國立政治大學金融學系研究所

碩士學位論文

# 金融壓力事件預警模型：
# 類神經網路、支援向量機與羅吉斯迴歸之比較

Financial Stress Model：

Comparison of Artificial Neural Network, Support Vector Machine and Logistic Regression

國立政治大學
National Chengchi University

指導教授：林士貴 博士

蔡炎龍 博士

研究生：朱君亞 撰

中 華 民 國 一 百 零 七 年 七 月

DOI:10.6814/THE.NCCU.MB.013.2018.F06

金融壓力事件預警模型：類神經網路、支援向量機與羅吉斯迴歸之比較

學生：朱君亞　　指導教授：林士貴　博士
蔡炎龍　博士
國立政治大學金融學系

# 摘要

全球化浪潮、科技進步與金融商品日趨複雜等原因，讓金融商品價格容易在短時間內發生大幅波動，也使得風險管理成為不可避免的議題。本文遂針對重大金融危機事件做出定義，稱之為「金融壓力事件」，以股市大跌作為「金融壓力事件」是否發生的判定標準，且分別針對 6 個月與 3 個月的兩種不同判定期間做出定義。再以利率、匯率、資產價格等變數，搭配羅吉斯迴歸模型、類神經網路模型與支援向量機，建立金融壓力事件預警模型，每日針對是否發生「金融壓力事件」進行預測。本研究以台灣加權股價指數實證，實證結果顯示，不論哪一種方法，判定期間為 6 個月的預測結果都優於判定期間為 3 個月；不論判定期間長短，都是支援向量機預測能力最好，其次為類神經網路模型，羅吉斯迴歸則較弱。

關鍵詞：金融情勢指數、金融壓力事件預警模型、羅吉斯迴歸模型、類神經網路模型、支援向量機



DOI:10.6814/THE.NCCU.MB.013.2018.F06

Financial Stress Model：Comparison of Artificial Neural Network, Support Vector Machine and Logistic Regression

Student: Chun-Ya Chu

Advisor: Dr. Shih-Kuei Lin
Dr. Yen-Lung Tsai

Submitted to Department of Money and Banking
Collage of Commerce National Chengchi University

# Abstract

With globalization, new technology and more complicated financial instruments, the financial market become more volatile, making risk management an inevitable issue. In this paper, we define a major financial crisis event as a "financial stress event." It uses the stock market crash as a criterion for the occurrence of a "financial stress event," and define two different judgment periods of 6-months and 3-months respectively. Using the variables such as interest rate, exchange rate, and asset price, together with the Logistic Regression model, Artificial Neural Network model, and Support Vector Machine, a financial stress model was established to predict the occurrence of "financial stress events" everyday. By using Taiwan Capitalization Weighted Stock Index as empirical evidence, the result shows that regardless of the method, the predictability of 6-months judgment period is better than the 3-months period. Regardless of the length of the judgment period, Support Vector Machine has the highest predictability, followed by Artificial Neural Network model, and Logistic Regression is the weakest.

Keywords：Financial Conditions Index, Financial Stress Model, Logistic Regression, Artificial Neural Network, Support Vector Machine



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 目 次

第一章 緒論.....1
第一節 研究背景與動機.....1
第二節 金融壓力事件定義.....2
第三節 研究目的.....3
第二章 文獻回顧.....4
第一節 金融情勢指數使用變數.....4
第二節 金融市場預測之研究方法.....5
第三章 研究方法.....8
第一節 迴歸模型.....8
第二節 類神經網路模型.....9
第三節 支援向量機.....12
第四章 實證資料.....17
第一節 資料期間.....17
第二節 變數選取.....17
第三節 模型建構.....19
第四節 超參數設定.....19
第五章 實證結果.....26
第一節 羅吉斯迴歸模型實證結果.....27
第二節 類神經網路模型實證結果.....29
第三節 支援向量機實證結果.....31
第四節 各模型結果比較.....31
第六章 結論.....33
參考文獻.....35
中文文獻.....35
英文文獻.....35

國立政治大學 National Chengchi University



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 表 次

表 2-1：FCI 變數之相關文獻整理 ..... 4

表 2-2：國內外金融市場預測方法之相關文獻整理 ..... 6

表 3-1：基本的核函數 ..... 16

表 4-1：金融變數列表 ..... 18

表 5-1：預測結果分類表 ..... 26

表 5-2：性能指標定義 ..... 26

表 5-3：羅吉斯迴歸模型實證結果-6 個月金融壓力事件預警模型 ..... 27

表 5-4：羅吉斯迴歸模型實證結果-3 個月金融壓力事件預警模型 ..... 28

表 5-5：類神經網路模型實證結果-6 個月金融壓力事件預警模型 ..... 29

表 5-6：類神經網路模型實證結果-3 個月金融壓力事件預警模型 ..... 30

表 5-7：支援向量機實證結果 ..... 31

國立政治大學
National Chengchi University



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 圖 次

圖 1-1：金融壓力事件圖示 ..... 2

圖 3-1：類神經網路架構 ..... 10

圖 3-2：支援向量機架構 ..... 13

圖 4-1：模型建構方式圖示 ..... 19

圖 4-2：6 個月金融壓力事件之誤差收斂圖-7 個神經元 ..... 20

圖 4-3：6 個月金融壓力事件之誤差收斂圖-13 個神經元 ..... 21

圖 4-4：6 個月金融壓力事件之誤差收斂圖-26 個神經元 ..... 21

圖 4-5：6 個月金融壓力事件之誤差收斂圖 ..... 22

圖 4-6：3 個月金融壓力事件之誤差收斂圖-7 個神經元 ..... 22

圖 4-7：3 個月金融壓力事件之誤差收斂圖-13 個神經元 ..... 23

圖 4-8：3 個月金融壓力事件之誤差收斂圖-26 個神經元 ..... 23

圖 4-9：3 個月金融壓力事件之誤差收斂圖 ..... 24

圖 4-10：6 個月金融壓力事件之 SVM 誤差圖 ..... 25

圖 4-11：3 個月金融壓力事件之 SVM 誤差圖 ..... 25

圖 5-1：各模型比較-6 個月金融壓力事件預警模型 ..... 32

圖 5-2：各模型比較-3 個月金融壓力事件預警模型 ..... 32

國立政治大學 National Chengchi University



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 第一章 緒論

## 第一節 研究背景與動機

隨著金融市場的發展，投資人對金融商品的需求越來越多樣化，金融商品也就越來越複雜，再加上日新月異的科技與全球化浪潮，更容易讓金融商品價格在短時間內發生大幅波動。尤其近幾年黑天鵝事件頻傳，引發金融市場波動加劇，風險管理已成為不可避免的議題。在 2008 年全球金融風暴後，Silvers and Slavkin (2009) 認為風險控制與監管機制的缺失是導致金融危機的主要原因，因此如何訂定風險管理機制成為一項重要課題。

為掌握整體金融情勢，加拿大銀行自 1990 年中期開始編製金融情勢指數（Financial Conditions Index, 簡稱 FCI），為融合利率、匯率、資產價格等多種金融變數的單一指標。然而 FCI 估計出來的是一數值，應用於金融市場前尚須衡量 FCI 的方向與數值意義。本研究希望訂定出一套更容易使用的金融市場風險判定系統，旨在針對重大金融危機事件做出預警。本文遂針對重大金融危機事件做出定義，並稱之為「金融壓力事件」，再根據 FCI 使用的利率、匯率、資產價格等變數，搭配分類模型，建立金融壓力事件預警模型。

金融壓力事件預警模型為二元分類模型，模型每日針對金融壓力事件做出「發生」或「未發生」的預測。傳統的統計分析中，多使用羅吉斯迴歸(Logistic Regression)做為二元分類模型。在金融科技的浪潮下，將人工智慧運用於金融是未來的趨勢，因此本研究除了傳統統計方法外亦使用人工智慧方法，常見的人工智慧分類方法有決策樹(Decision Tree)、類神經網路(Artificial Neural Network)與支援向量機(Support Vector Machine)。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

本研究欲探討人工智慧方法相較於傳統統計方法，是否能提高金融壓力事件預警模型的預測能力，因此分別使用傳統羅吉斯迴歸模型、人工智慧的類神經網路模型與支援向量機，建立金融壓力事件預警模型，並以台灣加權股價指數實證，比較三種模型的預測能力，決定哪一模型最為合適。

## 第二節 金融壓力事件定義

本文的「金融壓力事件」即代表發生重大金融危機事件，而通常都伴隨股市大跌，因此以股市大跌作為「金融壓力事件」是否發生的判定標準，判定期間越長，股市下跌的機率就越大，故分別針對兩種不同判定期間做出定義如下：

若判定期間為 6 個月，則定義為：自預警日到未來 6 個月內，任意一日發生股價跌破現值的 80%即定義為預警日發生金融壓力事件。

若判定期間為 3 個月，則定義為：自預警日到未來 3 個月內，任意一日發生股價跌破現值的 90%即定義為預警日發生金融壓力事件。

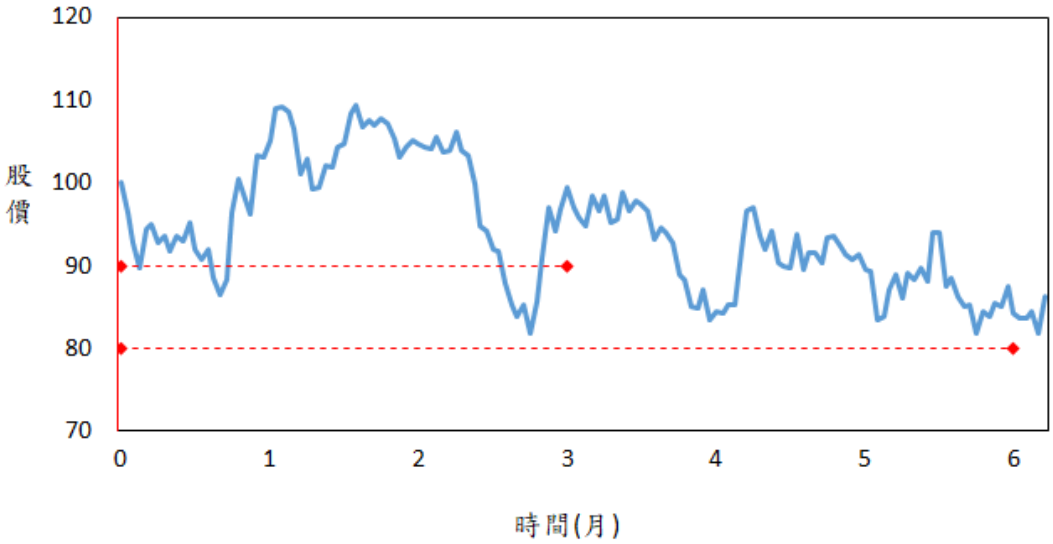


圖 1-1：金融壓力事件圖示



DOI:10.6814/THE.NCCU.MB.013.2018.F06

舉例說明，如圖 1-1 所示，假設今日為預警日，圖中實線代表股價，今日股價為 100 元整。自今日起至未來三個月內，股價曾經低於今日的 90%，也就是 90 元整，因此定義為今日「發生」3 個月的金融壓力事件；而自今日起至未來 6 個月內，不曾發生股價低於今日的 80%，也就是 80 元整，因此定義為今日「未發生」6 個月的金融壓力事件。

