**國 立 中 央 大 學**

資訊管理研究所

碩 士 論 文

類神經網路於投資策略的應用

Neural Network Applied in Investment Strategy

|  |  |  |
| --- | --- | --- |
|  | 研 究 生：林奇鋒  指導教授：李俊賢 博士 |  |

中 華 民 國 107 年 6 月

類神經網路於投資策略的應用

|  |  |
| --- | --- |
| 研究生：林奇鋒 | 指導教授：李俊賢博士 |

國立中央大學資訊管理學系碩士班

**摘要**

股票的波動是一種時間序列的資料。時間序列的預測是一個重要的研究議題，人工智慧計算模型目前正被廣泛使用於該議題，例如:類神經模糊系統等。本文提出複數型模糊類神經系統 (Complex neuro-fuzzy system)並應用於多目標時間序列預測，此模型具有多組複數型態輸出，其中，每一組複數型態的輸出，其實部和虛部可分別針對兩個不同實數型態目標進行預測。有關特徵挑選，本研究採用多目標特徵挑選，篩選出針對所有目標有利的特徵，並以此作為模型輸入，以降低模型整體運算負擔及提高資料運用效率。在模型方面，由輸入層、複數模糊集合神經層 (Complex fuzzy sets layer)、前提式神經層 (Premise neural layer)、T-S神經層 (Takagi-Sugeno neural layer)及輸出層建構出多層式類神經網路。在參數學習方面，訓練模型時我們採用分治原則(Divide-and-conquer principle)。複數模糊集合神經層的參數使用不同的演算法優化，像是粒子群演算法 (Particle swarm optimization, PSO)、人工蜂群演算法 (Artificial bee colony optimization, ABCO); T-S神經層的參數使用遞迴式最小平方演算法 (Recursive least-squares estimation, RLSE)進行優化; 其他的神經層沒有參數需要優化。在實驗方面，我們設計三個實驗檢驗模型的效能，將PSO-RLSE及ABCO-RLSE實驗結果結合投資策略，計算模型利潤互相比較也與不同的文獻方法比較。經由效能比較結果，本文提出多目標預測的研究方法表現出優秀效能。

|  |  |
| --- | --- |
| **關鍵字：** | 時間序列、複數模糊集、複數模糊類神經系統、人工蜂群演算法、粒子群演算法、遞迴最小平方演算法 |

Neural Network Applied in Investment Strategy

|  |  |
| --- | --- |
| Graduate student: Chi Feng Lin | Advisor: Dr. Chunshien Li |

Department of Information Management, National Central University, Taiwan

**Abstract**

Stock fluctuations are time series data. The prediction of time series is an important research topic. Artificial intelligence models are currently being widely used in this topic, such as neuro-fuzzy systems. This paper proposes a complex neuro-fuzzy system and applies it to multi-target time series prediction. This model has multiple complex-valued outputs, every output can have real and imaginary parts for two different real-valued targets, respectively. With regard to feature selection, this study uses multi-target feature selection to filter out features that are beneficial to all targets and use this as the model inputs to reduce the overall computational burden and improve data utilization efficiency. In terms of model, multi-layer neural network is constructed from input layer, Complex fuzzy set layer (CFS layer), Premise neural layer, Takagi-Sugeno neural layer (T-S neural layer), and output layer. For parameter learning, we use the divide-and-conquer principle when training the model. The parameters of the complex fuzzy set neural layer are optimized using different algorithm, such as particle swarm optimization (PSO), artificial bee colony optimization (ABCO); the parameters of the T-S neural layer are optimized using recursive least-squares estimation (RLSE), other neural layers have no parameters to optimize. In terms of experiments, we use three experiments to test the performance of the model. We combine investment strategy with PSO-RLSE and ABCO-RLSE experimental results, respectively, and calculate model profit to compare with each other and the different literature methods. Through the results of performance comparison, this paper presents a multi-target prediction method showing excellent performance.

|  |  |
| --- | --- |
| **Keywords:** | Time series; Complex fuzzy set; Complex neuro-fuzzy system; Artificial bee colony optimization; Particle swarm optimization; RLSE; |

# 致謝

# 目錄

|  |  |
| --- | --- |
| 章節 | 頁次 |

[類神經網路於投資策略的應用 i](#_Toc516354999)

[Neural Network Applied in Investment Strategy ii](#_Toc516355000)

[致謝 iii](#_Toc516355001)

[目錄 iv](#_Toc516355002)

[圖目錄 vi](#_Toc516355003)

[表目錄 vii](#_Toc516355004)

[第一章 緒論 1](#_Toc516355005)

[1.1 研究背景 1](#_Toc516355006)

[1.2 研究動機與目的 2](#_Toc516355007)

[1.3 研究方法概述 4](#_Toc516355008)

[1.4 論文架構 4](#_Toc516355009)

[第二章 文獻探討 6](#_Toc516355010)

[2.1 特徵選取 6](#_Toc516355011)

[2.2 複數模糊集合 7](#_Toc516355012)

[2.3 類神經網路 8](#_Toc516355013)

[2.4 複數類神經模糊系統 8](#_Toc516355014)

[第三章 系統設計與架構 10](#_Toc516355015)

[3.1 複數模糊集 10](#_Toc516355016)

[3.2 結構學習 11](#_Toc516355017)

[3.3 複數模糊類神經模型 14](#_Toc516355018)

[3.4 參數學習演算法 17](#_Toc516355019)

[3.4.1 粒子群演算法 17](#_Toc516355020)

[3.4.2 人工蜂群演算法 18](#_Toc516355021)

[3.4.3 遞迴最小平方演算法 19](#_Toc516355022)

[3.5 投資策略 21](#_Toc516355023)

[第四章 實驗 23](#_Toc516355024)

[4.1 實驗一：台灣股票加權指數單目標預測 24](#_Toc516355025)

[4.2 實驗二：台灣股票加權指數與恆生指數雙目標預測 28](#_Toc516355026)

[4.3 實驗三：台灣股票加權指數、道瓊工業指數、納斯達克和標準普爾500四目標預測 32](#_Toc516355027)

[第五章 討論 37](#_Toc516355028)

[5.1 利用複數類神經網路模型針對單目標資料進行預測 38](#_Toc516355029)

[5.2 利用複數型態輸出針對雙目標進行預測 38](#_Toc516355030)

[5.3 解構歸屬程度值進行多目標預測 39](#_Toc516355031)

[5.4 特徵選取之應用 39](#_Toc516355032)

[5.5 MGCACO-RLSE複合式演算法效能分析 40](#_Toc516355033)

[第六章 結論與未來研究方向 41](#_Toc516355034)

[6.1 結論 41](#_Toc516355035)

[6.2 未來研究方向 42](#_Toc516355036)

[參考文獻 44](#_Toc516355037)

# 圖目錄

|  |  |
| --- | --- |
| 圖 | 頁次 |

圖 1模糊集合輸入空間(2維) 12

圖 2資料密度總和 12

圖 3 複數模糊類神經系統模型 15

圖 4 PSO粒子更新位置示意圖 18

圖 5預測結果 27

圖 23 模型預測誤差值 27

圖 24 實驗一模糊類神經模型學習曲線 28

圖 5預測結果 31

圖 23 模型預測誤差值 31

圖 24 實驗一模糊類神經模型學習曲線 32

圖 5預測結果 35

圖 23 模型預測誤差值 36

圖 24 實驗一模糊類神經模型學習曲線 36

# 表目錄

|  |  |
| --- | --- |
| 表 | 頁次 |

表 1實驗一模型設定 24

表 2機器學習參數設定 25

表 3效能比較(TAIEX, 實驗一) 25

表 4十次重複實驗效能統計(TAIEX, 實驗一) 26

表 5模擬投資利潤表(實驗一) 26

表 1實驗一模型設定 29

表 2機器學習參數設定 29

表 3效能比較(TAIEX, 實驗一) 30

表 4十次重複實驗效能統計(TAIEX, 實驗一) 30

表 5模擬投資利潤表(實驗一) 30

表 1實驗一模型設定 33

表 2機器學習參數設定 33

表 3效能比較(TAIEX, 實驗一) 34

表 4十次重複實驗效能統計(TAIEX, 實驗一) 34

表 5模擬投資利潤表(實驗一) 34

# 緒論

## 研究背景

時間序列資料是以時間為依序所排列而成的資料序列，像是股市、匯率皆是一種時間序列資料。時間序列可視為依離散時間次序觀察的數據，可以數理方法或是統計模型進行分析，是現今社會進行常使用的方式。隨著網路快速發展，以及科技資訊的進步，導致資訊量的增加快速。利用機器學習或是資料探勘等人工智慧的方式，可以找出隱藏於資料背後所隱藏的關聯性或是規則，而這些知識是無法僅依靠人類智慧就可找出的。

時間序列資料的預測被廣泛的使用在各個領域上，已經有許多學術研究產出，如金融方面的匯率、股票漲跌等；除此之外，能源耗損、疾病預測等可運用在資源的分配，有助於國家或是機構做有效的資源配置。上述許多領域中，與我們自身最相關的為金融方面的預測，如何有效的運用金錢投資帶來利潤，是一個值得探討的議題。財務經濟方面的時間資料預測卻最為困難，因為當中包括許多影響因素，像是個別企業情況、全球情勢、總體經濟環境等。總總因素導致財務經濟的時間序列具有高度的變化。因此若是能夠有一預測效果良好的模型，配合有效的輸入資料以及正確的投資策略，便能協助投資者從中獲取利潤。

