

國立政治大學
企業管理研究所(MBA 學位學程)

碩士學位論文



指導教授：余清祥 博士

研究生：郭育丞 撰

中華民國 110 年 7 月

摘要

本文探討聯邦公開市場委員會（Federal Open Market Committee，簡稱 FOMC）會議紀要（minutes）的文字風格，尋找不同的利率調整結果（升息、降息、利率不變）時的 FOMC 會議紀要的用詞有哪些差異。本文研究資料為 1993~2020 年的 FOMC 會議紀要，以及同時期的目標聯邦基金利率（targeted federal funds rate）。透過探索式資料分析（exploratory data analysis）、文字分析（或稱文字探勘）的技術，比較 FOMC 會議紀要在升息、降息、利率不變時的會議紀要風格。分析發現，目標聯邦基金利率的調整並非隨機（亦即具有自相關性），經常出現連續幾期的升息、降息；文字使用，在不同利率調整時有不少差異，其中升息、降息的報告大多強調美國四大族裔的失業率。另外，由於分析會議紀要的原始文檔為高維度的文檔-詞頻矩陣（document-term matrix），考量多達 4102 個變數，除了具有稀疏矩陣（sparse matrix）的特質外，變數過多也會影響資料分析的效率。因此本文使用倍數指標篩選器、線性降維、非線性降維等方法，透過縮減特徵空間維度以提高執行效率，研究發現倍率指標的為度縮減效果最佳，配合羅吉斯迴歸得出之三分類準確率最高。

關鍵詞：文字分析、聯準會會議紀要、寫作風格、探索式資料分析、維度縮減

Abstract

In this study, our goal is to explore the writing style of FOMC (Federal Open Market Committee) minutes. In particular, we want to know if the style of minutes shows significant differences when the FOMC decided to raise, lower, or hold interest rates. We applied exploratory data analysis and text mining techniques to the FOMC 1993~2020. We found that the adjustments of targeted federal funds rates are not randomly distributed and they show signs of correlation. For example, among the 39 times of raising interest rate, there was one 17 consecutive intertest increase. Also, the minutes tend to emphasize on the unemployment of four major ethnicities when FOMC decided to raise or lower interest rates. On the other hand, there are 4102 variables involved in exploring the writing study of FOMC minutes. This means that the document-term matrix is a sparse matrix and high dimensionality requires a lot of computation time. Thus, adopted dimensionality reduction techniques: multiplication index, linear reduction and non-linear reduction methods. We found that the multiplication index has the best performance and, together with logistic regression, it has the highest accuracy in classifying the writing style of FOMC minutes in the cases of raising, lowering and holding interest rates.

Keywords: Text mining, FOMC minutes, Writing style, Exploratory data analysis, Dimension reduction

目次

第一章 緒論	1
第一節 研究動機	1
第一節 研究目的	1
第二章 文獻探討	4
第一節 聯邦準備系統與 FOMC 會議記錄	4
第二節 文獻回顧	5
第三節 資料介紹	9
第三章 研究方法	10
第一節 單一字詞、複合字、字詞	10
第二節 資料預處理	11
第三節 時間序列分析	15
第四節 標準化相異詞出現率與景氣指標	16
第五節 主題模型	17
第六節 詞嵌入	21
第七節 文本分類	22
第四章 時間序列分析與探索式資料分析	25
第一節 時間序列分析	25
第二節 探索式資料分析	28
第五章 主題模型與詞嵌入	31
第一節 主題模型	31
第二節 詞嵌入	33
第六章 文本分類	36
第七章 結論與討論	44
第一節 結論	44
第二節 討論	46

參考文獻	47
附錄一、名詞簡稱對照表	50
附錄二、詞嵌入完整圖表	52



表次

表 4-1、升息、降息、利率不變文檔的連續性.....	27
表 4-2、升息、降息文檔的調整幅度	27
表 5-1、LSA 與 LDA 前三個主題的前 20 大字詞.....	32
表 6-1、將 4 倍數指標得到的 8 個特徵分類.....	42
表 6-2、將 3 倍數指標得到的 36 個特徵分類.....	43
附表 1、四類文檔中，各自與 unemploy 的前 20 大近義詞	54



圖次

圖 1-1、分析流程圖	2
圖 3-1、1993/12/21 的 FOMC minutes 接近開頭的部分	12
圖 3-2、2004/12/14 的 FOMC minutes 接近結尾的部分	13
圖 4-1、實質聯邦基金利率變動、景氣循環、Fed 主席任期	25
圖 4-2、升息、降息、利率不變文檔各自的占比	26
圖 4-3、FOMC minutes 的總字數變化	28
圖 4-4、三類文檔的總字數箱形圖	29
圖 4-5、FOMC minutes 的 TTR 變化	29
圖 4-6、FOMC minutes 的 STTR 變化	30
圖 4-7、升息、降息、利率不變文檔的 STTR 箱型圖	30
圖 5-1、升息文檔的 t-SNE 視覺化	34
圖 5-2、美國四大族裔的失業率統計	35
圖 6-1、三種特徵篩選方式的測試資料集準確率	39
圖 6-2、倍數指標的測試資料集準確率	39
圖 6-3、線性降維方法的測試資料集準確率	40
圖 6-4、非線性降維方法的測試資料集準確率	40
圖 6-5、LR 分類器搭配倍數指標的混淆矩陣	41
附圖 1、升息文檔的 t-SNE 視覺化	52
附圖 2、降息文檔的 t-SNE 視覺化	53
附圖 3、利率不變文檔的 t-SNE 視覺化	54

第一章 緒論

第一節 研究動機

聯邦準備系統 (Federal Reserve System, 簡稱 Fed)，為美國的中央銀行體系，負責監管聯邦準備銀行、制訂貨幣政策，包括調控目標聯邦基金利率 (targeted federal funds rate)。聯邦公開市場委員會 (Federal Open Market Committee, 簡稱 FOMC)，主要任務為決策美國的貨幣政策，每年固定舉辦 8 次會議，並於會議結束後 3 週公布聯準會會議紀要 (Federal Open Market Committee minutes，簡稱 FOMC minutes)。

對於投資人而言，每次的 FOMC 會議後宣布的貨幣政策非常重要，影響市場甚大；財經新聞也會重點報導每次 FOMC 後的貨幣政策決策，尤其是目標聯邦基金利率的調整，也就是常聽到的聯準會升息、降息。Rosa (2013)、Huang and Kuan (2021) 都指出，FOMC minutes 對市場上的資產價格深具影響力。本研究透過文字分析 (text mining) 技術，分析 FOMC minutes 在升息、降息、利率不變這三類文檔 (document) 中的寫作風格 (writing style) 差異。然後，使用不同的特徵選取方法縮減維度，搭配不同的統計與機器學習分類器，對 FOMC minutes 做三分類，並觀察哪些字詞最具分類影響力。

第一節 研究目的

本研究以 1993 年 1 月 1 日到 2020 年 10 月 1 日間的目標聯邦基金利率、聯準會會議紀要為研究資料。本研究與先前文獻的差異，是本研究以資料驅動 (data driven) 原則，透過探索式資料分析 (exploratory data analysis) 找出目標聯邦基金利率的特性，找出 FOMC minutes 在升息、降息、利率不變這三類文檔上的寫作風格差異。

首先，先對目標聯邦基金利率，做探索式資料分析、時間序列分析模型 (使用 SARIMA)，了解其有什麼資料特性、了解其是否有自相關性。然後，利用目標聯邦基金利率的歷史資料，預測未來目標聯邦基金利率，並計算準確率。

接著，為了瞭解 FOMC minutes 的資料特性，每篇 FOMC minutes 的標準化相異詞出現率 (standardized type-token ratio) 等數值，並觀察這些數值在升息、降息、利率不變

這三類文檔下表現的差異。然後，用主題分析 (topic modeling)、詞嵌入 (word embedding) 以更深入了解 FOMC minutes 寫作風格在升息、降息、利率不變這三類文檔下的差異。

在對 FOMC minutes 做三分類 (升息、降息、利率不變) 時，為了因應文本資料常見的特徵維度數目遠高於資料點個數的問題，我們利用不同的特徵篩選方法來縮減特徵空間維度，再套用統計與機器學習模型。我們以初階降維方法 (如使用 TF (term frequency) 前 200 大字詞)、N 倍數指標篩選器、線性降維 (linear dimensionality reduction) 方法、非線性降維 (non-linear dimensionality reduction) 方法，來篩選 FOMC minutes 的特徵 (也就是相異的字詞)；觀察在不同特徵篩選方法下，分類準確率的差異，並找尋那些對分類最具影響力的字詞。

圖 1-1 中，展示了分析流程。在處理完 FOMC minutes 的文本後，首先對目標聯邦基金利率變動做時間序列分析，接著，以 FOMC minutes 處理過後的文本加上目標聯邦基金利率變動資料，並依序做探索式分析、主題模型分析、文本分類。

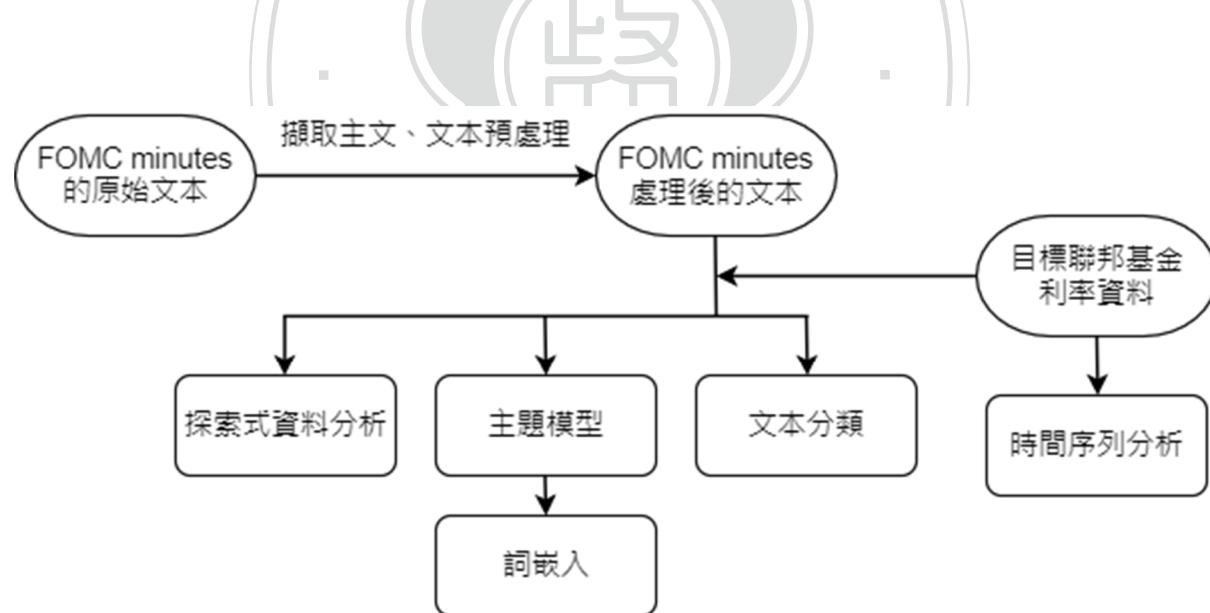


圖 1-1、分析流程圖

綜合以上，本文將於第二章說明文獻回顧、文本資料的取得、資料預處理流程，並於第三章研究方法介紹文字分析等技術，包括英文的文字預處理、主題模型、詞嵌入。第四章，介紹探索式資料分析、時間序列分析的研究結果。第五章，探討主題模型、詞嵌入的研究結果。第六章，以不同的特徵篩選方法來縮減維度，並套用統計與機器學習模型，對 FOMC minutes 進行三分類，然後觀察各情況下的分類準確率。最後，在第七章的結論總結上述所有分析。



第二章 文獻探討

本章節將介紹美國的聯邦準備系統、聯邦公開市場委員會會議的功用，以及聯準會的經濟使命、中央銀行的公共溝通。接著，介紹過去文獻中，如何使用 FOMC 會議紀錄的不同形式、如何分析 FOMC 不同形式的會議紀錄、如何進行文字預處理，以及本研究的特色。

第一節 聯邦準備系統與 FOMC 會議記錄

一、聯邦準備系統

聯邦準備系統 (Federal Reserve System, 簡稱 Fed) 為美國的中央銀行體系，目前 Fed 包含三大組織，有聯邦準備委員會 (Federal Reserve Board of Governors)、聯邦銀行 (Federal Reserve Banks)，以及聯邦公開市場委員會 (Federal Open Market Committee, 簡稱 FOMC)。Fed 的三大組織中，聯邦準備委員會底下有數個委員會，負責管理 Fed 的各業務面向；聯邦銀行分布於各地區，聽命於 Fed 的決策行事；而 FOMC 則主要負責開會決定貨幣政策，其中最重要的為決定目標聯邦基金利率調整，亦即財經新聞常看到的升息、降息或利率不變。而目標聯邦基金利率的變動，會引發一連串連鎖反應，影響其他短期利率、匯率、長期利率、貨幣流通量、融資，並最終影響到重要的經濟變數——失業率、經濟成長率、物價指數。

二、聯準會的三大經濟使命

聯準會身為美國的中央銀行，肩負三大經濟使命：高度可持續就業 (maximum employment)、穩定物價 (stable prices)、適度的長期利率 (moderate long-term interest rates)。¹

1

三、FOMC 會議紀錄的形式

¹ 本節參考了 Fed 網站上的介紹。

FOMC 每年舉辦 8 次例行會議，並於每次會議結束後立即發布 FOMC policy statements，於會議結束的 3 週後發布 FOMC minutes，於會議結束的 5~6 年後發布 FOMC transcripts；而其中，FOMC transcripts 是最詳細的會議紀錄，FOMC minutes 次詳細，FOMC policy statements 則最簡略。

這三種形式的會議紀錄，都有被使用過作為分析資料來源：Lucca and Trebbi (2009)、Doh et al. (2020) 使用了 FOMC policy statements；Boukus and Rosenberg (2006)、Hayo and Neuenkirch (2012)、Rosa (2013)、Jubinski and Tomljanovich (2013)、Stekler and Symington (2016)、Ericsson (2016)、黃于珊 (2017)、Huang and Kuan (2021) 使用了 FOMC minutes；Cannon (2015) 和 Shapiro and Wilson (2019) 使用了 FOMC transcripts。

四、中央銀行的公共溝通

當一國中央銀行進行清晰、透明的公共溝通 (public communication)，該國投資者就更能理解中央銀行的目標和計畫，因此能使貨幣政策更有效果；溝通有助於減少不確定性，增加貨幣政策在金融市場的效力。Blinder et al. (2008) 發現中央銀行的公共溝通，有效地影響了金融市場、增加了貨幣政策決定的可預測性，有助於中央銀行達到總體經濟的目標。Lucca and Trebbi (2009) 發現短期公債名目利率 (nominal Treasury yields) 的改變是根據 FOMC 會議後對利率調整的宣告，但長期公債的利率變動則主要根據 Fed 的政策溝通 (policy communication)。Huang and Kuan (2021) 指出 Fed 透過 FOMC minutes，影響了私部門的觀點與經濟數據。在前 Fed 主席 Ben Bernanke 的任期內 (2006/02~2014/01)，他極力推動聯準會決策的公開透明，盡力向投資人解釋清楚決策的原因和想法，這情況使得對 FOMC minutes 做文字分析在此之後更具價值。

第二節 文獻回顧

根據前人對 FOMC minutes 的研究，FOMC minutes 中蘊含著重要資訊，能影響市場資產價格和成交量，能幫助市場參與者更準確的預測長期公債利率。也有研究以主題模型挖掘 FOMC minutes 蘊含的價值。

