間	1,126	987	53.29	46.71
測試區間	311	210	59.69	40.31

註: 此表為兩日後修正後台股期貨價格的資料描述。全部區間為2007/07/03 到2018/02/21, 訓練區間為2007/07/03 到2015/12/31, 測試區間為2016/01/04 到2018/02/21。如同敘述統計, 台股期貨不論在哪個區間皆有上漲的現象, 除此之外, 測試區間的上漲比例與大盤績效遠超過全部區間。

表 4 修正後台股期貨各區間績效指標

	卡馬比率	Omega	索丁諾比率	潛在上漲比率	赫斯特指數
全部區間	0.84	1.06	0.03	0.49	0.55
訓練區間	0.32	1.02	0.01	0.48	0.58
測試區間	3.82	1.27	0.12	0.57	0.49

註: 此表為修正後台股期貨價格的績效, 區間與結論與上表相同。

透過表 3 與表 4 可以發現, 若於訓練區間全部做多, 其錯誤率為 46.71%, 卡馬比率為 0.32, Omega 為 1.02, 索丁諾比率為 0.01, 潛在上漲比率為 0.48, 赫斯特指數為 0.58, 這些結果顯示台股期貨在該區間有不錯的績效表現。測試區間的大盤表現更為良好, 有高達六成的時間是上漲, 且卡馬比率可以達到 3.82。

圖 1 橫軸為投入特徵數目 (m), 代表隨機森林每次建樹要抽取的特徵數目, 是模型最重要的參數之一, 另一則為決策樹總數量 ($ntree$), 其會影響模型收斂程度與運算時間, 在此設定為 10,000, 代表隨機森林內共存在 10,000 棵決策樹。圖 1 橫軸的投入特徵數目為 2 到 27 個, 縱軸部分則有兩軸, 分別為卡馬比率與袋外錯誤率。如同前述介紹, 袋外錯誤率為隨機森林獨有的錯誤率衡量方法, 其為模型誤差之不偏估計, 因此其為模型預測能力的衡量指標, 而非配適程度。透過圖 1 可以發現, 若投入特徵數目到達一定程度, 袋外錯誤率與卡馬比率就能維持在不錯的水準; 若投入特徵數目大於 6, 袋外錯誤率就能落在足夠低的區間。投入特徵數目為 23 時袋外錯誤率為全部參數中最

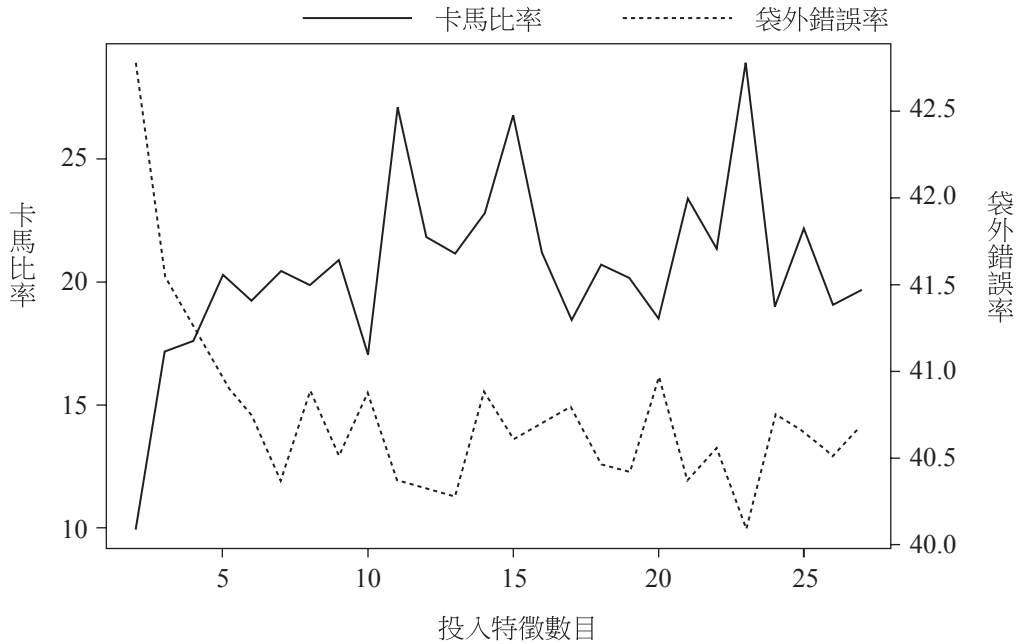


圖 1 各投入特徵數目之卡馬比率與袋外錯誤率

低, 卡馬比率則為最高, 因此, 我們選用參數 23 來進行策略的建構, 而其袋外錯誤率為 40.09%, 卡馬比率為 28.95, 代表其能夠壓低 6% 左右的錯誤率, 且樣本內策略績效能夠遠遠勝過大盤。這即是本文的 *Perturbed_R_F* 策略, 它的參數配置包含了絕多數的資料特徵。

圖 2 橫軸為第 1 棵到第 10,000 棵樹, 縱軸為該決策樹數量下的袋外錯誤率, 其代表在不同決策樹數量下 *Perturbed_R_F* 策略的收斂程度。由圖 2 結果可以發現, 當決策樹數量增長到大約 4,000 棵以上時, 袋外錯誤率就會收斂, 其也代表 *Perturbed_R_F* 策略之參數設定為 10,000 是足夠穩定的。

除了模型穩定收斂以外, 透過了解這 29 個特徵對模型的貢獻程度, 或許能夠使 *Perturbed_R_F* 策略去蕪存菁, 以更少的特徵達到同樣水準的分類能力與績效。在此構面上, 如同研究方法所述, 我們衡量隨機森林變數重要性的方法是透過平均準確度減少 (MDA) 與平均吉尼不純度減少 (MDG) 的求算, *Perturbed_R_F* 策略的變數重要性如圖 3 所示。

由圖 3 我們發現有些變數具有無庸置疑的重要性, 然而更有趣的是也有部分變數依兩種衡量方法得到了截然不同的評定結果。具重要性的變數有

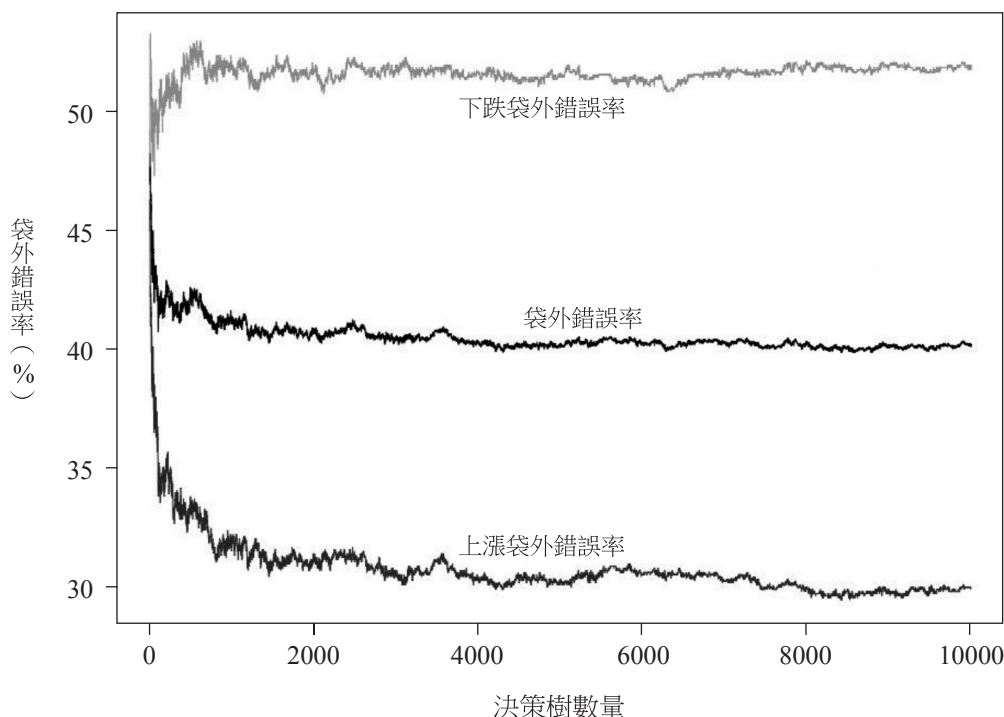


圖 2 Perturbed_R_F 策略於不同決策樹數量之下的袋外錯誤率

部分籌碼面指標, 像是買賣權未平倉比率、外資未平倉量等; 部分技術面指標, 如遠月交易量、部分天期的RSI與MA。籌碼面指標能夠衡量市場情緒與水位, 且台灣市場較不效率, 容易受市場情緒與熱錢影響, 因此加入籌碼面指標有很高的可能性可以提昇模型績效與預測能力。

此外, 圖 3 也顯示, 無論以哪種方法衡量, 買賣權未平倉比率皆為最重要的特徵, 而外資未平倉量的重要性也排在所有特徵中的前五名, 這便從變數重要性的角度驗證了籌碼面指標的不可或缺。另外, 本研究認為技術面指標能找出動能與市場趨勢所在, 並提昇資訊量的品質, 因此將其加入模型。從圖 3 可以發現, 遠月交易量與較長天期RSI都屬於高重要性的特徵, 而中等天期的MA也有一定程度的重要性。