## 第三節 研究目的

本文研究目的主要分為模型與實證兩部分。模型部分，介紹了傳統多變量分析中的羅吉斯迴歸模型、人工智慧方法的類神經網路模型與支援向量機，以及各個模型的超參數(Hyperparameter)選擇與設定；實證部分，使用 2002 年到 2017 年台灣加權股價指數資料，結合利率、匯率、資產價格變數，以上述模型建立金融壓力事件預警模型，並以準確率、F1-Score、Type I Error 與 Type II Error 比較三個模型的預測能力。

本文後續章節架構如下：第二章回顧金融情勢指數所使用的變數、金融市場預測之研究方法相關文獻；第三章介紹羅吉斯迴歸模型、逐步迴歸分析、類神經網路模型與支援向量機；第四章以台灣加權股價指數實證，描述資料期間、變數選取、模型建構方式、超參數設定；第五章分析實證結果；第六章為本文研究結論。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 第二章 文獻回顧

## 第一節 金融情勢指數使用變數

金融情勢指數（Financial Conditions Index, 簡稱FCI），起源於貨幣情勢指數（Monetary Conditions Index, 簡稱MCI），最早由加拿大銀行自1990年中期開始編製。因其為融合利率、匯率、資產價格等多種金融變數的單一指標，可以用來描述一個國家的整體金融情勢，並提供貨幣當局作為貨幣決策時之考量，故受到各界的重視，尤其2008年全球金融海嘯後，更彰顯出其重要性。

相較於MCI僅參考利率與匯率變數，FCI相關文獻之變數選取相當多元，彙整於表2-1中。除了利率與匯率變數外， Goodhart and Hofmann (2001) 以G7國家實證編制的FCI，使用的變數包含實質房價與實質股價；Gauthier et al. (2004) 編制的加拿大FCI中，包含貸款資料、利差；Skaarup et al. (2010) 編制的丹麥FCI中，包含油價、貨幣供給、信用放款等；Hatzius et al. (2010) 編制的美國FCI中，包含信用傳遞管道變數、二階動差（包含股票與債券報酬率相關性、股價波動率與VIX指數）等。

表2-1：FCI變數之相關文獻整理

| 學者 | 編制對象 | 使用變數 |
|---|---|---|
| Goodhart and Hofmann (2001) | G7國家 | 利率、實質有效匯率指數、股價與實質房價 |
| Gauthier et al. (2004) | 加拿大 | 貸款資料、利差、名目有效匯率指數、利率、股價與房價 |
| Skaarup et al. (2010) | 丹麥 | 利差、利率、匯率、股價、房價、油價、貨幣供給、信用放款 |
| Hatzius et al. (2010) | 美國 | 資產價格、利率、信用傳遞管道變數、放款調查資料與二階動差變數等 |
| 張天惠（2012） | 臺灣 | 匯率、利率、資產價格變數 |



DOI:10.6814/THE.NCCU.MB.013.2018.F06

綜合上述可以發現，FCI之變數選取沒有放諸四海的標準，因此本文以較多文獻選用且資料取得較容易之變數，並參考張天惠（2012）之作法，自股票市場、匯率市場與利率市場三個管道中選擇金融變數。

## 第二節 金融市場預測之研究方法

國內外針對金融市場預測的參考文獻眾多，研究方法亦不勝枚舉，以下針對使用統計迴歸、類神經網路與支援向量機等方法之文獻回顧，並將結論彙整於表2-2。

Kim (2003) 使用支援向量機、倒傳遞類神經網路、案例推論等方法，以12項技術指標為輸入變數，對韓國綜合股價指數之漲跌方向進行實證分析，結果發現支援向量機優於倒傳遞類神經網路與案例推論，在股市預測中為有前途的模型。

Huang et al. (2004) 使用倒傳遞類神經網路、支援向量機、線性分析、二次分析等方法，以對日本出口影響最大的美國之變數：S&P 500指數與日圓兌美元匯率作為輸入變數，預測日經225指數之一週漲跌方向，結果發現支援向量機優於其他模型，支援向量機結合其他分類方法的模型則有最好的預測能力。

黃華山與邱一薰 (2005) 運用倒傳遞類神經網路做為預測台灣50股票指數的研究，以台灣50中排名前20名之成份股及台灣50指數技術指標資料作為輸入變數，結果發現類神經網路模式具有良好的預測能力，且若配合該研究提出之交易策略可獲得較高的報酬。

呂奇傑等 (2009) 嘗試以支援向量機、支援向量迴歸為工具，建構財務時間序列預測模式，對台灣加權股價指數之漲跌方向預測進行實證研究，使用包含技術分析指標、外資每日淨買賣金額及美國加權股價指數的一階差分值共29個輸入



DOI:10.6814/THE.NCCU.MB.013.2018.F06

變數，結果發現不論在漲跌幅度的預測誤差或是趨勢預測準確度的表現上，支援向量機與支援向量迴歸均較倒傳遞類神經網路為佳。

表2-2：國內外金融市場預測方法之相關文獻整理

| 學者 | 研究方法 | 研究對象 | 研究結果 |
|---|---|---|---|
| Kim(2003) | 支援向量機、倒傳遞類神經網路、案例推論 | 韓國綜合股價指數之漲跌方向 | 支援向量機優於倒傳遞類神經網路與案例推論，在股市預測中為有前途的模型。 |
| Huang et al.(2004) | 倒傳遞類神經網路、支援向量機、線性分析、二次分析 | 日經225指數之一週漲跌方向 | 支援向量機優於其他模型，支援向量機結合其他分類方法的模型則有最好的預測能力。 |
| 黃華山與邱一薰(2005) | 倒傳遞類神經網路 | 台灣50股票指數 | 類神經網路模式具有良好的預測能力，且若配合該研究提出之交易策略可獲得較高的報酬。 |
| 呂奇傑等(2009) | 支援向量機、支援向量迴歸、倒傳遞類神經網路 | 台灣加權股價指數之漲跌方向 | 不論在漲跌幅度的預測誤差或是趨勢預測準確度的表現上，支援向量機與支援向量迴歸均較倒傳遞類神經網路為佳。 |
| Hsieh et al.(2011) | 結合DOE、田口方法、倒傳遞類神經網路 | 公司股票價值 | 結合DOE、田口方法、倒傳遞類神經網路可有效的改善股票價值的預測率。 |
| Kara et al.(2011) | 類神經網路、支援向量機 | 土耳其Nation-100指數之漲跌方向 | 類神經網路(75.74%)之預測能力優於支援向量機(71.52%)。 |
| 王翎聿(2015) | 倒傳遞類神經網路、支援向量機 | 台灣加權股價指數、電子類股指數、金融類股指數之隔日漲跌方向 | 以預測正確率做比較，倒傳遞類神經網路在各股價指數之預測能力優於支援向量機。 |



DOI:10.6814/THE.NCCU.MB.013.2018.F06

Hsieh et al. (2011) 以結合DOE、田口方法、倒傳遞類神經網路的方法，使用7項技術指標作為輸入變數，預測公司股票價值，發現結合DOE、田口方法、倒傳遞類神經網路可有效的改善股票價值的預測率。

Kara et al. (2011) 使用類神經網路、支援向量機，以10項技術指標作為輸入變數，預測土耳其Nation-100指數之漲跌方向，結果發現類神經網路(75.74%)之預測能力優於支援向量機(71.52%)。

王翎聿 (2015) 使用倒傳遞類神經網路、支援向量機，以6種技術指標，對台灣加權股價指數、電子類股指數、金融類股指數之隔日漲跌方向進行預測，結果發現，以預測正確率做比較，倒傳遞類神經網路在各股價指數之預測能力優於支援向量機。

國立政治大學 National Chengchi University



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 第三章 研究方法

## 第一節 迴歸模型

建立迴歸分析模型時，若有很多個自變數，則互相之間可能存在「共線性」問題，因此，本研究在建立羅吉斯迴歸模型（Logistic Regression Model）前，先以逐步迴歸分析（Stepwise Regression Model）篩選變數，再使用篩選過後的變數建立羅吉斯迴歸模型。以下將分別介紹羅吉斯迴歸模型與逐步迴歸分析。

### 一、羅吉斯迴歸模型

羅吉斯迴歸模型為多變量分析的一種，可以針對事件的發生機率進行預測，因此常用於依變數為二元類別資料的情形。

羅吉斯迴歸模型是由(3.1)式的線性機率模型（Linear Probability Model）衍生：

$$P_i = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i = \beta' X_i + \varepsilon_i \tag{3.1}$$

其中 $P_i$ 為依變數，也就是事件發生的機率，$X_i$為自變數，$\beta_i$為參數估計值，$\varepsilon_i$為隨機誤差。線性機率模型可以用來估計一個事件發生的機率，然而他的迴歸估計值 $E(P_i)$ 不一定會落在 [0,1] 之間，因此學者將依變數 $P_i$ 用Logistic累積機率密度函數（$L(x) = \frac{1}{1+e^{-x}}$）轉換，以確保估計值落在 [0,1] 之間，亦即：假設 $Z_i$ 為服從Logistic累積機率密度函數的隨機變數，令 $P_i = L(Z_i) = \frac{1}{1+e^{-z_i}}$，則 $Z_i = \ln\frac{P_i}{1-P_i}$，再以 $Z_i$ 作為依變數進行迴歸估計，即(3.2)式，

$$Z_i = \ln\frac{P_i}{1-P_i} = \beta_0 + \beta_1 X_{i1} + \beta_2 X_{i2} + \cdots + \beta_k X_{ik} + \varepsilon_i \tag{3.2}$$

可得到估計值 $\mathrm{E}(Z_i) = \beta' X_i$，轉換回 $P_i$ ，則 $E(P_i) = \frac{1}{1+e^{-z_i}}$ 。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

## 二、逐步迴歸分析

逐步迴歸分析為多元迴歸分析中一種挑選自變數的方法。開始時迴歸方程內除常數外無其他自變數，再透過增加或減少自變數，並同時進行檢定來決定是否要挑選該自變數，以下為逐步回歸之步驟：

步驟一：設定進入門檻與剔除門檻。

步驟二：針對所有迴歸方程外的自變數進行F檢定，選擇P-value小於進入門檻且最小的自變數放入迴歸方程。

步驟三：對所有迴歸方程內的自變數進行F檢定，剔除所有P-value大於剔除門檻的自變數。

步驟四：重複步驟二與步驟三，直到不再有變數進出迴歸方程為止。

# 第二節 類神經網路模型

類神經網路為模擬生物腦中神經元傳遞的人工智慧方法，依資訊的傳遞方向、神經元的連接方式可以分為許多種不同的神經網路，本研究選擇基本的前饋神經網路(Feedforward Neural Network)，亦即資訊輸入後僅經過單一方向的傳遞便輸出，架構則為全連接(Fully Connected)架構。此種類神經網路屬於監督式學習，意旨需要藉由訓練資料來訓練模型，而訓練資料須包含「輸入特徵(Input)」與「目標結果(Output)」，透過調整參數訓練模型，讓「輸入特徵(Input)」經由類神經網路模型的轉換後，得到「目標結果(Output)」。

類神經網路的架構如圖3-1所示，主要分為三大層：輸入層(Input Layer)、隱藏層(Hidden Layer)與輸出層(Output Layer)，其中隱藏層可設置一至多層，視資料需求自行決定層數，而各層神經元個數亦視需要自行決定神經元個數。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