一、FOMC minutes 的發布對市場的影響

Lucca and Trebbi (2009) 發現短期公債名目利率 (nominal Treasury yields) 的改變是根據 FOMC 會議後對利率調整的宣告，但長期公債的利率變動則主要根據 Fed 的政策溝通。Rosa (2013) 指出 FOMC minutes 的發布，顯著的影響美國各資產的價格和成交量；而其中，美國資產價格的變動在 2008 年後降低了，推測是由於 FOMC 的公開透明政策導致。Jubinski and Tomljanovich (2013) 調查了 2006 年至 2007 年間 2832 檔股票，觀察他們在 FOMC minutes 發布前與發佈後的差異，發現報酬率 (return) 並無差異，但波動率 (volatility) 在 FOMC minutes 發布後明顯變高。

二、FOMC minutes 的主題模型分析

早在 2006 年，來自紐約聯邦銀行的研究，Boukus and Rosenberg (2006) 就使用 FOMC minutes 做為資料來源，然後以主題模型中的潛在語義分析 (latent semantic analysis，簡稱 LSA)，找到不同的主題 (topic)，並發現這些主題的特性。黃于珊 (2017) 利用 LSA 萃取出升息、降息、利率不變文檔的潛在特徵，再以線性判別分析 (linear discriminant analysis) 進行分類，在三分類 (分為升息、降息、利率不變) 得到最高 77.72% 的準確率。Huang and Kuan (2021) 以 FOMC minutes 做為研究資料，使用 MAP-PLSA (一種基於 PLSA 的進階主題模型方法) 產生 Fed 對其三大經濟使命的情感指標 (sentiment indicators)，發現這個指標能有效預測美國經濟變數；此研究表明，Fed 除了在 FOMC 會議之後宣布貨幣政策來影響市場，也透過其發布的 FOMC minutes，影響了私部門的觀點與經濟。

三、FOMC minutes 的文本預處理

Boukus and Rosenberg (2006) 依序使用去除格式、移除停字詞 (stopwords)、詞幹提取 (stemming) (使用 Porter)、再次移除停字詞。黃于珊 (2017) 依序使用移除空白鍵與大小寫轉換、移除標點符號、移除停字詞、詞幹提取 (使用 Porter)。

Huang and Kuan (2020) 提及，先以程式找出 95 個 FOMC 常用的專門術語，將他們統一；接著，以 N-gram 技巧，挑出 572 個 Minutes 常用的複合字 (compound words)；然後，再以程式從 Wikipedia 中挑出 30 萬個常用的英文複合字，跟 FOMC minutes 比對，因此再從 FOMC minutes 中挑出 4,453 個複合字；然後，找出用連字符號 (hyphen) 或引號 (quotation mark) 產生的字詞共 848 個。整合上述所有複合字，得到了 5,770 個複合字。接著，使用史丹佛大學的自然語言處理 (natural language processing) 程式 coreNLP 標出了個字詞的詞性，然後使用字詞還原 (lemmatization)。最終，他們得到的文檔-詞頻矩陣 (document-term matrix) 為 192×9311 的稀疏矩陣，包含 192 個 FOMC minutes 文本，以及 9,311 個字詞。

四、本研究特色

本研究的特色，在於以下三點：

(一) 過往研究，在說明如何移除 FOMC minutes 的前後比較不重要的文字段落時，往往不夠明確，這可能導致研究結果因此受到影響。比方，Boukus and Rosenberg (2006) 提到他們選取每篇 FOMC minutes 資料的範圍是從 “The Committee then turned to discussion of the economic outlook...” 、“The information reviewed at this meeting suggested...” 或者類似的變形體開始，然後刪去末尾部分的委員會投票資訊；Huang and Kuan (2020) 描述了他們移除了開場敘述以及圖表的標註文字。而本研究詳細的描述不同時期 FOMC minutes 的開頭和結尾的差異，確切定義了選取資料範圍，使將 FOMC minutes 刪去不重要的文字段落後的分析資料，更有可比較性 (comparability)。

(二) 過往研究較少著墨於探索式資料分析，較多著重於驗證式資料分析 (confirmatory data analysis) 和機器學習模型，而本研究做了對目標聯邦基金利率、FOMC minutes 詳細的探索式資料分析。並且，研究目標聯邦基金利率的自身相關性，並且以某期的利率調整結果 (升息、降息、利率不變) 來預測下一期的利率調整結果。另外，發現在每次經濟衰退前，FOMC minutes 的標準化相異詞出現率 (standardized type-token ratio) 都會先出現上升趨勢。

(三) 在文本分類部分，黃于珊 (2017) 使用用了線性降維、LSA 兩種方法來篩選特徵，並搭配線性判別分析 (linear discrimination analysis) 分類器；本研究則使用 N 倍數指標篩選器、線性降維、非線性降維、等方法篩選特徵，羅吉斯迴歸 (logistic regression)、支持向量機 (supporting vector machine)、隨機森林 (random forest)、XGBoost 共 4 種分類器中。



第三節 資料介紹

一、目標聯邦基金利率

取自聯準會網站，本研究使用 1993 年 1 月 1 日到 2020 年 10 月 1 日間的目標聯邦基金利率 (targeted federal funds rate)，其全名為 ”FOMC's target federal funds rate or range, change (basis points) and level”。² 聯準會所設定的目標聯邦基金利率，與實質聯邦基金利率 (effective federal funds rate) 有非常些微的差異，原因是聯準會並不能直接控制實質聯邦基金利率，而是使用數個貨幣政策措施，去影響實質聯邦基金利率到其所設定的目標值。³ 目標聯邦基金利率為 Fed 的貨幣政策中的核心面向之一，每年例行的 8 次 FOMC 會議負責決定如何調整目標聯邦基金利率，選項有升息、降息、利率不變，而升息、降息的幅度的基本單位是一碼 (quarter of a percentage point)，一碼相當於 0.25%。一般而言，目標聯邦基金利率的調整，都是在 FOMC 會議結束後馬上進行；但在極少數特殊情況下也能不藉由 FOMC 會議的臨時調整，比如美東時間 2020 年 3 月 3 日，Fed 在道瓊開盤後，無預警直接降息了 0.5%。

二、聯準會會議紀要

本研究使用 1993 年 1 月 1 日到 2020 年 10 月 1 日間的 FOMC minutes，取自聯準會網站，共 222 個文檔。⁴ FOMC 會議中將決定 Fed 的貨幣政策，會議一結束就會發布 FOMC policy statement，而在會議結束後 3 週後會發布更詳細的 FOMC minutes。相較之下，黃于珊 (2017) 使用了 1993 年至 2017 年 3 月間的 FOMC minutes，共 193 個文檔；Huang and Kuan (2021) 使用了 1993 年至 2016 年的 FOMC minutes，共 192 個文檔。

² 見 <https://www.federalreserve.gov/monetarypolicy/openmarket.htm>

³ 見 <https://fred.stlouisfed.org/series/FEDFUNDS>

⁴ 見 https://www.fedsearch.org/fomc-docs/?advanced_search=true

第三章 研究方法

本研究使用 Python 執行，並主要用到了幾個套件：nltk、pmdarima、gensim、scikit-learn。並且，本研究的研究資料、程式碼、圖片都可以在 GitHub 上可找到。⁵ 第四章時間序列分析與探索式資料分析、第五章的主題模型，是在本機上運行；第五章的詞嵌入和第六章的文本分類，是在 Google Colaboratory 上運行。在開始閱讀本章前，可再次參考圖 1-1 的分析流程圖，快速了解分析流程。

第一節 單一字詞、複合字、字詞

英文文法中的單一字詞 (single word)、複合字 (compound words)、字詞 (word)，這三者將在本文出現多次，因此有必要特別說明其定義。單一字詞指的是 apple、mongo 這種，並不包含 skin-deep 這種複合字；字詞則包含 apple、mongo、skin-deep，亦即包含單一字詞與複合字。

複合字有三種型式：第一種是開放式複合字 (open compounds)，由兩個字詞組成，比方 ice cream；第二種是封閉式複合字 (closed compounds)，由兩個字詞組成一個字詞，比方 doorknob；第三種是連字號複合字 (hyphenated compounds)，由幾個字詞中間以連字號 (hyphen) 連接而成，比方 short-term、father-in-law。為求敘述方便，本文中所指稱的複合字，包含開放式複合字、連字號複合字，但不包含封閉式複合字；Huang and Kuan (2021) 所指稱的複合字，推測應該也是包含開放式複合字、連字號複合字，但不包含封閉式複合字。

另外，還有一個相異字詞 (unique words)，在 “I have an apple, and my brother have a kiwifruit.” 這句話中，共有 10 個字詞，但只有 9 個相異字詞，因為 have 出現了 2 次。

相較於 Huang and Kuan (2021) 考慮到了複合字，本研究只考慮了單一字詞，因此本研究內所稱的字詞，其實都是單一字詞。

⁵ 見 https://github.com/mortonkuo/text_mining_on_FOMC_minutes

第二節 資料預處理

在資料預處理 (data preprocessing) 階段，必須要刪減掉每篇 FOMC minutes 的前後較不重要的段落，也就是定義和擷取 FOMC minutes 的本文部分。接著下一步，是文本預處理 (text preprocessing)，包含一系列的步驟使文本成為可以轉化為矩陣形式。最後，是文本向量化 (text vectorization)，將經過文本預處理的文本，轉化成矩陣形式。⁶

一、定義和擷取 FOMC minutes 的主文

每篇 FOMC minutes 的前後都會有較不重要的文字段落，我們需要將其刪除並保留其他文字段落，在此將這些保留的文字段落稱為主文。因為每段時間的 FOMC minutes 都具不同特性，無法將定義和擷取主文的流程自動化，因此本研究是手動觀察每一篇 FOMC minutes，歸納出各時期 FOMC minutes 結構的不同處，然後定義和擷取主文。在進行此步驟時，Boukus and Rosenberg (2006) 只提到他們選擇的資料範圍從 “The Committee then turned to discussion of the economic outlook...”、“The information reviewed at this meeting suggested...” 或類似的變形開始，然後刪除了 FOMC minutes 末尾的委員會投票行為紀錄。黃于珊 (2017) 基於 Boukus and Rosenberg (2006) 的上述手法，更詳細的解構了 FOMC minutes 的架構，但讀者還是難以了解其確切刪除了那些段落。Huang and Kuan (2020) 只提到他們刪除了開頭的陳述、圖表的註解。如果不確切定義研究使用的刪除頭尾後的 FOMC minutes，那很可能造成每篇研究其實是使用了不一樣的分析資料，導致研究結果的差異可能是來自於分析資料有所差異。因此，本研究將確切描述如何刪除每篇 FOMC minutes 的頭尾。

本研究刪除了每篇 FOMC minutes 參與者名單 (participants)、投票行為紀錄 (voting actions)、投票者的評論 (voters' reviews)，剩下的段落，都是關於未來經濟展望、經濟成長、失業率、聯邦基金利率、國債利率等等的經濟和財金相關的敘述。因為投票者的評論並不是每一篇 FOMC minutes 都有，因此本研究刪除了它們。

⁶ 本節參考了 Joshi (2016) 和 Sarkar (2019)。

接著，說明我們如何刪除 FOMC minutes 的頭、尾較不重要的文字段落，如何定義主文。

(一) 主文的開始點

The Committee then turned to a discussion of the economic and financial outlook and the implementation of monetary policy over the intermeeting period ahead. A summary of the economic and financial information available at the time of the meeting and of the Committee's discussion is provided below, followed by the domestic policy directive that was approved by the Committee and issued to the Federal Reserve Bank of New York.

The information reviewed at this meeting suggested that economic activity had recorded a strong advance in recent months. Consumer spending had picked up, and business purchases of durable equipment had remained on a marked upward trend. Residential construction was rising rapidly, and nonresidential construction had turned up from depressed levels. Industrial production had been boosted by developments in the motor vehicle industry, and employment had continued to post solid gains. Most indexes of prices pointed to little change in inflation trends despite the recent acceleration of economic activity.

Total nonfarm payroll employment rose appreciably further in November. Another substantial increase in jobs was recorded in the services industries, notably in health and business services. Construction employment was up significantly further after registering modest gains on balance over the first three quarters of 1993. In manufacturing, there were back-to-back increases in jobs in October and November following seven consecutive monthly declines, and both overtime hours and the average workweek remained at a high level. Most of the November expansion in factory jobs occurred in the motor vehicle and capital goods industries. The civilian unemployment rate fell

圖 3-1、1993/12/21 的 FOMC minutes 接近開頭的部分

我們選用圖 3-1 作為範例，圖 3-1 中有三個段落，我們從上到下依序稱呼其為段落 A、段落 B、段落 C。這三個段落有以下的特性：

1. 段落 A：通常以 “The Committee then turned to a discussion...” 為開頭，但自 2009/01/28 起，此段落消失了。
2. 段落 B：通常以 “The information reviewed at this meeting suggested that...” or “The information reviewed by the Committee at this meeting suggested that...” 為開頭，但自 2009/01/28 起，此段落上方有了一個小標題 “Staff Review of the Economic and Financial Situation” 或 “Staff Review of the Economic Situation”。
3. 段落 C：此段落有幾種變形，但這些變形都攜帶著同樣的資訊，包括著非農業就業人數 (nonfarm payroll employment)、失業 (unemployment)、就業 (employment)。

有時候段落 A、段落 B、段落 C 變形的太多，導致我們無法輕易認出他們；但是，我們還是能用上述段落 A、段落 B、段落 C 的特徵，來迅速辨識這三個段落。而本研究刪除了段落 A、段落 A 之前的段落（通常是參與者名單），保留了段落 B 和段落 B 之後的段落。也就是說，段落 B 是本研究定義的主文的開始段落。接著，我們來看主文的結束點。

（二）主文的結束點

At the conclusion of the discussion, the Committee voted to authorize and direct the Federal Reserve Bank of New York, until it was instructed otherwise, to execute transactions in the System Account in accordance with the following domestic policy directive.

“The Federal Open Market Committee seeks monetary and financial conditions that will foster price stability and promote sustainable growth in output. To further its long-run objectives, the Committee in the immediate future seeks conditions in reserve markets consistent with increasing the federal funds rate to an average of around 2¼ percent.”

The vote encompassed approval of the paragraph below for inclusion in the statement to be released shortly after the meeting:

“The Committee perceives the upside and downside risks to the attainment of both sustainable growth and price stability for the next few quarters to be roughly equal. With underlying inflation expected to be relatively low, the Committee believes that policy accommodation can be removed at a pace that is likely to be measured. Nonetheless, the Committee will respond to changes in economic prospects as needed to fulfill its obligation to maintain price stability.”