其中當然也有重要性較低的變數, 如長天期MA與部分天期RSI, 但使用隨機森林而非決策樹建構模型, 便是希望讓一些資訊含量較低的變數也能發

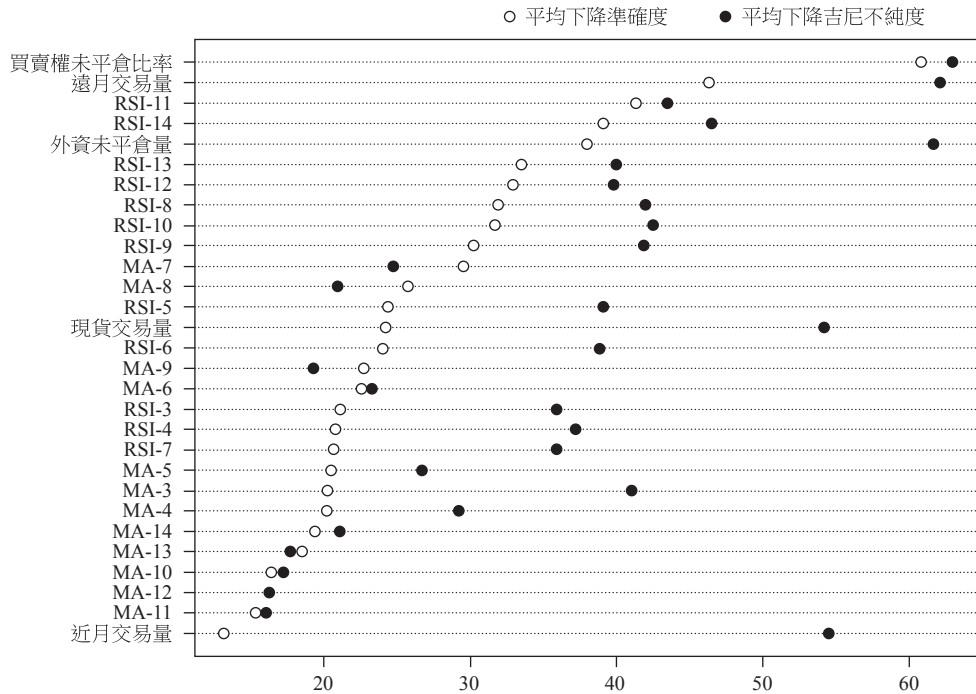


圖3 Perturbed_R_F 策略變數重要性

揮穩定模型的功用(依照吉尼不純度, 資訊含量低的變數便不會成為節點, 也因此少了部分資訊), 因此並非變數重要性低就要將其拿出模型。

本研究中, 兩種衡量方法判斷結果一致的變數就會留在模型內, 而會產生爭議的是判斷結果相差極大的變數, 尤其是MDG 高但MDA 低的變數。MDG 高, 代表其在隨機森林內建構單一決策樹時, 只要選入該特徵, 便會成為上層節點。隨機森林最大的特點是建構模型過程的隨機性, 因此計算變數重要性時, 除了有吉尼不純度可以當做標準以外, 還能夠將建模過程中計算的袋外錯誤率轉為另一種標準, 而這種標準更偏向全域解, 因此當MDA 非常低, 代表特徵對模型來說貢獻度不大, 但若其MDG 非常高, 代表此特徵可能只是區域解, 它會是建構決策樹的過程中優先被考慮的節點, 但可能對模型整體沒有太大幫助, 因此放在模型內反而會使個別決策樹的建構過程有偏誤, 導致模型預測能力下降。

為了找出這種現象特別明顯的特徵, 本研究首先將全部特徵的MDA 與MDG 轉換成值域為0 到1 的調整後MDA 與MDG, 再取其差做為盒鬚圖(box-

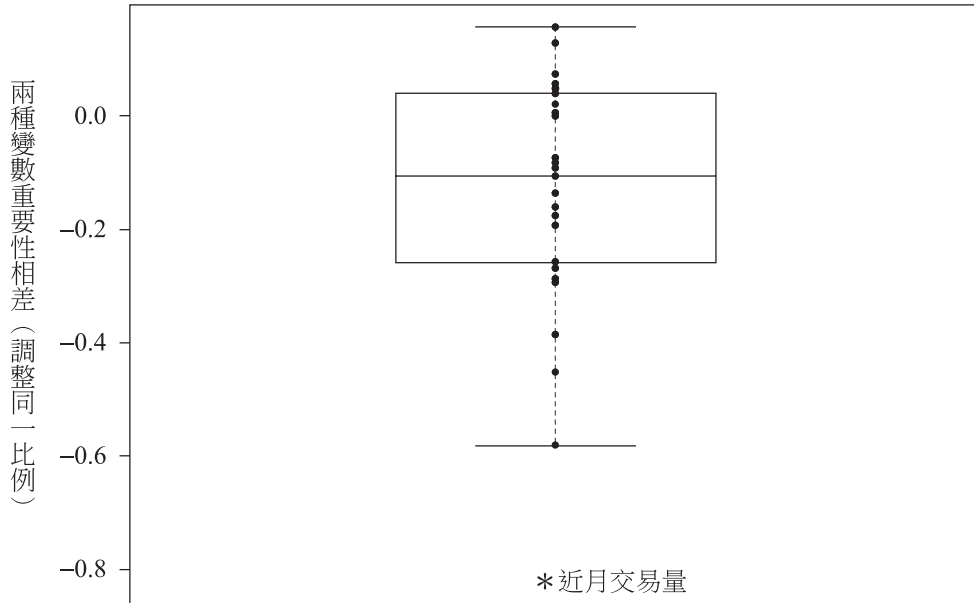


圖 4 Perturbed_R_F 策略不同變數重要性的差之盒鬚圖

plot), 如圖 4 所示。⁶ 透過圖 4 可以發現, 只有近月交易量落於盒鬚圖之內籬 (inner fence) 與外籬 (outer fence) 間, 代表它屬於溫和離群值 (mild outlier), 因此判斷其為兩種變數重要性差異過大的特徵。我們因而辨認近月交易量是為本模型的爭議變數。

設定決策樹總數量為 10,000, 重新建構之模型與原始模型在不同投入特徵數目下之卡馬比率與袋外錯誤率如圖 5 所示。首先, 若將焦點放在袋外錯誤率, 可以發現移除近月交易量後確實能將錯誤率控制在更低水位, 最低甚至能到 39.85%, 確實解決了上述提出之問題。從績效角度出發, 又可以發現重新建構的交易策略之卡馬比率在各參數之下, 明顯比有放入近月交易量的策略低, 代表近月交易量被拿出模型後, 雖然能達到更低的錯誤率, 但卻會更難找到關鍵節點, 導致策略績效降低。如前述, 交易策略之焦點應為績效衡量, 而分類能力的重要性來自於其與策略績效的相關性, 因此提昇分類能力便能間接提昇績效。當上述情況發生時, 可以合理推測近月交易量雖不能讓

⁶ 本文另有嘗試另外兩種離群值偵測方法, 分別為 Carling (2000) 之盒鬚圖與 Schwertman and de Silva (2007) 之盒鬚圖, 此兩種方法得出內籬上下界較寬, 對離群值敏感度較低, 然而與本文的盒鬚圖所辨認出的爭議變數一致, 此兩種方法所偵測出的離群值仍為近月交易量。

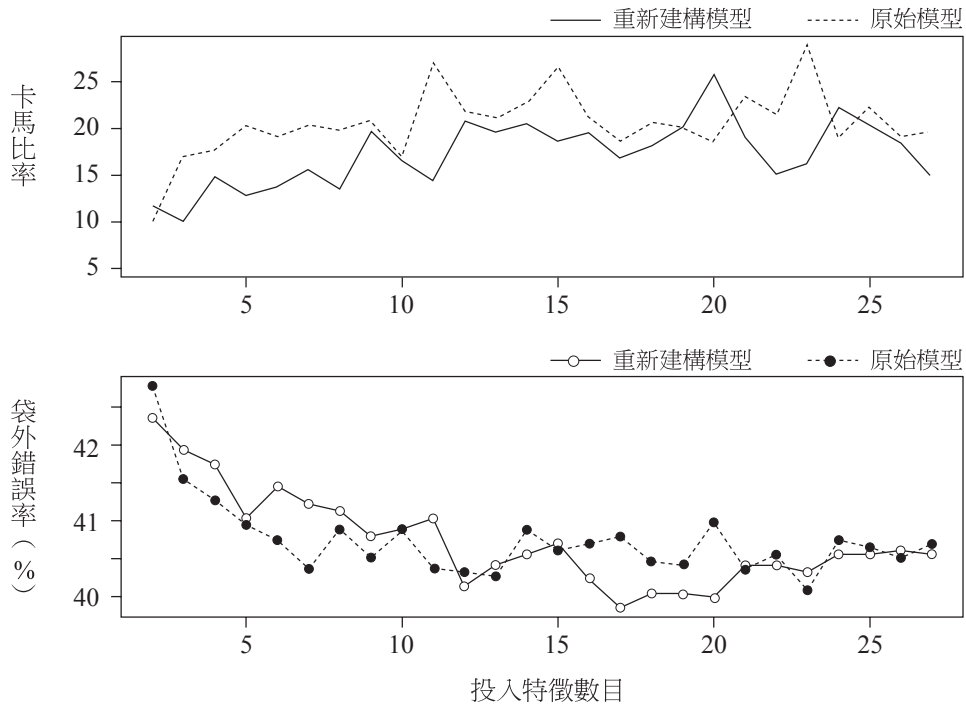


圖 5 各投入特徵數目下重新建構與原始模型之卡馬比率與袋外錯誤率比較

漲跌更容易被分類, 但可以更容易抓出波段特徵, 因此有利於績效成長。由於投入特徵數目為 20 時績效高且錯誤率低, 故選擇其來建構 unPerturbed_R_F 策略。

圖 6 為 unPerturbed_R_F 策略於第 1 棵到第 10,000 棵樹下的袋外錯誤率, 橫軸為決策樹數量, 縱軸為該數量下的袋外錯誤率, 代表在不同決策樹數量下隨機森林模型的收斂程度。由圖 6 結果可以發現, 當決策樹數量增長到大約 6,000 時, 袋外錯誤率就會收斂, 也代表 unPerturbed_R_F 策略中參數設定為 10,000 足以讓分類結果穩定。

總結上述的分析, 本文所辨認出的爭議變數為近月交易量, 因而所建構之 Perturbed_R_F 策略的具體特徵為: 3-14 日 MA 與 RSI 指標、近遠月期貨交易量、現貨交易量、期貨外資未平倉量與買賣權未平倉比率。在另一方面, 所建構之 unPerturbed_R_F 策略的具體特徵則為: 3-14 日 MA 與 RSI 指標、遠月期貨交易量、現貨交易量、期貨外資未平倉量與買賣權未平倉比率。本文將比較 Perturbed_R_F 和 unPerturbed_R_F 兩策略與大盤之間的績效差異, 再