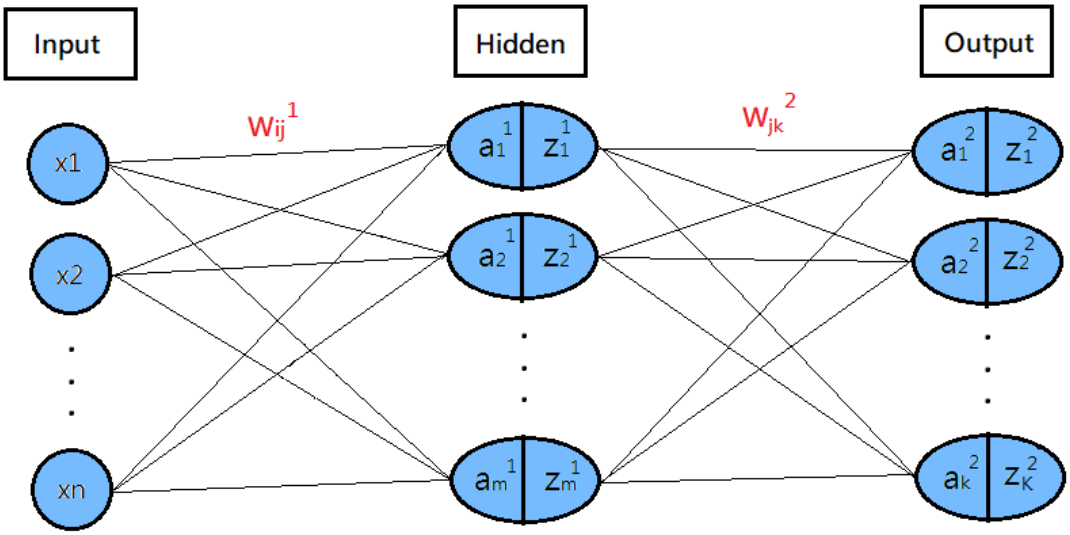


圖 3-1：類神經網路架構

輸入層即為訓練資料之特徵值$x_i$；隱藏層分為接收值$a_j^1$與輸出值$z_j^1$兩部分，接收值$a_j^1$為輸入層乘上權重加總並加上Bias，輸出值$z_j^1$為$a_j^1$經由活化函數(Activation Function)轉換後，傳向下一層；輸出層同樣分為接收值$a_k^2$與輸出值$z_k^2$兩部分，接收值$a_k^2$為上一層輸出值$z_j^1$乘上權重加總並加上Bias，輸出值$z_k^2$為$a_k^2$經由活化函數轉換後，即為與輸入特徵相對應的目標結果。其中活化函數是為解決非線性問題，因此須為非線性函數，且因為計算權重時須對活化函數微分，因此須為可微函數，本研究中選擇Logistic Function作為活化函數。整體架構以數學式表達如下：

$$a_j^1 = \sum_i^n x_i \times w_{ij}^1 + b_j^1 \quad ,j = 1,2,\dots,\mathrm{m}$$

$$a_k^2 = \sum_j^m z_j^1 \times w_{jk}^2 + b_k^2 \quad ,k = 1,2,\dots,K$$

$$z_j^h = \varphi(a_j^h) \ \text{, where} \ \ h = 1,2 \ \ , \ \ \varphi(x) = \frac{1}{1+e^{-x}}$$

為使輸入特徵經由模型轉換得到目標結果，需要找到最適合的權重，此過程



DOI:10.6814/THE.NCCU.MB.013.2018.F06

即為訓練模型。類神經網路是以「誤差倒傳遞」的方式訓練模型，首先隨機給定一組初始權重，輸入特徵後可算出一個預測值，預測值與目標結果之差為誤差(Loss Function)，接著利用誤差往回傳的方式計算每一個權重對誤差的影響力，繼而決定權重調整的比例，透過不斷計算誤差並調整權重，最後達到誤差最小化，即可得到最適合的權重，過程敘述如下：

目　標：Minimize $L$

各神經元誤差 $E_{O_k} = Y_k - z_k^H$，其中$Y_k$為目標結果，$z_k^H$為預測值，總誤差(Loss Function)為 $L = \frac{1}{2}\sum(Y_k - z_k^H)^2$，平方是為確保誤差為正，係數$\frac{1}{2}$則為之後微分計算方便。

步驟一：給定初始權重

首先給定一組隨機的初始權重。

步驟二：計算 $\frac{\partial L}{\partial w_{ij}^h}$ ，$h = 1,2,\ldots,H$

計算每一個權重對 $L$ 的影響力：$\frac{\partial L}{\partial w_{ij}^h}$，微分的計算則透過微積分的連鎖率，從最後一層往前一層一層微分計算，以數學式表達如下：

For $h = H$(隱藏層與輸出層中間的權重)：

$$\frac{\partial L}{\partial w_{jk}^H} = \frac{\partial L}{\partial z_k^H} \times \frac{\partial z_k^H}{\partial a_k^H} \times \frac{\partial a_k^H}{\partial w_{jk}^H} = -(Y_k - z_k^H) \times \varphi'(a_k^H) \times z_k^{H-1} = \delta_k^H \times z_k^{H-1}$$

where $\delta_k^H = \frac{\partial L}{\partial z_k^H} \times \frac{\partial z_k^H}{\partial a_k^H} = -(Y_k - z_k^H) \times \varphi'(a_k^H)$

For $b_k^H$(隱藏層與輸出層中間的bias )：

$$\frac{\partial L}{\partial b_k^H} = \frac{\partial L}{\partial z_k^H} \times \frac{\partial z_k^H}{\partial a_k^H} \times \frac{\partial a_k^H}{\partial b_k^H} = -(Y_k - z_k^H) \times \varphi'(a_k^H) \times 1 = \delta_k^H$$



DOI:10.6814/THE.NCCU.MB.013.2018.F06

For $h = 1,2,\dots,H-1$(隱藏層與輸入層、隱藏層中間的權重)：

$$\frac{\partial L}{\partial w_{ij}^{h}} = \frac{\partial L}{\partial z_{j}^{h}} \times \frac{\partial z_{j}^{h}}{\partial a_{j}^{h}} \times \frac{\partial a_{j}^{h}}{\partial w_{ij}^{h}}$$

$$\frac{\partial L}{\partial z_{j}^{h}} = \sum_{k} \frac{\partial L}{\partial z_{k}^{h+1}} \times \frac{\partial z_{k}^{h+1}}{\partial a_{k}^{h+1}} \times \frac{\partial a_{k}^{h+1}}{\partial z_{j}^{h}}$$

$$\frac{\partial L}{\partial w_{ij}^{h}} = \left( \sum_{k} \frac{\partial L}{\partial z_{k}^{h+1}} \times \frac{\partial z_{k}^{h+1}}{\partial a_{k}^{h+1}} \times \frac{\partial a_{k}^{h+1}}{\partial z_{j}^{h}} \right) \times \varphi'\left(a_{j}^{h}\right) \times z_{k}^{h-1} = \delta_{j}^{h} \times z_{k}^{h-1}$$

where $\delta_{j}^{h} = \left( \sum_{k} \frac{\partial L}{\partial z_{k}^{h+1}} \times \frac{\partial z_{k}^{h+1}}{\partial a_{k}^{h+1}} \times \frac{\partial a_{k}^{h+1}}{\partial z_{j}^{h}} \right) \times \varphi'\left(a_{j}^{h}\right)$

For $b_{j}^{h}(h = 1,2,\dots,H-1)$(隱藏層與輸入層、隱藏層中間的bias)：

$$\frac{\partial L}{\partial b_{j}^{h}} = \frac{\partial L}{\partial z_{j}^{h}} \times \frac{\partial z_{j}^{h}}{\partial a_{j}^{h}} \times \frac{\partial a_{j}^{h}}{\partial b_{j}^{h}} = \delta_{j}^{h} \times 1 = \delta_{j}^{h}$$

步驟三：利用梯度下降的方式調整權重

新權重為舊權重減 $\frac{\partial L}{\partial w_{ij}^{h}}$ 乘上學習率 $\eta$，學習率越大則收斂速度越快。

$w^{new} = w - \eta \times \frac{\partial L}{\partial w}$ , where $\eta$ is Learning Rate

步驟四：重複步驟二與步驟三直到 $L$ 收斂

在重複的過程中，利用梯度下降的方式找到最小值，讓 $L$ 收斂。

## 第三節 支援向量機

支援向量機(Support Vector Machine)是一種監督式機器學習方法，將資料放入d維空間中，欲找到一個超平面(Hyperplane)，可將資料區分為兩類。但是許多資料並非線性可分，因此發展出使用非線性的核函數(Kernel Function)來處理非線性可分的資料。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

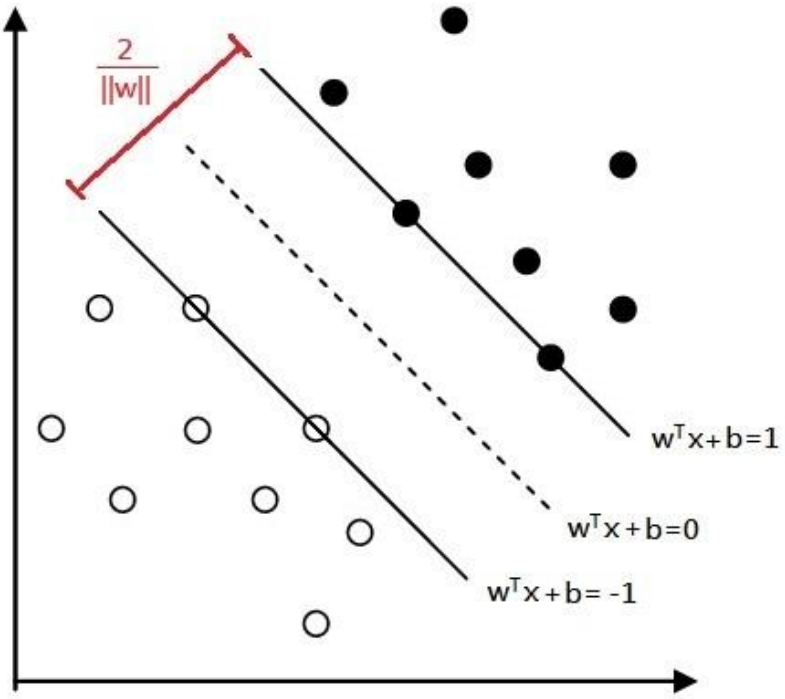


圖 3-2：支援向量機架構

支援向量機的架構如圖3-2所示，圖上每一個點代表一筆資料，資料中包含向量 $x_i$與 $y_i$。$x_i$為資料的特徵，因此特徵的個數即為 $x_i$的長度；$y_i$代表分類的類別，依照 $y_i$的類別可以將資料在d維空間中分為兩邊，找出可以區分兩邊的超平面(Hyperplane)即為目的。而在多個可區分的超平面中，距離兩邊資料臨界點最遠的超平面為最佳的超平面，也是我們所欲尋找的。以數學式表達如下：

$$x_i \in R^d,\ y_i \in \{-1,1\} \quad \text{where } i \in N, N = \{1,2,\ldots,n\}$$

$$\text{Hyperplane}: w^T x + b = 0 \tag{3.3}$$

$$\text{condition on } \begin{cases} w^T x_i + b \geq 1 & if\ y_i = 1 \\ w^T x_i + b \leq -1 & if\ y_i = -1 \end{cases} \tag{3.4}$$

$$\text{or equivalently,}\ \ y_i(w^T x_i + b) \geq 1\ \ \forall\, i \in N \tag{3.5}$$