Votes for this action: Messrs. Greenspan, Geithner, Bernanke, Ms. Bies, Messrs. Ferguson,
圖 3-2、2004/12/14 的 FOMC minutes 接近結尾的部分

圖 3-2 中劃上底線的段落 “The vote encompassed approval of the paragraph...”，以及其後的段落，都被我們刪除了，不包含在主文內。也就是在圖 3-2 中劃上底線段落的上一個段落，是主文的結束段落。然而，在 2000/12/01 以前的 FOMC mintes 中沒有 “The vote encompassed approval of the paragraph...” 段落，因此我們改為刪掉 “Votes for this action:...“ 段落以及其後的段落。

二、文本預處理

在定義和擷取 FOMC minutes 的主文之後，下一步，即為將文本處理好至可以轉化為矩陣形式，本研究的文本預處理 (text preprocessing) 步驟為：

- (一) 還原縮寫 (expanding contractions)
- (二) 將大寫字母轉為小寫 (lower case transformation)
- (三) 移除非英文的符號 (removing non-alphabet characters)
- (四) 斷詞 (tokenization)。
- (五) 移除停字詞 (removing stopwords)。
- (六) 詞幹提取 (Stemming)。

在斷詞、移除停字詞、詞幹提取這三個步驟，都使用 Python 的 nltk 套件。在詞幹提取步驟，選用了 Porter stemmer，而 Boukus and Rosenberg (2006) 也是使用 Porter stemmer。相較於 Huang and Kuan (2021) 考慮到了複合字，本研究只考慮了單一字詞，因此本研究內所稱的字詞，其實都是單一字詞。

三、文本向量化

文本向量化 (text vectorization)，有時候又被稱為文本表達式 (text representation) 的特徵工程 (feature engineering)，即為將文本轉換成矩陣，常見的模型有 one-hot encoding、詞袋模型 (bag-of-words model, BOW) (以下簡稱 BOW)、TF-IDF (term frequency-inverse document frequency)、word2vec。本研究使用了 TF-IDF，參數設定為 ngram_range=(1,1)，亦即不考慮複合字。本研究的原始文檔-詞頻矩陣 (document-term matrix, 簡稱 DTM) 為 222×4102 的矩陣 (222 個 FOMC minutes 文檔、4102 個字詞)，並未設定最低詞頻門檻來刪除特徵 (參數設定：CountVectorizer(binary=False, min_df=0.0, max_df=1.0, ngram_range=(1, 1)))；相較之下，Huang and Kuan (2021) 考慮到了複合字，其 DTM 為 192×9311 的矩陣，雖然本研究的研究資料比 Huang and Kuan (2021) 還多並且包含他們所有的文檔，但他們卻有更高的特徵空間維度。

本研究使用的原始 DTM 為 222×4102 的矩陣，由於特徵數量 4102 遠大於資料點數量 222，這會有維度災難 (curse of dimensionality) 的問題。該原始 DTM 為稀疏矩陣 (sparse matrix)，這會導致過擬合 (overfitting)，使模型的泛化 (generalization) 能力較差。因此，本研究在不同的分析部分，都使用了維度縮減 (dimensionality reduction) 方法。

為了縮減維度，在主題模型分析時，本研究使用的參數設定為：

```
CountVectorizer(analyzer='word', max_features=5000, max_df=1.0, min_df=0.0)。
```

為了縮減維度，本研究在詞嵌入時，使用的參數設定為：

```
feature_size = 100 # Word vector dimensionality  
window_context = 30 # Context window size  
min_word_count = 5 # Minimum word count  
sample = 1e-3 # Downsample setting for frequent words  
w2v_model = word2vec.Word2Vec(tokenized_corpus, size=feature_size,  
                                window=window_context, min_count=min_word_count,  
                                sample=sample, iter=50)
```

為了縮減維度，本研究在文本分類時，使用了 TF 前 200 大字詞、TF-IDF 前 200 大字詞、N 倍數指標篩選器、線性降維方法 PCA、非線性降維方法 kernel PCA，來使特徵維度降至 222 個以下。注意到在文本分類的全部字詞時，使用的 4102 個字詞，並沒有設定最低詞頻門檻來刪除特徵。

第三節 時間序列分析

一、季節差分整合移動平均自迴歸模型

季節差分整合移動平均自迴歸模型 (seasonal autoregressive integrated moving average model, SARIMA) (以下簡稱 SARIMA)，可拆分為 S、AR、I、MA 四項，分別代表季節性、自迴歸、差分、移動平均。此模型的結果形式為 $SARIMA(p,d,q)(P,D,Q)m$ ，其中 p 代表 AR (自迴歸)、d 代表 I (差分)、q 代表 MA (移動平均)，而 $(P,D,Q)m$ 則是季節性的項。本研究使用 Python 的 pmdarima 中的 auto_arima 函數來實踐 SARIMA。

二、Ljung-Box test

Ljung-Box test 用於檢測某個時間率列模型，跟資料實際上的時間序列模式，是否相符，以確保某個時間序列模型配適程度。在使用 Python 的 pmdarima 中的 auto_arima 函數跑出 SARIMA 的最佳結果後，我們接著使用 Ljung-Box test 來確認模型是否對資料配適。

三、天真預測法

對於分類問題，天真預測法 (naïve method) 使用 t 期的類別，來作為對於 $t + 1$ 期類別的預測值。亦即，若 t 期 FOMC minutes 文檔為降息，則天真預測法將預測 $t + 1$ 期的 FOMC minutes 文檔為降息。

第四節 標準化相異詞出現率與景氣指標

一、標準化相異詞出現率

相異詞出現率 (type-token ratio, 簡稱 TTR) 計算了一個文檔中的相異字詞 (unique words, 亦即 types) 的數目，除以全部字詞的數目 (total words, 亦即 tokens)，得到的 TTR 代表該文檔中字詞的豐富度。⁷ 但是，當文檔字詞數增大時，相異字詞數通常不會一起線性增大，這導致字詞數相異的文檔間，會失去可比較性。因此，標準化相異詞出現率 (standardized type-token ratio, 簡稱 STTR)，以拔靴法 (bootstrap) 對文檔重複取後放回抽樣 1000 次，每次抽取 500 個字詞，然後計算 1000 個 TTR 結果的平均，得出的平均數更具有可比較性。在第四章研究結果的圖 4-5、圖 4-6，可以看到 FOMC minutes 的 TTR 和 STTR 的結果比較，了解到為何 STTR 比 TTR 更具可比較性。

二、景氣指標

⁷ 見 <https://carla.umn.edu/learnerlanguage/spn/comp/activity4.html>

景氣循環 (business cycle) 衡量國家總體經濟活動的波動，一個完整景氣循環週期，包含一個擴張期 (expansion)，一個收縮期 (contraction)。景氣谷底 (trough) 至景氣高峰 (peak) 之期間，為擴張期；景氣高峰至景氣谷底之期間，為收縮期。

景氣指標，用一些足以反映經濟活動、景氣變化的重要總體經濟變數，綜合計算出經濟景氣的狀況。景氣指標有三種，包含領先指標 (leading indicator)、同時指標 (coincident indicator)、落後指標 (lagging indicator)，而美國經濟諮商局 (The Conference Board) 負責界定各重要總體經濟變數，屬於上述三種景氣指標之一，或是都不屬於這三種。⁸

(一) 領先指標：具領先景氣波動性質，可用來預測未來景氣之變動，包含住宅投資 (residential investment)、存貨投資 (inventory investment)、平均勞動生產力 (average labor productivity)、貨幣供應量 (money supply)、股價 (stock prices)。

(二) 同時指標：代表當前景氣狀況，可以衡量當時景氣之波動，包含工業生產 (industrial production)、消費 (consumption)、企業固定投資 (business fixed investment)、就業 (employment)。

(三) 落後指標：用以驗證過去之景氣波動，包含通貨膨脹 (inflation)、名目利率 (nominal interest rates)。

若能找到一個變數，其為領先指標，則就能對經濟景氣具預測力，也有可能對股價具預測力。在計算了每一份 FOMC minutes 文檔的字詞數、相異字詞數、標準化相異詞出現率之後，我們加上景氣循環的資訊，繪製成折線圖；我們將美國景氣循環的衰退期，用淺藍色的陰影區域在折線圖上繪出，以觀察某變數是否為領先指標。⁹

第五節 主題模型

主題模型 (topic modeling) 的目的，是以數學和統計技術發現與語料庫 (corpus) 中隱藏的語義結構。主題模型從 DTM 中提取特徵，並使用矩陣分解和奇異值分解 (singular

⁸ 本點景氣循環參考了 Abel et al. (2013)。

⁹ 見 <https://www.nber.org/research/data/us-business-cycle-expansions-and-contractions>

value decomposition，簡稱 SVD) 等數學架構，來生成彼此不同的詞組 (cluster of terms)，每一組詞組代表著一個主題 (topic)，每個主題都代表個一個語義聚類 (semantic clusters)。簡而言之，每份文檔都會包含多個主題，而每個主題也會與多個字詞有關。主題模型有 4 種常見的實現方法：LSA (latent semantic analysis)、NMF (non-negative matrix factorization)、PLSA (probabilistic latent semantic analysis)、LDA (latent Dirichlet allocation)，從左到右大略為從簡單到複雜，而本研究使用了 LSA 與 LDA。LSA、NMF、PLSA、LDA 都屬於低秩 (low-rank) 估計方法，都可看作矩陣分解 (matrix factorization) 方法。雖然本研究只使用 LSA 與 LDA，但為求連貫性，本節將 4 種方法都介紹。¹⁰

一、Latent semantic analysis (LSA)

LSA，有時又被稱為 LSI (latent semantic indexing) 或 SVD，為一種降維 (dimensionality reduction) 並產生正交 (orthogonal) 主題的方法。其優劣如下：

(一) 優點：LSA 具正交特徵向量的幾何優勢，對於不包含在原始資料矩陣中的新文檔進行融入 (fold in) 時，也就是對樣本外 (out-of-sample) 的文檔進行融入時，LSA 表現優異。這種樣本外的嵌入對其他矩陣分解方法而言，是非常難處理的。

(二) 缺點：得出的主題中的字詞們重疊率很高，因此各主題並不真正的正交。大多數的 NMF 類別方法 (包含 NMF、PLSA、LDA)，並不使用正交基底向量 (orthogonal basis vectors)，也因此更能真實呈現資料中的聚類 (cluster) 結構。LSA 因為各主題會都含正向、負向字詞，這將導致 LSA 在語義 (semantic) 上不可解釋。

二、Non-negative matrix factorization (NMF)

NMF 的矩陣裡只允許非負元素。

¹⁰ 本節參考了 Aggarwal (2018) 與 Sarkar (2019)。

(一) 優點：由於允許基底向量裡出現非正交，各個主題能更正確的被 NMF 捕捉到。這很合理，因為各個語義聚類（或稱主題）本身就會有相互的關聯。語義上的解釋度非常高，比 LSA 更具解釋性。有些研究指出，一般而言 NMF 表現會比 LDA 和 LSA。

(二) 缺點：在計算原始資料矩陣之外的新文檔，NMF 表現比 LSA 更差。

三、Probabilistic latent semantic analysis (PLSA)

PLSA 單純只是 NMF 的變形，可以將 PLSA 當成機率模型 (probabilistic model) 而分解模型 (factorization model)，因為 PLSA 不對一個文檔中的主題們的分布做任何事前假設。Huang and Kuan (2021) 使用的 MAP-PLSA，即為基於 PLSA 發展出來的模型。

(一) 優點：此方法簡單直覺、易理解。PLSA 不對主題在文檔中的分布做任何假設，因此在大資料集上有很好的一般性 (generality)。語義上的解釋度非常高，比 LSA 更具解釋性。

(二) 缺點：PLSA 不是完全生成式模型 (fully generative model)，所以很難融入新的文檔。PLSA 的估計程序，隨著資料集大小而線性成長，所以在語料庫大小非常大時，需要花費大量運算資源。

四、Latent Dirichlet allocation (LDA)

LDA 是 PLSA 的一般化，可以將 LDA 看作為 PLSA 的進階版本。LDA 利用狄利克雷分布 (Dirichlet distribution) 提前確定各文檔中主題的組成，然後一次性生成 DTM (document-term matrix) 的某一行的所有元素。因此，狄利克雷分布對每個文檔都引入了先驗結構，在文檔的主題分布上使用了 Dirichlet 先驗分布，以創建一個有效地融入新文檔的完全生成式模型 (fully generative model)。

(一) 優點：語義上的解釋度非常高，比 LSA 更具解釋性。可能比 PLSA 更節省一點運算資源。

(二) 缺點：有人指出，LDA 表現不比 NMF 更好。LDA 是 PLSA 的一般化，所以，LDA 花費的運算資源，也會隨著語料庫增大而線性增大。

NMF 與 PLSA 幾乎是等價的，他們在語義上都可解釋，並可以很好地處理同義詞和多義詞的問題；另一方面，NMF 和 PLSA 都不能融入樣本外的文檔。而本研究使用 LSA 與 LDA。



第六節 詞嵌入

詞嵌入 (word embedding) 技術中，以 word2vec 為代表，並具有兩大特色。第一，word2vec 技術運用單詞的前後文，在沒有人為干預下來學習單詞的語義，運用神經網絡方法來計算每個單詞的近義詞 (synonymous words)。第二，相較於 one-hot encoding 產生的稀疏矩陣，詞嵌入產生密集矩陣 (dense matrix)，有效縮減了維度。許多自然語言處理的方法，都必須依靠良好的單詞特徵表達方式才能運作，而這種單詞表達方式，不只要保留單詞的語義，也要保留下單詞前後文的相關訊息。¹¹

一、詞嵌入

相較於 one-hot encoding 得到的詞向量是二進位制的、稀疏的、維度高的 (維度大小等於文本中的字詞數)，詞嵌入使用密集的詞向量，是維度低的密集向量。與 one-hot encoding 得到的詞向量不同，詞嵌入是從資料中學習得到的。常見的詞嵌入的詞向量維度，有 256、512、1,024；與此相對，BOW 的詞向量維度通常高達 20,000 (對應到包含個 20,000 字詞的文檔) 或更高。所以，詞嵌入是有效縮減維度的方法。常見的詞嵌入實現方法，有 2013 年 Google 提出的 word2vec、2014 年 Stanford 自然語言處理小組提出的 GLoVe，與 2016 年 Facebook 提出的 FastText。

二、word2vec

2003 年，詞嵌入方法 NPLM (neural probabilistic language model) 被提出，NPLM 可將詞語轉換成維度可控制的向量，並且能表示語義訊息。比方，兩個近義詞在 NPLM 的向量空間中的位置也是相近的。但因為 NPLM 的算法運算速度較慢，因此在當時並未帶起風潮。2013 年 Mikolov et al. (2013) 提出了 word2vec 算法，word2vec 是 NPLM 的升級版，word2vec 進行了改進，大幅提升了運算速度與運算精確，也帶起了詞嵌入的風潮。word2vec 算法包含兩個模型：continuous bag-of-words (CBOW) model 和 skip-gram

¹¹ 本節參考了 VanderPlas (2017)、Chollet (2018) 與 Ganegedara (2018)。

model。CBOW 模型是使用當前詞的前 n 個詞和後 n 個詞(即上下文)來預測當前詞。而 skip-gram 則相反，使用當前詞來預測前 n 個詞和後 n 個詞(即上下文)。

三、t-SNE

t-SNE (t-distributed stochastic neighbor embedding) 是一種非線性降維技術，把高維度資料投影到二維空間。降維方法有線性降維、非線性降維兩大類別，常見的線性降維方法有主成分分析 (principal component analysis, 簡稱 PCA) 與因素分析 (factor analysis)，而常見的非線性降維方法有 kernel PCA、t-SNE、Isomap (isometric mapping)、MDS (multidimensional scaling)、LLE (locally linear embedding)。由於我們很難以視覺化方式呈現三維以上的資料，因此 t-SNE 有效的幫助我們解讀資料點之間的相關性。然而，這些將資料點的分布，從高維度降維到二維平面的非線性方法，都不是很穩健 (robust)：在使用 t-SNE 分析 FOMC minutes 時，偶爾會發現，加上某個資料點 X 的二維平面上的資料點分布，和不加入某個資料點 X 的二維平面上的資料點分布，會有滿大的差異。而觀察這些資料點在二維平面上的分佈，涉及聚類 (clustering) 分析，在解釋上需要參考這些資料點聚類的語義，使解釋結果涉及部分主觀成分。

第七節 文本分類

一、特徵選取方法：初階特徵選取方法、N 倍數指標篩選器、線性降維方法、非線性降維方法

(一) 初階特徵選取方法：全部字詞、TF 前 200 大字詞、TF-IDF 前 200 大字詞。全部字詞，是使用全部相異字詞作為特徵。後兩種方法，都是把 222 個 FOMC minutes 文檔合併後，篩選出詞頻出現比率前 200 大的字詞。