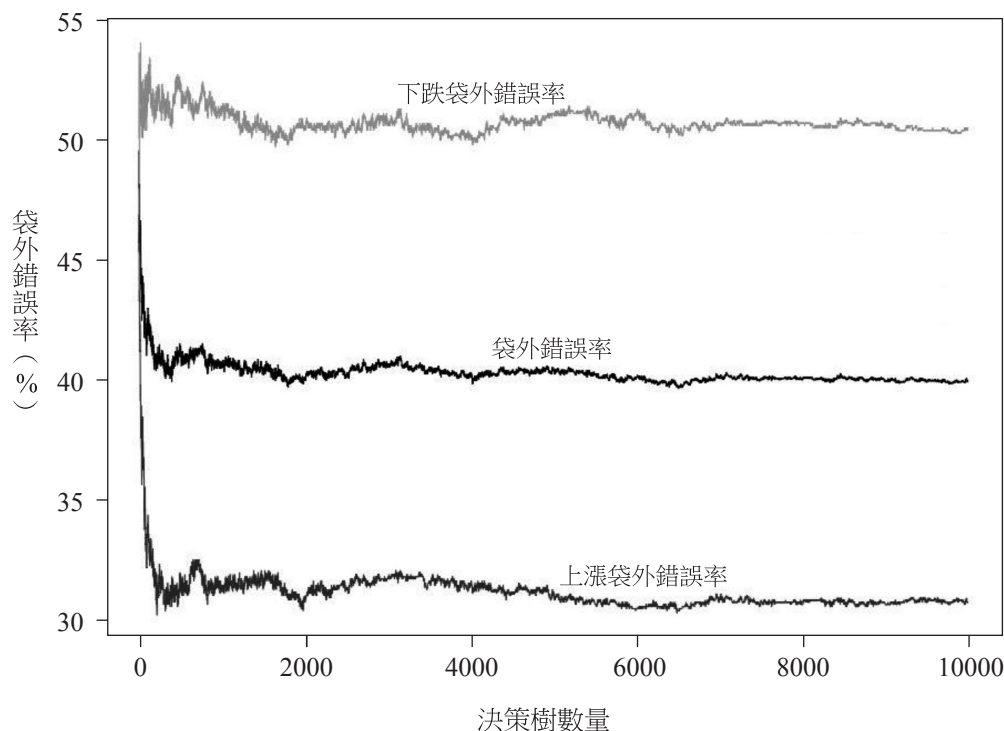


圖 6 unPerturbed_R_F 策略於不同決策樹數量之下的袋外錯誤率

去分析兩策略的參數配置之於模型預測能力與實際策略績效衡量結果上的影響。

4.3 訓練區間的模型預測能力和策略績效檢視

透過袋外分類結果的混淆矩陣與各比率, 可以更詳細地衡量模型的預測能力。首先, 盛行率即代表訓練區間上漲比率, 而準確率是袋外錯誤率的相對值。表 5 的結果中所示, 本文之隨機森林模型的預測能力呈現了 53.29% 盛行率和 59.91% 準確率, 透過這兩個比率可以得知模型預測能力提昇了大約 6%。佐以其餘比率, 即可更精確地判斷整體模型預測能力的各個面向。精確率與召回率(陰性預測率與特異率)之間存在互抵沖銷的關係, 當模型擁有高精確率代表其正確預測的比率很高, 但若其預測的結果佔全部資料極低比例, 反而代表模型預測能力不足。反之, 當一個模型擁有高召回率, 即代表多

表 5 Perturbed_R_F 策略於訓練區間之袋外混淆矩陣

	真實為上漲	真實為下跌	
預測為上漲	790	511	精確率 = 60.72%
預測為下跌	336	476	陰性預測率 = 58.62%
	召回率 = 70.16%	特異率 = 48.23%	
	盛行率 = 53.29%	準確率 = 59.91%	$F_1P = 65.10\%$ $F_1N = 52.92\%$

註: 混淆矩陣於研究方法有詳細說明。訓練區間為 2007/07/03 到 2015/12/31。Perturbed_R_F 策略於訓練區間內的預測結果中, 精確率與陰性預測率皆位於大約六成的水準, 召回率能達到七成, 而特異率僅能達到五成左右, 故 Perturbed_R_F 策略預測上漲能力優於預測下跌能力。

表 6 unPerturbed_R_F 策略於訓練區間之袋外混淆矩陣

	真實為上漲	真實為下跌	
預測為上漲	779	498	精確率 = 61.00%
預測為下跌	347	489	陰性預測率 = 58.49%
	召回率 = 69.18%	特異率 = 49.54%	
	盛行率 = 53.29%	準確率 = 60.09%	$F_1P = 64.84\%$ $F_1N = 53.65\%$

註: 混淆矩陣於研究方法有詳細說明。訓練區間為 2007/07/03 到 2015/12/31。unPerturbed_R_F 策略於訓練區間的預測結果中, 精確率與陰性預測率皆位於大約六成的水準, 但召回率能達到七成, 而特異率僅能達到五成左右, 故預測上漲能力優於預測下跌, 但若與 Perturbed_R_F 策略比對, 可以發現 unPerturbed_R_F 策略預測下跌能力較好, 且準確率亦略高。

數的上漲樣本有被預測到, 但若某模型預測全部資料為上漲, 則其不存在預測能力。因此, 在評量模型預測能力上, 這兩種比率必須一併被參考。

透過表 5 可以發現, 雖然精確率與陰性預測率都落於六成左右的水準, 但預測上漲方面召回率能達到七成, 而預測下跌方面, 特異率僅能達到五成左右的水準。上述結果代表模型預測上漲與下跌的準確程度接近, 但被預測為下跌資料的比例較低, 此即代表 Perturbed_R_F 策略預測結果以上漲為多數。另外, 我們更進一步採用本文之於陽性(上漲)資料預測和陰性(下跌)資料預測所建立的 F_1 值綜合評量標準, 即本文的第(22)式及(23)式中所定

義的 F_1P 值與 F_1N 值，來個別衡量模型在預測上漲與下跌的能力。表 5 的結果顯示，模型在 F_1P 與 F_1N 值的表現都不錯，但 F_1P 明顯高於 F_1N ，顯示 Perturbed_R_F 策略中所採用的隨機森林模型其預測上漲的能力優於預測下跌。就我們敘述統計所觀察到的台股期貨長期多頭市場特性，在此市場做空往往較為不利，而依上述結果顯示，預測上漲的能力優於預測下跌的 Perturbed_R_F 策略於訓練區間更能捕捉到市場真正的方向。

從績效角度出發，透過表 7 可以發現訓練區間 Perturbed_R_F 與 unPerturbed_R_F 策略之卡馬比率分別為 28.95 與 25.99，而 unPerturbed_R_F 策略不論是獲利或是風險控管皆劣於 Perturbed_R_F 策略，代表 unPerturbed_R_F 策略雖然錯誤率較低，但降低的幅度不足以抵消近月交易量的重要性。本文在敘述統計有提到，台股期貨長期屬於多頭市場，考慮到這樣的市場特性，便可以解釋為何 unPerturbed_R_F 策略的錯誤率下降，但績效也跟著下降。unPerturbed_R_F 策略錯誤率下降主要貢獻於預測下跌的能力提昇，但同時也犧牲些許上漲預測能力，代表策略較不順應於市場特性，故導致 unPerturbed_R_F 策略整體績效下降。除此之外，相較於 Perturbed_R_F 策略，其更容易預測下跌，且其陰性預測率並沒有提昇，故預測下跌錯誤的絕對次數會上升，使空單績效更差。

另一方面，台股期貨之卡馬比率僅為 0.32，且兩策略在其餘績效指標的部分皆勝過台股期貨。與混淆矩陣得出之結果相異的是，若將策略績效拆為多單與空單兩部分，兩策略之空單卡馬比率與其他績效指標更勝於多單，但若以袋外混淆矩陣的角度來看，上漲類別之預測結果應當優於下跌，因此我們必須進一步拆解卡馬比率才能更了解多空單在績效與性質上的差異。

將卡馬比率拆解為累積損益與最大歷史跌幅兩部分，其結果如表 7「多單」與「空單」兩欄所示。可以發現，Perturbed_R_F 策略於訓練區間的多單累積損益能達到 17,970 點，超越空單的 16,072 點，但多單最大跌幅為 1,537 點，遠大於空單的 837 點；unPerturbed_R_F 策略之多單累積損益能達到 16,631 點，超越空單的 14,733 點，但多單最大跌幅為 1,537 點，遠大於空單的 837 點。上述結果代表本研究建構之兩策略，其多單風險較空單高但報酬也高，也隱含了此兩策略空單交易風險管控較好，且策略全部交易的卡馬比率能分別達到 28.95 與 25.99，代表多空單有達到互補的作用。

由於兩策略之訓練區間多空單交易損益跌幅圖相近，故僅以 Perturbed_R_F 策略為例，如圖 7 所示。上面的部分為多空單於訓練區間之損益曲線；中間的

表 7 Perturbed_R_F 和 unPerturbed_R_F 策略與大盤於訓練區間內績效比較

		所有交易	多單	空單
累積損益 (點數)	Perturbed_R_F	34,042	17,970	16,072
	unPerturbed_R_F	31,364	16,631	14,733
最大跌幅 (點數)	Perturbed_R_F	1,176	1,537	837
	unPerturbed_R_F	1,207	1,614	1,155
卡馬比率	Perturbed_R_F	28.95	11.69	19.20
	unPerturbed_R_F	25.99	10.30	12.76
	台股期貨	0.32	N/A	N/A
Omega	Perturbed_R_F	1.57	1.46	1.69
	unPerturbed_R_F	1.51	1.43	1.59
	台股期貨	1.02	N/A	N/A
索丁諾比率	Perturbed_R_F	0.25	0.16	0.22
	unPerturbed_R_F	0.23	0.15	0.19
	台股期貨	0.01	N/A	N/A
潛在上漲比率	Perturbed_R_F	0.70	0.52	0.54
	unPerturbed_R_F	0.68	0.50	0.51
	台股期貨	0.48	N/A	N/A
赫斯特指數	Perturbed_R_F	0.65	0.57	0.72
	unPerturbed_R_F	0.65	0.55	0.68
	台股期貨	0.58	N/A	N/A

註: 訓練區間為 2007/07/03 到 2015/12/31。上述結果代表兩策略的多單風險較空單高, 但報酬也高, 也隱含此兩策略空單交易風險管控較好。另外, unPerturbed_R_F 策略於訓練區間不論獲利能力或下檔風險控制皆劣於 Perturbed_R_F 策略, 若將卡馬比率與其他績效指標分為多單與空單, 結果仍然相同。

部分為多單於訓練區間之跌幅與最大跌幅; 下面的部分為空單於訓練區間之跌幅與最大跌幅。此外, 灰色區塊代表國家發展委員會之台灣景氣循環峰谷認定中的景氣循環衰退期間, 其中, 第一次衰退期間為 2008 年 3 月至 2009 年 2 月; 第二次為 2011 年 2 月至 2012 年 1 月; 第三次為 2014 年 10 月至 2016 年 2 月。