財經的預測中，類神經網路是最被廣泛使用的方法，從2005年開始有非常多針對真實世界中財經問題的研究產出，像是Xiong et al. 〔27〕使用了適應性神經模糊推理系統(Adaptive neuro-fuzzy inference system, ANFIS)與人工類神經網路(Artificial neural network, ANN)預測年度超額收益並比較效能，發現兩者預測效果都非常好；2006年，Patel et al. 〔28〕使用ANFIS、模糊推理系統和ANFIS協助投資者做經濟上的決策；Yao〔29〕於2007年提出一個方法針對外幣做買賣策略，從預測匯率中獲利；Li et al. 〔30〕學者使用群體智慧與自我組織的模糊類神經系統預測人民幣與美金的匯率；除了匯率之外，股票預測也是有許多的預測產出，如Abbasi於2008使用ANFIS配合四個獨立變數，交易量、本益比和每股盈餘預測股票收盤價〔31〕；於2017，Koijen et al. 〔34〕透過股票和債券預測彼此的數據；Pan et al. 〔35〕透過預測股票的波動性，比較短期和長期的槓桿效應。

## 研究動機與目的

在股市裡面，有著幾個著名的理論，其一為Kendall 1953年提出的隨機漫步理論 (random walk theory) 〔1〕，意即股價的變動是獨立的，其間並無規律性或週期性存在；另一延伸為有效市場假說 (efficient market hypothesis, EMH)〔2〕，是學者Fama於1970年所提出，認為買方和賣方的價錢是雙方願意的，且假設所有市場之參與者均可無償取得資訊，因此投資人所收集的資訊並不能使其獲得超額利潤。但另有一派的專家認為股票是可以預測的，因此不斷的研究，於1990年Kimoto et al. 〔3〕利用倒傳遞類神經網路，搭配兩種指標來預測日經指數的漲跌和買賣時機；1999年Yao et al. 採用多項技術指標以倒傳遞類神經網路來預測股票市場 〔4〕；Kim et al. 於2000年將類神經網路中加入基因演算法 〔5〕，不但改進了類神經網路的學習速度，且降低了特徵空間的複雜度，實驗結果相對於倒傳遞類神經網路顯得更加優秀，並發現非線性類神經網路預測能力較好； 2012年，Wei 提出了基於經驗模態分解 (Empirical Mode Decomposition, EMD)的混合式適應性類神經模糊推論系統 (Adaptive network-based fuzzy inference system, ANFIS) 〔6〕，並配合投資策略計算利潤，在虛擬投資中，有著不錯的獲利效果；同年，Li et al. 〔32〕以差分整合移動平均自迴歸(Autoregressive integrated moving average, ARIMA)模型結合類神經模糊系統(Neuro-fuzzy system, NFSs)預測股票的收盤價，隔年，Li et al. 〔33〕使用複數模糊類神經模糊系統對雙目標同時預測。

從上述可得知，模糊類神經系統對財經方面的預測非常的有用，且一個經過良好訓練的模型預測效能會精準許多，因此本研究將會以模糊類神經為基準，結合其他人工智慧技術，並比較訓練前後對模型預測的影響。在進行模型訓練的過程中，經常發生預測偏差或是模型運算過久等問題。因此如何挑選有效的輸入資料以及模型整體的大小是大多研究的主要議題。模型在進行預測的過程中，參數需要經過學習以及最佳化的過程來增加效能表現，目前有許多機器學習演算法被使用於尋找參數最佳解，像是粒子群最佳化演算法 (Particle swarm optimization, PSO) 〔7〕、人工蜂群演算法 (Artificial bee colony optimization, ABCO) 〔8〕、基因演算法 (Genetic algorithm, GA) 〔9〕等。目前文獻大多是以Takagi-Sugeno 模糊系統建立架構 〔10〕，因此所需最佳化的參數較多，會讓PSO等演算法面臨求解的參數維度過大，使得效果不如預期，發生收斂過於緩慢或落入區域最佳解的問題。

本研究將使用混合的演算法優化參數，分別以PSO及ABCO結合遞迴最小平方誤差 (Recursive least-square estimation, RLSE)演算法 〔11〕，形成PSO-RLSE〔12〕和ABCO-RLSE。以分治法使用不同的演算法訓練不同部分的參數，減低機器學習在尋求最佳解的過程中，落入區域最佳解的機率。如上述所提，除了股票目標本身，股票的波動是由諸多因素所形成，因此本實驗將從多組股票中進行特徵挑選，也同時一次預測多個目標，從大量的資料中萃取對每個目標都有效益的內容。為了使預測效果提升，模型的建置上應用了類神經網路系統結合T-S (Takagi-Sugeno)的模糊規則型態組成模糊類神經模型。並且將模糊類神經模型的參數分為兩部分，分別為前鑑部參數 (If-part parameters)與後鑑部參數 (Then-part parameters)。

傳統上，在建立模糊集的設計多是以高斯函數建立模糊集〔13〕，資料輸入後可得到介於0到1的隸屬程度值。Ramot在2002年提出了Complex Fuzzy Sets 〔14〕，將原本隸屬程度加上一個虛部值，將隸屬程度的值從傳統的一維區間投影到實數軸x與虛數軸y兩個維度上的單位圓盤，使其相較於原本的實數型隸屬函數，具有較豐富的資料量。

## 研究方法概述

本研究提出混合型的演算法，PSO-RLSE以及ABCO-RLSE，用以最佳化模型參數。模型中有兩部分參數，分別為前鑑部參數與後鑑部參數，PSO與ABCO負責優化前鑑部參數，RLSE則優化後鑑部參數，以降低各個演算法的搜尋維度。

在本研究中，首先會利用夏農資訊熵〔15〕的概念對資料分析，並做多目標特徵選取 (Multi-target feature selection) 〔16〕，藉由計算原始資料對目標所提供的資訊量作為標準，並且加入冗餘資訊的計算，選出對目標有良好影響的資料作為輸入至模型的依據，可以降低資料複雜度與計算成本。之後透過PSO-RLSE與ABCO-RLSE演算法，調整前鑑部參數及更新後鑑部參數。最後將結果與投資策略結合，計算出模型帶來的利潤。

## 論文架構

本論文一共分為六章，第一章為緒論，介紹本研究的背景、動機及研究之目的、概述本研究所要使用的方法及流程。第二章為研究理論方法的介紹及概述，包括過去特徵選取的一些方式，除此之外，將會介紹模糊集的起源以及複數模糊集的概念。接著介紹類神經網路的理論，以及相關的研究，最後提及被廣泛使用的模糊類神經系統。第三章為系統設計與架構，概述如何進行多目標特徵挑選，以及機器學習決定模型大小的方法；複數模糊類神經網路中各神經層的介紹也會在此章節敘述。接著會詳細說明參數學習的演算法，包括PSO、ABCO、RLSE，概述演算法彼此混合後進行參數學習與訓練的過程，最後會提及投資策略的操作及概念。第四章是實驗及結果，本研究用了3個實驗進行時間序列預測之研究。實驗一利用台股加權指數進行單一目標的預測，實驗二是同時進行台股加權指數與上海恆生指數的預測，實驗三則是一次針對四個目標預測，分別是台股加權指數、道瓊工業指數、納斯達克指數以及標準普爾500。以上實驗皆會做兩種比較，其一為不同混合演算法的效能比較，包含訓練前後的差異，其比較標準包含配合投資策略後的利潤；除此之外也會與過去文獻所提出的方法進行效能比較，以證實本研究方法的優劣性。第五章為針對實驗的結果進行討論。第六章則是本研究的結論以及未來的研究方向。

# 文獻探討

## 特徵選取

資料計算模型是透過機器學習或資料探勘後所建構，像是分類、回歸、分群等。資料前處理的目的有兩個，其一是降低資料集合的大小，使得分析更具效果；其二是可針對特定的資料集合選出較適當的資料。在現今龐大的資料中，前者對整個分析方法來說是很重要的，因此特徵挑選在資料探勘、網路異常檢測、文字分類、基因分析等許多領域中扮演著不可或缺的角色。特徵集合的縮減是根據特徵的關聯性，以及他們對目標所造成的冗餘性〔17〕。特徵挑選的方法可以被分類為數種〔18〕，最常見的就是過濾法(Filter approach) 〔19〕、包裝法(Wrapper approach) 〔20〕、嵌入法(Embedded) 〔19〕、混合法(Hybrid approach) 〔21〕。以下為各方法的簡略說明。