(二) N 倍數指標篩選器：必須對兩個成對的文檔使用，假設某字詞 A 在文檔一的詞頻為 x 個，而某字詞 A 在文檔二的詞頻為 p 個。另外，假設全部文檔一的字詞共有 y 個，而文檔二的字詞共有 q 個。則 N 倍數指標篩選器，將計算某字詞 A 是否符合下列兩項條件之一：

$$\frac{x}{y} \geq N \frac{p}{q} \text{ 或 } \frac{p}{q} \geq N \frac{x}{y}$$

若符合上述兩項條件之一，則該字詞 A 通過 N 倍數指標篩選器的篩選。我們使用此方法，分別對「全部升息文檔 / 全部降息文檔」、「全部升息文檔 / 全部降息文檔加上全部利率不變文檔」、「全部降息 / 全部升息文檔加上全部利率不變文檔」這三個成對的文檔，分別以 N 倍數指標篩選器進行上述的特徵篩選流程，找出通過 N 倍數指標篩選器的字詞。在第六章表..... 可看到.....，更清楚了解 N 倍數指標篩選器。

(三) 線性降維方法：使用 PCA。

(四) 非線性降維方法：使用 kernel PCA。

透過上述四大類特徵選取方法，我們可以將特徵空間維度數目，降至資料點數目以下，從而避免維度災難。

二、分類器

本研究選用了 4 種文本分類器，大約從簡單到複雜依序排列，分別是羅吉斯迴歸 (logistic regression，簡稱 LR)、支持向量機 (supporting vector machine, 簡稱 SVM)、隨機森林 (random forest，簡稱 RF)、XGBoost。羅吉斯迴歸是傳統統計方法，支持向量機是統計學習 (statistical learning) 方法，隨機森林為集成學習 (ensemble learning) 的 bagging 算法，而 XGBoost 則為集成學習 (ensemble learning) 的 boosting 算法。¹²

集成學習的核心思想，是建構多個不同的模型，並將這些模型聚合起來，從而提高模型的性能。在集成學習中，建構起來的一系列學習器稱作基學習器 (base learner)。集成學習可分為兩類，一類是併行方法 (parallel method)，以 bagging 為代表；另一種是順序方法 (sequential method)，以 boosting 為代表。

LR 為簡單的傳統統計方法。SVM 比 LR 更複雜，但在分類高度重疊的資料點時表現較差。RF 為 bagging 方法，透過 bootstrap 取樣，然後併行建構並運算所有基學習器，此方法能減少雜訊 (noise) 的影響，且更容易分類高度重疊的資料點。XGBoost 為 boosting 方法，依照順序建構並運算所有基學習器，並且在建構新的基學習器時，會重

¹² 本節參考了孫亮、黃倩（2017）。

點學習以前的基學習器錯誤分類的訓練樣本。相較於 RF 的參數數量少非常容易優化，XGBoost 的參數數量眾多優化不易；因此，如果使用 XGBoost 但沒有花時間調整參數，經常得不到比 RF 更好的結果。



第四章 時間序列分析與探索式資料分析

第一節 時間序列分析

圖 4-1 為實質聯邦基金利率的變動，一個資料點代表一個月，可發現實質聯邦基金利率是經濟衰退的領先指標，而股市也是經濟衰退的領先指標。並且，每一任 Fed 主席任期內，聯邦基金利率的變動況狀差距都很大。從圖 4-2，可發現這是一個不平衡 (unbalanced) 資料集，222 個資料點中，有 153 個資料點 (68.9%) 都是利率不變。

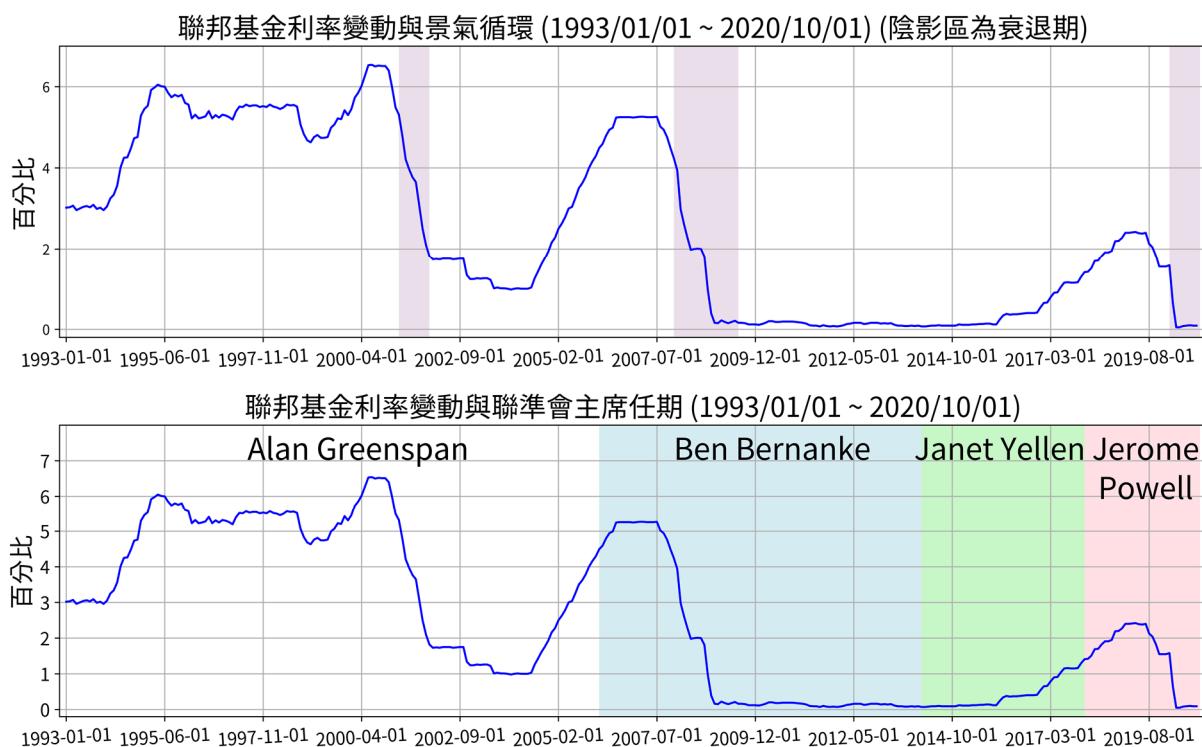


圖 4-1、實質聯邦基金利率變動、景氣循環、Fed 主席任期

FOMC Minutes 中升息、降息、利率不變文檔的占比

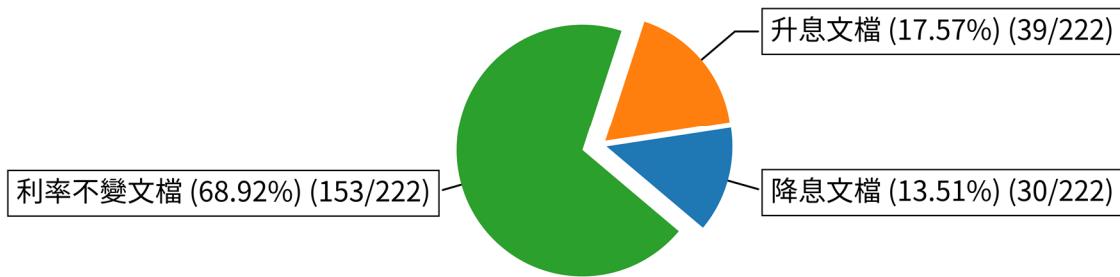


圖 4-2、升息、降息、利率不變文檔各自的占比

表 4-1 展示了目標聯邦基金利率調整結果的連續性，以升息文檔為例，連續次數為 1 代表該點前、後兩個點都不是升息，連續次數為 2 即為連續兩個升息文檔；因此， $1 \times 14 + 2 \times 1 + 3 \times 2 + 17 \times 1 = 39$ 。表 4-1 中，64.1% (25/39) 的升息是連續兩次以上、90.0% (27/30) 的降息是連續兩次以上、92.2% (131/153) 的利率不變是連續兩次以上，可以發現在 Fed 在降息上的決策是比升息更果斷的。從表 4-2，進一步觀察升息、降息的調整幅度，只有 15.4% (6/39) 的升息決策調整幅度為 2 碼 ($2 \times 0.25\%$) 以上，但高達 60.0% (18/30) 的降息決策調整幅度為 2 碼以上，顯示 Fed 對升息決策的保守性、對降息決策的果斷性。

表 4-1、升息、降息、利率不變文檔的連續性

升息文檔 (39 個)		降息文檔 (30 個)		利率不變文檔 (153 個)	
連續次數	次數	連續次數	次數	連續次數	次數
1 (無連續)	14	1 (無連續)	3	1 (無連續)	12
2	1	2	3	2	2
3	2	3	2	3	2
17	1	6	1	4	5
		9	1	6	1
				7	2
				8	2
				9	1
				11	1
				55	1

表 4-2、升息、降息文檔的調整幅度

升息文檔 (39 個)		降息文檔 (30 個)	
利率改變幅度	次數	利率改變幅度	次數
0.25 %	33	0.25 %	12
0.50 %	5	0.50 %	12
0.75 %	1	0.75 %	2
		1.00 %	4

接著，使用 Python 的 pmdarima 套件的 AutoARIMA 函式，訊息準則 (information criterion) 選擇 AIC，得到 SARIMA(1,1,2)(0,0,0)[12]，並以 Ljung-Box test 確認此時間序列模型的配適是否貼切。亦即，我們得到一個 AR(p) = 1、MA(q) = 2 的模型，代表目標聯邦基金利率的變動，具自身相關性。既然目標聯邦基金利率有自相關性，也從上述分析看到升息、降息、利率不變，這三者通常是各自連續出現的，因此很自然地我們想到使用天真預測法來預測，以前一期的利率調整結果 (升息、降息、利率不變) 來預測某

一期的利率調整結果，並得到 74.7% 的預測準確率。因此，使用任何方法來預測目標或實質聯邦基金利率時，所參照的基準線 (baseline) 都應設定為天真預測法產生的數字。

第二節 探索式資料分析

對 FOMC minutes 做探索式資料分析，從圖 4-3 發現從 2008 年金融海嘯之後，FOMC minutes 總字數出現顯著上升了一個水平，而此段上升的時期也是前 Fed 主席 Ben Bernanke 的任期期間。在去掉了圖 4-3 中的離群值 (outlier) 後，繪成圖 4-4，並在此後都將此離群值刪除。圖 4-4 中，三類文檔中的中位數，從大到小排序，是利率不變、降息、升息；但如果資料為多峰，則此箱形圖就較不具參考價值。

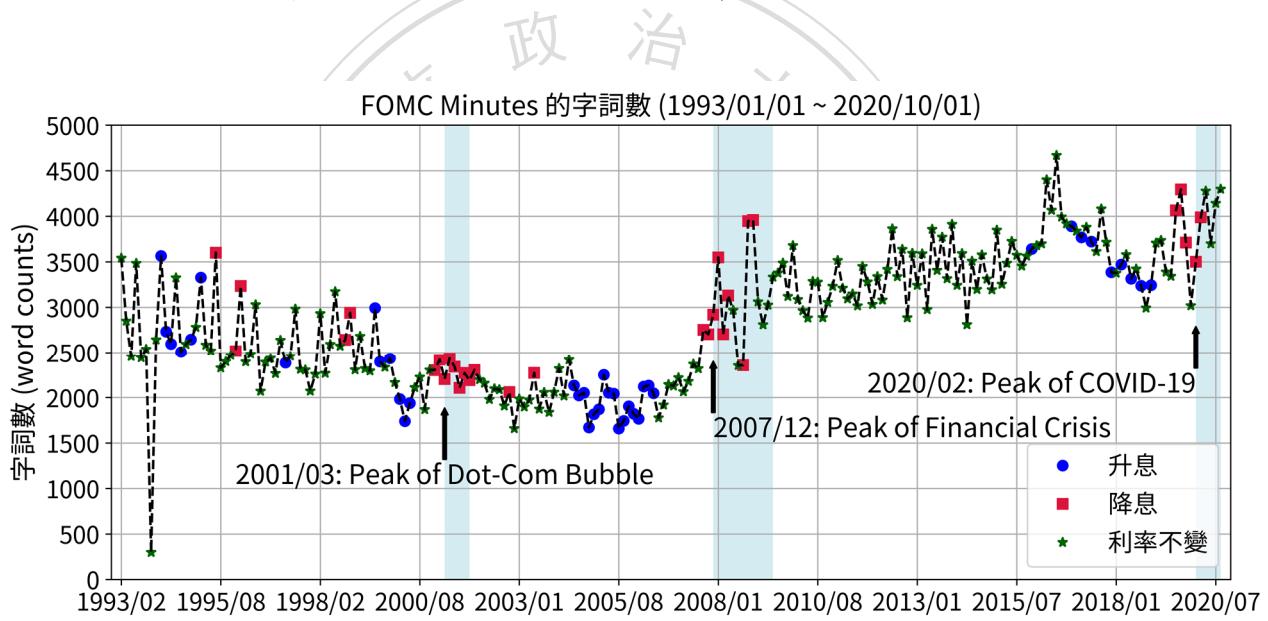


圖 4-3、FOMC minutes 的總字數變化

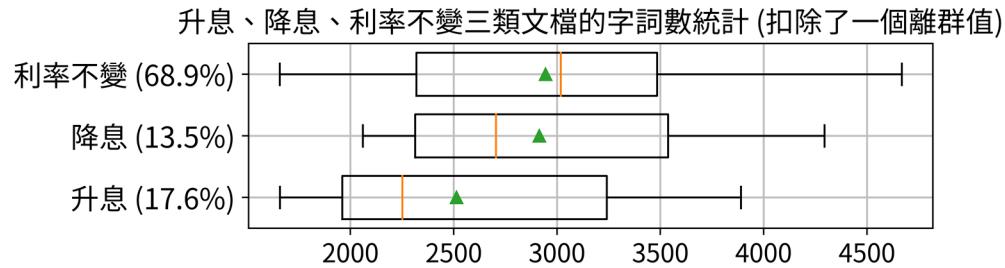


圖 4-4、三類文檔的總字數箱形圖

接著，以 STTR 來分析。在此 STTR 採用拔靴法抽取 1000 次，每次抽取 500 個字詞。圖 4-6 中，可發現 STTR 大約是景氣循環的同時指標，亦即，在經濟衰退期間的 STTR 較高。比較圖 4-5 的 TTR 結果和圖 4-6 的 STTR 結果，可理解到在第三章研究方法提到為何 STTR 比 TTR 更具有可比較性的原因。FOMC minutes 的總字詞數在圖 4-5 的右半部明顯上升了，導致圖 4-5 右半部的 TTR 明顯低於左半部的 TTR；但是，在圖 4-6 中的右半部，STTR 並沒有因為總字詞數的上升而明顯降低。圖 4-7 的箱型圖，展示了降息文檔的 STTR 的中位數和平均數都大於升息文檔的，並且是三類中最高的。