透過圖 7 可以發現, 空單損益於每個衰退期間都有明顯地增長, 跌幅則有縮小的現象, 而在衰退期間以外的部分, 空單損益便沒有太大變動; 多單

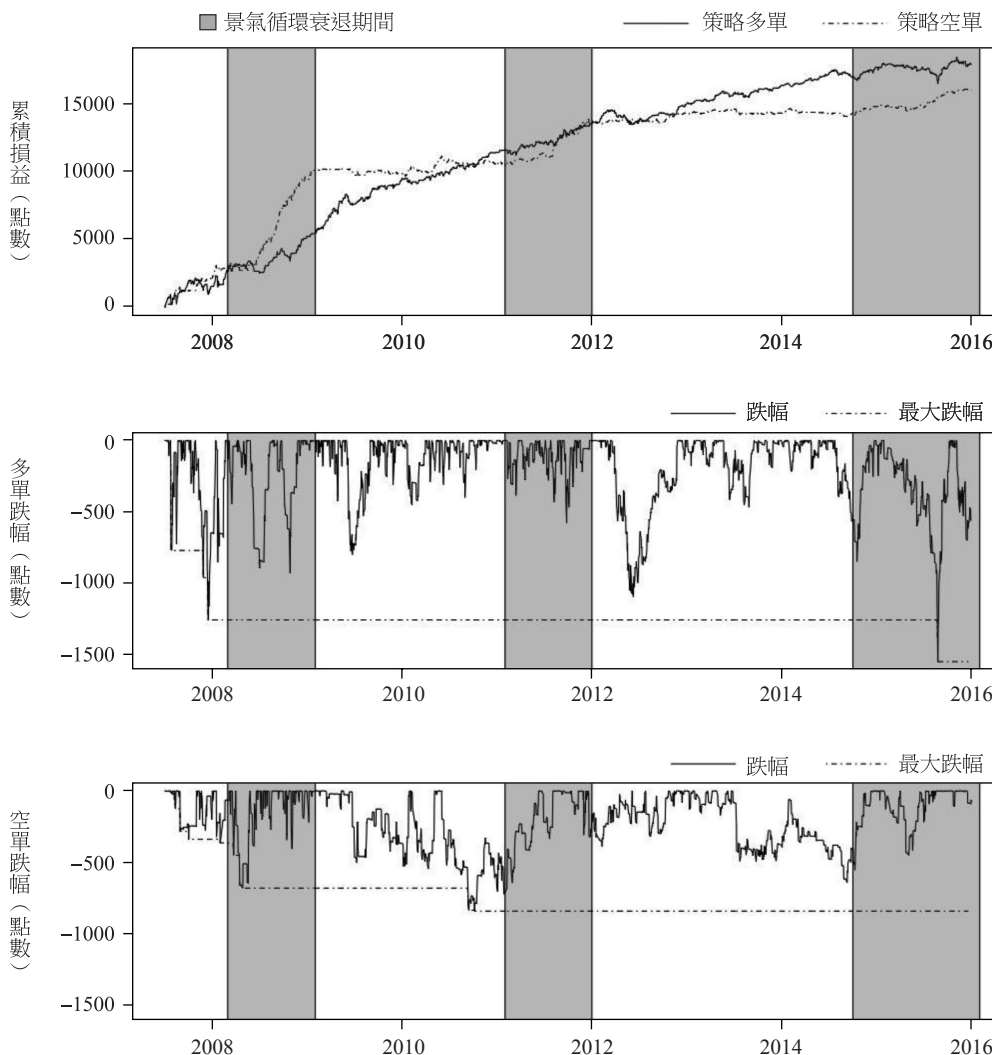


圖 7 Perturbed_RF 策略多空單交易於訓練區間內損益跌幅比較

損益的部分則在訓練區間持續增長,但在第一次與第三次衰退期間內其跌幅波動極大,第二次衰退期間則沒有明顯改變。若以事件來解釋,第一次衰退主因為美國次貸風暴引起的全球系統性衰退;第二次為歐債危機衝擊全球景氣,再加上民間投資意願降低導致;第三次為中國大陸供應鏈在地化導致的經濟衰退,可以發現三次中以第二次與台灣金融市場的關係程度最低,再加上台股期貨動盪程度最小,因此第二次衰退對策略影響程度可能不大。

綜合多單與空單的互補性後, Perturbed_R_F 策略與台股期貨之損益與跌幅如圖 8 所示。可以發現原先衰退期間的極大波動因多空互補而抵消, 使得策略卡馬比率得以上升到 28.95, 並完全在獲利與風險控管上擊敗大盤。

4.4 測試區間的模型預測能力及策略績效檢視

本部分將利用 2016/01/04 到 2018/02/21 共 521 個樣本, 探究爭議變數所產生的屬性擾動對於模型的預測能力與交易策略績效的影響。第一部分會先對交易策略進行測試區間的績效驗證, 第二部分則以 Perturbed_R_F 與 unPerturbed_R_F 策略於全部區間的績效為主軸, 另外也會以不同的角度去檢視策略績效的可行性。

表 8 為 Perturbed_R_F 策略於測試區間的袋外混淆矩陣, 可以發現在此段區間內盛行率高達 59.69%, 代表台股期貨有六成的時間是在上漲的。以全球角度來看, MSCI 編制的世界指數 (MSCI all-country world equity index) 從 2016 年 11 月份至 2018 年 1 月份連漲了 15 個月, 為全球股市 30 年來最長多頭; 在國內, 台股期貨於此區段內屢創站上萬點最長記錄, 代表測試區間內台股確實跟隨著全球的腳步, 故對於策略來說更具挑戰性。以表 8 來說, 測試區間內的準確率為 59.12%, 略低於盛行率 0.57%, 但訓練區間內準確率為 59.91%, 這樣的結果與訓練區間並沒有太大差別, 因此本文認為 Perturbed_R_F 策略具有預測能力。

此外, 測試區間召回率達到 89.39%, 而精確率與訓練區間差不多, 讓 F_1P 值比訓練區間還高出 7.10%, 代表此區間預測上漲能力提昇; 相反來看, 測試區間特異率僅有 14.29%, 陰性預測率也低於訓練區間的水準, 故 F_1N 僅有 21.98%, 代表測試區間預測下跌的能力下降。測試區間與訓練區間存在這種程度的差別便是因為市場行情的改變, 由於測試區間的市場行情為強勁多頭, 故模型預測做多的能力會適應市場行情而傾向做多, 但這段區間便會更難去預測下跌。

同樣的, 我們對 unPerturbed_R_F 策略做測試區間的模型驗證, 如表 9 所示。將其與 Perturbed_R_F 策略比較, 可以發現 F_1N 比 Perturbed_R_F 策略高 3.46%; F_1P 僅低於 Perturbed_R_F 策略 0.1%; 準確率略高於 Perturbed_R_F 策略之結果, 但仍然低於測試區間的盛行率。從上述結果可以得知, 預測類別這

表 8 Perturbed_R_F 策略於測試區間之袋外混淆矩陣

	真實為上漲	真實為下跌	
預測為上漲	278	180	精確率 = 60.70%
預測為下跌	33	30	陰性預測率 = 47.62%
	召回率 = 89.39%	特異率 = 14.29%	
	盛行率 = 59.69%	準確率 = 59.12%	$F_1P = 72.30\%$ $F_1N = 21.98\%$

註：混淆矩陣於研究方法有詳細的計算方法。測試區間為 2016/01/04 到 2018/02/21。測試區間的預測結果中，準確率為 59.12%，略低於盛行率，但由於其與訓練區間準確率位於同樣水準，故本研究認為 Perturbed_R_F 策略具有足夠預測能力。

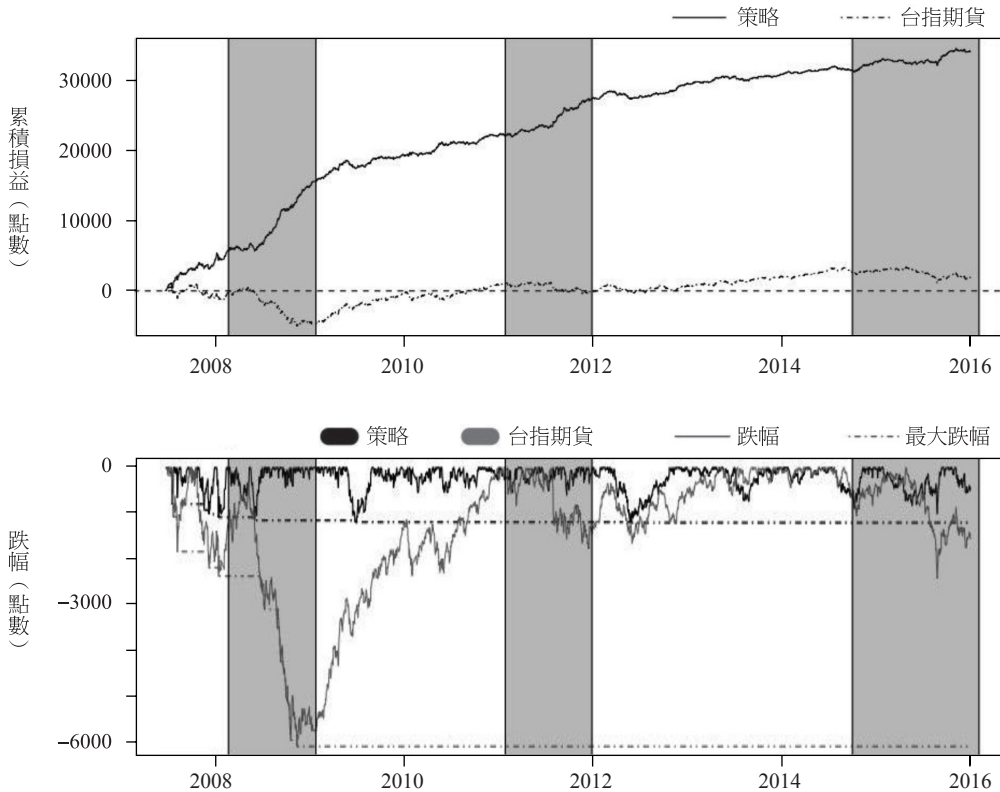


圖 8 Perturbed_R_F 策略與台股期貨於訓練區間內損益跌幅比較

表 9 unPerturbed_R_F 策略於測試區間之袋外混淆矩陣

	真實為上漲	真實為下跌	
預測為上漲	274	174	精確率 = 61.16%
預測為下跌	37	36	陰性預測率 = 49.32%
	召回率 = 88.10%	特異率 = 17.14%	
	盛行率 = 59.69%	準確率 = 59.50%	$F_1P = 72.20\%$ $F_1N = 25.44\%$

註: 混淆矩陣於研究方法有詳細的計算方法。測試區間為 2016/01/04 到 2018/02/21。測試區間的預測結果中, 準確率略低於盛行率, 但由於其與訓練區間準確率位於同樣水準, 故本研究認為 unPerturbed_R_F 策略具有足夠預測能力。與 Perturbed_R_F 策略比較, 可以發現 F_1N 的部分較其高 3.46%, 而 F_1P 略低 0.1%。