但是並非所有資料都可以完美被區分成兩側，因此需要在原本的條件(3.5)式中加入允許誤差的寬鬆變數(Slack Variables) $\xi_i$ 及控制誤差的常數$C$，改寫為(3.6)式。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

$$y_i(w^T x_i + b) \geq 1 - \xi_i \quad \forall \mathrm{i} \in \mathrm{N} \tag{3.6}$$

而許多資料並非線性可分，此時可以將$x_i$透過非線性的映射函數$\Phi: R^d \longrightarrow F$轉換到另外一個較高維度的特徵空間(Feature Space)中，使得資料在高維度特徵空間中線性可分，故將(3.6)式改為(3.7)式

$$y_i(w^T \Phi(x_i) + b) \geq 1 - \xi_i \quad \forall \mathrm{i} \in \mathrm{N} \tag{3.7}$$

為找出距離兩邊資料臨界點最遠的超平面，需要估計其中的係數$w$，步驟如下：

目　標：Maximize $\dfrac{2}{||w||^2} - C\sum_{i=1}^{n}\xi_i \equiv$ Minimize $\dfrac{||w||^2}{2} + C\sum_{i=1}^{n}\xi_i$

$$subject\ to \begin{cases} y_i(w^T\Phi(x_i) + b) + \xi_i \geq 1 \\ \xi_i \geq 0 \end{cases} \quad \forall \mathrm{i} \in \mathrm{N}$$

步驟一：轉換為Lagrange最佳化問題

Minimize $L_p(w, b, \xi, \alpha, \mu)$

$$= \frac{||w||^2}{2} + C\sum_{i=1}^{n}\xi_i - \sum_{i=1}^{n}\alpha_i[\, y_i(w^T\Phi(x_i) + b) + \xi_i - 1] - \sum_{i=1}^{n}\mu_i\xi_i$$

$$subject\ to\ \alpha_i, \mu_i \geq 0 \quad \forall \mathrm{i} \in \mathrm{N}$$

步驟二：列出KKT Condition(Karush-Kuhn-Tucker Condition)

$$\frac{\partial L_p}{\partial w} = w - \sum_{i=1}^{n}\alpha_i\, y_i\Phi(x_i) = 0 \tag{3.8}$$

$$\frac{\partial L_p}{\partial b} = -\sum_{i=1}^{n}\alpha_i y_i = 0 \tag{3.9}$$

$$\frac{\partial L_p}{\partial \xi_i} = C - \alpha_i - \mu_i = 0 \tag{3.10}$$

$$y_i(w^T\Phi(x_\mathrm{i}) + b) \geq 1 - \xi_i \tag{3.11}$$



DOI:10.6814/THE.NCCU.MB.013.2018.F06

$$\xi_i \ , \ \alpha_i \ , \ \mu_i \ \geq 0 \tag{3.12}$$

$$\alpha_i[\ y_i(w^T\Phi(x_\mathrm{i}) + b) - 1 + \xi_i] = 0 \tag{3.13}$$

$$\mu_i\xi_i = 0 \tag{3.14}$$

步驟三：求解

解(3.8)式可得$w = \sum_{i=1}^{n} \alpha_i y_i \Phi(x_i)$ 。

解(3.13)式，加上(3.10)式與(3.12)式的條件，可得

$b = \ y_j - w^T\Phi(x_j)\ , \forall 0 < \alpha_j < C$，

取平均值得$b = \frac{1}{N_s} \sum_{0<\alpha_j<C} y_j - w^T\Phi(x_j)$ 。

因此Hyperplane可改寫為：

$$w^T\Phi(x_i) + b$$

$$= \sum_{i=1}^{n} \alpha_i y_i \Phi(x_i)^T\Phi(x_j) + \frac{1}{N_s} \sum_{0<\alpha_j<C} \left( y_j - \sum_{i=1}^{n} \alpha_i y_i\ \Phi(x_i)^T\Phi(x_j) \right)$$

$$= \sum_{i=1}^{n} \alpha_i y_i K(x_i, x_j) + \frac{1}{N_s} \sum_{0<\alpha_j<C} \left( y_j - \sum_{i=1}^{n} \alpha_i y_i\ K(x_i, x_j) \right) = 0$$

其中$K(x_i, x_j) = \Phi(x_i)^T\Phi(x_j)$，稱之為核函數(Kernel Function)。

核函數(Kernel Function)有很多種，亦可自己定義符合需求的函數，基本且常使用的核函數整理於表3-1中。Hsu et al.(2003)提到，在選擇核函數時，RBF Kernel (Radial Basis Function Kernel) 為合理的選擇，原因有三：第一，它可以處理非線性資料，而Linear Kernel是其特例之一，Sigmoid Kernel在特定參數下，亦和RBF



DOI:10.6814/THE.NCCU.MB.013.2018.F06

Kernel有相似的表現；第二，相較於Polynomial Kernel，RBF Kernel參數較少，模型相對來說就沒有那麼複雜；最後，RBF Kernel在數值計算上較簡單，其Kernel Function $K(x_i, x_j)$ 介於 (0,1] 之間，而 Polynomial Kernel 的 Kernel Function $K(x_i, x_j)$ 介於 $(0, \infty)$，至於Sigmoid Kernel則在某些參數下是無法內積計算的。因此本研究採用RBF Kernel作為支援向量機之Kernel Function。

表3-1：基本的核函數

| Kernel Function | 數學式 |
|---|---|
| Radial Basis Function (簡稱RBF) | $K(x_i, x_j) = \exp\left(-\gamma \lVert x_i - x_j \rVert^2\right)$, $\gamma > 0$ |
| Polynomial | $K(x_i, x_j) = (\gamma {x_i}^T x_j + r)^d$, $\gamma > 0$ |
| Linear | $K(x_i, x_j) = {x_i}^T x_j$ |
| Sigmoid | $K(x_i, x_j) = \tanh(\gamma {x_i}^T x_j + r)$ |

國立政治大學 National Chengchi University



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 第四章 實證資料

## 第一節 資料期間

本研究的資料期間自 2002/04/30 至 2017/12/31，除了台幣實質有效匯率為月資料外，其餘資料皆為日資料，資料來自台灣經濟新報資料庫、Datastream 資料庫、政府資料開放平台以及國際清算銀行(Bank for International Settlement)。

本研究以台灣加權股價指數作為實證對象，針對每日收盤價判定是否發生金融壓力事件。總共 3525 個交易日中，發生 6 個月金融壓力事件共 456 次，發生 3 個月金融壓力事件共 750 次。模型輸入的金融變數於下一章節詳細介紹。

## 第二節 變數選取

自文獻中可發現，FCI 之變數選取沒有放諸四海的標準，因此本文以較多文獻選用且資料取得較容易之變數，並參考張天惠(2012)之作法，自股票市場、匯率市場與利率市場三個管道中選擇金融變數。

除台灣市場之金融變數外，考量到近幾年資金全球化的影響，及台灣貿易總額占 GDP 比重超過 95%、外資持有股票占總市值超過 30%等經濟型態，亦納入重要的國際金融市場變數，所有變數列於表 4-1，放入模型前皆經過標準化處理。

國內股票市場變數部分，選擇能代表整體台灣股市的台灣加權股價指數；台灣加權股價指數 CMAX，也就是台灣加權股價指數的現值除以過去一年的最大值，可以表現出目前股價與過去的相對價格水準；以及考量到模型是針對未來做出預測，而放入近月台指期貨。國外股票市場部分，選擇對於全球金融市場都有較大影響力的 S&P 500 指數；以及俗稱「恐慌指數」的 CBOE VIX 指數，代表



DOI:10.6814/THE.NCCU.MB.013.2018.F06

投資人對未來 30 天市場波動性預期。

國內利率市場變數部分，選擇長期的台灣 10 年期公債殖利率、短期的 90 天期商業本票利率以及金融業隔夜拆款利率。國外利率市場部分，選擇可以反映資金流動性和投資者風險偏好變化的 Treasury-Eurodollar Spread，以及代表美國整體金融政策的 Fed Fund Rate。

國內匯率市場匯率採用可以衡量一國對外價格競爭能力的台幣實質有效匯率，台幣實質有效匯率為月資料，因此以線性內插方式填補為日資料後使用；國外部分考量到美元目前仍是最重要的國際貨幣，許多原物料等貿易結算多以美元計價，選擇新台幣兌美元匯率。

表4-1：金融變數列表

| 類別 | 國內/外 | 變數名稱 | 變數發布頻率 |
|---|---|---|---|
| 股票市場 | 國內 | 台灣加權股價指數 | 日 |
| | | 台灣加權股價指數 CMAX | 日 |
| | | 近月台指期貨 | 日 |
| | 國外 | S&P 500 指數 | 日 |
| | | CBOE VIX 指數 | 日 |
| 利率市場 | 國內 | 台灣 10 年期公債殖利率 | 日 |
| | | 90 天期商業本票利率 | 日 |
| | | 金融業隔夜拆款利率 | 日 |
| | 國外 | Treasury-Eurodollar Spread | 日 |
| | | Fed Fund Rate | 日 |
| 匯率市場 | 國內 | 台幣實質有效匯率 | 月 |
| | 國外 | 新台幣兌美元匯率 | 日 |



DOI:10.6814/THE.NCCU.MB.013.2018.F06

## 第三節 模型建構

本研究使用的三種模型皆需要先以歷史資料建置模型，再將預測日資料輸入建置好的模型進行預測。本研究選擇將所有預測日之前的歷史資料皆放入來建構模型，故模型建置方式為：以過去 1～n-1 日資料作為訓練資料(Training Data)建置模型，以第 n 日資料作為測試資料(Testing Data)，預測第 n 日是否發生金融壓力事件。這意味著每一日模型都會因為新加入的資料而有些微調整，能更貼近市場現況，也能更精準的預測。

本研究資料自 2002/04/30 開始蒐集，為避免資料不足導致模型失準，第一個模型以頭兩年的資料來建置，也就是以 2002/04/30 至 2004/04/29 資料作為訓練資料(Training Data)，以 2004/04/30 作為測試資料(Testing Data)預測是否發生金融壓力事件，以此類推，第二個模型以 2002/04/30 至 2004/04/30 資料作為訓練資料，以 2004/05/01 作為測試資料預測是否發生金融壓力事件；第三個模型以 2002/04/30 至 2004/05/01 資料作為訓練資料，以 2004/05/02 作為測試資料預測是否發生金融壓力事件，如圖 4-1 所示，以此模式持續進行預測直到 2017/06/30。

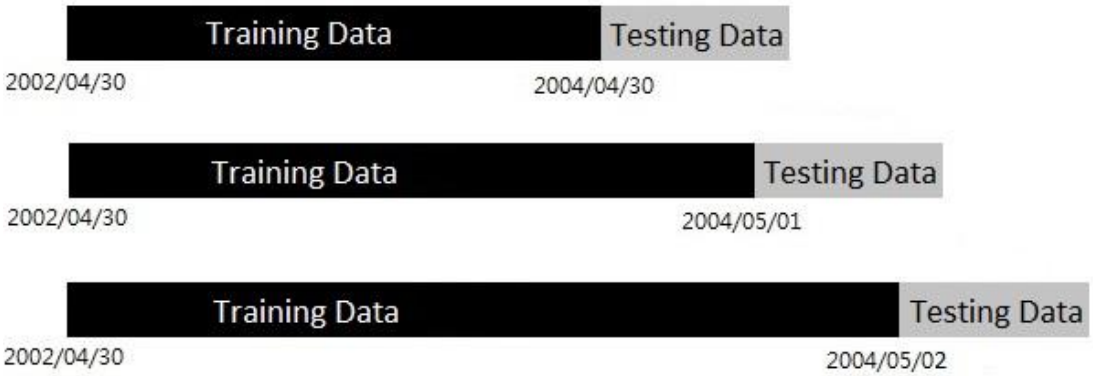


圖 4-1：模型建構方式圖示

## 第四節 超參數設定



DOI:10.6814/THE.NCCU.MB.013.2018.F06

## 一、類神經網路超參數設定

類神經網路有若干超參數(Hyperparameter)需設定，包含隱藏層層數、每一層神經元個數、學習率(Learning Rate)及迭代次數，以下詳述本研究之超參數設定。