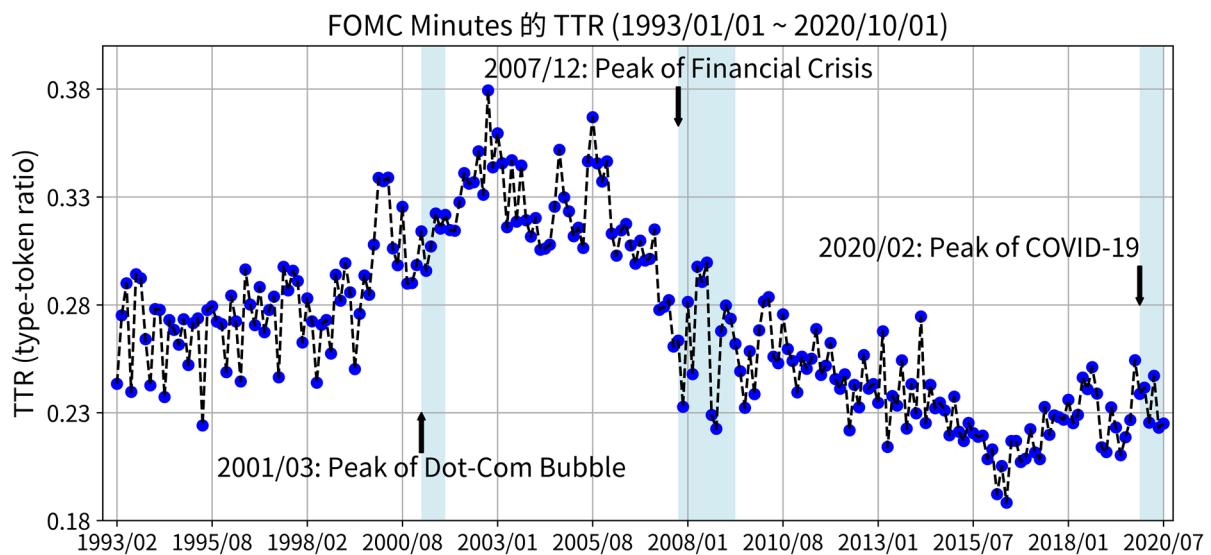


圖 4-5、FOMC minutes 的 TTR 變化

FOMC Minutes 的 STTR (拔靴法抽樣 1000 次，每次抽取 500 個字詞) (1993/01/01 ~ 2020/10/01)

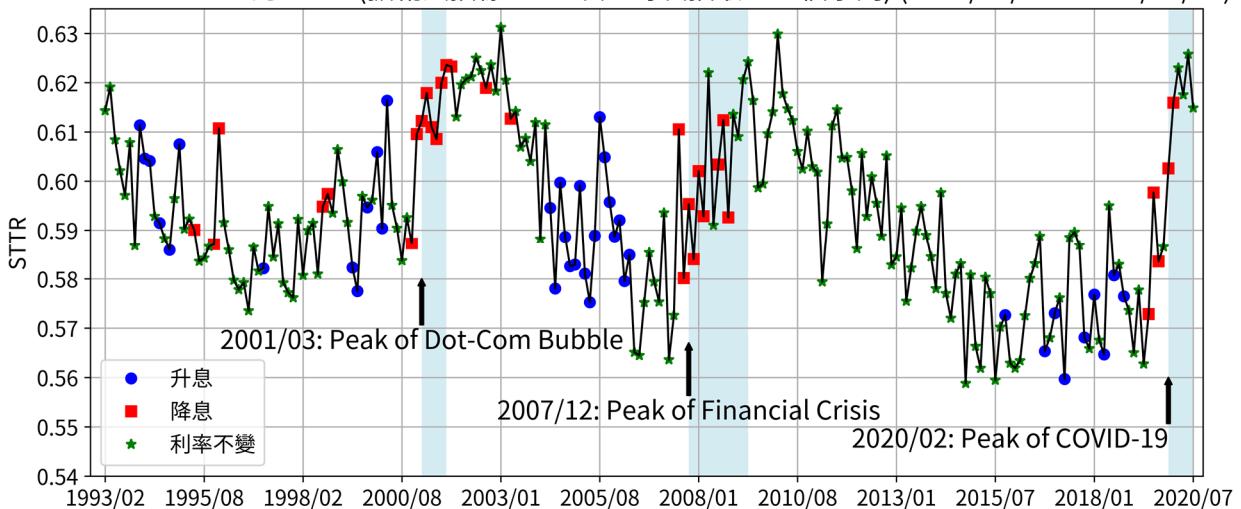


圖 4-6、FOMC minutes 的 STTR 變化

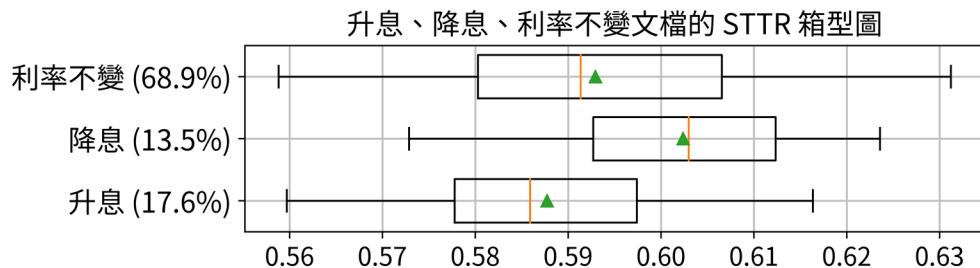


圖 4-7、升息、降息、利率不變文檔的 STTR 箱型圖

第五章 主題模型與詞嵌入

第一節 主題模型

主題模型可以幫助我們找到文檔中的主題，並列出前幾大的主題。我們擷取了 LSA 與 LDA 的前三大主題的各自的前 20 大字詞，並表 5-1 以不同顏色標註出了幾個在各主題之間重複的字詞，原本是期望能看到前三大主題對應到升息、降息、利率不變，或是對應到 Fed 的三大經濟使命 (高度可持續就業、穩定物價、適度的長期利率)，但在表 5-1 中，沒有找到與上述兩者的對應。推測應該是因為每一篇 FOMC minutes 都會提到 Fed 的三大經濟使命，所以各主題的字詞會互相重疊。

因為本研究沒有考慮到複合字，只使用了字詞，所以在此 LSA、LDA 充其量只能當作是抓取重要字詞的工具，黃于珊 (2017) 即使用 LSA 選取重要字詞，用以對 FOMC minutes 分類成升息、降息、利率不變三種分類。主題模型分析的失敗，讓我們轉向使用下一節的詞嵌入，希望能進一步發現 FOMC minutes 在升息、降息、利率不變三類文檔下不同的寫作風格。

表 5-1、LSA 與 LDA 前三個主題的前 20 大字詞

LSA	主題 1	主題 2	主題 3	LDA	主題 1	主題 2	主題 3
字詞 1	rate	growth	octob	字詞 1	april	april	octob
字詞 2	price	member	septemb	字詞 2	particip	juli	novemb
字詞 3	inflat	expans	novemb	字詞 3	march	august	august
字詞 4	market	inventori	rang	字詞 4	juli	march	septemb
字詞 5	econom	price	august	字詞 5	june	februari	decemb
字詞 6	growth	quarter	committe	字詞 6	hurrican	particip	januari
字詞 7	increas	sale	third	字詞 7	novemb	januari	juli
字詞 8	quarter	year	percent	字詞 8	may	decemb	particip
字詞 9	particip	product	rate	字詞 9	octob	octob	februari
字詞 10	continu	demand	juli	字詞 10	august	novemb	third
字詞 11	committe	juli	total	字詞 11	pce	june	fourth
字詞 12	year	restraint	restraint	字詞 12	agenc	guidanc	agenc
字詞 13	would	consider	monitor	字詞 13	medium	threshold	june
字詞 14	expect	economi	reserv	字詞 14	septemb	restraint	twelv
字詞 15	remain	aggreg	veloc	字詞 15	fomc	agenc	loan
字詞 16	polici	industri	object	字詞 16	loan	program	restraint
字詞 17	recent	moder	lower	字詞 17	februari	mb	march
字詞 18	declin	rang	monetari	字詞 18	decemb	loan	al
字詞 19	month	twelv	financi	字詞 19	district	recoveri	fomc
字詞 20	consum	direct	develop	字詞 20	gdp	fourth	pce



第二節 詞嵌入

參考了 Hunag and Kuan (2021) 整理的 FOMC minutes 中 Fed 三大經濟使命對應到的重要字詞，我們挑選出了 Fed 三大經濟使命中的 9 個關鍵字：

- (一) 高度可持續就業 (maximizing employment) : unemployment、labor、growth
- (二) 穩定物價 (stable prices) : inflation、price
- (三) 適度的長期利率 (moderate long-term interest rates) : fund、interest、exchange、money

其中，fund 對應到的是 federal fund rate，interest 則是對應到 interest rate，因為本研究僅使用單一字詞，沒有考慮到複合字，因此經處理過後的文本並沒有複合字存在。接著，我們對這 9 個關鍵字，以 word2vec 取各這 9 個關鍵字各自的前 5 大近義詞，然後將這 9 個關鍵字和 45 個近義詞，加起來總共 54 個字詞，以 t-SNE 技術繪製成圖，如圖 5-1。對 FOMC minutes 中的升息文檔、降息文檔、利率不變文檔，都做了上述流程後，我們觀察到這 9 大關鍵字在上述 3 類文檔中的聚類，以主觀的方式判斷聚類，發現在升息文檔中，inflation、price 屬於同一個聚類，且 unemployment、labor 屬於同一個聚類，其他關鍵字則都屬於不同的聚類；在降息文檔中，inflation、price 屬於同一個聚類，其他關鍵字則都屬於不同的聚類；在利率不變文檔中，9 大關鍵字都屬不同的聚類。這顯示 Fed 在升息、降息時 inflation 和 price 的語義較相近 (上下文距離比較近)；而且在升息時，unemployment 和 labor 的語義較相近。

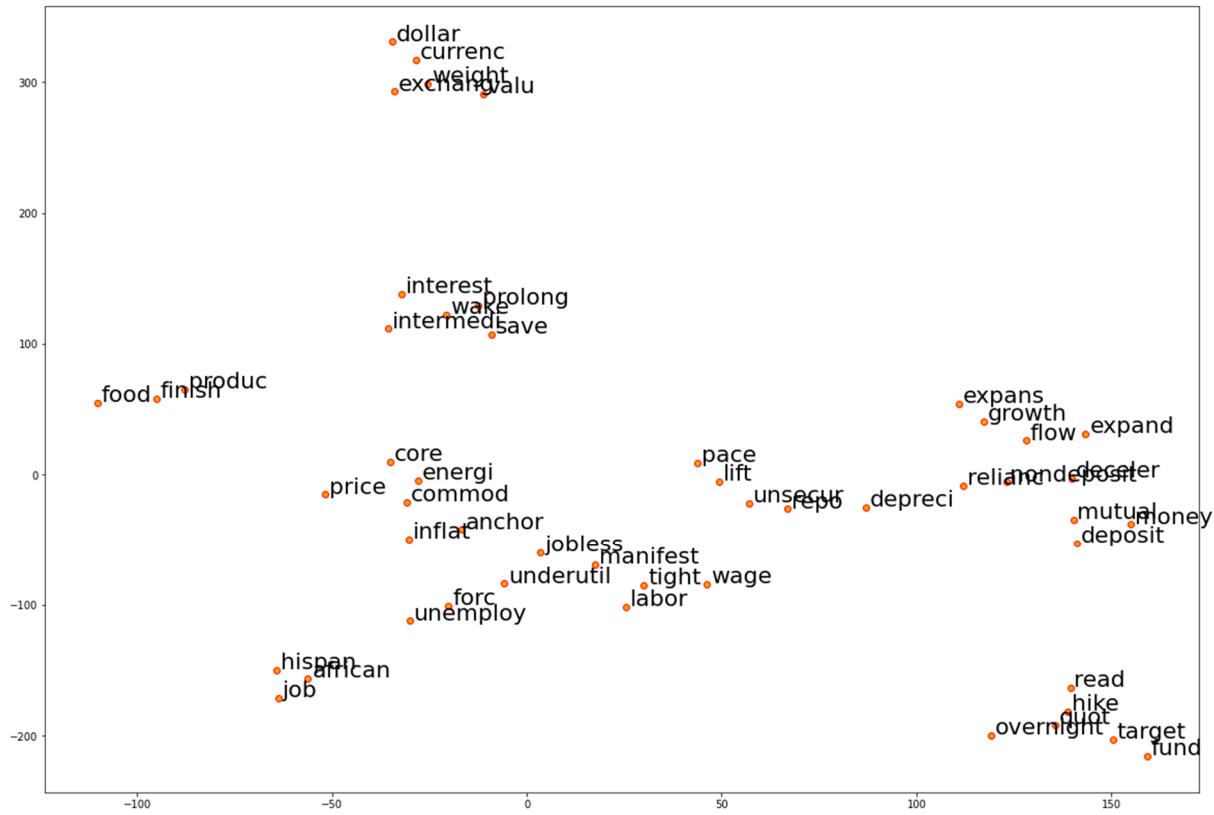


圖 5-1、升息文檔的 t-SNE 視覺化

接著，一一檢視 9 大關鍵字與各自的前 20 大近義詞，發現 *unemploy* 與其對應的前 20 大近義詞，在升息、降息、利率不變這三類文檔中差異特別大：在升息、降息的文檔中，會特別出現 *african*、*hispan*、*white*、*american*、*civilian*。其中，*civilian* 應該是對應到 *civilian unemployment rate*，而 *african*、*hispan*、*white*、*american* 對應到的是在 FRED (federal reserve economic data) 上一個非常重要的圖表，也是 Fed 重點看待的數據，那就是美國四大族裔的失業率統計，如圖 5-2。也就是說，在升息、降息的文檔中，Fed 會特別強調各族裔的失業率。¹³

¹³ 見 <https://fred.stlouisfed.org/graph/?g=houW>

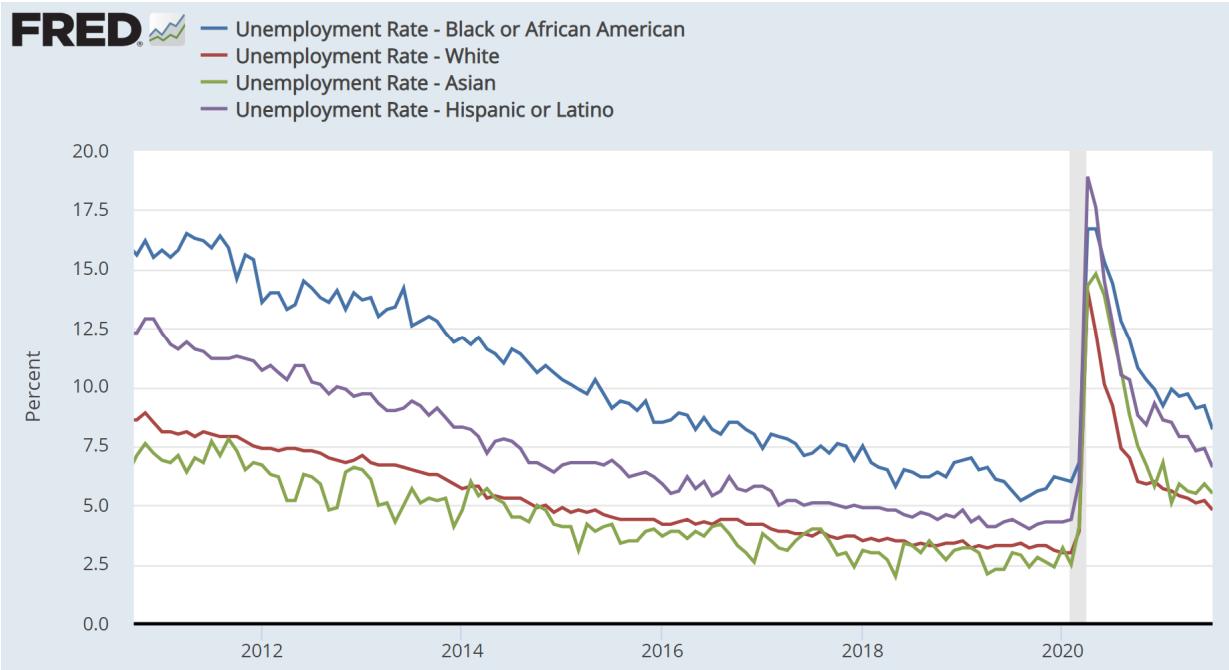


圖 5-2、美國四大族裔的失業率統計



第六章 文本分類

將 222 個 FOMC minutes 文檔，針對不同的特徵選取方法，分別做了 10 次、20 次、100 次的 80-20 資料集切分，來隨機分成 80% 訓練資料集、20% 測試資料集，並以準確率 (accuracy) 作為評價分類模型的指標，計算測試資料集的平均準確率。¹⁴ 圖 6-1 的三種初階的特徵選取方法中，做了 10 次的 80-20 資料集切分；圖 6-2 的 N 倍數指標篩選器中，做了 20 次的 80-20 資料集切分；圖 6-3 的線性降維方法 PCA 中，做了 100 次的 80-20 資料集切分；圖 6-4 的非線性降維方法 kernel PCA 中，做了 100 次的 80-20 資料集切分。