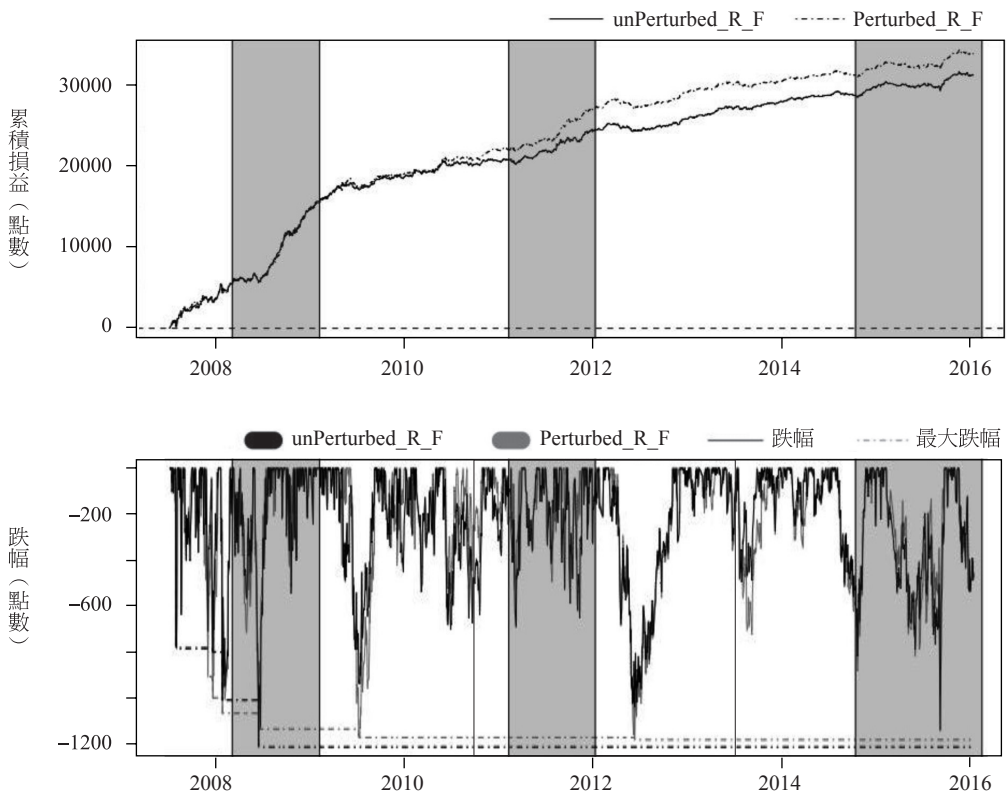


圖 9 Perturbed_R_F 與 unPerturbed_R_F 策略於訓練區間內損益跌幅比較

部分結論與訓練區間相同，推測 unPerturbed_R_F 策略辨別下跌類別的能力優於 Perturbed_R_F 策略，但辨別上漲類別的能力差不多，因此 unPerturbed_R_F 策略整體錯誤率略低於 Perturbed_R_F 策略。

測試區間內，Perturbed_R_F 策略、unPerturbed_R_F 策略與大盤之績效比較如表 10，可以發現台股期貨兩年兩個月的時間內卡馬比率可以達到 3.82（卡馬比率分子為累積報酬，故績效為正斜率之策略其卡馬比率應當隨時間增加），除了赫斯特指數以外，其餘績效指標表現亦為突出。然而，就算身處於強勁牛市，Perturbed_R_F 與 unPerturbed_R_F 策略之卡馬比率依然可以分別達到更好的 4.01 與 4.32，其他績效指標皆優於大盤，因此，此策略之績效不僅存在於樣本內，即使到了測試區間也經得起考驗。

有趣的是，unPerturbed_R_F 策略於測試區間的策略、多單與空單卡馬比率皆略優於 Perturbed_R_F 策略，且其他績效指標也有產生類似結果，但訓練區間的績效結果中，unPerturbed_R_F 策略之績效明顯低於 Perturbed_R_F 策略。本部分對此有一推論：依照前述種種數據所示，測試區間市場行情為強勁多頭市場，而本研究建構之模型會順應市場特性，讓做多比例大幅上升，此外，這樣的市場狀態亦會使不同交易策略之間相異度降低，使決策趨於一致。上述推論可依據幾組實證數據佐證，其一為 Perturbed_R_F 與 unPerturbed_R_F 策略在訓練區間與測試區間內預測上漲的比例，其二為 Perturbed_R_F 與 unPerturbed_R_F 策略在訓練區間與測試區間的決策相似程度。

訓練區間內，Perturbed_R_F 與 unPerturbed_R_F 策略預測上漲的比例分別為 61.57% 與 60.44%；測試區間內，Perturbed_R_F 與 unPerturbed_R_F 策略預測上漲的比例分別為 87.91% 與 85.99%，代表測試區間內確實會使決策做多比例大幅上升。此外，訓練區間內 Perturbed_R_F 與 unPerturbed_R_F 策略之決策相同比例為 94.51%；測試區間此比例提升到 97.31%，代表兩種策略決策確實趨於一致。綜上所述，由於測試區間的市場行情，導致決策做多比例上升，並使 Perturbed_R_F 與 unPerturbed_R_F 策略決策趨於一致，再加上 unPerturbed_R_F 策略之下跌預測能力較好，更能在此區間內正確地做出做空決策，故其測試區間之卡馬比率高於 Perturbed_R_F 策略。

另外，透過表 10 亦可發現此區間確實屬於多頭市場。一般而言，策略卡馬比率會高於多單與空單之卡馬比率，因為其會有互補效果使策略風險降

表 10 Perturbed_R_F、unPerturbed_R_F 策略與大盤於測試區間內之績效比較

		所有交易	多單	空單
累積損益(點數)	Perturbed_R_F	3,638	3,483	155
	unPerturbed_R_F	3,914	3,621	293
最大跌幅(點數)	Perturbed_R_F	907	843	334
	unPerturbed_R_F	907	843	338
卡馬比率	Perturbed_R_F	4.01	4.13	0.46
	unPerturbed_R_F	4.32	4.30	0.87
	台股期貨	3.82	N/A	N/A
Omega	Perturbed_R_F	1.32	1.32	1.24
	unPerturbed_R_F	1.35	1.34	1.31
	台股期貨	1.27	N/A	N/A
索丁諾比率	Perturbed_R_F	0.14	0.13	0.05
	unPerturbed_R_F	0.15	0.14	0.07
	台股期貨	0.12	N/A	N/A
潛在上漲比率	Perturbed_R_F	0.59	0.55	0.27
	unPerturbed_R_F	0.60	0.55	0.31
	台股期貨	0.57	N/A	N/A
赫斯特指數	Perturbed_R_F	0.52	0.50	0.52
	unPerturbed_R_F	0.52	0.48	0.59
	台股期貨	0.49	N/A	N/A

註: 卡馬比率分子為累積報酬, 故不同區間長度之卡馬比率不能相互比較。測試區間為 2016/01/04 到 2018/02/21, 在此區間內 Perturbed_R_F 與 unPerturbed_R_F 策略績效皆優於大盤, 策略獲利來源幾乎來自於多單, 但相反地多單亦承擔較多下檔風險。此外, unPerturbed_R_F 策略於測試區間的策略、多單與空單卡馬比率以及其他績效指標皆優於 Perturbed_R_F 策略。

低, 但 Perturbed_R_F 策略在測試區間的多單卡馬比率達到了 4.13, 略高於策略卡馬比率 4.01。拆解卡馬比率後可以發現, 多單部分的獲利佔了 95.74%, 且多單之最大跌幅比所有交易還少, 代表這段區間內做多非常容易獲利且風險偏低, 再次驗證此段區間的強勁多頭力道。可以觀察到的是, 雖然此區間大盤屬於強勁牛市, 但其赫斯特指數卻為 0.49, 代表在此區間內大盤的每日損益有些微反持續性的現象, 亦接近隨機漫步, 此數值代表大盤以不同頻率的

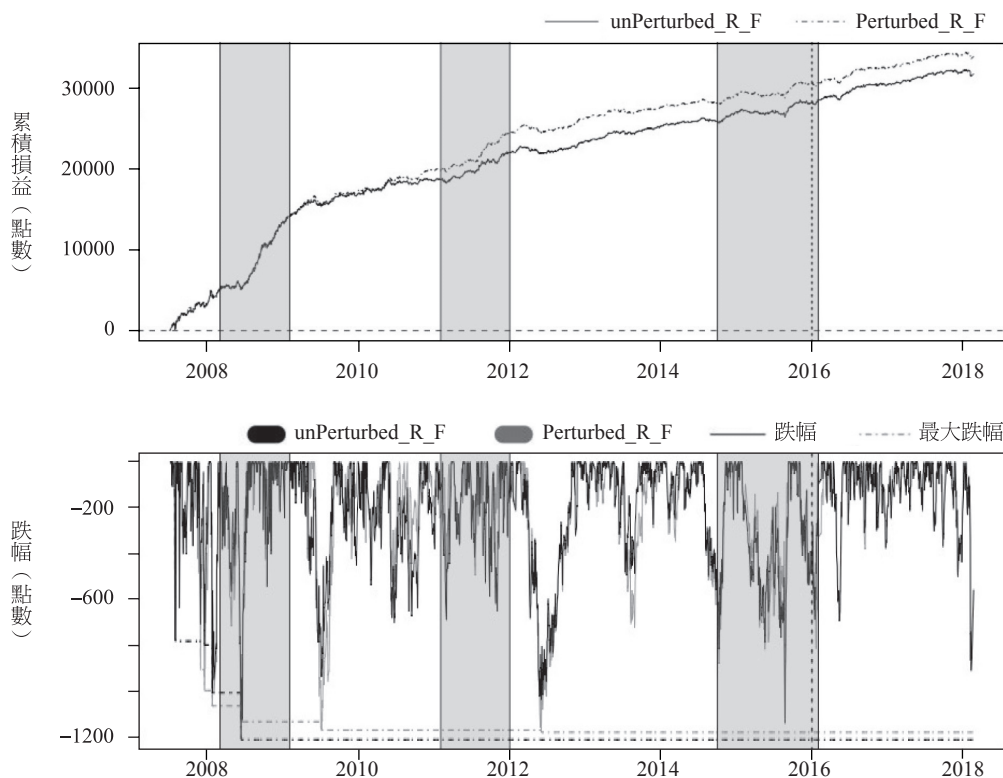


圖 10 Perturbed_R_F 與 unPerturbed_R_F 策略於全部區間之損益跌幅比較

時間去回溯績效，其績效游走於均值的運動無法掌握，但此一均值損益為正故績效良好，而類似的情況也出現在策略多單的部分。

4.5 全部測試區間模型預測能力與策略績效檢視

圖 10 為 Perturbed_R_F 與 unPerturbed_R_F 策略在全部區間的損益跌幅，垂直虛線左側為訓練區間，右側為測試區間。可以發現，兩策略於測試區間的累積損益非常穩定，並未突破最大跌幅，代表策略的最大跌幅足以代表策略風險，模型預測穩定度足夠，且兩策略在測試區間的累積損益並沒有太大差異。

兩策略與大盤於全部區間的績效指標比較如表 11，可以發現兩策略之績效皆遠優於大盤，⁷ 而 Perturbed_R_F 策略略優於 unPerturbed_R_F 策略。此

⁷ 除了以各種績效指標驗證 Perturbed_R_F、unPerturbed_R_F 與台股期貨的績效差異之外，本文亦有嘗試透過累積損益突破記錄高點之比較圖來再次驗證績效之差異。其中，圖的橫軸表累積損益突

表 11 Perturbed_R_F 和 unPerturbed_R_F 策略與大盤於全部區間內之績效比較

		所有交易	多單	空單
累積損益 (點數)	Perturbed_R_F	37,871	21,447	16,424
	unPerturbed_R_F	35,469	20,246	15,223
最大跌幅 (點數)	Perturbed_R_F	1,176	1,537	837
	unPerturbed_R_F	1,207	1,614	1,155
卡馬比率	Perturbed_R_F	32.20	13.95	19.62
	unPerturbed_R_F	29.39	12.54	13.18
	台股期貨	0.84	N/A	N/A
Omega	Perturbed_R_F	1.53	1.43	1.66
	unPerturbed_R_F	1.49	1.41	1.57
	台股期貨	1.06	N/A	N/A
索丁諾比率	Perturbed_R_F	0.24	0.16	0.20
	unPerturbed_R_F	0.22	0.15	0.17
	台股期貨	0.03	N/A	N/A
潛在上漲比率	Perturbed_R_F	0.68	0.52	0.50
	unPerturbed_R_F	0.66	0.51	0.47
	台股期貨	0.49	N/A	N/A
赫斯特指數	Perturbed_R_F	0.64	0.54	0.71
	unPerturbed_R_F	0.64	0.53	0.71
	台股期貨	0.55	N/A	N/A