隱藏層層數與神經元個數部分，根據葉怡成（2003）之建議，設定隱藏層一層，隱藏層之神經元個數則依照輸入層與輸出層之神經元個數總和來決定，可設置為輸入層與輸出層之神經元個數總和的 1/2、1、2 倍，以本研究輸入層 12 個神經元，輸出層 1 個神經元來說，分別為 7、13、26 個神經元。學習率之設定通常介於 0 與 1 之間，故本研究選擇 0.01、0.1、0.5、1 四種不同的學習率來比較。迭代次數部分，選擇 5000 次。

本研究將以上各參數共 12 種組合，分別以全部資料建置模型，並計算出在不同迭代次數之下的誤差平方和(Sum of the Squared Errors, 以下簡稱 SSE)，也就是第三章第二節提到的總誤差(Loss Function) $L = \sum(Y_k - z_k^H)^2$，繪製出誤差收斂圖，再從 12 個組合中選擇 SSE 小且穩定的組合作為本研究之超參數設定。

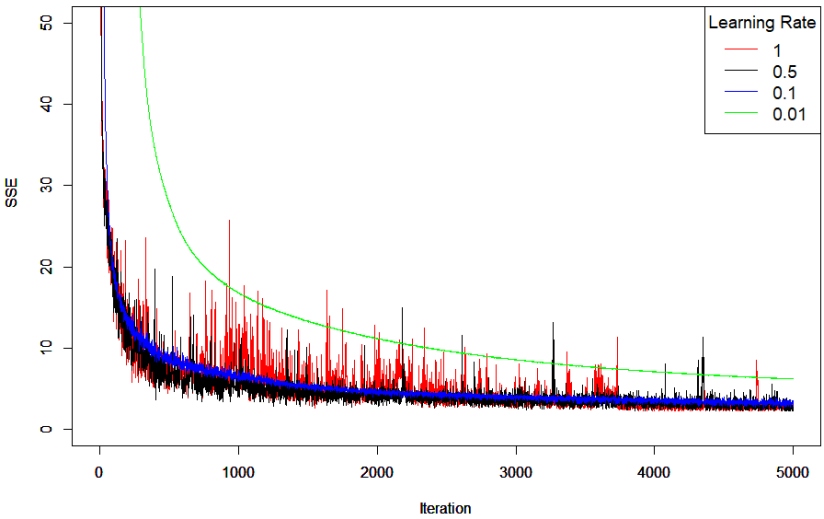


圖 4-2：6 個月金融壓力事件之誤差收斂圖-7 個神經元

6 個月金融壓力事件預警模型部分，先將神經元個數固定為 7，並比較不同學習率下之 SSE，如圖 4-2，可以發現學習率為 0.01 時 SSE 較大，而學習率為 1、0.5 與 0.1 時 SSE 相近，但三者之中以學習率為 0.1 時 SSE 最穩定。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

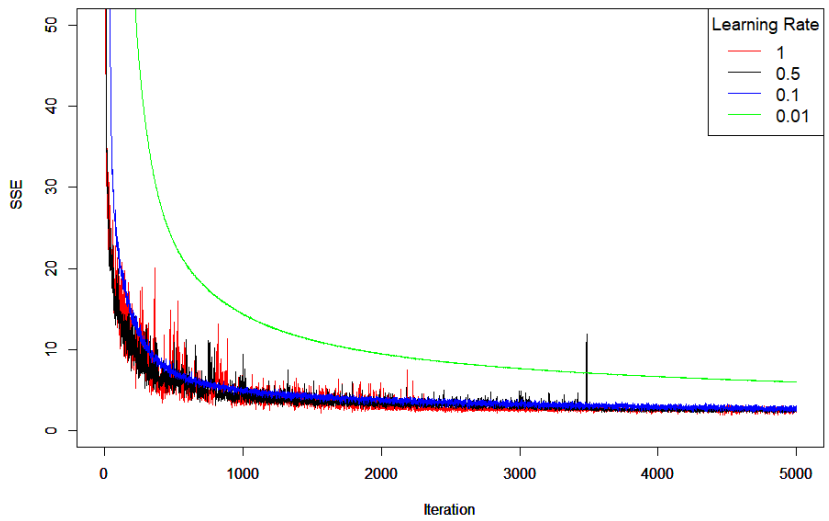


圖 4-3：6 個月金融壓力事件之誤差收斂圖-13 個神經元

再來將神經元個數固定為 13，同樣比較不同學習率下之 SSE，如圖 4-3，發現學習率為 0.01 時 SSE 較大，而學習率為 1、0.5 與 0.1 時 SSE 相近，但三者之中以學習率為 0.1 時 SSE 最穩定。

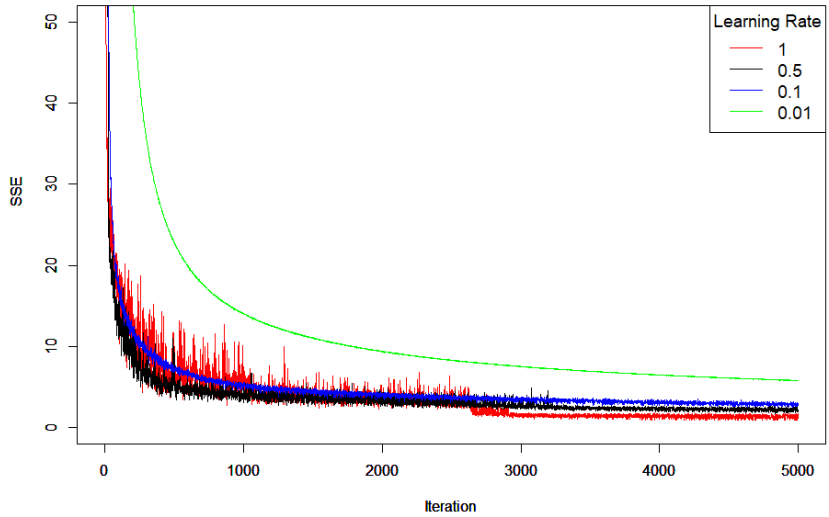


圖 4-4：6 個月金融壓力事件之誤差收斂圖-26 個神經元

再將神經元個數固定為 26，同樣比較不同學習率下之 SSE，如圖 4-4，發現迭代次數超過大約 3000 次後 SSE 皆較穩定，SSE 排序由大到小為學習率 0.01、0.1、0.5、1 時。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

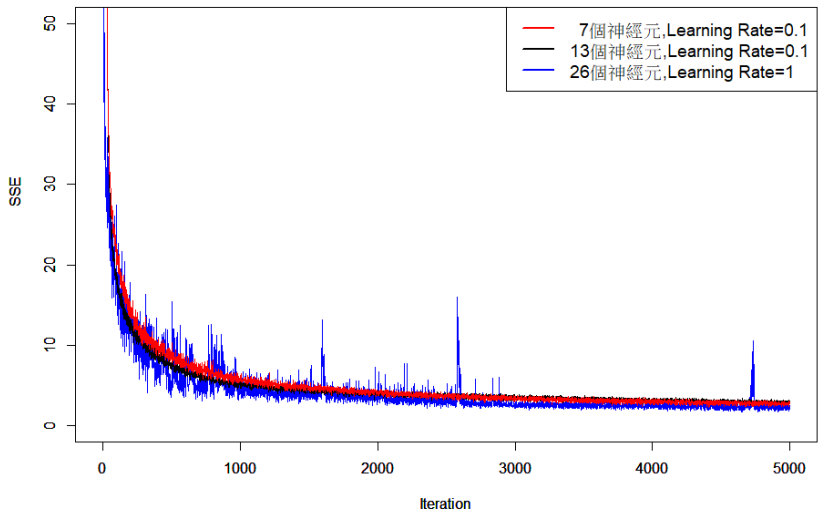


圖 4-5：6 個月金融壓力事件之誤差收斂圖

最後將三種不同神經元數下表現最好的超參數做比較，如圖 4-5，分別為 7 個神經元搭配學習率 0.1、13 個神經元搭配學習率 0.1、26 個神經元搭配學習率 1，發現 26 個神經元時 SSE 較不穩定，7 個神經元與 13 個神經元則 SSE 相近，以訓練速度來說 7 個神經元表現較好。故最終將 6 個月金融壓力事件預警模型的超參數設定為：1 層隱藏層、隱藏層包含 7 個神經元、學習率 0.1、迭代 5000 次。

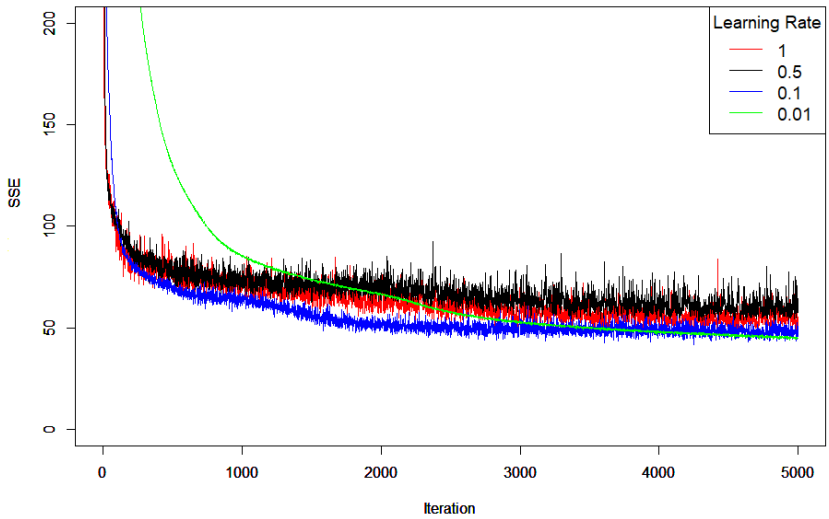


圖 4-6：3 個月金融壓力事件之誤差收斂圖-7 個神經元

3 個月金融壓力事件預警模型部分，一樣先將神經元個數固定為 7，比較不同學習率下之 SSE，如圖 4-6，可以發現當迭代次數夠大時，SSE 皆相近，但四者之中以學習率為 0.01 時 SSE 最穩定。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