過濾法(Filter method)是由Guyon於2003年所提出〔19〕，此方法首先對資料集進行特徵選擇，接著將選擇出來的特徵子集合進行相關性評估並評分，最後設定門檻值個數，藉此選擇特徵，特徵選擇的過程與後續的模型無關，其優點是比包裝法還要快速。於1997年，Blum所提出的包裝法(Wrapper approach) 〔20〕，與過濾法最大的差別在於特徵子集合的挑選評分標準，其評分標準是由模型的運算結果所決定，因此在挑選過程中，會不斷的使用到模型計算，運算時間長且計算複雜，對於某些特定的分類器有著較高的分類準確度。嵌入法(Embedded approach) 〔19〕是將特徵挑選與模型的訓練一同進行，因此在模型訓練結束後也會完成特徵挑選，時間與模型運算介於上述兩種方法之間。混合法(Hybrid approach) 〔21〕是把過濾法及包裝法的概念結合，透過過濾法挑選出適當的特徵子集合，接著將這些特徵子集合透過模型運算選出最終的特徵，此方法提高了過濾法的精準度，也降低了包裝法的運算時間，結合了兩者的優點。

如上述，特徵集合的縮減，是根據特徵的關聯性與冗餘性，故本文採用熵的概念計算特徵彼此間提供的訊息量。熵一詞最早是由德國物理學家Rudolph Clausius 於 1854 年提出〔22〕，是一種對物理系統之無秩序或亂度的量度，當熵值越高代表混亂程度也越高。關於資訊的亂度，夏農提出了夏農資訊熵(Shannon information entropy)的理論〔15〕，若資訊的隨機性越高，則資訊熵值會越高，對於某一個隨機變數，資訊熵的定義如下。

|  |  |
| --- | --- |
|  |  |

其中，是隨機變數的資訊熵；則是事件的發生機率；則被視為的資訊混亂度。由公式可以發現發生機率越大，資訊熵值越小。反之，若事件出現機率趨於平均，則熵值會趨近最大值。如前所敘，熵值越大代表所提供的資訊隨機性越高。本研究基於夏農資訊熵(Shannon information entropy)的理論〔15〕，進行多目標特徵挑選〔16〕，此方法中考慮到資訊量以及冗餘資訊量，最後配合使用者設定篩選出數個特徵當作模型的輸入。

## 複數模糊集合

過去集合的概念只有「屬於」或「非屬於」，但此概念卻無法應用於真實世界中，好比溫度若大於30度時稱為「熱」，但29.9度與30度並無太大差異卻不屬於「熱」，此情況與人類的思維有所衝突，因此在1965年，Zadeh學者提出了模糊集的概念〔23〕，藉由此概念，我們可以透過一歸屬函數計算出元素屬於集合的隸屬程度，隸屬程度則介於0到1之間，隸屬程度越高，則該元素屬於集合的程度越高。之後Buckley提出了模糊複數(Complex number)的概念〔24〕-〔26〕，2002年，Ramot et al. 〔43〕更進一步提出了複數模糊集合(Complex fuzzy sets, CFSs)，意即歸屬程度可以透過函式得到一個複數型態的值，這使得歸屬程度可以呈現在一個半徑為1的複數單位圓盤中。這個概念使原本能表示的歸屬程度更加的豐富。

## 類神經網路

類神經網路，是一種模擬人類大腦的技術，人類的大腦包含著大量的神經元，這些神經元主要功用是處理資訊和記憶，神經元之間是由「突觸」所連接，而「突觸」的功用如同權重，控制著資訊的流通，於1943年，McCulloch et al. 〔37〕學者使用一種稱為閥值邏輯(Threshold logic)的演算法模擬出神經元的概念，當一個神經元接收到資訊，會判斷是否要產生興奮反應的機制。而後心理學家Hebb提出了海伯理論(Hebb’s law) 〔39〕，意即突觸前的神經元與突觸後的神經元同時活化時，則這兩個神精元之間連結的強度會增加；此想法於1956年Rochester et al. 〔38〕將之轉化成數學模型創造了感知機，以模擬人類大腦，可以說是當今深度學習(Deep learning)的始祖，但相較來說感知機神經元較少，且傳遞訊號權重為0或1，仍然無法處理異或問題(非線性問題)，且當時電腦沒有足夠能力運算。直到1975年Werbos 〔40〕提出反向傳播演算法 (Back propagation, BP)，這個演算法有效地解決了異或問題以及訓練多層神經網路的問題。日後有許多其他的人工智慧方法被提出，像是支援向量機〔41〕、貝氏分類器〔42〕等，直到近年來，因資訊科技快速的發展，電腦所能運算的速度倍增，在某些專長領域已可超越人類水平，神經網路才再度被大量研究。

## 複數類神經模糊系統

在過去有許多研究方法被提出如ARIMA、模糊理論、神經網路運算等。其中，最常被提出的就是類神經模糊系統 (Neuro-fuzzy systems, NFSs)。類神經模糊系統是一種模糊系統，它可以說是類神經網絡理論〔37〕與模糊理論〔23〕的結合，有文獻將NFS分類成三種型態〔36〕分別是合作型類神經模糊系統(Cooperative NFS)、併發型類神經模糊系統(Concurrent NFS)和混合型類神經模糊系統(Hybrid NFS)。

***合作型類神經模糊系統****:* 是先用類神經網路計算規則等所需參數，在將參數帶入模糊系統中運算。

***併發型類神經模糊系統****:* 是數據經過類神經網路運算後，視為模糊系統的輸入再次運算獲得最後結果。

***混合型類神經模糊系統****:* 是將模糊系統中的規則等元素融入類神經網路中，並透過學習演算法修正系統中的參數，是目前三者中最被廣泛研究的系統。

由於類神經模糊系統的特性使他對時間序列的預測有著不錯的效果。所以現在有關時間預測這方面的研究，大多採用類神經模糊系統為模型架構。而本研究使用混合式類神經網路，為了使模糊系統所涵蓋的資訊量更加豐富，本研究使用複數型態的模糊集取代傳統的模糊集，使得模型可以同時預測多個目標，細節將會在第三章第一小節中探討。

# 系統設計與架構

本章節將逐一說明運用於研究中之方法設計與模型架構。本研究使用機器學習決定模型結構的大小，模型實作使用複數類神經模糊系統，分別利用不同的演算法(PSO、ABC)優化模型前鑑部參數，遞迴式最小平方演算法最佳化後鑑部參數。在資料進入模型之前，透過多目標特徵挑選〔16〕，挑選出對所有目標較為有效之特徵資料集合，減少龐大資料對模型的負擔。最後將結果配合投資策略做不同演算法的比較。

## 複數模糊集

傳統的模糊集合概念〔23〕，元素對集合的歸屬程度為一對一的關係。複數歸屬度型態的模糊集合，可以擁有更豐富的歸屬程度，透過此概念可以計算出一筆資料於集合中複數型態的歸屬程度，以便之後模型可以一次預測多個目標。歸屬程度計算流程如下，假設有一複數模糊集合，可以表示如下。

|  |  |
| --- | --- |
|  |  |

其中，為元素的歸屬程度，表示如下。

|  |  |
| --- | --- |
|  |  |

其中，為宇集合的數值變數；為振幅函數，是一實數數值介於[0,1]間；為相位函數，是一實數數值；為。

本實驗採用高斯複數模糊集，此概念由Li et al.提出〔27〕，為複數模糊集合與高斯函數的結合，使其可以進入模型並分析資料。高斯複數模糊集(Complex Gaussian membership function, cGMF)可以表示如下：

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

其中， 分別為輸入資料、中心值以及模糊集合的延展度，值為此函數的相位頻率參數，此函數會進入參數學習過程，以增加模型整體的彈性。則採用高斯函數的一次微分，目的在於可重複使用高斯原有的參數，降低運算時參數的複雜度。透過複數高斯型態的模糊集，可得出一複數歸屬程度。我們可以透過拆解，得出一組歸屬程度向量，成分表示如下。

|  |  |
| --- | --- |
|  |  |
|  |  |
|  |  |

其中，為擷取該值的實數部位置數值；為擷取該值的實數部位置數值；為公式(5)所提及的高斯函數；為公式(6)所提及的高斯函數一次微分。透過上述的拆解可以在不增加參數的情況下，得到有別於傳統模糊集的歸屬程度，此方法提供豐富的資訊量，方便日後做更多元的應用。

## 結構學習

結構學習是為了透過訓練資料，建造出更適當的模型架構，此外結構學習中的結果，也會成為之後參數學習的一部分。在本研究採用高斯型態的模糊集合，需要中心以及標準差兩個參數，因此會透過減數分群(Subtractive cluster, SC)演算法〔44〕分群不同輸入維度的訓練資料。並將分群後的群中心配合每個維度的標準差形成模糊集，各個維度的模糊集個數總和，即為第1層神經元的數量，基於各個輸入維度的模糊集，共可以組成個區塊。

|  |  |
| --- | --- |
| , |  |

其中，為第個輸入維度的分群個數，亦即模糊集個數。以第個區塊的組成為例，如下。

|  |  |
| --- | --- |
| , |  |

，其中，為第個輸入的語意變數;為第個輸入變數，;為第個區塊中第個輸入語意變數的模糊集合，使用高斯函數建構，如公式(5)。



圖 1模糊集合輸入空間(2維)

兩個輸入維度，各分出3群則會形成共9塊區域，其中z軸為該區的資料密度。



圖 2資料密度總和

可看出部分區塊資料密度總和較高，代表用於建造模型較有利。

為了降低模型的運算負擔，以提升效率，我們將會篩選出若干個較重要的區塊成為前鑑部神經元，也就是第2層中的神經元。若以兩個輸入維度為例子可得到如圖1的區塊，而後透過資料密度的概念，將資料灑入區塊中，計算資料密度量，累加每個區域的資料密度量後可得到如圖2，從中可挑選資料密度較高的區塊，作為第2層神經元。詳細步驟如下:

1. 從各個輸入維度與模糊集可以得到每個區塊的資料密度量，以第個區塊為例，資料密度量公式如下。

|  |  |
| --- | --- |
| , |  |

其中，為第個輸入維度的第筆資料;為第個區塊中第個輸入維度的模糊集。

1. 將第個區塊的資料密度量累加標記為，公式如下。

|  |  |
| --- | --- |
| , |  |

其中，為資料總筆數。計算平均值標記為，標準差標記為。

1. 查看每個區塊，若，則將累加。設定上下界，標記為和，透過上下界找出，表示最後選取的區塊數目。若介於上下界間，則將設定成;若小於下界，則將設定成;若大於上界則將設定成。
2. 將排序，並保留前個區塊，當作之後模型第2層的神經元。

位於第4層的後鑑部神經元個數，在本研究中與前鑑部神經元數目相同。後鑑部神經元為T–S神經元，由T–S function構成，T–S function公式如下。

|  |  |
| --- | --- |
| , |  |

其中，{}是第個T–S神經元的參數，是第個輸入。

在結構學習後，我們可以透過分群確認第1層的神經元個數，藉由區塊挑選得到個第2層神經元以及第4層神經元個數，藉此創建模型，模型詳細說明將在下個小節探討。

## 複數模糊類神經模型

本研究利用Takagi-Sugeno (T-S) fuzzy 建立模糊系統。T-S模糊模型最早是由Takagi與Sugeno於1985年提出〔10〕，以一個複合式非線性系統並藉由一系列的If-Then模糊規則組合而成。本研究是透過複數高斯模糊集非線性的系統，結合線性的T-S function，形成非線性的If-Then模糊規則網路架構。If-Then模糊規則類似人類的經驗法則，因此更容易被人類所理解。本章節將探討模型各類神經層的輸入、計算方法與輸出。

本研究的模型為一個複合式非線性系統。訓練資料集合標記為，為資料總筆數，是-by-1的輸入向量，為輸入維度數量;為-by-1的目標向量，為複數型態目標的數量。透過模型可以得到輸出。

模型是透過條T-S模糊規則組成，每條規則是由前鑑部(If part)與後鑑部(Then part)所結合，其規則形式如下：

|  |  |
| --- | --- |
| , |  |

規則數；為模糊系統的輸入變數；為第條規則的複數模糊集；則是輸入的語意變數。複數模糊集的參數即是前鑑部(If part)參數，為後鑑部(Then part)參數。此複數模糊類神經模型可轉為一個六層架構的類神經網路模型，如圖3，以下將對各層進行說明。

Layer 0

(輸入層)

Layer 1

(複數模糊集層)

Layer 2

(前鑑部層)

Layer 4

(後鑑部層)

Layer 5

(輸出層)

Layer 3

(正規化層)

圖 3 複數模糊類神經系統模型

**Layer 0**: 此層為輸入層，是原始資料透過多目標特徵選取後，將最後挑出的特徵當作訓練資料，我們將時間序列第個點的輸入向量標記如下。

|  |  |
| --- | --- |
| , |  |

**Layer 1**: 此層為複數模糊集神經層，透過前面結構學習的分群，可在不同維度上建構數個模糊集，每個不同維度的輸入都可經由模糊集得到歸屬程度。透過複數模糊集合可得到多組複數型態的歸屬程度，不同的歸屬程度可以給不同的模型輸出做應用，以達到多目標預測的效果，透過複數模糊集的公式(2)-(9)，可得到歸屬度向量，如下。

|  |  |
| --- | --- |
| , |  |

**Layer 2**: 此層為前鑑部神經層，經過前面的結構學習，我們可以篩選出個對模型較有用的區塊並建構神經元，由於神經元的輸入為上一層的歸屬程度，且輸出為每個輸入維度的歸屬程度相乘結果，故稱之為神經元，每個神經元的輸出為該神經元的啟動強度，由於本研究採用複數模糊集合，因此每個神經元的輸入會是向量型態，輸出亦然如此。

|  |  |
| --- | --- |
| , |  |
| , |  |

，為的乘積，為第個神經元中第個維度歸屬度向量的第項歸屬程度，。

**Layer 3**: 此層為正規化層，會將個輸入向量中的各個元素正規化，之後將結果以向量型態輸出，輸出如下。

|  |  |
| --- | --- |
| , |  |
|  |  |

其中，為第個神經元中第個元素正規化後的值，。本研究中使用的是複數模糊集，因此輸入是複數型態，故輸出也是複數型態。

**Layer 4:** 此層為後鑑部層，經過此層的運算可以得到個模型輸出，公式如下。

|  |  |
| --- | --- |
| , |  |

，為第個T–S神經元輸出；{}是第個T–S神經元的參數。

**Layer 5:** 此層為輸出層，將上一層得到的個神經元輸出加總，即為我們的模型輸出。

|  |  |
| --- | --- |
|  |  |

## 參數學習演算法

根據分治法(Divide-and-conquer)的概念，我們將使用不同的機器學習演算法，對各層的參數優化，以便更容易找到最佳解。對於第1層複數模糊集的參數優化，我們使用兩種不同的演算法，其中包含粒子群演算法(Particle swarm optimization, PSO) 〔7〕以及人工蜂群演算法(Artificial bee colony optimization, ABCO) 〔8〕。第2層與第3層沒有參數需要優化。在第4層的部分，採用遞迴最小平方演算法(Recursive least squares estimation, RLSE) 〔11〕最佳化後鑑部參數，希望透過不同演算法的結合，降低所需搜尋的參數維度，為模型帶來更好的效能表現。以下將會詳細敘述三種演算法的運作原理以及公式等細節。

## 粒子群演算法

粒子群演算法是由J. Kennedy et al.〔7〕於1995年開發的一種演化計算技術，來源於對一個簡化社會模型的模擬。其原理類似鳥群在尋找食物，除了自身提供的資訊，慣性以及自身最佳位置，也運用到群體智慧中全群最佳位置，用以調節速度，如圖4，此演算法特性為收斂快速，公式如下。

|  |  |
| --- | --- |
| , |  |
| , |  |

其中，為第回合時第個粒子的位置，為第回合時第個粒子的速度，為第回合時第個粒子的最好位置，為第回合時全部粒子中最好的位置，為PSO的參數，、為介於0到1的隨機數。在本實驗中，粒子的位置代表前鑑部的參數，其中包含了每個維度的分群中心、標準差以及相位頻率參數。

目前位置

新位置

透過自我最佳位置

調整方向

透過全群最佳位置

調整方向

全群最佳位置

(Gbest)

自我最佳位置

(Pbest)

慣性方向

第顆粒子

圖 4 PSO粒子更新位置示意圖

## 人工蜂群演算法

人工蜂群演算法為Karaboga學者所提出〔7〕，原理類似蜜蜂尋找食物來源的概念，其特性包含群體智慧與隨機性，如蜜蜂傳達食物位置時會彼此透過搖擺舞溝通，而搖擺舞所指示食物位置會有偏差帶有隨機性。此演算法中，總共有三種蜜蜂，包括工蜂(Employed bee)、觀察蜂(Onlooker bee)以及偵查蜂(Scout bee)，工蜂負責對每個食物源的附近進行開發，然後以跳舞傳達食物源的收益訊息，每隻工蜂代表著一食物源，代表著大維度的搜尋；觀察蜂觀察各個食物收益度後，以輪盤法選擇一個食物源，到食物源的鄰近周遭進行開發的行為，代表著小維度的搜尋；偵查蜂作用為當食物源經過一定的開發次數後，食物源的收益度沒有改善時，會派出偵查蜂取代食物源，意味著隨機的探索一個全新的食物源，可以避免演算法陷入區域最佳解中。其步驟與真實的蜜蜂找尋食物不盡相同，演算法步驟如下:

1. 隨機形成工蜂位置，並更新工蜂位置，公式如下。

|  |  |
| --- | --- |
| , |  |

其中，為第隻工蜂在移動後的第個維度;為第隻工蜂移動前的第個維度;為其他隨機蜜蜂的第個維度的值。

1. 使用輪盤法(Roulette method)選一個工蜂的位置，輪盤機率公式如下。

|  |  |
| --- | --- |
|  |  |

其中，為第隻工蜂被選中的機率;為適應值，本實驗將成本函數的倒數視為適應值;為工蜂的總數目

1. 所有的觀察蜂在剛剛被選中的食物源位置附近搜尋，公式如下。

|  |  |
| --- | --- |
|  |  |

其中，為第隻工蜂在移動後的第個維度;為第隻工蜂移動前的第個維度;為被選中蜜蜂的第個維度的值。

1. 判斷每隻工蜂是否已經達到限制回合都未更新，若為真則派出偵查蜂取代，偵查蜂位置產生公式如下。

|  |  |
| --- | --- |
|  |  |

其中，為第隻工蜂的第個維度;為所有工蜂第個維度的最大值;為所有工蜂第個維度的最小值。

1. 重複step 2~ step 4，直到反覆運算結束。

## 遞迴最小平方演算法

本研究使用遞迴式最小平方演算法(Recursive least square estimation, RLSE) 〔11〕更新T–S神經元參數，RLSE方法在更新參數時是利用每筆資料，不斷的更新比起一次接收所有資料的LSE方法更加有效，一般來說LSE問題可以被視為一個線性的問題，如下。

|  |  |
| --- | --- |
| , |  |

其中，y是目標;u是模型的輸出;{}是u已知的方程式;{, =1,2,…,m}是我們估計的未知參數，則是整個模型的誤差LSE的問題也可以被寫成矩陣的方式表達，如下。

|  |  |
| --- | --- |
| , |  |

其中，

|  |  |
| --- | --- |
| , |  |
| , |  |
| , |  |
| , |  |

是輸入的矩陣，是我們估計的未知參數矩陣，是目標矩陣，是誤差的向量。要最佳化，可透過RLSE的等式[17]運算。

|  |  |
| --- | --- |
| , |  |
| , |  |

其中，是遞迴次數，{}，為資料總筆數，是的第行，再開始RLSE演算法時，會設定為0，則設定為，為一極大整數，為單位矩陣。

混合型演算法中各部分參數優化會交由不同演算法負責，本研究使用到PSO-RLSE以及ABCO-RLSE兩種混合型演算法，PSO和ABCO負責前鑑部的參數學習，RLSE則負責線性的T-S function 參數學習。混合演算法與模型計算的流程如下:

1. 準備訓練資料及測試資料。
2. 以前鑑部演算法(PSO或ABCO)粒子位置作為模糊集參數，將訓練資料帶入模型，並計算每個神經元啟動強度。
3. 用RLSE更新T-S神經元的參數，RLSE算式中的和向量如下。

|  |  |
| --- | --- |
| , |  |
| , |  |
| , |  |
| , |  |

其中。因為多目標預測中，正規化後的啟動強度為一向量，使為一矩陣，因此在原本的公式(36)中，利用單位矩陣取代原本的常數項1，改良後公式如下。

|  |  |
| --- | --- |
| , |  |

1. 更新完所有參數後，計算出模型的輸出。
2. 計算成本，更新前鑑部演算法粒子的位置及相關數據。
3. 重複Step 2~Step 5，直到迭代結束。

## 投資策略

為了評估模型是否對投資有實質上的幫助，使用成本函數是不夠的，因為從中無法看出是否有利潤，只能瞭解模型的配市率，而配市率高並不代表投資效益高，因此本實驗將預測出來的收盤價配合投資策略〔45〕，進一步決定要買進或是賣出，買進與賣出公式如下。

|  |  |
| --- | --- |
| 買進: if , |  |
| 賣出: if , |  |

其中，為股票的波動與門檻參數，為模型的輸出，意即預測日的收盤價格;為日實際的收盤價格。若預測明天的收盤價高於今天實際收盤價，代表必須買進;若預測明天的收盤價低於今天實際收盤價，代表模型預測明天會跌，所以要儘快賣出。

計算利潤的方式，則透過今天實際的收盤價與隔天實際的收盤價去做運算，公式如下。

|  |  |
| --- | --- |
|  |  |

其中，為利潤，為策略為買的總天數;為策略為賣的總天數;代表第天的收盤價。

透過上述投資策略以及利潤公式，我們可以計算出整個模型的所帶來的利潤值，並大致模擬出此模型運用到真實世界的效果。本研究將會在每個實驗中秀出利潤值與其他參數。

# 實驗

本研究總共有3個實驗，實驗一為單目標的預測，使用的目標為2001年台灣股票加權指數(The Taiwan stock exchange capitalization weighted stock index, TAIEX)；實驗二為雙目標的預測，檢驗模型透過複數型態的歸屬程度可行性，使用第一組複數輸出預測兩個目標，實數部分負責第一個目標，虛數部分負責第二個目標，同時預測TAIEX以及恆生指數(Hang seng index, HSI)；實驗三為四個目標的多目標預測，使用到兩組複數型態的輸出，比實驗二複雜度更高，其中包含了2001年的台股指數、道瓊工業指數(Dow Jones industrial average index, DJI)、納斯達克(National association of securities dealers automated quotation, NASDAQ)以及標準普爾500(Standard and Poor’s 500, S&P 500)。上述目標皆為非常著名的股票指標，像是TAIEX為台灣上市的股票中經過加權計算出的指標，代表著台灣上市股票的波動；HSI是以反映香港股市行情的重要指標，指數由五十隻恆指成份股的市值計算出來的，相當於香港交易所所有上市公司十二個月平均市值涵蓋率的63%；DJI涵蓋著財務等9大產業，為一股價加權指標；NASDAQ為超過三千檔股票所組合成的市值加權指標，大多以科技產業為例；S&P500為美國前500大公司的市值加權，當中包含IT等11個產業。這些指標代表著一個國家甚至全球股票的趨勢，故能精準預測可為投資者帶來不少的幫助。

為了與其他文獻比較模型的好壞，我們將透過誤差指標對模型評估以及計算模擬投資後的利潤。成本函數(Cost function)與評估指標皆使用均方根誤差(Root mean square errors, RMSE)，公式如下。

|  |  |
| --- | --- |
| , |  |
| , |  |

，為資料總筆數；為模型第筆資料的誤差向量;為第筆目標向量；為第筆模型輸出向量；為埃爾米特共軛(Hermitian transpose)，意即轉置矩陣後，並對矩陣元素做共軛運算。

在結構學習部分，本研究中所有實驗的上限皆設定為15，下限皆設定為4，故模型的規則數會介於4~15之間。

## 實驗一：台灣股票加權指數單目標預測

本實驗使用真實世界的時間序列數據來驗證模型的效能，目標為TAIEX。實驗目標是建立一適當的模型，經過訓練後預測每日股價指數。模型輸出為一複數值，取其實數部分預測目標。 使用的資料為TAIEX 2001年的每日收盤價，此實驗的原始資料為278筆，經過多目標特徵挑選後選出特徵做為模型輸入資料。結構學習部分，會透過減數分群演算法分群，在進行區塊挑選如第三章第二小節所介紹，本實驗各個輸入維度分別有X個複數模糊集，進行區塊挑選後，從原本的XX神經元篩選到剩下XX條規則，藉由資料驅動(Data driven)概念適當地建構模型，每個複數模糊集有3個參數，其中包含中心、標準差以及相位頻率參數，總共有X個複數模糊集，故前鑑部參數數量為XX，後鑑部型態為T-S function，因此總共會有個參數，為後鑑部數目，為模型輸入維度數目。整體模型在結構學習後的參數，如表1所示。機器學習參數設定，如表2所示。本次範例的結果除了兩個演算法比較也與其他論文〔45〕所提的方法做比較，像是Chen (1996)、Yu (2005)、SVR-RBF等。比較結果如表3所示。為了測試模型的穩定性，我們總共跑了十次的實驗，效能統計結果如表4所示。模擬投資所賺的利潤，以及買賣次數如表5所示。目標與模型輸出的結果，如圖5所示；模型的機器學習曲線，如圖6所示;預測誤差如圖7所示，誤差呈現亂數狀態，範圍介於-50至50，代表模型預測能力穩定。

表 1實驗一模型設定

|  |  |
| --- | --- |
| **參數** | **值** |
| 特徵變數(輸入變數) | {} |
| 模糊集數量 | {2, 3, 3, 3} |
| 模糊集型態 |  |
| 複數型態目標數目 |  |
| 神經元數目 |  |
| 前鑑部參數數目 |  |
| T–S神經元數目 |  |
| 後鑑部參數數目 |  |

\* 每一個複數目標的實部與虛部，分別包含兩個實數目標

表 2機器學習參數設定

|  |  |
| --- | --- |
| **PSO** | |
| 粒子群大小 | 50 |
| 迭代次數 | 100 |
|  | {0.8, 2.0, 2.0} |
|  | Random in [0,1] |
| 粒子初始化位置 | 藉由第三章所提SC演算法決定 |
| 粒子初始化速度 | 0 |
| **ABCO** |  |
| 工蜂數量 | 40 |
| 觀察蜂數量 | 10 |
| 迭代次數 | 100 |
| 限制回合數 | 20 |
| **RLSE** | |
|  |  |
|  | 25-by-1 全零向量 |
|  | **I** |
| **I** | 25-by-25 單位矩陣 |

表 3效能比較(TAIEX, 實驗一)

|  |  |
| --- | --- |
| 方法 | RMSE |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| PSO-RLSE |  |
| ABCO-RLSE |  |

表 4十次重複實驗效能統計(TAIEX, 實驗一)

|  |  |  |
| --- | --- | --- |
| Trials | PSO-RLSE | ABCO-RLSE |
| 1 |  |  |
| 2 |  |  |
| 3 |  |  |
| 4 |  |  |
| 5 |  |  |
| 6 |  |  |
| 7 |  |  |
| 8 |  |  |
| 9 |  |  |
| 10 |  |  |

表 5模擬投資利潤表(實驗一)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Trials | PSO-RLSE | | ABCO-RLSE | |
| 利潤 | 買/賣 | 利潤 | 買/賣 |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |
| 5 |  |  |  |  |
| 6 |  |  |  |  |
| 7 |  |  |  |  |
| 8 |  |  |  |  |
| 9 |  |  |  |  |
| 10 |  |  |  |  |