因為 FOMC minutes 中的升息文檔 (佔 17.57%)、降息文檔 (佔 13.51%)、利率不變文檔 (佔 68.92%)，這三者是不平衡資料，因此有使用分層抽樣，依照這三類文檔的比例去切分出訓練資料集、測試資料集；並以將文檔全部分類為利率不變文檔的準確率 68.92% 設定為基準。對於不平衡資料的進階處理，並沒有使用上採樣 (up-sampling)、下採樣 (down-sampling)，原因是考量到文檔稀少，只有 222 個，並且是要做三分類而不是二元分類，使用上採樣會放大雜訊、下採樣會損失很多文檔訊息，故沒有使用。而沒有使用 F1-score 而採用準確率作為評價分類模型的指標，是因為在三分類下 F1-score 的算法有權重選擇方面的問題。

在研究方法章節，我們介紹過了 N 倍數指標篩選器。在特徵篩選部份，我們使用總字詞數 (共 4102 個特徵)、全部文檔合併後的 TF 前 200 大字詞、全部文檔合併後的 TF-IDF 前 200 大字詞、N 倍數指標篩選器、線性降維方法 PCA、非線性降維 kernel PCA，來篩選特徵並縮減特徵維度；因為本研究的原始 DTM 為 222×4102 的稀疏矩陣，有必要將特徵空間維度縮減至 222 以下，以避免維度災難。在分類器模型部份，我們使用了 LR、SVM、RF、XGBoost，共 4 種分類器。對於在不同特徵選取方法下的 LR、SVM、RF、XGBoost 四種分類器的參數調整，是用網格搜索法 (grid search) 初步找尋模

¹⁴ 原本預計跑 500 次 80-20 資料集切分，但因為使用 Google Colaboratory 跑，並限定跑 N 次 80-20 資料集切分的時間小於 1 小時，所以最終分別在不同資料選取方法下跑了 10 次、20 次、100 次 80-20 資料集切分。

型最佳化的參數的大致範圍，然後在其範圍開一個區間填入幾個數值，再使用網格搜索法讓其自動尋找最佳的參數。

圖 6-1 的初階特徵篩選方法中，x 軸有三種初階的特徵選取方法——全部字詞、TF 前 200 大字詞、TF-IDF 前 200 大字詞——並搭配四種分類器；y 軸則為準確率。圖 6-1 的初階特徵篩選方法中，準確率最高的情況，為使用全部字詞，搭配上 XGBoost 分類器，得到 85.56% 的準確率，而基準線是 68.92%。另外，在圖 6-1 的初階特徵篩選方法中，LR 與 XGBoost 分類器在全部字詞的準確率表現，都比 TF 前 200 大字詞、TF-IDF 前 200 大字詞還高。圖 6-2 的 N 倍數指標篩選器中，使用 N 倍數指標篩選器，搭配上四種分類器；其中，得到最高準確率的情況，為使用 1.7 倍數指標篩選器，搭配上 LR 分類器，得到 88.11% 的準確率，比圖 6-1 的初階特徵篩選方法中最高的準確率 85.56% 還高。另外，圖 6-2 的 N 倍數指標篩選器中，大致上而言，對四個分類器而言，最佳的準確率出現在 1.7 倍數指標篩選器、2 倍數指標篩選器、3 倍數指標篩選器中。

圖 6-3 的線性降維方法 PCA 中，使用線性降維方法 PCA，搭配四種分類器，得到最高 78.47% 的準確率，低於圖 6-1 的初階特徵選取方法中最高的準確率 85.56%，也低於圖 6-2 的 N 倍數指標篩選器中最高的準確率 88.11%。在圖 6-3 的線性降維方法 PCA 中，對於四種分類器而言，在可解釋變數率 (explained variance ratio) 於 95%~80% 區間時，隨著可解釋變異率下降，四種分類器的準確率大致上呈現上升，這可能代表資料中存在不少雜訊。圖 6-4 的非線性降維方法 kernel PCA 中，使用非線性降維方法 kernel PCA，搭配四種分類器，得到最高 73.18% 的準確率；因為在此 kernel PCA 作用下最高只能得到 177 個特徵，所以圖 6-4 中最高只有 177 個特徵。整體而言，非線性降維方法 kernel PCA 的準確率表現，都差於線性降維方法 PCA。

接著，我們固定好分類器為 LR，來觀察不同的特徵選取方法、特徵降維方法的差異。在圖 6-1 的初階特徵選取方法中，LR 分類器在使用全部字詞（使用了 4102 個特徵）有最高的準確率 83.22%。在圖 6-2 的 N 倍數指標篩選器中，LR 分類器在 1.7 倍數指標篩選器（使用了 210 個特徵）有最高的準確率 88.11%；而 LR 分類器在 3 倍數指標篩選器（使用了 36 個特徵）僅僅使用了 36 個特徵，就得到 83.49% 的準確率，這個準確率高於了 LR 分類器在使用全部字詞（使用了 4102 個特徵）得到的 83.22%。在圖 6-3 的線性

降維方法 PCA 中，LR 分類器在 80% 可解釋變異率（約 31 個特徵）時得到了最高的準確率 76.78%，並且隨著可解釋變異率降低，準確率些微升高。在圖 6-4 的非線性降維方法 kernel PCA 中，LR 分類器在使用 100 個特徵時，得到最高的準確率 71.27%；並且，考慮到特徵數量後，LR 分類器的線性降維方法得到的每個準確率，都高於 LR 分類器在非線性降維方法得到的每個準確率。

將總結起來，在本研究使用的不同的特徵選取方法下，LR、SVM、XGBoost 分類器有較好的表現，RF 分類器表現則較差。而使用 LR、SVM、RF、XGBoos 共四種分類器，搭配 N 倍數指標篩選器、線性降維方法 PCA、非線性降維方法 kernel PCA 共四種特徵選取方法、特徵降維方法下的準確率排名，是 N 倍數指標篩選器表現最好、線性降維方法 PCA 表現次之，非線性降維方法 kernel PCA 表現最差。線性降維方法 PCA 的準確率，在可解釋變數率於 95%~80% 區間時，隨著可解釋變異率下降，四個分類器的準確率大致上呈現上升，暗示著此資料集可能含有不少雜訊。最後，LR 分類器在 3 倍數指標篩選器（使用了 36 個特徵）僅僅使用了 36 個特徵，就得到 83.49% 的準確率，這個準確率高於了使用 LR 分類器搭配全部字詞（使用了 4102 個特徵）得到的 83.22%。