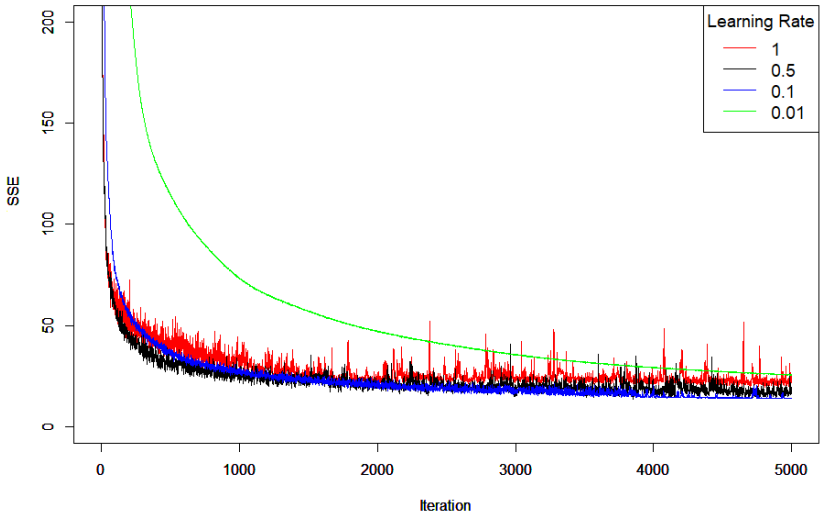


圖 4-7：3 個月金融壓力事件之誤差收斂圖-13 個神經元

再來將神經元個數固定為 13，同樣比較不同學習率下之 SSE，如圖 4-7，發現學習率為 0.01 時 SSE 較大，而學習率為 1、0.5 與 0.1 時 SSE 相近，但三者之中以學習率為 0.1 時 SSE 最穩定。

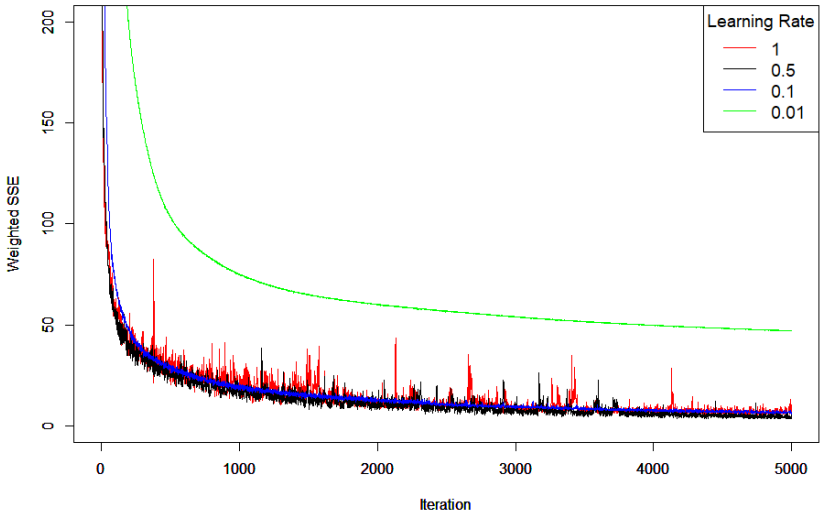


圖 4-8：3 個月金融壓力事件之誤差收斂圖-26 個神經元

再將神經元個數固定為 26，同樣比較不同學習率下之 SSE，如圖 4-8，發現學習率為 0.01 時 SSE 較大，而學習率為 1、0.5 與 0.1 時 SSE 相近，但三者之中以學習率為 0.1 時 SSE 最穩定。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

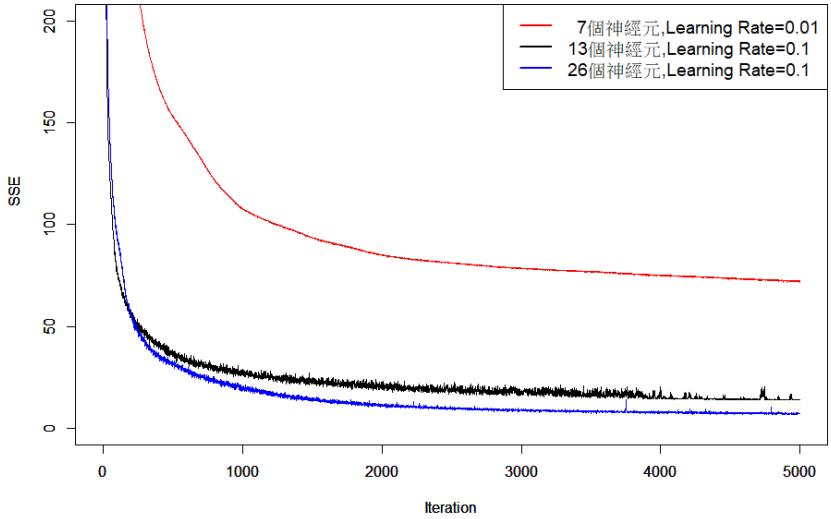


圖 4-9：3 個月金融壓力事件之誤差收斂圖

最後將三種不同神經元個數之下表現最好的超參數做比較，如圖 4-9，分別為 7 個神經元搭配學習率 0.01、13 個神經元搭配學習率 0.1、26 個神經元搭配學習率 0.1，結果發現具有 26 個神經元的模型 SSE 最小。故最終將 3 個月金融壓力事件預警模型的超參數設定為：1 層隱藏層、隱藏層包含 26 個神經元、學習率 0.1、迭代 5000 次。

## 二、支援向量機超參數設定

支援向量機中需要估計的超參數包含控制誤差的常數$C$以及 radial basis function 中的常數$\gamma$。控制誤差的常數$C$越大，越能容忍誤差，但也可能因此造成過度配適的結果。而常數$\gamma$是用來控制輸入資料$x_i$轉換到高維度的特徵空間中時縮放的比例，$\gamma$越大，資料在特徵空間中就越分散。

為找出最適合的超參數，本研究以全部資料建置模型，嘗試$C$介於$2^{-2}$到$2^{8}$，$\gamma$介於$2^{-4}$到$2^{4}$，所有組合共 99 種，選擇 SSE 最小的組合作為支援向量機的超參數設定。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

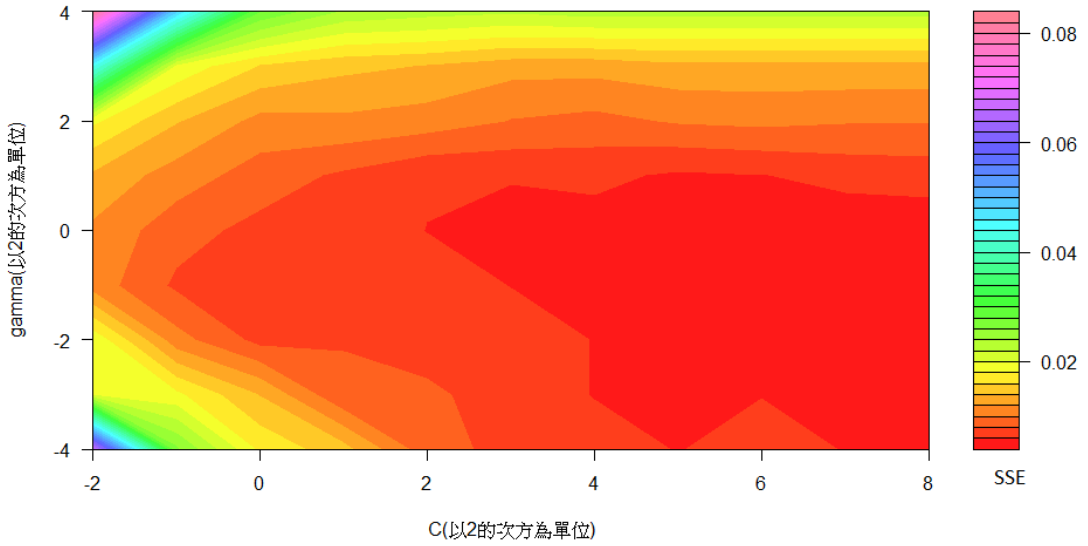


圖 4-10：6 個月金融壓力事件之 SVM 誤差圖

6 個月金融壓力事件預警模型部分，誤差如圖 4-10 所示，最小誤差之超參數為$C = 2^6 = 64$，$\gamma = 2^{-1} = 0.5$，此時 SSE 約等於 0.0045。

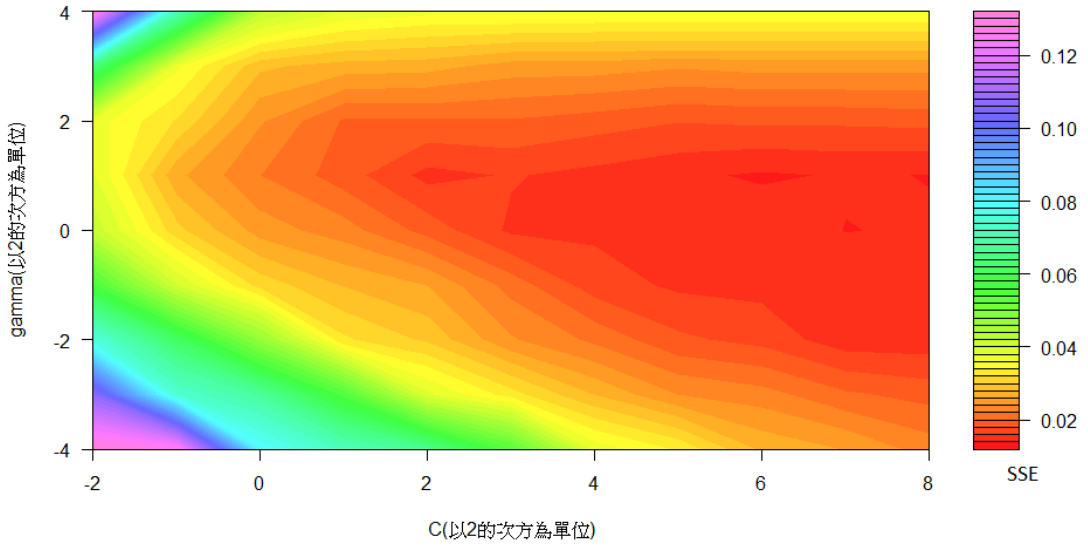


圖 4-11：3 個月金融壓力事件之 SVM 誤差圖

3 個月金融壓力事件預警模型部分，誤差如圖 4-11 所示，最小誤差之超參數為$C = 2^7 = 128$，$\gamma = 2^0 = 1$，此時 SSE 約等於 0.014。



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 第五章 實證結果

本章將以準確率、F1-Score、Type I Error 與 Type II Error 等四種性能指標評斷各模型的實證結果，並比較分析其優劣。如表 5-1 所示，預測結果可分為四類，分別是真實市場發生金融壓力事件且金融壓力事件預警模型也正確預警的 True Positive (以下簡稱 TP)，真實市場發生金融壓力事件但金融壓力事件預警模型卻沒有預警的 False Negative (以下簡稱 FN)，真實市場未發生金融壓力事件但金融壓力事件預警模型卻預警了的 False Positive (以下簡稱 FP)，以及真實市場未發生金融壓力事件而金融壓力事件預警模型也沒有預警的 True Negative (以下簡稱 TN)。以下搭配表 5-1 之名詞，於表 5-2 中詳述各性能指標之定義。

表5-1：預測結果分類表

| | | 金融壓力事件預警模型 | |
|---|---|---|---|
| | | 預警 | 無預警 |
| 真實市場 | 金融壓力事件發生 | True Positive (以下簡稱 TP) | False Negative (以下簡稱 FN) |
| | 金融壓力事件未發生 | False Positive (以下簡稱 FP) | True Negative (以下簡稱 TN) |