註: 卡馬比率分子為累積報酬, 故不同區間長度之卡馬比率不能相互比較。全部區間為 2007/07/03 到 2018/02/21, 訓練區間為 2007/07/03 到 2015/12/31, 測試區間為 2016/01/04 到 2018/02/21。

外, 表 12 為兩策略於全部區間之年週期績效分析, 可以發現除了 2018 年, 兩策略累積損益皆為正, 且 2018 年因僅累積不滿兩個月的損益, 兩策略之實際交易次數皆僅 4 筆, 故其為負的意義不大, 代表 Perturbed_R_F 與 unPerturbed_R_F 策略的獲利能力穩定且持續, 而兩策略之最低歷年獲利也分別能達到 1,409

破記錄高點次數, 縱軸則為策略累積損益突破記錄高點的相隔天數之排序。所得的結果呈現, 無論訓練與測試區間, 兩策略之績效表現皆優於台股指數, 此結果呼應表 11 與前述之結論。細節之索取可聯繫本文作者。

表 12 Perturbed_R_F 與 unPerturbed_R_F 策略於全部區間之年週期分析

		累積損益 (點數)	交易 次數	平均持有天數 (天)	多單持有比例 (%)
Perturbed_R_F	2007	3,494	38	3.34	69.29
	2008	11,400	86	2.90	46.18
	2009	4,524	75	3.35	76.10
	2010	2,587	113	2.22	54.18
	2011	5,269	78	3.17	57.49
	2012	2,119	101	2.48	61.20
	2013	1,427	82	3.00	55.28
	2014	1,813	87	2.85	70.56
	2015	1,409	80	3.05	67.62
	2016	2,207	33	7.39	88.52
	2017	1,934	30	8.20	89.02
	2018	-312	4	7.75	74.19
		平均	3,550	76	3.26
	加總	37,871	807	N/A	N/A
unPerturbed_R_F	2007	3,616	40	3.18	67.72
	2008	11,442	86	2.90	46.18
	2009	3,960	87	2.89	74.10
	2010	1,623	115	2.18	52.59
	2011	3,949	86	2.87	52.63
	2012	1,533	105	2.38	61.60
	2013	2,071	78	3.15	56.10
	2014	1,967	89	2.79	70.97
	2015	1,203	80	3.05	65.57
	2016	2,605	41	5.95	84.84
	2017	1,812	32	7.69	88.62
	2018	-312	4	7.75	74.19
		平均	3,325	79	3.12
	加總	35,469	843	N/A	N/A

註: 2007 年績效僅記錄 2007 年 7 月到年底; 2018 年績效僅記錄 1、2 月份, 因此其不具太大意義。除了 2018 年, 其他年度累積損益皆為正, 且 Perturbed_R_F 與 unPerturbed_R_F 策略最低獲利分別能達到 1,409 點與 1,203 點(皆為 2015 年), 此外, 兩策略歷年累積損益前三高皆為台灣景氣循環衰退期間。最後, unPerturbed_R_F 策略交易次數普遍高於 Perturbed_R_F 策略, 而 Perturbed_R_F 策略平均持有天數與多單持有比率普遍高於 unPerturbed_R_F 策略。

點與 1,203 點之水準(皆為2015年)。值得注意的是,兩策略的累積損益前三高年度皆為台灣景氣循環的衰退期間,代表 Perturbed_R_F 與 unPerturbed_R_F 策略對於風險具有足夠良好的抑制能力,甚至能以之獲利。然而,兩個策略前三高的累積損益除了 2008 年沒有太大差距以外,2009 年與 2011 年仍存在些許差距。

本研究之策略執行若連續預測都為同一類別,則會繼續持有部位,而交易次數欄即是記錄該年度中變動持有部位的次數。可以發現,兩策略於訓練區間內的交易次數明顯高於測試區間,且平均持有天數(僅計算交易日)與多單持有比例測試區間明顯高於訓練區間,原因為測試區間屬於強勢多頭,所以交易策略預測上漲的比率也大幅上升,導致大部分時間都持有多單,也間接使交易次數減少許多。另外,unPerturbed_R_F 策略交易次數普遍高於 Perturbed_R_F 策略,其原因為 unPerturbed_R_F 策略更容易做空,導致交易頻率提高。最後,依照表12,交易頻率最低的年度平均一個月也會交易 2 次,因此投資人並不會長期持有相同部位,在實際執行交易時,投資人較不容易因短期波動而對策略失去信心。

圖 11 與圖 12 分別為 Perturbed_R_F 與 unPerturbed_R_F 策略於全部區間的每月損益,除了確認績效穩定性以外,此圖還能以交易實務的角度了解執行策略時投資人的下方風險胃納。透過此兩圖可以發現,除了 2008 年金融海嘯的極高異常報酬以外,兩策略的每月損益非常穩定。另外,兩策略皆僅有在 2012 年初與 2013 年中有連續三個月虧損的情況發生,代表策略波段虧損情況的發生率不高,因而在現實市場中有一定可被執行的價值。

從另一個角度來看,策略在歷史中曾發生連續三個月虧損的情況也可以給投資人做為另類風險指標。另外,更確切的風險揭露可以發現 Perturbed_R_F 策略的每月虧損極端情況較嚴重,兩策略的最大虧損月份都為 2009 年 6 月份,但 Perturbed_R_F 策略該月虧損 929 點,而 unPerturbed_R_F 策略該月虧損僅 597 點。

錯誤率為衡量模型預測能力的重要標準,而勝率則是檢視策略績效的重要指標。全部類別中有多少比例沒有被預測到即為錯誤率,而勝率的概念與之類似,然而其計算的是策略的「每筆交易」最後損益是否為正,因此勝率是最攸關投資人的績效衡量。在此層面上,由表 13 可發現,若去除掉交易次數僅 4 筆且涵蓋期間僅為兩個月的 2018 年, Perturbed_R_F 與 unPerturbed_R_F

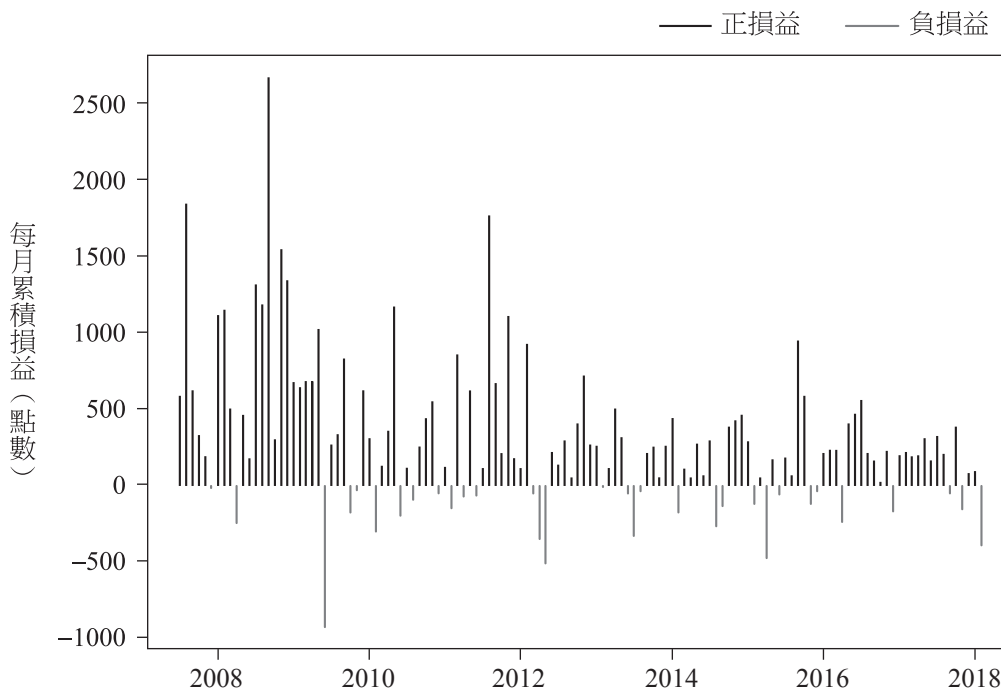


圖 11 Perturbed_R_F 策略於全部區間之每月損益

策略的勝率皆極高, 且其各自最低的勝率分別也有 57.32% 與 54.78%。此外, Perturbed_R_F 策略之多空單勝率差距不大, 並與市場情勢相符, 除了 2008 年 (第一次衰退期間)、2014 年到 2015 年 (第三次衰退期間) 以外, 多單勝率皆大於空單勝率; unPerturbed_R_F 策略亦有相同現象, 多空單勝率差距不大並與市場情勢相符, 除了 2007 到 2008 年 (第一次衰退期間) 與 2015 年 (第三次衰退期間) 以外, 多單勝率皆大於空單勝率。表 13 的結果呈現, 本研究建構之交易策略皆具有一定預測市場走勢能力, 並具有相當的獲利能力來為投資人創造穩定收益。

進一步比較兩策略的勝率, 可以發現 unPerturbed_R_F 策略平均勝率略高於 Perturbed_R_F 策略, 且其為多單勝率較高所導致的。除此之外, 2013 年及 2016 年 unPerturbed_R_F 策略勝率都比 Perturbed_R_F 策略高 5% 以上, 且 2013 年 unPerturbed_R_F 策略的多單勝率甚至高於 Perturbed_R_F 策略 13%, 而該年度累積損益的確是 unPerturbed_R_F 策略優於 Perturbed_R_F 策略最多的一年。