對不同特徵選取方法下，四種分類器各自在網格搜索下的參數，讀者有興趣想知道的話，可以參見第三章研究方法的開頭第一段，去本研究的 GitHub 查看程式碼。

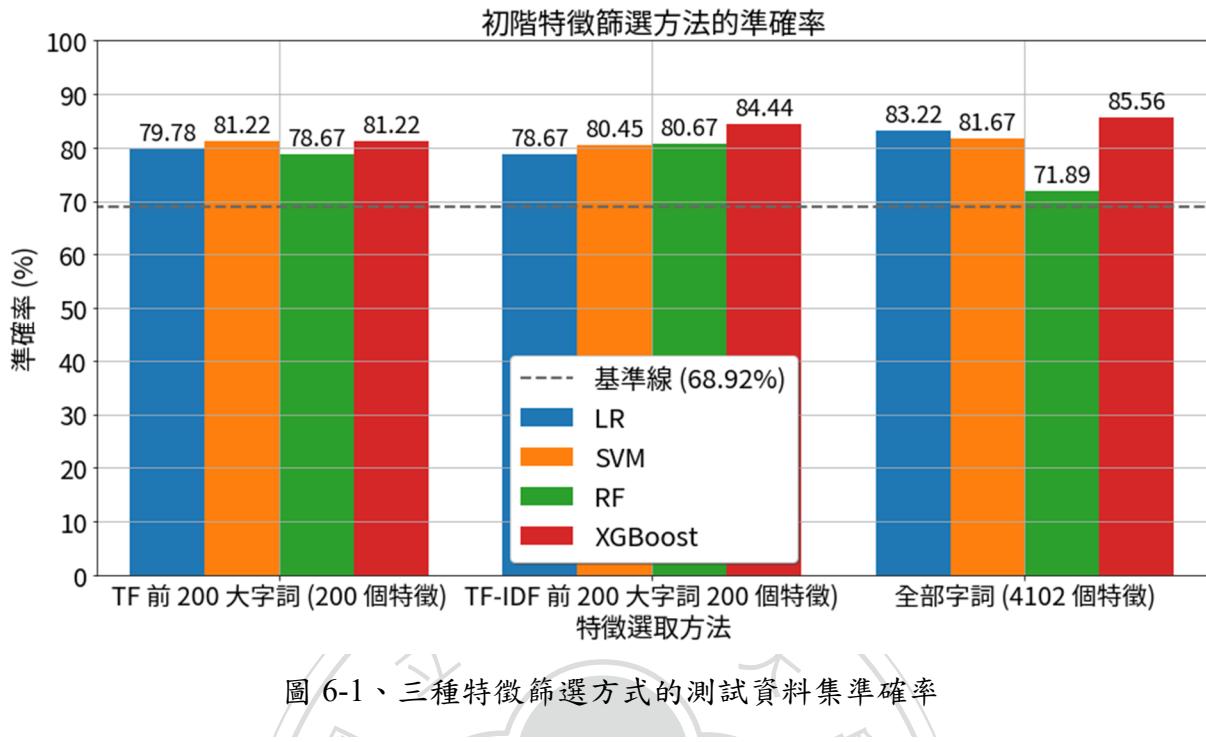


圖 6-1、三種特徵篩選方式的測試資料集準確率

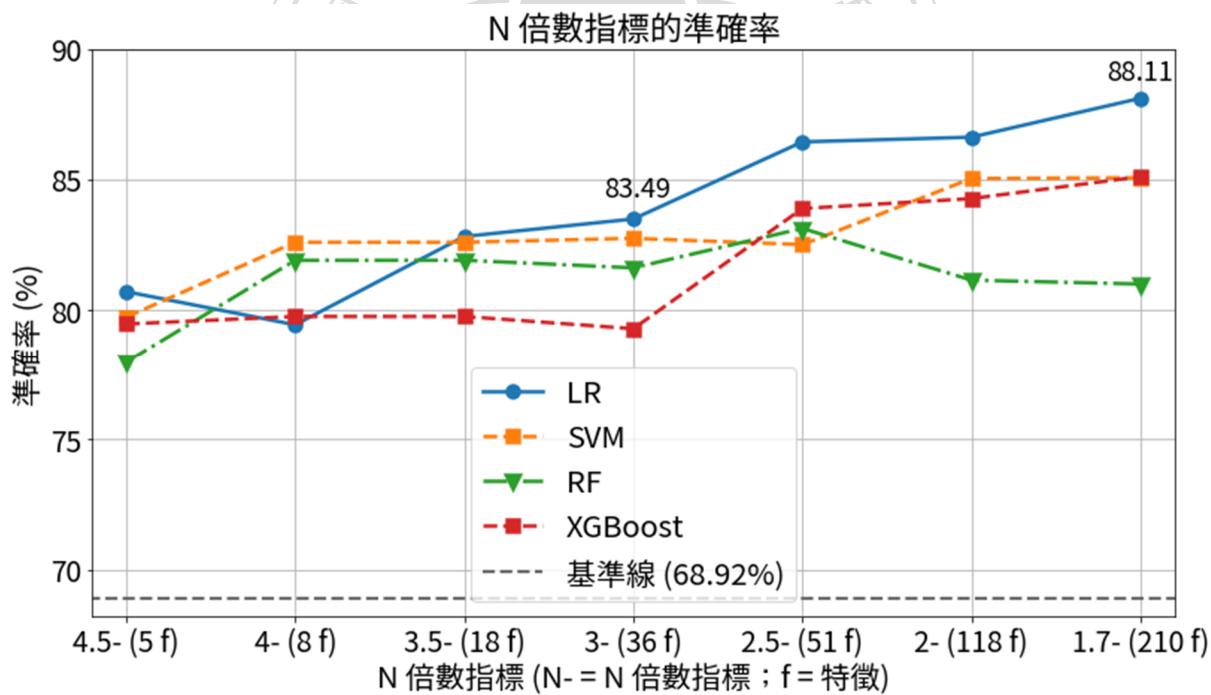


圖 6-2、倍數指標的測試資料集準確率

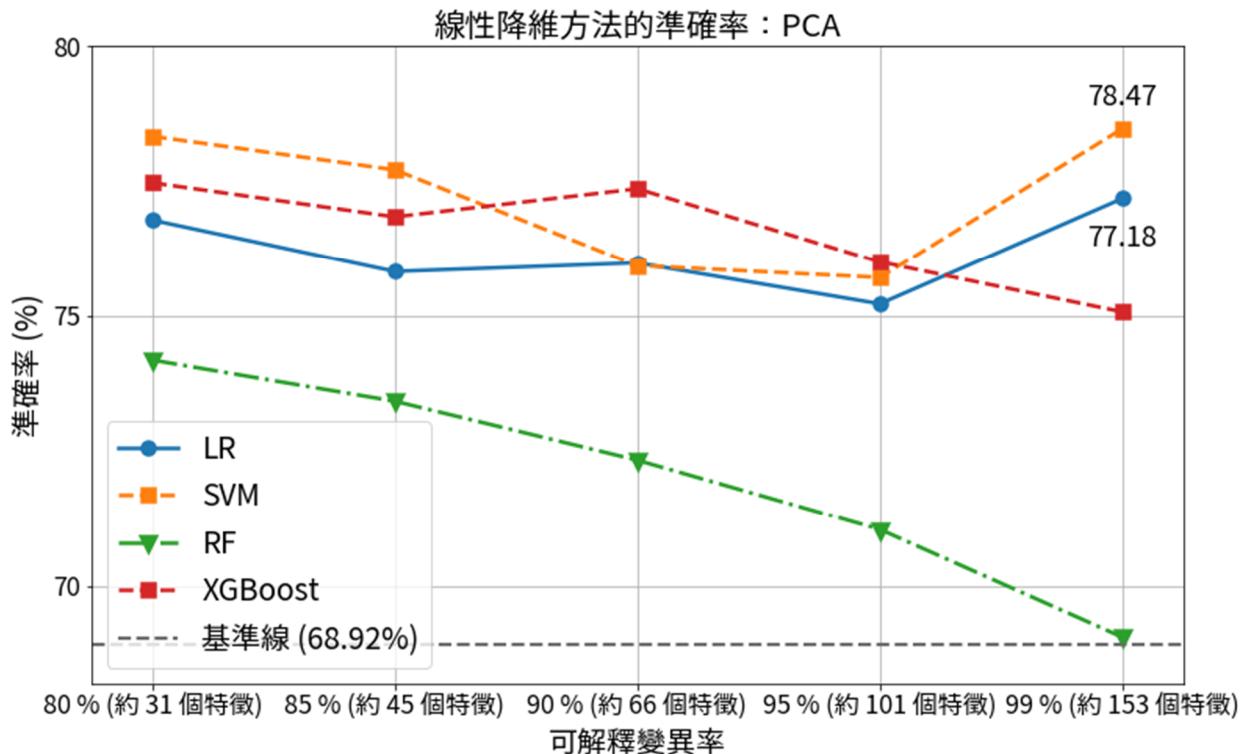


圖 6-3、線性降維方法的測試資料集準確率

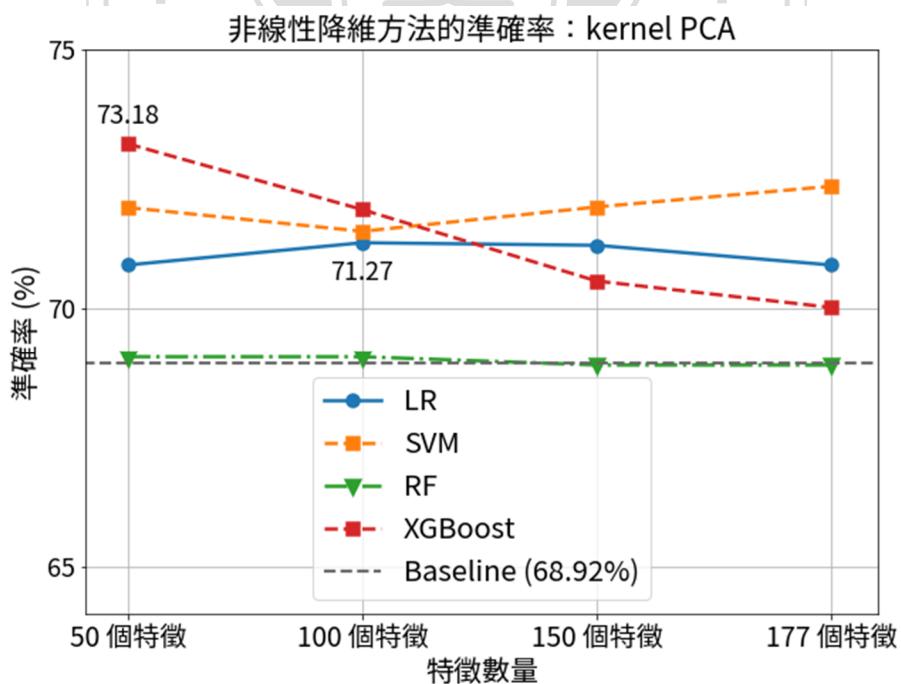


圖 6-4、非線性降維方法的測試資料集準確率

因為使用 LR 模型、1.7 倍數指標篩選器(使用了 210 個特徵)在測試資料集上得到了最高的 88.11% 準確率，所以我們隨機選擇某一個 80-20 訓練與測試資料集切分下的結果，進一步觀察其混淆矩陣(confusion matrix)，如圖 6-5(圖中 +1 為升息、0 為利率不變，-1 為降息)。圖 6-5 中的左半邊，發現錯誤分類的 6 個文檔中，5 個是將升息錯誤分類為利率不變，剩下 1 個是將不變錯誤分類為升息。這代表了降息文檔跟其他兩類文檔差異過大，導致降息文檔不會被錯誤分類為升息、利率不變，以及升息、利率不變文檔不會被錯誤分類為降息。從圖 6-5 的右半邊的數值區，也能看到降息文檔的 F1-score 高達 1.0，利率不變文檔的 F1-score 也有 0.91，但是升息文檔的 F1-score 却只有 0.50，代表升息樣本是最難分類正確的類別。

		Predicted			precision	recall	f1-score	support
		-1	0	1				
Actual	-1	6	0	0				
	0	0	30	1				
	1	0	5	3				

	accuracy	0.87	0.87	45	
	macro avg	0.87	0.78	0.80	45
	weighted avg	0.86	0.87	0.85	45

圖 6-5、LR 分類器搭配倍數指標的混淆矩陣

既然 N 倍數篩選器對文本分類幫助這麼大，很自然的我們想更深入觀察其篩選出來的字詞。因此，我們重點觀察使用 3 倍數篩選器(篩選出 36 個特徵)、4 倍數篩選器(篩選出 8 個特徵)。先從 4 倍數篩選器(篩選出 8 個特徵)開始，從表 6-1 中，可看到我們將篩選出的 8 個特徵，用兩種方式分類：第一種分類方式，是以 N 倍數篩選器的篩選過程使用的三對文檔來分類；第二種分類方式，是以字詞特性分類，但這個分類方法牽涉到了部分主觀意見。表 6-1 中，發現降息文檔中特有的字詞共有 contract、correct、eas、institut、weaken、turmoil，共 6 個；而升息文檔中特有的字詞有 hurrican、remov，共 2 個。在表 6-2 的 3 倍數指標篩選器篩選出的 36 個特徵中，降息文檔特有且與總體經濟較相關的字詞有 credit、cutback、deterior、discount、downturn、drop、ease、lend、

liquid、reduct、repo、soften、strain、turmoil、weak、weaken、weaker，共 17 個；降息文檔特有且與總體經濟較相關的字詞有 expand、medium、owe、rebuild、unseason、upsid、warm，共 7 個。

表 6-1、將 4 倍數指標得到的 8 個特徵分類

4 倍數指標篩選器，篩選出的 8 個特徵						
依倍數指 標篩選時 使用的文 檔分類 (會有重複)	全部升息文檔 / 全部降息文檔		全部升息文檔 / 全部降息文檔加 上全部利率不變文檔		全部降息文檔 / 全部升息文檔加 上全部利率不變文檔	
	在升息的詞 頻 $\geq 4 \times$ 在降 息的詞頻	在降息的詞頻 $\geq 4 \times$ 在升息的 詞頻	在升息的詞頻 $\geq 4 \times$ 在降息加 上利率不變的 詞頻	在降息加上利 率不變的詞頻 $\geq 4 \times$ 在升息的 詞頻	在降息的詞頻 $\geq 4 \times$ 在升息加 上利率不變的 詞頻	在升息加上利 率不變的詞頻 ≥ 4 \times 在降息的詞頻
		['contract', 'eas', 'institut', 'weaken']	['hurrican', 'remov']	['eas']	['correct', 'eas', 'turmoil']	[]
依字詞特 性分類	與總體經濟較相關的字詞 eas、turmoil、weaken，共 3 個		突發事件的字詞 hurrican，共 1 個		不包含在前面兩類的字詞 contract、correct、institut、 remov，共 4 個	



表 6-2、將 3 倍數指標得到的 36 個特徵分類

3 倍數指標篩選器，篩選出的 36 個特徵					
	全部升息文檔 / 全部降息文檔		全部升息文檔 / 全部降息文檔加上全部利率不變文檔		全部降息文檔 / 全部升息文檔加上全部利率不變文檔
依倍數指標篩選時使用的文檔分類 (會有重複)	在升息的詞頻 $\geq 3 \times$ 在降息的詞頻	在降息的詞頻 $\geq 3 \times$ 在升息的詞頻	在升息的詞頻 $\geq 3 \times$ 在降息加上利率不變的詞頻	在降息加上利率不變的詞頻 $\geq 3 \times$ 在升息的詞頻	在降息的詞頻 $\geq 3 \times$ 在升息加上利率不變的詞頻
	['expand', 'gradual', 'medium', 'upsid']	['august', 'contract', 'credit', 'drop', 'eas', 'institut', 'lend', 'liquid', 'reduct', 'soften', 'weak', 'weaken', 'weaker']	['contain', 'hurrican', 'owe', 'pass', 'rebuild', 'remov', 'unseason', 'warm']	['eas', 'recoveri']	['eas', 'deterior', 'weaken', 'strain', 'soften', 'downturn', 'turmoil', 'coronaviru', 'outbreak', 'correct', 'discount', 'repo', 'cutback']
依字詞特性分類	與總體經濟較相關的字詞		突發事件的字詞		不包含在前面兩類的字詞
	credit、cutback、deterior、discount、downturn、drop、eas、expand、lend、liquid、medium、owe、rebuild、recoveri、reduct、repo、soften、strain、turmoil、unseason、upsid、warm、weak、weaken、weaker，共 25 個		coronoviru、hurrican、outbreak，共 3 個		august、contain、contract、correct、gradual、institut、pass、remov，共 8 個

第七章 結論與討論

第一節 結論

本研究以 FOMC minutes 為出發點展開，首先分析了目標聯邦基金利率的變動，發現 Fed 在降息上比較果斷乾脆，在升息上比較保守，而 Fed 在升息、降息上通常都是連續為之。並且，目標聯邦基金利率的變動有自身相關性，若以天真預測法來預測其變動（升息、降息、利率不變），得到高達 74.7% 的準確率。接著，透過探索式資料分析，發現 FOMC minutes 的 STTR 與景氣循環為同時指標。

為了進一步探索 FOMC minutes 升息、降息、利率不變文檔的寫作風格差異，使用主題模型但沒有具價值的發現，因此進一步使用詞嵌入模型，以跟 Fed 三大經濟使命相關的 9 大關鍵字為核心，發現在升息文檔裡，inflation、price 在聚類分析上距離較近（語義較近，以字詞間的距離衡量），unemployment、labor 在聚類分析上距離較近；在降息文檔裡，inflation、price 在聚類分析上距離較近。進一步觀察這 9 大關鍵字各自的近義詞，發現在 9 大關鍵字中的 unemploy 與其對應的前 20 大近義詞，在升息、降息、利率不變這三類文檔中差異特別大，在升息、降息時會特別出現 african、hispan、white、american，代表在升息、降息時 Fed 會特別強調各族裔的失業率。

在文本分類部分，因為本研究的原始 DTM 為 222×4102 的稀疏矩陣，有必要將特徵空間維度縮減至 222 個以下，以避免維度災難。因此我們使用幾種特徵篩選方法來縮減維度至 222 個以下，包含 N 倍數指標篩選器、線性降維、非線性降維，接著再套用到羅吉斯迴歸、支持向量機、隨機森林、XGBoost 共 4 種分類器中，進行文本分類。以上述方法得到的所有準確率數字中，我們得到的最高準確率是 88.11%，遠大於 68.92% 的基準線，而此最高的準確率發生於使用 1.7 倍數指標篩選器（共 210 個特徵）搭配上 LR 分類器時。並且，LR 分類器在 3 倍數指標篩選器（使用了 36 個特徵）僅僅使用了 36 個特徵，就得到 83.49% 的準確率，這個準確率高於了使用 LR 分類器搭配全部字詞（使用了 4102 個特徵）得到的 83.22%。

在文本分類部分，使用的不同的特徵選取方法時，LR、SVM、XGBoost 分類器有較好的表現，RF 分類器表現則較差。而使用 LR、SVM、RF、XGBoos 共四種分類器，

搭配 N 倍數指標篩選器、線性降維方法 PCA、非線性降維方法 kernel PCA 共四種特徵選取方法、特徵降維方法下的準確率排名，是 N 倍數指標篩選器表現最好、線性降維方法 PCA 表現次之，非線性降維方法 kernel PCA 表現最差。線性降維方法 PCA 的準確率，在可解釋變數率於 95%~80% 區間時，隨著可解釋變異率下降，四個分類器的準確率大致上呈現上升，暗示著此資料集可能含有不少雜訊。因為使用 LR 模型、1.7 倍數指標篩選器（使用了 210 個特徵）在測試資料集上得到了最高的 88.11% 準確率，所以我們隨機選擇某一個 80-20 訓練與測試資料集切分下的結果，進一步觀察其混淆矩陣，發現降息文檔的 F1-score 高達 1.0，利率不變文檔的 F1-score 也有 0.91，但是升息文檔的 F1-score 却只有 0.50，代表升息樣本是最難分類正確的類別。進一步觀察 3 倍數篩選器（篩選出 36 個特徵）、4 倍數篩選器（篩選出 8 個特徵）篩選出來的特徵，發現 FOMC minutes 在升息、降息文檔，各自用有特有的字詞。

本研究使用目標聯邦基金利率、FOMC minutes 做為資料，先後進行時間序列分析、探索式分析、主題模型、詞嵌入、文本分類，並發現了 FOMC minutes 在升息、降息、利率不變文檔，有著不一樣的寫作風格，找到在升息、降息、利率不變文檔間具較大差異的字詞。雖然在文本分類，我們使用倍數指標得到了令人滿意的準確率，但仔細探究會發現，存在一些與總體經濟不是非常相關的字詞，大幅的增加了在文本分類的準確率。

本研究的特色，在於以下三點：第一，相較於過去文獻，更明確地界定了如何定義和擷取 FOMC minutes 的主文（亦即刪除每篇 FOMC minutes 前後較不重要的文字段落），使 FOMC minutes 之間更具可比較性。第二，過往文獻較少墨於探索式資料分析，而本研究對目標聯邦基金利率、FOMC minutes 做了詳細的探索式分析。第三，在文本分類部分，黃于珊（2017）使用用了線性降維、LSA 兩種方法來篩選特徵，並搭配線性判別分析（linear discrimination analysis）分類器；本研究則使用 N 倍數指標篩選器、線性降維、非線性降維、等方法篩選特徵，接著套用到羅吉斯迴歸、支持向量機、隨機森林、XGBoost 共 4 種分類器中。

第二節 討論

本研究的 DTM 並沒有在各分析部分都重點刪除大量詞頻較低的字詞，因此，未來可以設定一個合理的最低詞頻門檻來篩掉大量詞頻較低的字詞。本研究沒有考慮到複合字，只考慮了單一字詞，因此本研究的 DTM 為 222×4102 的矩陣 (222 個 FOMC minutes 文檔、4102 個字詞)；相較之下，Huang and Kuan (2021) 考慮到了複合字，其 DTM 為 192×9311 的矩陣。雖然本研究的研究資料比 Huang and Kuan (2021) 還多，並且包含他們所有的文檔，但他們卻有更高的特徵空間維度。未來可以參考 Huang and Kuan (2021)，在資料預處理階段，找出專門術語、使用 N-gram、使用語料庫，來辨識出複合字。

本研究在資料集切分上使用 80-20 資料集切分，並在 Google Colaboratory 上運行詞嵌入、文本分類。在文本分類時，原本預計跑 500 次 80-20 資料集切分，但是，考量到運行時間後，針對不同的資料選取方法，分別只跑了 10 次、20 次、100 次 80-20 資料集切分。因此，未來可以試著跑 500 次 80-20 資料集切分，得出來的分類準確率，將更準確。

本研究在文本分類時，使用 N 倍數指標篩選後，發現一些對分類準確率影響力較大的字詞，但這些字詞當中有不少是與總體經濟無明顯直接相關的字詞，比方突發事件相關的 hurricane 和 coronavirus。未來可以排除與總體經濟無明顯直接相關的字詞，針對跟 Fed 三大經濟使命有關的字詞、其他重要的經濟字詞，探討這些字詞對 FOMC minutes 三分類 (升息、降息、利率不變) 準確率的影響。

參考文獻

一、中文文獻

- 柏南克（2013）。《柏南克的四堂課：聯準會與金融危機》。臺北：財信。
- 孫亮、黃倩（2017）。《實用機器學習》。北京：人民電郵。
- 黃于珊（2017）。「文字探勘在總體經濟上之應用—以美國聯準會會議紀錄為例」，政治大學金融學系碩士論文。

二、英文文獻

- Abel, A. B., Bernanke, B., Croushore, D. (2013). Macroeconomics (8nd ed.). New Jersey, NJ: Pearson.
- Aggarwal, C. C. (2018). Machine Learning for Text. Cham, CH: Springer International Publishing. <https://doi.org/10.1007/978-3-319-73531-3>
- Bernanke, B. (2012). "The Federal Reserve and the Financial Crisis: The Aftermath of the Crisis, Lecture 4." George Washington University School of Business.
<https://www.federalreserve.gov/mediacenter/files/chairman-bernanke-lecture4-20120329.pdf>
- Blinder, A. S., Ehrmann, M., de Haan, J., Fratzscher, M., & Jansen, D.-J. (2008). "Central bank communication and monetary policy: A survey of theory and evidence," *Journal of Economic Literature*, 46, 910–945. <https://doi.org/10.1257/jel.46.4.910>
- Board of Governors of the Federal Reserve System. (2021, January 14). "Federal Open Market Committee: Transcripts and other historical materials." Board of Governors of the Federal Reserve System. https://www.federalreserve.gov/monetarypolicy/fomc_historical.htm
- Boukus, E., & Rosenberg, J. V. (2006). "The information content of FOMC minutes." Federal Reserve Bank of New York. <https://doi.org/10.2139/ssrn.922312>
- Cannon, S. (2015). "Sentiment of the FOMC: Unscripted," *Economic Review [Federal Reserve Bank of Kansas City]*, Fall 2015, pp. 55.
- Chollet, F. (2018). Deep learning with Python. Shelter Island, NY: Manning Publications.