表5-2：性能指標定義

| 名稱 | 定義 | 說明 |
|---|---|---|
| 準確率 (Accuracy) | $\frac{(TP+TN)}{(TP+TN+FP+FN)}$ | 為整體模型預測的正確率，亦即所有預測正確的數量除以總數量 |
| Precision | $\frac{TP}{(TP+FP)}$ | 在金融壓力事件預警模型預警的情況下之正確率 |
| Recall | $\frac{TP}{(TP+FN)}$ | 在真實市場發生金融壓力事件的情況下之正確率 |
| F1 Score | $\frac{P \times R}{2(P+R)}$ | 為 Precision 與 Recall 的調和平均 |
| Type I Error | $\frac{FN}{(TP+FN)}$ | 在真實市場發生金融壓力事件的情況下之錯誤率 |
| Type II Error | $\frac{FP}{(FP+TN)}$ | 在真實市場未發生金融壓力事件的情況下之錯誤率 |



DOI:10.6814/THE.NCCU.MB.013.2018.F06

## 第一節 羅吉斯迴歸模型實證結果

首先是 6 個月金融壓力事件預警模型，以逐步迴歸分析分析後，使用台灣加權股價指數、台灣加權股價指數 CMAX、S&P 500 指數、CBOE VIX 指數、台灣 10 年期公債殖利率、90 天期商業本票利率、金融業隔夜拆款利率、Treasury-Eurodollar Spread、Fed Fund Rate、台幣實質有效匯率等十項變數建立羅吉斯迴歸模型，實證結果列於表 5-3。

表5-3：羅吉斯迴歸模型實證結果-6個月金融壓力事件預警模型

| 性能指標 / 門檻值 | 準確率 | F1 Score | Type I Error | Type II Error |
|---|---|---|---|---|
| 5 % | 92.13 % | 73.26 % | 6.86 % | 8.00 % |
| 10 % | 94.18 % | 78.11 % | 10.29 % | 5.23 % |
| 15 % | 95.27 % | 81.11 % | 12.29 % | 3.74 % |
| 20 % | 95.57 % | 81.84 % | 13.71 % | 3.21 % |
| 25 % | 95.47 % | 80.95 % | 16.86 % | 2.92 % |
| 30 % | 95.47 % | 80.51 % | 19.14 % | 2.62 % |
| 35 % | 95.54 % | 80.06 % | 22.57 % | 2.09 % |
| 40 % | 95.50 % | 79.27 % | 25.71 % | 1.72 % |
| 45 % | 95.54 % | 78.87 % | 28.00 % | 1.38 % |
| 50 % | 95.34 % | 77.29 % | 31.43 % | 1.16 % |
| 55 % | 94.71 % | 73.33 % | 37.14 % | 1.12 % |
| 60 % | 94.12 % | 69.31 % | 42.57 % | 1.08 % |
| 65 % | 93.72 % | 66.55 % | 46.00 % | 1.08 % |
| 70 % | 93.62 % | 65.60 % | 47.43 % | 1.01 % |
| 75 % | 93.72 % | 65.33 % | 48.86 % | 0.71 % |
| 80 % | 93.79 % | 65.06 % | 50.00 % | 0.49 % |
| 85 % | 93.52 % | 62.31 % | 53.71 % | 0.30 % |
| 90 % | 93.29 % | 59.96 % | 56.57 % | 0.19 % |
| 95 % | 92.83 % | 55.80 % | 60.86 % | 0.15 % |

門檻值意指，當估計出來的數值大於等於門檻值時，模型「會預警」，小於門檻值時模型則「不會預警」，因此可以發現當門檻值越低，Type I Error 亦隨之



DOI:10.6814/THE.NCCU.MB.013.2018.F06

降低，反之，Type II Error 則升高。此模型之下，準確率至少有 92.13%，F1 Score 至少有 55.8 %，Type I Error 最高為 60.86%，Type II Error 最高為 8%。

3 個月金融壓力事件預警模型部分，以逐步迴歸分析分析後，使用台灣加權股價指數、台灣加權股價指數 CMAX、S&P 500 指數、CBOE VIX 指數、台灣 10 年期公債殖利率、90 天期商業本票利率、金融業隔夜拆款利率、Treasury-Eurodollar Spread、Fed Fund Rate、台幣實質有效匯率等十項變數建立羅吉斯迴歸模型，實證結果列於表 5-4。

表5-4：羅吉斯迴歸模型實證結果-3個月金融壓力事件預警模型

| 性能指標 / 門檻值 | 準確率 | F1 Score | Type I Error | Type II Error |
|---|---|---|---|---|
| 5 % | 48.88 % | 40.12 % | 5.56 % | 61.18 % |
| 10 % | 59.97 % | 44.58 % | 10.93 % | 46.46 % |
| 15 % | 70.37 % | 48.97 % | 21.33 % | 31.46 % |
| 20 % | 78.80 % | 56.52 % | 23.66 % | 20.66 % |
| 25 % | 81.88 % | 58.50 % | 29.21 % | 15.67 % |
| 30 % | 83.76 % | 58.97 % | 35.30 % | 12.03 % |
| 35 % | 83.92 % | 55.64 % | 44.09 % | 9.89 % |
| 40 % | 83.73 % | 52.06 % | 51.25 % | 8.55 % |
| 45 % | 84.18 % | 50.76 % | 55.02 % | 7.16 % |
| 50 % | 84.60 % | 49.03 % | 59.14 % | 5.74 % |
| 55 % | 84.73 % | 47.54 % | 61.83 % | 4.99 % |
| 60 % | 84.76 % | 46.77 % | 63.08 % | 4.67 % |
| 65 % | 85.02 % | 46.08 % | 64.70 % | 4.00 % |
| 70 % | 85.32 % | 45.54 % | 66.13 % | 3.32 % |
| 75 % | 85.45 % | 44.14 % | 68.28 % | 2.69 % |
| 80 % | 85.58 % | 42.64 % | 70.43 % | 2.06 % |
| 85 % | 85.41 % | 39.57 % | 73.66 % | 1.54 % |
| 90 % | 85.09 % | 35.26 % | 77.60 % | 1.07 % |
| 95 % | 84.64 % | 27.57 % | 83.87 % | 0.24 % |

與判定期間為 6 個月相比，此處 Type II Error 最高為 61.18%，代表模型較難



DOI:10.6814/THE.NCCU.MB.013.2018.F06

區分要預警或不要預警，導致預測出來的機率中間值較多，Type II Error 也因此隨門檻值的降低提升較多。

## 第二節 類神經網路模型實證結果

6 個月金融壓力事件預警模型部分，結果列於表 5-5。此模型下 Type II Error 最高為 1.98%，代表在真實市場未發生金融壓力事件時，模型可以非常良好的透過自變數分析出不要預警；同樣的，Type I Error 在門檻值最高為 18%，較羅吉斯迴歸模型的 60.86%低非常多，代表在真實市場發生金融壓力事件時，此模型較容易分析出要預警。

表5-5：類神經網路模型實證結果-6個月金融壓力事件預警模型

| 性能指標 / 門檻值 | 準確率 | F1 Score | Type I Error | Type II Error |
|---|---|---|---|---|
| 5 % | 97.26 % | 88.52 % | 8.57 % | 1.98 % |
| 10 % | 97.82 % | 90.60 % | 9.14 % | 1.27 % |
| 15 % | 97.85 % | 90.65 % | 10.00 % | 1.12 % |
| 20 % | 97.88 % | 90.72 % | 10.57 % | 1.01 % |
| 25 % | 97.95 % | 90.96 % | 10.86 % | 0.90 % |
| 30 % | 98.02 % | 91.23 % | 10.86 % | 0.82 % |
| 35 % | 98.02 % | 91.23 % | 10.86 % | 0.82 % |
| 40 % | 98.08 % | 91.47 % | 11.14 % | 0.71 % |
| 45 % | 98.15 % | 91.74 % | 11.14 % | 0.64 % |
| 50 % | 98.15 % | 91.72 % | 11.43 % | 0.60 % |
| 55 % | 98.12 % | 91.53 % | 12.00 % | 0.56 % |
| 60 % | 98.12 % | 91.53 % | 12.00 % | 0.56 % |
| 65 % | 98.08 % | 91.34 % | 12.57 % | 0.52 % |
| 70 % | 98.08 % | 91.32 % | 12.86 % | 0.49 % |
| 75 % | 98.05 % | 91.13 % | 13.43 % | 0.45 % |
| 80 % | 98.02 % | 90.96 % | 13.71 % | 0.45 % |
| 85 % | 98.05 % | 91.07 % | 14.00 % | 0.37 % |
| 90 % | 97.95 % | 90.52 % | 15.43 % | 0.30 % |
| 95 % | 97.72 % | 89.27 % | 18.00 % | 0.22 % |



DOI:10.6814/THE.NCCU.MB.013.2018.F06

3 個月金融壓力事件預警模型部分，結果列於表 5-6。此模型下 Type I Error 在最高為 20.25%，較羅吉斯迴歸模型的 83.87%低非常多，代表在真實市場發生金融壓力事件時，此模型較容易分析出要預警。

表5-6：類神經網路模型實證結果-3個月金融壓力事件預警模型

| 性能指標 / 門檻值 | 準確率 | F1 Score | Type I Error | Type II Error |
|---|---|---|---|---|
| 5 % | 93.94 % | 85.06 % | 5.91 % | 6.09 % |
| 10 % | 94.94 % | 87.15 % | 6.81 % | 4.67 % |
| 15 % | 95.49 % | 88.38 % | 7.17 % | 3.92 % |
| 20 % | 95.95 % | 89.30 % | 7.89 % | 3.21 % |
| 25 % | 96.11 % | 89.63 % | 8.42 % | 2.89 % |
| 30 % | 96.11 % | 89.59 % | 8.78 % | 2.81 % |
| 35 % | 96.11 % | 89.54 % | 9.32 % | 2.69 % |
| 40 % | 96.14 % | 89.54 % | 10.04 % | 2.49 % |
| 45 % | 96.27 % | 89.87 % | 10.04 % | 2.33 % |
| 50 % | 96.50 % | 90.42 % | 10.22 % | 2.02 % |
| 55 % | 96.69 % | 90.79 % | 10.57 % | 1.70 % |
| 60 % | 96.69 % | 90.73 % | 11.29 % | 1.54 % |
| 65 % | 96.73 % | 90.77 % | 11.65 % | 1.42 % |
| 70 % | 96.69 % | 90.62 % | 12.37 % | 1.31 % |
| 75 % | 96.82 % | 90.96 % | 12.37 % | 1.15 % |
| 80 % | 96.73 % | 90.62 % | 13.26 % | 1.07 % |
| 85 % | 96.63 % | 90.22 % | 14.70 % | 0.87 % |
| 90 % | 96.53 % | 89.83 % | 15.95 % | 0.71 % |
| 95 % | 95.95 % | 87.75 % | 20.25 % | 0.47 % |



DOI:10.6814/THE.NCCU.MB.013.2018.F06

## 第三節 支援向量機實證結果

表5-7：支援向量機實證結果

| 性能指標 / 判定期間 | 準確率 | F1 Score | Type I Error | Type II Error |
|---|---|---|---|---|
| 6 個月 | 99.24 % | 96.67 % | 4.57 % | 0.26 % |
| 3 個月 | 97.37 % | 92.70 % | 7.89 % | 1.46 % |

支援向量機實證結果列於表 5-7。從以上三節可以發現，不論哪一種模型，在預測能力方面，相較於判定未來 3 個月內是否有發生股價低於判定日的 90%，判定未來 6 個月內是否有發生股價低於判定日的 80%的結果都較好。可能是因為股價要低於 80%的發生機率較低，發生時金融變數變動的幅度也比較明顯，讓模型能夠比較清楚的判斷。從實證資料也可以發現總共 3525 個交易日中， 6 個月金融壓力事件共發生 456 次，確實低於發生 3 個月金融壓力事件的 750 次。