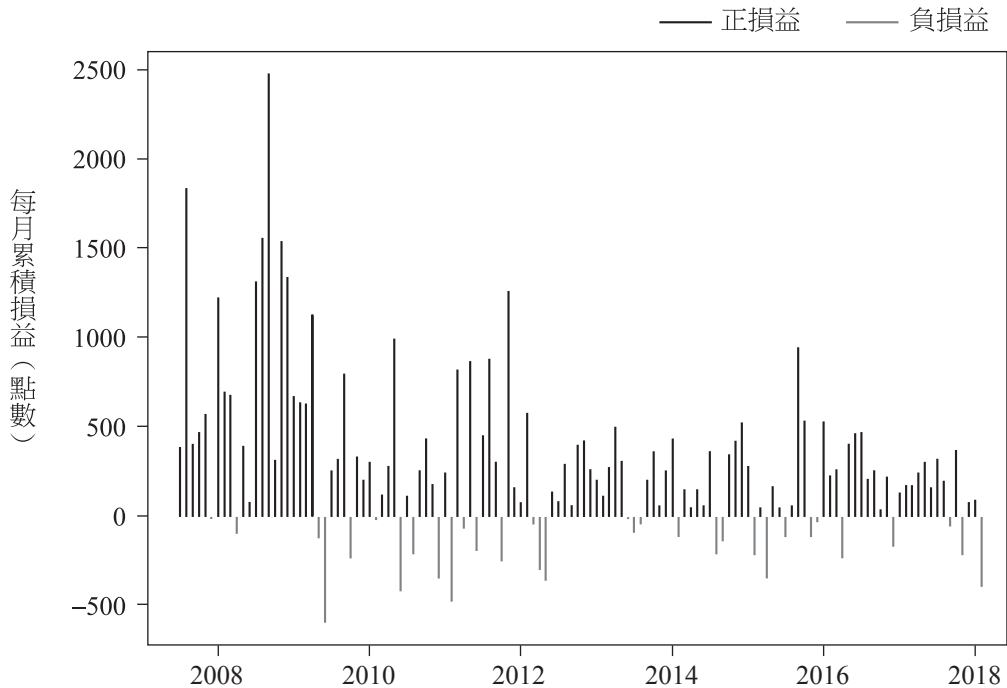


圖 12 unPerturbed_R_F 策略於全部區間之每月損益

2016 年則是因為空單勝率優於 Perturbed_R_F 策略 12.5% 導致, 而的確該年度累積損益也為 unPerturbed_R_F 策略優於 Perturbed_R_F 策略次多的年度。

表 14 為 Perturbed_R_F、unPerturbed_R_F 與台指期貨於全部區間之不同信心水準的歷史風險值與期望短缺, 其選定了 95% 與 99% 兩種常見的信心水準來計算風險值與期望短缺。其中, 若單純比較歷史風險值與期望短缺, Perturbed_R_F 策略相較於 unPerturbed_R_F 策略來的更小一些, 亦指若以每日損益的角度出發, Perturbed_R_F 在判斷錯誤時損失幅度較小, 但兩者之間僅有些微差距。若將台指期貨加入比較, 則可明顯地看出兩策略的風險皆遠小於單純持有台指期貨, 其風險指標相較於大盤, 皆有 10% 以上的改善, 最高甚至有達 18.81%, 代表本研究所建構之多數決學習台股指數期貨交易策略所取得的「風險承擔之下」的績效, 相較於大盤, 確實呈現著有效的下方風險沖銷。

表 13 Perturbed_R_F 與 unPerturbed_R_F 策略於全部區間之每年勝率分析

		勝率 (%)	多單勝率 (%)	空單勝率 (%)
Perturbed_R_F	2007	71.05	68.42	73.68
	2008	70.93	62.79	79.07
	2009	66.67	78.95	54.05
	2010	59.29	64.29	54.39
	2011	66.67	69.23	64.10
	2012	63.37	64.71	62.00
	2013	57.32	58.54	56.10
	2014	58.62	58.14	59.09
	2015	62.50	55.00	70.00
	2016	69.70	76.47	62.50
	2017	76.67	86.67	66.67
	2018	25.00	50.00	0.00
	平均	63.94	65.35	62.53
unPerturbed_R_F	2007	72.50	70.00	75.00
	2008	70.93	67.44	74.42
	2009	67.82	77.27	58.14
	2010	54.78	59.65	50.00
	2011	61.63	65.12	58.14
	2012	61.90	64.15	59.62
	2013	66.67	71.79	61.54
	2014	62.92	65.91	60.00
	2015	60.00	50.00	70.00
	2016	75.61	76.19	75.00
	2017	71.88	87.50	56.25
	2018	25.00	50.00	0.00
	平均	64.18	66.59	61.76

註：區間限制與上表相同。除卻 2018 年，Perturbed_R_F 與 unPerturbed_R_F 策略皆擁有極高勝率，且其各自最低的勝率分別也有 57.32%（2013 年）與 54.78%（2010 年）。

表 14 策略與台指期貨之歷史風險值與期望短缺

	歷史風險值		期望短缺	
	95%	99%	95%	99%
Perturbed_R_F	139	246	212	337
unPerturbed_R_F	140	263	215	344
台指期貨	166	303	253	387

5. 結論與後續研究建議

視金融資產價格走勢的預測問題為分類問題，並以隨機森林為基礎架構，本文建構輔以多數決機器學習的台灣加權指數期貨交易策略。相較於過去文獻，我們更著重於策略執行面，對於績效的討論與驗證倍加著墨。除此之外，本研究亦對變數重要性有深入的探討，並依照不同變數重要性衡量方法的差異，找出爭議變數，探討其對模型預測能力及策略投資績效的影響。

在模型預測能力上，於訓練區間內本文所建構的 Perturbed_R_F 與 unPerturbed_R_F 兩策略，其錯誤率分別為 40.09% 與 39.91%，而績效指標方面其卡馬比率分別為 28.95 與 25.99，而大盤在該區間僅有 0.32，因此兩策略皆大幅勝出。經由比較兩策略的參數配置，會發現取出近月交易量後的模型預測能力確實能夠提昇，但其績效卻會下降。進一步的分析結果發現，unPerturbed_R_F 策略之預測下跌能力確實有提昇，但是其所付出的代價是其預測上漲的能力。在測試區間，Perturbed_R_F 與 unPerturbed_R_F 策略仍然能分別以 4.01 與 4.32 的卡馬比率擊敗大盤的 3.82。此外，比較測試區間內兩策略之卡馬比率，可以發現 unPerturbed_R_F 策略勝過 Perturbed_R_F 策略。實證結果驗證，當策略所處環境為短期強勁多頭，其多單比例便會大幅上升，而導致策略鑒別度降低，但由於 unPerturbed_R_F 策略預測下跌能力較 Perturbed_R_F 策略為優，因而得以在測試區間內取得勝出。

另外，本文亦對交易策略的績效做全盤的檢視。首先，Perturbed_R_F 與 unPerturbed_R_F 策略於全部區間內之卡馬比率分別能達到 32.20 與 29.39，遠遠勝過大盤的 0.84，而為了更詳細地了解策略績效，本研究從年週期分析與

月損益分析的雙重時間框架來檢視策略績效。年週期分析部分，除去時間過短的2018年，兩策略每年都有穩定獲利，此外，其在衰退期間還能維持金流，甚至達到更好的獲利，代表本研究建構之策略對風險有良好的抑制能力。另一方面，兩策略於全部區間的平均勝率皆能達到60%以上，且Perturbed.R.F與unPerturbed.R.F策略各年度最低勝率也分別能達到57.32%與54.78%，代表本研究建構之交易策略皆穩定且具有預測市場走勢的能力。月損益分析的部分則可以發現，兩策略皆僅在2012年初與2013年中有連續三個月虧損情況發生，代表投資人在執行策略時能保持足夠信心，不至於被波段虧損擊垮。

在本文所建立的多數決學習指數期貨交易策略中，交易訊號的生成是攸關模型預測能力好壞的重要環節。此重要環節，與適應於市場資料的特徵參數之間所能取得的最適配置息息相關。基於多數決的漲跌分類預測結果，即決定了本文的多單抑或空單的持有訊號。在本文中，我們同時受益且受限於採用了直覺的多空分類預測來進行模型訓練，然而在後續的研究上，若能進一步延拓此分類預測問題的界定，則可以產生更具彈性的交易訊號生成機制。如此的分類預測問題，的確是可被界定且實驗的，然而它的分類預測能力，和適應於市場資料的特徵參數最適配置之間的權衡關係為何，不失為一個有趣的後續研究方向。此外，本研究的測試區間僅包含景氣循環成長期，因此在測試時涵蓋的場景並不足夠，故為了使交易策略能更全面地被檢驗，未來或許可以嘗試不同訓練區間與測試區間，或以移動窗格(rolling window)的方法來進行模型訓練。

最後，本文雖然以台灣加權股價指數近月期貨單一商品來呈現所建構之交易策略，然而我們所提供的建構流程可被應用至其他商品或市場，並以個別市場的特徵參數來代替本研究所使用的特徵參數，進而產生適應於其市場特性的多數決交易策略。本文在執行面上的思索，及對於模型預測能力和策略投資績效的詳細闡述，應可提供投資人及市場交易者在建構適應於台灣金融市場獨特性的人工智慧交易策略時所參考的依據。

附錄 1

1-1 Omega 指標

Omega 原為 Keating and Shadwick 於 2002 年提出 (Keating and Shadwick, 2002), 其代表的意義為一超過報酬率基準與低於報酬率基準的機率加權報酬率比率, 並將更高階動差資訊納入計算, 且僅衡量下方風險。由於報酬率並非連續, 許多文獻會將其表現為下方型式:

$$\text{Omega}_i(r_\tau) = \frac{\text{UPM}_{1,i}(r_\tau)}{\text{LPM}_{1,i}(r_\tau)}. \quad (\text{A1})$$

$$\text{UPM}_{n,i}(r_\tau) = \frac{1}{T} \sum_{t=1}^T \max[r_{i,t} - r_\tau, 0]^n. \quad (\text{A2})$$

$$\text{LPM}_{n,i}(r_\tau) = \frac{1}{T} \sum_{t=1}^T \max[r_\tau - r_{i,t}, 0]^n. \quad (\text{A3})$$

其中, UPM 與 LPM 分別為上偏動差 (upper partial moment) 與低偏動差 (lower partial moment), n 代表動差階數, i 代表第 i 個金融資產, r_τ 代表報酬率基準 (普遍設定為 0), r 代表報酬率, $0 < t < T$, t 代表天數, T 代表最後一日, \max 為最大值運算符號。

由於本研究交易標的為期貨, 若不假設本金則無法計算報酬率, 而本研究漲跌幅皆以點數表示, 故將 Omega 公式轉換為點數型式, 如下:

$$\text{Omega}_i^*(\tau) = \frac{\text{UPM}_{1,i}^*(\tau)}{\text{LPM}_{1,i}^*(\tau)}. \quad (\text{A4})$$

$$\text{UPM}_{n,i}^*(\tau) = \frac{1}{T} \sum_{t=1}^T \max[\text{DProfit}_{i,t} - \tau, 0]^n. \quad (\text{A5})$$

$$\text{LPM}_{n,i}^*(\tau) = \frac{1}{T} \sum_{t=1}^T \max[\tau - \text{DProfit}_{i,t}, 0]^n. \quad (\text{A6})$$

其中, n 代表動差階數, i 代表第 i 個金融資產, τ 代表日報酬基準, 而本研究會使用普遍設定的 0 為日報酬基準, DProfit 代表當日報酬, $0 < t < T$, t 代表天數, T 代表最後一日, \max 為最大值運算符號。

1-2 索丁諾比率 (Sortino Ratio)