圖 5預測結果

藍色實線為實際目標值，紅色虛線為模型預測值



圖 23 模型預測誤差值



圖 24 實驗一模糊類神經模型學習曲線

## 實驗二：台灣股票加權指數與恆生指數雙目標預測

本實驗使用真實世界的時間序列數據來驗證模型的效能，目標為TAIEX。實驗目標是建立一適當的模型，經過訓練後預測每日股價指數。模型輸出為一複數值，取其實數部分預測目標。 使用的資料為TAIEX 2001年的每日收盤價，此實驗的原始資料為278筆，經過多目標特徵挑選後選出特徵做為模型輸入資料。結構學習部分，會透過減數分群演算法分群，在進行區塊挑選如第三章第二小節所介紹，本實驗各個輸入維度分別有X個複數模糊集，進行區塊挑選後，從原本的XX神經元篩選到剩下XX條規則，藉由資料驅動(Data driven)概念適當地建構模型，每個複數模糊集有3個參數，其中包含中心、標準差以及相位頻率參數，總共有X個複數模糊集，故前鑑部參數數量為XX，後鑑部型態為T-S function，因此總共會有個參數，為後鑑部數目，為模型輸入維度數目。整體模型在結構學習後的參數，如表1所示。機器學習參數設定，如表2所示。本次範例的結果除了兩個演算法比較也與其他論文〔45〕所提的方法做比較，像是Chen (1996)、Yu (2005)、SVR-RBF等。比較結果如表3所示。為了測試模型的穩定性，我們總共跑了十次的實驗，效能統計結果如表4所示。模擬投資所賺的利潤，以及買賣次數如表5所示。目標與模型輸出的結果，如圖5所示；模型的機器學習曲線，如圖6所示;預測誤差如圖7所示，誤差呈現亂數狀態，範圍介於-50至50，代表模型預測能力穩定。

表 1實驗一模型設定

|  |  |
| --- | --- |
| **參數** | **值** |
| 特徵變數(輸入變數) | {} |
| 模糊集數量 | {2, 3, 3, 3} |
| 模糊集型態 |  |
| 複數型態目標數目 |  |
| 神經元數目 |  |
| 前鑑部參數數目 |  |
| T–S神經元數目 |  |
| 後鑑部參數數目 |  |

\* 每一個複數目標的實部與虛部，分別包含兩個實數目標

表 2機器學習參數設定

|  |  |
| --- | --- |
| **PSO** | |
| 粒子群大小 | 50 |
| 迭代次數 | 100 |
|  | {0.8, 2.0, 2.0} |
|  | Random in [0,1] |
| 粒子初始化位置 | 藉由第三章所提SC演算法決定 |
| 粒子初始化速度 | 0 |
| **ABCO** |  |
| 工蜂數量 | 40 |
| 觀察蜂數量 | 10 |
| 迭代次數 | 100 |
| 限制回合數 | 20 |
| **RLSE** | |
|  |  |
|  | 25-by-1 全零向量 |
|  | **I** |
| **I** | 25-by-25 單位矩陣 |

表 3效能比較(TAIEX, 實驗一)

|  |  |
| --- | --- |
| 方法 | RMSE |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| PSO-RLSE |  |
| ABCO-RLSE |  |

表 4十次重複實驗效能統計(TAIEX, 實驗一)

|  |  |  |
| --- | --- | --- |
| Trials | PSO-RLSE | ABCO-RLSE |
| 1 |  |  |
| 2 |  |  |
| 3 |  |  |
| 4 |  |  |
| 5 |  |  |
| 6 |  |  |
| 7 |  |  |
| 8 |  |  |
| 9 |  |  |
| 10 |  |  |

表 5模擬投資利潤表(實驗一)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Trials | PSO-RLSE | | ABCO-RLSE | |
| 利潤 | 買/賣 | 利潤 | 買/賣 |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |
| 5 |  |  |  |  |
| 6 |  |  |  |  |
| 7 |  |  |  |  |
| 8 |  |  |  |  |
| 9 |  |  |  |  |
| 10 |  |  |  |  |



圖 5預測結果

藍色實線為實際目標值，紅色虛線為模型預測值



圖 23 模型預測誤差值



圖 24 實驗一模糊類神經模型學習曲線

## 實驗三：台灣股票加權指數、道瓊工業指數、納斯達克和標準普爾500四目標預測

本實驗使用真實世界的時間序列數據來驗證模型的效能，目標為TAIEX。實驗目標是建立一適當的模型，經過訓練後預測每日股價指數。模型輸出為一複數值，取其實數部分預測目標。 使用的資料為TAIEX 2001年的每日收盤價，此實驗的原始資料為278筆，經過多目標特徵挑選後選出特徵做為模型輸入資料。結構學習部分，會透過減數分群演算法分群，在進行區塊挑選如第三章第二小節所介紹，本實驗各個輸入維度分別有X個複數模糊集，進行區塊挑選後，從原本的XX神經元篩選到剩下XX條規則，藉由資料驅動(Data driven)概念適當地建構模型，每個複數模糊集有3個參數，其中包含中心、標準差以及相位頻率參數，總共有X個複數模糊集，故前鑑部參數數量為XX，後鑑部型態為T-S function，因此總共會有個參數，為後鑑部數目，為模型輸入維度數目。整體模型在結構學習後的參數，如表1所示。機器學習參數設定，如表2所示。本次範例的結果除了兩個演算法比較也與其他論文〔45〕所提的方法做比較，像是Chen (1996)、Yu (2005)、SVR-RBF等。比較結果如表3所示。為了測試模型的穩定性，我們總共跑了十次的實驗，效能統計結果如表4所示。模擬投資所賺的利潤，以及買賣次數如表5所示。目標與模型輸出的結果，如圖5所示；模型的機器學習曲線，如圖6所示;預測誤差如圖7所示，誤差呈現亂數狀態，範圍介於-50至50，代表模型預測能力穩定。

表 1實驗一模型設定

|  |  |
| --- | --- |
| **參數** | **值** |
| 特徵變數(輸入變數) | {} |
| 模糊集數量 | {2, 3, 3, 3} |
| 模糊集型態 |  |
| 複數型態目標數目 |  |
| 神經元數目 |  |
| 前鑑部參數數目 |  |
| T–S神經元數目 |  |
| 後鑑部參數數目 |  |

\* 每一個複數目標的實部與虛部，分別包含兩個實數目標

表 2機器學習參數設定

|  |  |
| --- | --- |
| **PSO** | |
| 粒子群大小 | 50 |
| 迭代次數 | 100 |
|  | {0.8, 2.0, 2.0} |
|  | Random in [0,1] |
| 粒子初始化位置 | 藉由第三章所提SC演算法決定 |
| 粒子初始化速度 | 0 |
| **ABCO** |  |
| 工蜂數量 | 40 |
| 觀察蜂數量 | 10 |
| 迭代次數 | 100 |
| 限制回合數 | 20 |
| **RLSE** | |
|  |  |
|  | 25-by-1 全零向量 |
|  | **I** |
| **I** | 25-by-25 單位矩陣 |

表 3效能比較(TAIEX, 實驗一)

|  |  |
| --- | --- |
| 方法 | RMSE |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
|  |  |
| PSO-RLSE |  |
| ABCO-RLSE |  |

表 4十次重複實驗效能統計(TAIEX, 實驗一)

|  |  |  |
| --- | --- | --- |
| Trials | PSO-RLSE | ABCO-RLSE |
| 1 |  |  |
| 2 |  |  |
| 3 |  |  |
| 4 |  |  |
| 5 |  |  |
| 6 |  |  |
| 7 |  |  |
| 8 |  |  |
| 9 |  |  |
| 10 |  |  |

表 5模擬投資利潤表(實驗一)

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Trials | PSO-RLSE | | ABCO-RLSE | |
| 利潤 | 買/賣 | 利潤 | 買/賣 |
| 1 |  |  |  |  |
| 2 |  |  |  |  |
| 3 |  |  |  |  |
| 4 |  |  |  |  |
| 5 |  |  |  |  |
| 6 |  |  |  |  |
| 7 |  |  |  |  |
| 8 |  |  |  |  |
| 9 |  |  |  |  |
| 10 |  |  |  |  |



圖 5預測結果

藍色實線為實際目標值，紅色虛線為模型預測值



圖 23 模型預測誤差值



圖 24 實驗一模糊類神經模型學習曲線

# 討論

本研究提出以複合型演算法MGCACO-RLSE優化模型參數，並且在訓練資料進入模型前，事先進行特徵選取，選出對預測目標較為有利之特徵資料進行時間序列之預測。根據前3個實驗，顯示本研究的一系列方法對於時間序列的預測擁有較佳的預測性能。

本研究在資料的特徵挑選方針，取自夏農資訊熵的理論。利用資料彼此之間的資訊熵以及資訊熵差量，計算劃分出的資料候選特徵對於目標資料所能提供的資訊量。並考慮與已被選取的特徵的冗餘資訊量，以選取增益代表最後該特徵被選取後能夠提供給目標的資訊量。為了避免選入太多的特徵輸入至模型，針對單目標以及多目標的資料，還會進行二次的特徵篩選，選出所有有正向效益的特徵中，選取增益較為大量的部分特徵，以減少模型對於太多目標所產生的預測負擔。