## 第四節 各模型結果比較

本章將羅吉斯迴歸模型、類神經網路模型與支援向量機三種模型的準確率與 F1 Score 數值繪製成圖，來比較三種模型。6 個月金融壓力事件預警模型如圖 5-1 所示，3 個月金融壓力事件預警模型如圖 5-2 所示，

不論判定期間長短及下跌幅度，從比較結果中都可以發現，在準確率與 F1 Score 上，皆是支援向量機表現最好，類神經網路模型次之，羅吉斯迴歸模型則明顯較差。支援向量機與類神經網路模型的預測能力相差不多，以準確率來說，6 個月金融壓力事件預警模型：類神經網路最高為 98.15%，支援向量機為 99.24%；3 個月金融壓力事件預警模型：類神經網路最高為 96.73%，支援向量機為 97.37%。差異都低於 1.1%，若再進一步改善模型的超參數設置，應該可以增進兩個模型



DOI:10.6814/THE.NCCU.MB.013.2018.F06

的預測能力，也有可能讓類神經網路的結果優於支援向量機。

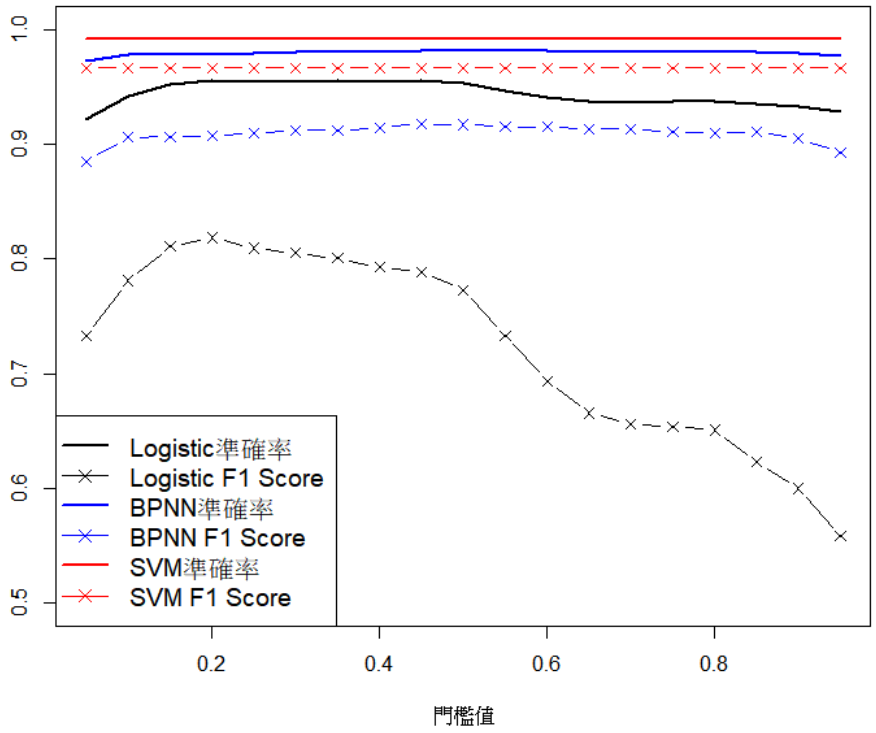


圖 5-1：各模型比較-6 個月金融壓力事件預警模型

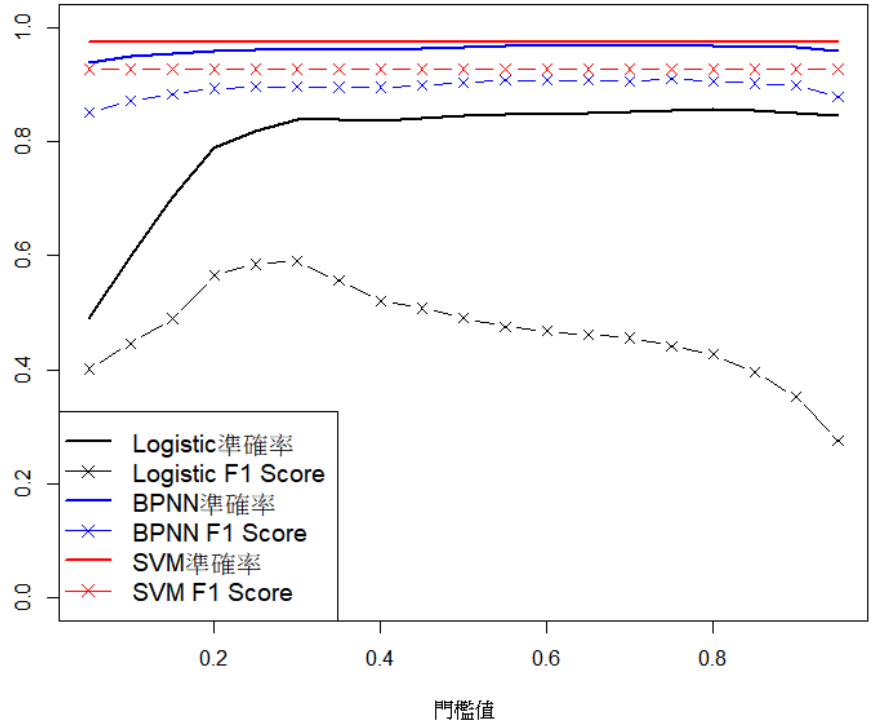


圖 5-2：各模型比較-3 個月金融壓力事件預警模型



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 第六章 結論

本研究以 2002 年至 2017 年之台灣加權股價指數實證，並分別使用羅吉斯迴歸模型、類神經網路模型與支援向量機，建立金融壓力事件預警模型。實證結果顯示，不論判定期間長短及下跌幅度，都是支援向量機預測能力最好，其次為類神經網路模型，羅吉斯迴歸模型則較弱，推測原因為台灣加權股價指數與各自變數間非單純線性關係。

雖然支援向量機預測能力較類神經網路模型好，但結果相差不多，以兩個模型之最高準確率來比較，不論是在預測未來 6 個月內是否有發生股價低於判定日的 80%，或是 3 個月內是否有發生股價低於判定日的 90%，準確率都相差不超過 1.1%。

綜觀過去以類神經網路模型與支援向量機預測金融市場之文獻，實證結果有的支援向量機預測能力較好，有的類神經網路模型預測能力較好，本研究認為是因為市場不同，因此資料特性也不同，另一原因則為超參數設置不同。故未來可以朝向超參數設置部分進一步研究，包括類神經網路模型中的活化函數(Activation Function)、隱藏層層數與神經元個數、學習率，以及支援向量機中的核函數(Kernel Function)、核函數中的參數、控制誤差的參數 C 等。而在這些超參數的選擇方法上除了以本研究使用的格子點方式逐一計算其誤差，並選擇誤差最小之超參數外，亦可以嘗試使用粒子群演算法，來加速整體計算過程。

由於本研究定義之金融壓力事件為判斷未來數個月內是否發生股價大跌，故在實務使用方面與輸入資料相對應的輸出結果需要等到判定期間結束才有結果，也就是要 3 或 6 個月後才能知道判定日是否發生金融壓力事件，進而導致模型出現延遲預警的情況。為改善此一情形，未來可以嘗試針對較短期間進行預測，例如 3 天至一週，或是更進一步以本文之研究方法來計算 VaR 風險值衡量(Value at



DOI:10.6814/THE.NCCU.MB.013.2018.F06

Risk)，讓投資人能夠對金融市場短期之風險有衡量的標準，增加此研究方法於金融市場之實用性。

另一未來研究方向可以朝向深度學習方面發展，其中時間遞歸神經網路(Recurrent Neural Network)因可以記憶過去資料的特性而常用於分析時間序列資料，卷積神經網路(Convolutional Neural Network)則有需考量之參數較少的優勢，都是可以進一步研究的方向。

國立政治大學
National Chengchi University



DOI:10.6814/THE.NCCU.MB.013.2018.F06

# 參考文獻

## 中文文獻


[1] 王翎聿（2015），應用倒傳遞類神經網路與支援向量機預測加權股價指數，國防大學管理學院財務管理學系碩士班碩士論文。

[2] 呂奇傑、李天行、高人龍、黃敏菁（2009），支援向量機與支援向量迴歸於財務時間序列預測之應用，數據分析，第4卷第2期，35-56。

[3] 張天惠（2012），我國金融情勢指數與總體經濟預測，〈中央銀行季刊〉，第34卷第2期，11-42。

[4] 黃華山與邱一薰（2005）類神經網路預測台灣50 股價指數之研究，資訊、科技與社會學報，第5卷第2期，19-42。

[5] 葉怡成（2003），類神經網路模式應用與實作，臺北市：儒林。


## 英文文獻


[1] Cortes, C. and Vapnik, V. (1995). Support-Vector Networks. *Machine Learning*, 20, 273-297.

[2] Gauthier, C., Graham, C., and Liu, Y. (2004). Financial Conditions Indexes for Canada. *Bank of Canada Working Paper* 2004, 22.

[3] Goodhart, C. and Hofmann, B. (2001). Asset Prices, Financial Conditions, and the Transmission of Monetary Policy. *Paper prepared for the conference on Asset Prices, Exchange rates, and Monetary Policy, Stanford University, March* 2-3.

[4] Hatzius, J., Hooper, P., Mishkin, F. S., Schoenholtz, K. L., and Watson, M. W. (2010). Financial Conditions Index: A Fresh Look after the Financial Crisis. *NBER Working Paper* 16150.

[5] Hsieh, L. F., Hsieh, S. C., and Tai, P. H. (2011). Enhanced Stock Price Variation Prediction via DOE and BPNN-based Optimization. *Expert Systems with Applications* 38, 14178-14184.


國立政治大學 National Chengchi University



DOI:10.6814/THE.NCCU.MB.013.2018.F06

[6] Hsu, C. W., Chang, C. C., and Lin, C. J. (2003). A Practical Guide to Support Vector Classification. Available at http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf.

[7] Huang, W., Nakamoria, Y., and Wang, S. Y. (2004). Forecasting Stock Market Movement Direction with Support Vector Machine. *Computers & Operations Research* 32, 2513-2522.

[8] Kara, Y., Boyacioglu, M. A., and Baykan, Ö. K. (2011). Predicting Direction of Stock Price Index Movement using Artificial Neural Networks and Support Vector Machines: The Sample of the Istanbul Stock Exchange. *Expert Systems with Applications* 38, 5311-5319.

[9] Kim, K. J. (2003). Financial Time Series Forecasting Using Support Vector Machines. *Neurocomputing* 55, 307-319.

[10] Rumelhart, D. E., Hinton, G. E., and Williams, R. J. (1986). Learning Representations by Back-propagating Errors. *Nature*, 323, 533-536.

[11] Silvers, D. and Slavkin, H. (2009). The Legacy of Deregulation and the Financial Crisis: Linkages Between Deregulation in Labor Markets, Housing Finance Markets, and the Broader Financial Markets. *Journal of Business & Technology Law* 4, 2, 301.

[12] Skaarup, M., Duschek-Hansen, C., and Nielsen, S. (2010). A Financial Conditions Index for Denmark. *Working Paper no* 23/2010, *The Danish Ministry of Finance*.

[13] Svozil, D., KvasniEka, V., and Pospichal, J. (1997). Introduction to Multi-layer Feed-Forward Neural Networks. *Chemometrics and Intelligent Laboratory Systems* 39, 43-62.

國立政治大學 National Chengchi University



DOI:10.6814/THE.NCCU.MB.013.2018.F06