索丁諾比率原為 Sortino and van der Meer 於 1991 年首先提出 (Sortino and van der Meer, 1991), 其為報酬率資料和報酬率基準的差與下方標準差 (二階低偏動差開根號) 的比率, 其型式與 Omega 類似, 但分母更接近於標準差, 此外, 其分子為策略績效報酬率減去基準的報酬率, 故兩者相比, Omega 是從報酬率分配的角度計算整體績效, 而索丁諾比率是從最終獲利能力的角度來衡量整體績效。

如同前一段落所述, 本文將索丁諾比率公式轉換為點數型式, 如下:

$$\text{Sortino}_i^*(\tau) = \frac{E(\text{Profit}_{i,t}) - \tau}{\sqrt{\text{LPM}_{2,i}^*(\tau)}}. \quad (\text{A7})$$

其中, i 代表第 i 個金融資產, τ 代表日報酬基準, 而本研究會使用普遍設定的 0 為日報酬基準, Profit 為累積報酬, $0 < t < T$, t 為天數, T 為最後一日, E 代表期望值運算符號, LPM* 的定義如同上一段落所述。

1-3 潛在上漲比率 (Upside Potential Ratio)

潛在上漲比率原為 Sortino, van der Meer and Plantinga 於 1999 年提出 (Sortino et al., 1999), 其值與 Omega 概念也非常相似, 即為報酬率資料的上偏動差與低偏動差的比率, 差別只在於潛在上漲比率是利用二階低偏動差, 而 Omega 是利用一階低偏動差計算, 因此潛在上漲比率的風險衡量部分會更接近於標準差的概念。

如同前一段落所述, 本文將潛在上漲比率公式轉換為點數型式, 如下:

$$\text{UPR}_i^*(\tau) = \frac{\text{UPM}_{1,i}^*(\tau)}{\sqrt{\text{LPM}_{2,i}^*(\tau)}}. \quad (\text{A8})$$

其中, i 代表第 i 個金融資產, τ 代表日報酬基準, 而本研究會使用普遍設定的 0 為日報酬基準, UPM* 與 LPM* 的定義如同 Omega 段落所述。

1-4 赫斯特指數 (Hurst Exponent)

赫斯特指數的值域為 0 到 1 之間, 而一般赫斯特指數會以 0.5 為分界, 當值落於 0.5 時顯現此時間序列游走方式為標準布朗運動; 當值大於 0.5 則代表該序列具有持續性; 當值小於 0.5 時代表時間序列具有反持續性。但在投資層面上, 赫斯特指數的解讀方式較不直覺, 策略績效具有持續性與否並不具有褒義或貶義, 而是中性的形容詞, 因為赫斯特指數僅在描述時間序列資料在不同頻率之下, 其各頻率段落之損益分配之間的相似程度。

赫斯特指數的原始定義如下:

$$E \left[\frac{R(n)}{S(n)} \right] = Cn^H, \quad \text{as } n \rightarrow \infty. \quad (\text{A9})$$

其中, n 為一段時間區間內的資料數量, $R(n)$ 為 n 個值的全距, $S(n)$ 為 n 個值的標準差, C 為某常數, E 代表期望值運算符號, 而 H 即為赫斯特指數。

赫斯特指數有相當多種估計方法, 而本研究使用最廣為人知的 R/S 分析法 (rescaled range analysis), 這個方法是於 1968 年由 Mandelbrot and Wallis 根據先前 Hurst 作出之水文研究提出, 本文的計算流程亦參考於該篇文獻, 故不在此贅述 (Mandelbrot and Wallis, 1968; Hurst, 1951)。

附錄 2

本文為了使策略績效符合真實情況, 計算獲利時會扣除手續費與滑價 (合稱費用), 其扣除方法為一口期貨契約開倉、平倉皆會扣除一次費用, 而部位反轉 (多單轉空單, 抑或相反) 則會扣除兩次費用, 單次費用為三點期貨點數。期交所的交易所手續費試算每次期貨交易手續費的大約金額, 其規定如附表 1:

附表 1 期交所期貨交易手續費

交易所手續費(新台幣元/口)			期貨交易稅率(%)
交易經手費	結算手續費	交割手續費	
12	8	8	0.002

以一口小台為例, 假設台灣股價指數期貨一萬點, 其交易成本為新台幣 38 元, 而各家期貨商與交易人議定的手續費一般都會較此價格低, 故以此手續費做為比較基準, 此篇文章設定的費用約為此手續費的四倍。

參考文獻

- 巫和懋與許智翔 (2010), 「交易量在預測內部交易機率與技術分析的訊息價值」, 《經濟論文》, 38(2), 211–244. [Wu, H.-M. and C.-H. Hsu (2010), “The Informational Role of Trading Volume in the Probability of Informed Trading and Technical Analysis,” *Academia Economic Papers*, 38(2), 211–244.]
- Blume, L., D. Easley, and M. O’Hara (1994), “Market Statistics and Technical Analysis: The Role of Volume,” *Journal of Finance*, 49(1), 153–181.
- Borovkova, S. and I. Tsiamas (2018), “An Ensemble of LSTM Neural Networks for High-Frequency Stock Market Classification,” *Quantitative Finance*, forthcoming.
- Breiman, L. (2001), “Random Forests,” *Machine Learning*, 45(1), 5–32.
- Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone (1984), *Classification and Regression Trees*, Belmont, CA: Wadsworth International Group.
- Brooks, C., A. Rew, and S. Ritson (2001), “A Trading Strategy Based on the Lead-Lag Relationship between the Spot Index and Futures Contract for the FTSE 100,” *International Journal of Forecasting*, 17, 31–44.
- Carling, K. (2000), “Resistant Outlier Rules and the Non-Gaussian Case,” *Computational Statistics and Data Analysis*, 33, 249–258.
- Dutta, J., D. Bandopadhyay, and S. Sengupta (2012), “Prediction of Stock Performance in the Indian Stock Market Using Logistic Regression,” *International Journal of Business and Information*, 7(1), 105–136.
- Easley, D. and M. O’Hara (1992), “Time and the Process of Security Price Adjustment,” *Journal of Finance*, 47(2), 577–605.
- Giacomel, F., R. Galante, and A. Pereira (2015), “An Algorithmic Trading Agent Based on a Neural Network Ensemble: A Case of Study in North American and Brazilian Stock Markets,” Working Paper, The 2015 IEEE/WIC/ACM International Joint Conferences on Web Intelligence and Intelligent Agent Technologies.
- Ho, T. K. (1995), “Random Decision Forests,” Working Paper, The 3rd International Conference on Document Analysis and Recognition.

- Ho, T. K., J. J. Hull, and S. N. Sridhari (1994), “Decision Combination in Multiple Classifier Systems,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 16(1), 66–75.
- Hurst, H. E. (1951), “Long-Term Storage Capacity of Reservoirs,” *Transactions of the American Society of Civil Engineers*, 116, 770–799.
- Keating, C. and W. F. Shadwick (2002), “A Universal Performance Measure,” *Journal of Performance Measurement*, 6(3), 59–84.
- Khaidem, L., S. Saha, and S. R. Dey (2016), “Predicting the Direction of Stock Market Prices Using Random Forest,” arXiv preprint arXiv:160500003.
- Kim, K. J. (2003), “Financial Time Series Forecasting Using Support Vector Machines,” *Neurocomputing*, 55(1–2), 307–319.
- Kumar, M. and M. Thenmozhi (2006), “Forecasting Stock Index Movement: A Comparison of Support Vector Machines and Random Forest,” Working Paper, The Ninth Indian Institute of Capital Markets Conference.
- Lo, A. W., H. Mamaysky, and J. Wang (2000), “Foundations of Technical Analysis: Computational Algorithms, Statistical Inference, and Empirical Implementation,” *Journal of Finance*, 55(4), 1705–1765.
- Mandelbrot, B. B. and J. Wallis (1968), “Noah, Joseph and Operational Hydrology,” *Water Resources Research*, 4, 909–918.
- Ren, N., M. Zargham, and S. Rahimi (2006), “A Decision Tree-Based Classification Approach to Rule Extraction for Security Analysis,” *International Journal of Information Technology and Decision Making*, 5(1), 227–240.
- Richards, A. (2005), “Big Fish in Small Ponds: The Trading Behavior and Price Impact of Foreign Investors in Asian Emerging Equity Markets,” *Journal of Financial and Quantitative Analysis*, 40(1), 1–27.
- Schwertman, N. C. and R. de Silva (2007), “Identifying Outliers with Sequential Fences,” *Computational Statistics and Data Analysis*, 51, 3800–3810.
- Sen, J. and T. Chaudhuri (2017), “A Robust Predictive Model for Stock Price Forecasting,” Working Paper, The 5th International Conference on Business Analytics and Intelligence.
- Shannon, C. E. (1948), “A Mathematical Theory of Communication,” *Bell System Technical Journal*, 27(3), 379–423; 623–656.

Simon, D. P. and R. A. Wiggins (2001), "S&P Futures Returns and Contrary Sentiment Indicators," *Journal of Futures Markets*, 21(5), 447–462.

Sortino, F. A. and R. van der Meer (1991), "Downside Risk," *The Journal of Portfolio Management*, 17(4), 27–31.

Sortino, F. A., R. van der Meer, and A. Plantinga (1999), "The Dutch Triangle," *The Journal of Portfolio Management*, 26(1), 50–57.

漫步於隨機森林－輔以多數決學習的台股指數期貨交易策略（鄭仁杰 江彌修）

A RANDOM WALK DOWN RANDOM FOREST: ENSEMBLE-LEARNING-ASSISTED TRADING STRATEGIES FOR TAIEX FUTURES

Jen-Chieh Cheng

Department of Quantitative Trading
Yuanta Securities

Mi-Hsiu Chiang*

Department of Money and Banking
National Chengchi University

Keywords: Ensemble machine learning, Random forest, Trading strategies, TAIEX futures, Calmar ratio

JEL classification: G11, G17

* Correspondence: Mi-Hsiu Chiang, Department of Money and Banking, National Chengchi University, Taipei 116, Taiwan. Tel: (02) 2939-3091 ext. 81265; Fax: (02) 2939-8004; E-mail: mhchiang@nccu.edu.tw. Mi-Hsiu Chiang gratefully acknowledges research support from the Ministry of Science and Technology of Taiwan (MOST 107-2410-H-004-071).

ABSTRACT

With the ensemble learning of specific TAIEX market characteristics drawn from technical analysis data, in this paper we construct futures trading strategies where price directional forecasts are generated by Random Forest classification models. By quantifying the model attributes' extent of contribution to the overall prediction outcomes, we identify attributes-in-dispute and explore their perturbative effects on the predictive ability of Random Forest and thus the risk-reward performance of the proposed strategies. Using 2007–2018 TAIEX futures data, our in-sample and out-of-sample test results show that, after transaction costs, risk-adjusted outperformance over the market is consistently observable when the Random Forest models adapt the 3-14 days MA and RSI indicators, far-month futures trading volume, spot transaction volume, foreign capital open interest in futures, and open interest ratio in options.