複數模糊類神經模型以複數模糊集、T-S模糊系統等方式建立。T-S模糊系統能夠處理較為模糊的資訊，以一種非線性的方式描述輸入資料的強度，並且以線性規則代表模型輸出，使其系統可以較為人類所理解。複數模糊集相較於傳統的模糊集合，其歸屬程度從一維實數空間延伸到一個二維單位圓盤的平面空間，此方式使歸屬程度能夠容納更多的資訊，有助於提升模糊系統的推理能力與應用效能。而利用複數歸屬程度可以使模型系統產生雙輸出，提供系統同時預測兩組時間序列資料。除此之外，其複數歸屬程度亦能解構，分別擷取其實數部以及虛數部，讓整體模糊集產生多個歸屬程度，使原本的雙輸出預測能夠預測更多目標，增強模型解釋及預測能力。

進行複數模糊類神經模型的參數學習，利用多群連續型蟻群演算法結合遞迴最小平方演算法分別對模型的前鑑部以及後鑑部參數優化。多群連續型蟻群演算法採用多群搜尋的方式，避免演算法結果過早收斂，並且能同時兼顧蟻群解的多樣性。而本研究在蟻群演算法的演化上，再加上資訊流通、淘汰、繼承等流程。使本研究之多群演算法可以共享每個群體的搜尋結果，並且將表現最差的群體淘汰。但是保留該群體的較佳解再加上其他群體的較佳解，產生新的群體繼續搜索。結合此三流程，加速搜尋的能力以及演化收斂的速度。

後鑑部是採用遞迴式最小平方演算法，利用輸入資料點，以及前次的計算結果，尋找一線性函數，使資料點與該函數的平方誤差達到最小值，藉此找到符合資料關係的函數。並且利用不斷遞迴計算的方式，最佳化其後鑑部參數。

本研究在實驗部分以三個不同資料的實驗呈現建構之模糊類神經網路模型，並且以MGCACO-RLSE優化其參數以及將輸入資料進行特徵選取的實驗結果。針對以上理論及方法將進行以下探討：

## 利用複數類神經網路模型針對單目標資料進行預測

本節將討論本研究所提出全新的多群複合演算法用於複數模型對於單一目標的計算。目的是利用較為容易的單一目標預測，驗證本研究理論在基礎上的可行性。實驗一利用複數類神經網路模型預測效能，並且與多個模型方式進行比較。根據表 6，可以發現本研究測試結果的RMSE為517。相較於其他方式，如最佳之PHM模型，提升約八成的預測效能。比較MAPE，即誤差百分比，也從原本的三成誤差下降到只有3個百分比。該實驗結果如圖 21、圖 22所示，皆有極近似的預測效能。顯示出利用複數模糊集，其歸屬程度包含實數與虛數部分，大幅增加歸屬程度的描述能力，提升模型的預測性能。

## 利用複數型態輸出針對雙目標進行預測

本節將探討採用複數模糊集使模型達到複數型態的輸出，並且針對雙目標進行預測應用，增強模型的應用能力。於實驗二，模型藉由輸入資料，並且利用複數模糊集，在模型輸出，個別擷取其實數部與虛數部，讓兩個不同的數值個別預測不同的目標。實驗的雙目標預測結果如圖 26與圖 28所示，在測試資料集的部份，模型的預測輸出與實際資料極為接近。證實了此複數型態輸出以及模型的預測能力。此模型的輸出校能比較，本實驗的MAPE分別為1.2612%(高盛)與1.0297%(微軟)，皆優於過去的文獻。而且本實驗的預測方式是一次進行雙目標的預測，與以往一次僅進行單一目標預測，困難度更高。本實驗更再次證明加入複數模糊集合的優勢，利用其可以讓模型進行複數型態的輸出，預測兩個不同的時間序列。而且表現皆優於只進行單一目標預測的模型，讓複數模糊集的重要性不言可喻。

## 解構歸屬程度值進行多目標預測

如3.2所敘，利用複數模糊集所得到的複數型態歸屬程度，可以個別擷取其實數部及虛數部，使歸屬函數可以獲得多個不同值，以便進行多個目標的預測應用。於實驗三，利用了歸屬程度的原始值以及其實數部數值，加上複數的模型輸出型態，進行三個不同目標的輸出。利用歸屬程度原始值調整及預測一複數型態的模型輸出，預測道瓊工業指數以及日經平均指數的資料。再將歸屬程度解構，取其歸屬程度實數值用以預測調整及另一模型輸出的實數值，預測巴西股市指數的資料。相關的實驗預測結果如圖 34、圖 36、圖 38所示，在測試資料的部分皆有極為優秀的預測校能。而在校能比較方面，相較於其他的模型預測，雖然並沒有達到最佳效果，但是仍在同一校能階層，證明了利用解構歸屬程度達到多個模型輸出及多目標預測的可行性。未來在數據資料極為龐大之下，能夠以單一模型參數達到多個目標的預測會是一個極大的優勢。

## 特徵選取之應用

於實驗中，雖然利用特徵選取的優勢並無法明顯的從實驗預測結果中得知。但是根據輸入資料產生的候選特徵數量而言，如實驗一，產生了30個候選特徵。而實驗二與實驗三更是多達60個候選特徵。若是將這些候選特徵全數作為輸入資料進入模型，會給模型帶來相當大的資料負擔。但是經由特徵選取，將實驗一與實驗二的輸入資料的特徵降到兩個，實驗三只剩三個。而在預測校能仍然能與對比文獻更加或是在同一效能階層，間接證明了採用特徵選取的優勢所在。

## MGCACO-RLSE複合式演算法效能分析

本研究所採用的模型，由於其前鑑部以及後鑑部的參數量過於龐大，並無法以人工或是直觀的方式直接計算參數值，因此利用此複合式演算法優化參數。但是當求解的參數維度增加，其落入區域最佳解的可能性也增高。於是本研究在優化前鑑部的參數時，採用了多群蟻群演算法。利用其多群搜尋的能力，減少落入區域最佳解的機率。並且加入淘汰的流程，加速尋找最佳解。對於後鑑部參數的部分，採用了遞迴最小平方演算法，利用其可以迅速處理並調整大量後鑑部線性規則參數的優勢，使後鑑部參數不須利用蟻群演算法即可達到優化。

透過三個實驗的學習曲線：圖 23、圖 30、圖 40所示，在少量的時間迭代次數即可達到較佳的效能，並且於學習過程中不斷往全域最佳解搜尋。由此可證明MGCACO-RLSE複合式演算法具有快速的演化速度，並且能夠在落入區域最佳解時可以很快脫離。

# 結論與未來研究方向

## 結論

本研究提出一新型態複合演算法，結合多群連續型蟻群演算法以及遞迴最小平方法，用於優化複數型態模糊類神經模型系統的參數集合。而模型結合複數型模糊集合、T-S模型系統以及類神經網路。於資料輸入前輔以特徵選取，減少過多資料對模型的負擔。經由實驗證明以上方式並統整本研究貢獻如下：

1. 特徵選取處理輸入資料降低模型負擔  
    本研究根據夏農資訊熵的原理與方法，開發出一種新的特徵選取方式。讓輸入資料在進入模型前，針對該資料所產生的候選特徵進行篩選。避免冗餘的輸入資料進入模型，耗費運算效能。
2. 採用複數模糊集使模型能夠有多個複數型態輸出  
    本研究採用的複數模糊集，其複數型態的歸屬程度，讓模型可以有複數值的輸出。而利用解構，將歸屬程度的實數部與虛數部取出各自成為另一歸屬程度，此舉使模型可產生多個輸出。讓模型能夠有同時進行多目標預測的能力。於實驗二與實驗三亦證明此方式的貢獻。
3. MGCACO-RLSE複合式演算法使系統快速學習及避免落入區域解  
    模型系統中的參數好壞可以說是決定預測結果的優劣。本研究採用了多群連續型蟻群演算法結合遞迴最小平方演算法形成一複合式演算法，用於優化模糊模型的前鑑部以及後鑑部參數。蟻群演算法用於優化前鑑部參數，遞迴最小平方演算法用於最佳化後鑑部的參數。此方式可以分配不同類別參數給不同演算法，增加優化的效率。  
    在多群連續型蟻群演算法，加入資訊流通的特性，使不同群體間的搜尋結果可以互相流通，加速各群體搜尋最佳解的速度；加入淘汰的特性，讓演化中效能表現最差之群體被淘汰。由於其效能差，淘汰該群可以減少計算資源不必要的浪費；加入繼承的特性，使前述被淘汰的群體，其較為優秀的蟻群解可以留下。使其即使遭排除，該群體中較為良好的蟻群解可以保留下來，並融入其他群體的較佳解，建立新的一群體。該群體保有各個群體的較佳解，使其表現更為良好，增加預測能力。結合以上三個特性，使預測效能收斂速度加快，並且減少落入區域最佳解的機會。  
    而在遞迴最小平方法，估計的部分將常數項更改為單位矩陣，符合多目標輸出預測的應用。讓原本只能進行單一目標預測應用的演算法，增強到進行多目標的預測及參數調整，增強模型預測及應用能力。

## 未來研究方向

本研究提出之複數模糊類神經模型，並用複合式演算法優化其參數。效能不管是在單一目標或是多目標的預測，根據實驗結果所示，皆能有良好的預測性能。但是仍然有部分細節能夠在未來有延伸。以下提出未來可延續的研究方向：

1. 模型輸入維度之模糊集建立  
    目前研究所建立各輸入維度模糊集的方式是直觀在每個維度建立三個模糊集，並以輸入資料的平均值等初始化模糊集參數，期望能有良好的覆蓋性。但是實際情況中，資料並不會在每個維度上平均分配。若能根據資料在維度上分布的情況，利用群聚演算法，如減法集群演算法(Sub clustering algorithm)或是模糊 c-means演算法(Fuzzy C-means clustering)便能明確的了解資料的分布情況，進而利用其分布情形建立模糊集，讓模糊集能夠更有效率的覆蓋資料點。
2. 前鑑部區域選擇  
    目前在前鑑部區域的劃分是採用各輸入資料維度上的模糊集數目以網格分區法劃分出，但是並非所有劃分出的前鑑部區域都有足夠的資料點分布。若是為了部分僅涵蓋極為少數的資料點而使用該前鑑部，亦會造成不必要的資源消耗。因此未來的研究方向可以朝向利用各個前鑑部當中資料點的密集程度，選取部分的前鑑部於模型中使用。
3. T-S模糊系統  
    本研究中，採用的規則是源自於T-S模糊系統。該方式是以一前鑑部建立一後鑑部。但是若前鑑部數目過多，隨之建立的後鑑部規則數也會相當龐大。加上需要優化的參數數量，對模型是龐大的預測負擔。因此未來可以朝向將前鑑部與後鑑部的建立方式分離，使後鑑部數量不會與前鑑部有太大關係，如此便能使模型建立更為彈性。

# 參考文獻

1. R.F. Engle, “Autoregressive Conditional Heteroscedasticity with Estimates of the Variance of United Kingdom Inflation,” *Econometrica*, vol. 50, iss. 4, p.p. 987-1007, 1982

[ ] T. Bollerslev, ”Generalized Autoregressive Conditional Heteroscedasticity,” *Journal of Econometrics*, vol.31, pp. 307-27, 1986

[ ] T. Kimoto and K. Asakawa and M. Yoda and M. Takeoka,” Stock market prediction system with modular neural networks,” *International Joint Conference* , USA, 1990

[ ] K. Kim and I. Han, “Genetic algorithms approach to feature discretization in artificial neural networks for prediction of stock index,” *Expert System with Applications*, vol. 19, pp. 125–132, 2000

[ ] T.H. Roh, “Forecasting the volatility of stock price index,” *Expert Systems with Applications*, vol.33, pp. 916–922, 2007

[ ] J. Kennedy and R.C. Eberhart, “Particle swarm optimization,” Proceedings *IEEE International Conferencnce on Neural Networks (Perth, Australia)*, vol. 4, pp. 1942-1948, 1995

[ ] R. C. Eberhart and J. Kennedy, “A new optimizer using particle swarm theory,” *Proceedings IEEE International Symposium on Micro Machine and Human Science (Nagoya, Japan)*, pp. 39-43, October 1995

[ ] A. Colorni and M. Dorigo and V. Maniezzo, “Distributed optimization by ant colonies,” *Proceedings of the 1st European Conference on Artificial Life*, pp. 134-142, Paris, 1991

[ ] C.Juang and T.Jeng and Y.Chang, “An Interpretable Fuzzy System Learned Through Online Rule Generation and Multiobjective ACO With a Mobile Robot Control Application,” *IEEE Transactions on Cybernetics*, vol. 46, iss. 12, pp.2706-2718, December 2015

[ ] H.J. Sadaei and R. Enayatifar and M.H. Lee and M. Mahmud “A hybrid model based on differential fuzzy logic relationships and imperialist competitive algorithm for stock price forecasting,” *Applied Soft Computing*, vol.40, pp. 132-149, March 2016

[ ] C.E. Shannon and W. Weaver, “*The Mathematical Theory of Communication*,” Univ of Illinois Press, 1949.

[ ] D. Ramot and R. Milo and M.Friedman and A. Kandel , “Complex fuzzy sets,” *IEEE Transactions on Fuzzy Systems*, vol. 10, no. 2, pp. 171-186, April 2002

[ ] D.Pedro ,” A few useful things to know about machine learning,” *Communications of the ACM*, vol.55, iss. 10, pp.78-87, October 2012

[ ] M. Dash and H. Liu, “Feature selection for classification,” *Intelligent Data Analysis*, vol.1, iss. 1-4, pp. 131-156, 1997

[ ] L. Yu and H. Liu, “Efficient Feature Selection via Analysis of Relevance and Redundancy,” *The Journal of Machine Learning Research*, vol. 5, pp. 1205-1224, Oct 1 2004

[ ] K. Kira and L.Rendell, “ A practical approach to feature selection,” In: *Proc. of the Ninth International Conference on Machine Learning*, pp. 249–256, 1992

[ ] M.A. Hall, “Correlation-based Feature Selection for Machine Learning,” University of Waikato, Hamilton, New Zealand, PhD thesis, 1999

[ ] H. Yu and J. Oh and W.-S. Han, “Efficient feature weighting methods for ranking,” in:*Proc. of the 18th ACM Conference on Information and Knowledge Management* *Hong Kong, China*, pp. 1157-1166, 2009

[ ] Z. Sun and T. Qin and Q. Tao and J. Wang, ”Robust sparse rank learning for non-smooth ranking measures,” in:*Proc. of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2009) 20 Boston, MA*, pp. 259-266, 2009

[ ] R.Clausius, " Ueber eine veränderte Form des zweiten Hauptsatzes der mechanischen Wärmetheorie," *Annalen der Physik und Chemie*, vol.93, issue.12, pp. 481–506, 1854

[ ] G. Cantor, “Ueber eine Eigenschaft des Inbegriffs aller reellen algebraischen Zahlen,” *Journal für die reine und angewandte Mathematik*, no. 77, pp. 258-262, 1874

[ ] B. Russell, “Vagueness,” *Australasian Journal of Philosophy*, vol.1, iss. 2, pp.84 – 92, 1923

[ ] L.A. Zadeh, ”Fuzzy sets,” *Information and Control*, vol.8, issue.3, pp.338-353, June 1965

[ ] D. Ramot and R. Milo and M. Friedman and A. Kandel, “Complex fuzzy sets,” *IEEE Transactions on Fuzzy Systems*, vol. 10, iss. 2, pp. 171-186, April 2002

[ ] C. Li and T. Chiang, “Complex Neurofuzzy ARIMA Forecasting—A New Approach Using Complex Fuzzy Sets,” *IEEE Transactions on Fuzzy Systems*, vol. 21, iss. 3, pp. 567-584, June 2013

[ ] F. Rosenblatt, “The Perceptron: A Probabilistic Model For Information Storage And Organization in the Brain,” *Psychological Review*, vol.65, no.6, pp. 386–408, December 1958

[ ] D.E. Rumelhart and J. L. McClelland,” *Parallel distributed processing: explorations in the microstructure of cognition, vol. 1: foundations*”, January 1986

[ ] J.H. Holland, “Adaptation in Natural and Artificial Systems,” *Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology*, Control and Artificial Intelligence, MIT Press Cambridge, MA, USA, ISBN:0262082136

[ ] A.Colorni and M. Dorigo and V. aniezzo,“Distributed optimization by antcolonies,”*Proceedings of the 1st European Conference on Artificial Life*,pp.134-142, Paris, 1991.

[ ] K. Socha and M. Dorigo,” Ant colony optimization for continuous domains,” *European Journal of Operational Research*, vol.185, iss. 3, pp. 1155–1173, 16 March 2008

[ ] T. Takagi and M. Sugeno, “Fuzzy identification of systems and its applications to modeling and control,” *IEEE Trans. Syst. Man Cybern*, vol.15, pp. 116–132, 1985

[ ] P. Husbands and F. Mill, “Simulated Co-Evolution as the Mechanism for Emergent Planning and Scheduling,” Proceedings of the 4th International Conference on Genetic Algorithms, San Diego, CA, USA, pp. 264-270, January 1991

[ ] F. van den Bergh and A.P. Engelbrecht, “A Cooperative approach to particle swarm optimization,” IEEE Transactions on Evolutionary Computation, vol.8 , iss.3, pp.225-239, June 2004

[ ] J.J. Wang and J.Z. Wang and Z.G. Zhang and S.P. Guo, “Stock index forecasting based on a hybrid model,” Omega, vol. 40, iss. 6, pp. 758-766, December 2012

[ ] JL Ticknor, “A Bayesian regularized artificial neural network for stock market forecasting, ”Expert Systems with Applications,” Expert Systems with Applications, vol. 44, iss. 14, pp. 5501-5506, 15 October 2013

[ ] L.J. Kao and C.C. Chiu and C.J. Lu and C.H. Chang, “A hybrid approach by integrating wavelet-based feature extraction with MARS and SVR for stock index forecasting ,” *Decision Support Systems*, vol. 54, iss. 3, pp. 1228-1244, February 2013

[ ] 國立中央大學資訊管理所李俊賢教授, 研究生訓練課程內容2015-2017, 紀錄筆記。 (未發表)