- Doh, T., Song, D., & Yang, S. K. (2020). "Deciphering Federal Reserve Communication via Text Analysis of Alternative FOMC Statements." Federal Reserve Bank of Kansas City, Research Working Paper no. 20-14, October. <https://doi.org/10.18651/RWP2020-14>
- Ericsson, N. R. (2016). "Eliciting GDP forecasts from the FOMC's minutes around the financial crisis," *International Journal of Forecasting*, 32, 571–583. <https://doi.org/10.1016/j.ijforecast.2015.09.007>
- Ganegedara, T. (2018). Natural Language Processing with TensorFlow. Birmingham, UK: Packt Publishing.
- Hayo, B., & Neuenkirch, M. (2013). "Do Federal Reserve Presidents Communicate with a Regional Bias?" *Journal of Macroeconomics*, 35(4), 62–72. <https://doi.org/10.1016/j.jmacro.2012.10.002>
- Huang, Y. L., & Kuan, C. M. (2021). "Economic Prediction with the FOMC Minutes: An Application of Text Mining," *International Review of Economics & Finance*, 71, 751-761. <https://doi.org/10.1016/j.iref.2020.09.020>
- Hyndman, R. J., & Athanasopoulos, G. (2018). Forecasting: principles and practice (2nd ed.). Melbourne, AU: OTexts. <https://otexts.com/fpp2/>
- Joshi, P. (2016). Python Machine Learning Cookbook. Birmingham, UK: Packt Publishing.
- Jubinskia, D., & Tomljanovich, M. (2013). "Do FOMC Minutes Matter to Markets? An Intraday Analysis of FOMC Minutes Releases on Individual Equity Volatility and Returns," *Review of Financial Economics*, 22(3), 86–97. <https://doi.org/10.1016/j.rfe.2013.01.002>
- Kliesen, K. L., Levine, B., & Waller, C. J. (2019). "Gauging Market Responses to Monetary Policy Communication," *Federal Reserve Bank of St. Louis Review*, pp. 69-91. <https://doi.org/10.20955/r.101.69-91>
- Lucca, D. O., & Trebbi F. (2009). "Measuring Central Bank Communication: An Automated Approach with Application to FOMC Statements." NBER Working Paper, No. 15367. <https://doi.org/10.2139/ssrn.1470443>
- Mikolov, T., Sutskever, I., Chen, K., Corrado, G. and Dean, J. (2013). "Distributed Representations of Words and Phrases and their Compositionality." Proceedings of the 26th International Conference on Neural Information Processing Systems, NIPS'13, Curran Associates, Red Hook, NY, Vol. 2, pp. 3111-3119.

- Rosa, C. (2013). “The Financial Market Effect of FOMC Minutes,” *FRBNY Economic Policy Review*, 19(2), 67–81. <https://ssrn.com/abstract=2378398>
- Sarkar, D. (2019). Text Analytics with Python (2nd ed.). Bangalore, India: Apress.
- Shapiro, A.H., & Wilson, D.J. (2019). “Taking the Fed at its Word: A New Approach to Estimating Central Bank Objectives using Text Analysis.” Federal Reserve Bank of San Francisco Working Paper 2019-02. <https://doi.org/10.24148/wp2019-02>
- Stekler, H., & Symington, H. (2016). “Evaluating qualitative forecasts: The FOMC minutes, 2006–2010,” *International Journal of Forecasting*, 32, 559–570.
<https://doi.org/10.1016/j.ijforecast.2015.02.003>
- VanderPlas, J. (2017). Python Data Science Handbook. California, CA: O'Reilly Media.



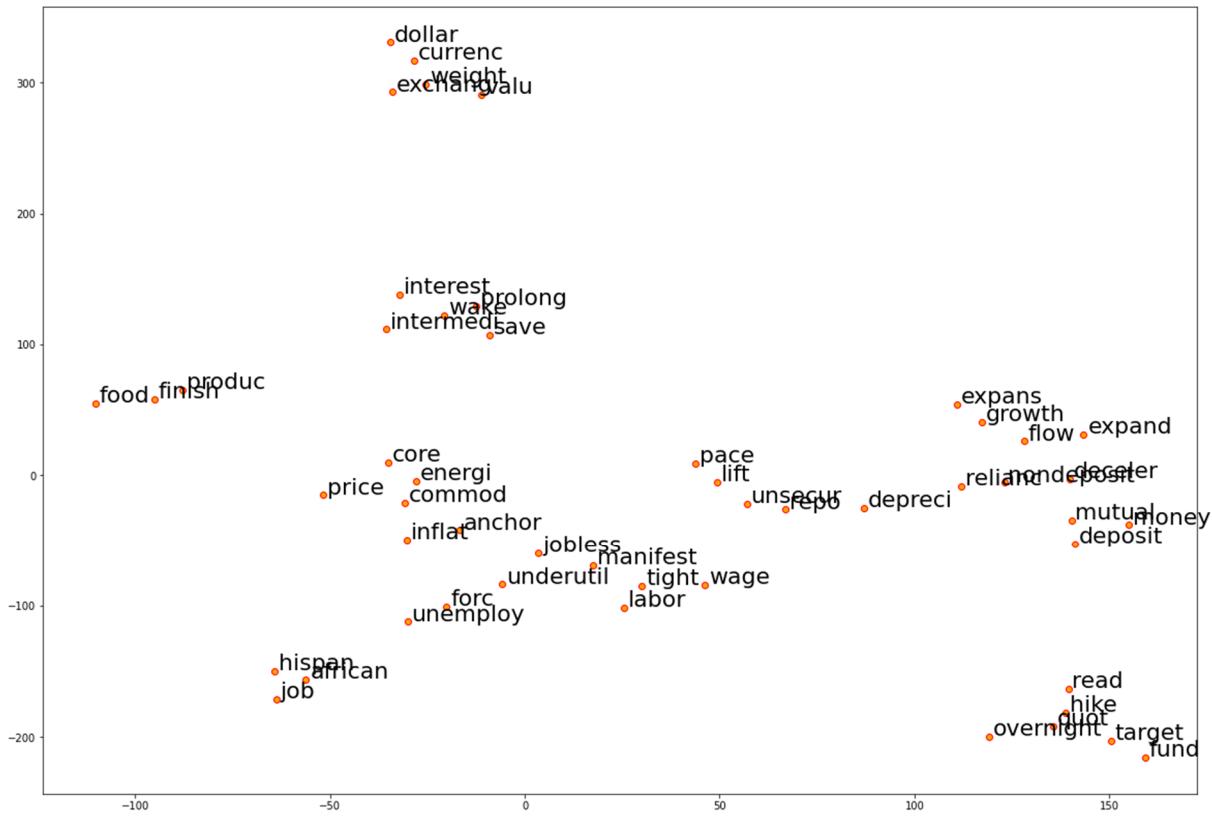
附錄一、名詞簡稱對照表

簡稱	中文名稱	英文名稱
BOW	詞袋模型	bag-of-words model
CBOW	連續詞袋模型	continuous bag-of-words
DTM	文檔-詞頻矩陣	document-term matrix
Fed	聯邦準備系統	Federal Reserve System
FOMC	聯邦公開市場委員會	Federal Open Market Committee
FOMC minutes	聯邦公開市場委員會會議紀要	Federal Open Market Committee minutes
kernel PCA	核主成分分析	kernel principal component analysis
LDA	隱含狄利克雷分布	latent Dirichlet allocation
LR	羅吉斯迴歸	logistic regression
LSA	潛在語義分析	latent semantic analysis
NMF	非負矩陣分解	non-negative matrix factorization
NPLM	神經統計語言模型	neural probabilistic language model
PCA	主成分分析	principal component analysis
PLSA	機率潛在語義分析	probabilistic latent semantic analysis
STTR	標準化相異詞出現率	standardized type-token ratio
SVD	奇異值分解	singular value decomposition
SVM	支持向量機	supporting vector machine
TF	詞頻	term frequency

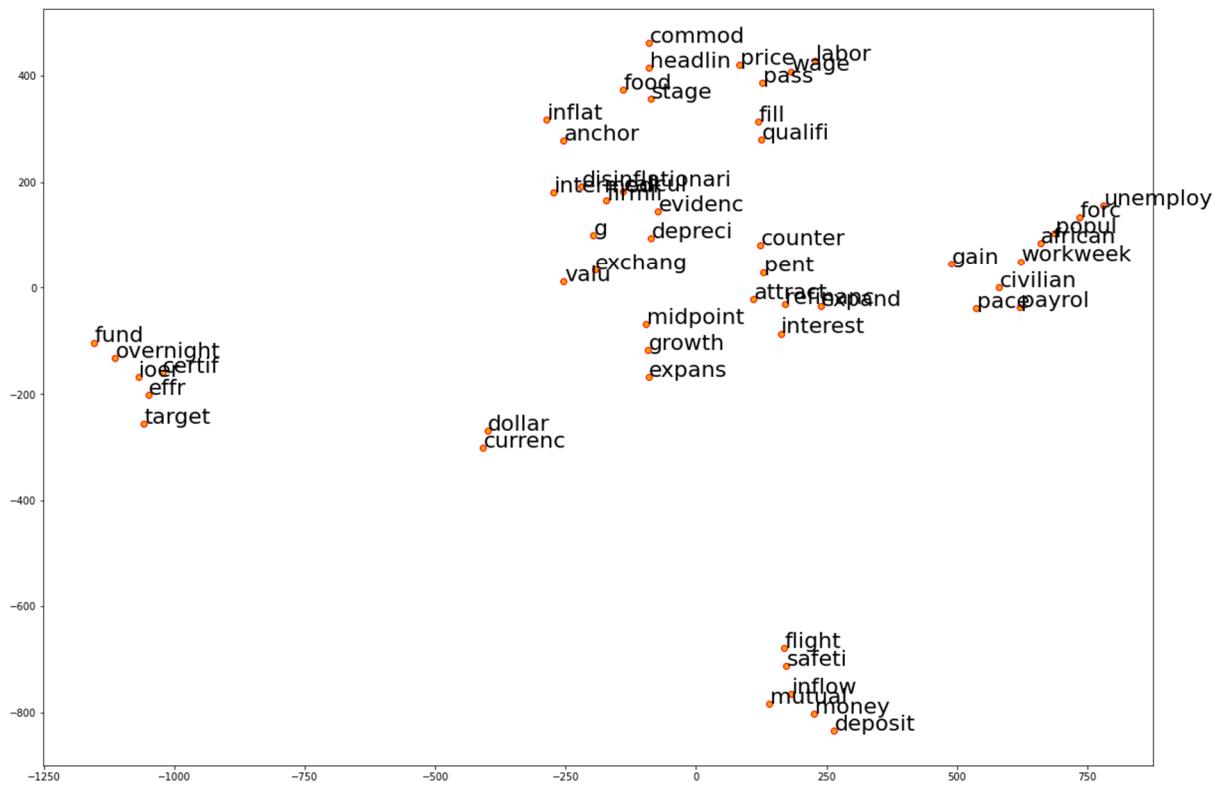
TF-IDF	詞頻-逆向檔案頻率	term frequency-inverse document frequency
t-SNE	t-隨機鄰近嵌入法	t-distributed stochastic neighbor embedding
TTR	相異詞出現率	type-token ratio



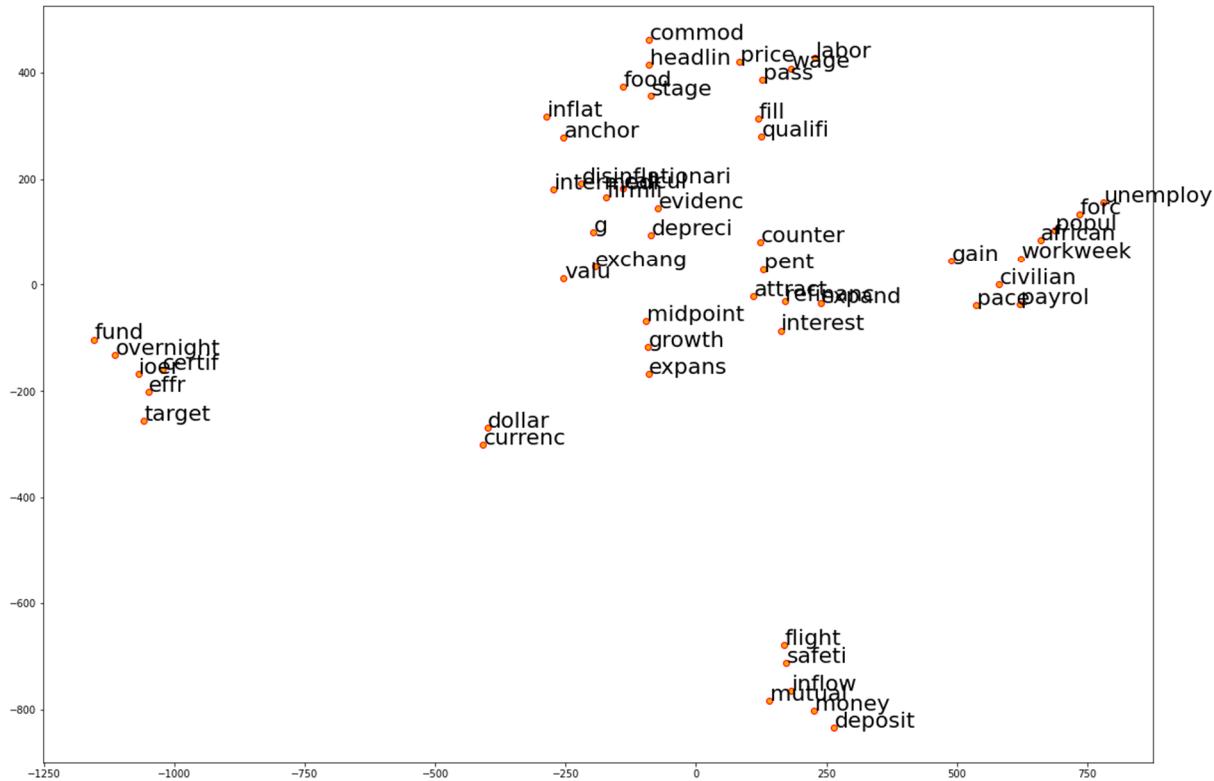
附錄二、詞嵌入完整圖表



附圖1、升息文檔的 t-SNE 視覺化



附圖 2、降息文檔的 t-SNE 視覺化



附圖 3、利率不變文檔的 t-SNE 視覺化

附表 1、四類文檔中，各自與 *unemploy* 的前 20 大近義詞

(1) 全部文檔	unemploy : [' <i>forc</i> ', ' <i>workweek</i> ', ' <i>hire</i> ', ' <i>age</i> ', ' <i>popul</i> ', ' <i>reason</i> ', ' <i>jobless</i> ', ' <i>save</i> ', ' <i>underutil</i> ', ' <i>interest</i> ', ' <i>hour</i> ', ' <i>sideway</i> ', ' <i>summari</i> ', ' <i>quit</i> ', ' <i>group</i> ', ' <i>delinqu</i> ', ' <i>addendum</i> ', ' <i>converg</i> ', ' <i>work</i> ', ' <i>fill</i> ']
(2) 升息文檔	unemploy : [' <i>forc</i> ', ' african ', ' hispan ', ' workweek ', ' american ', ' <i>hour</i> ', ' <i>underutil</i> ', ' <i>men</i> ', ' <i>job</i> ', ' white ', ' <i>jobless</i> ', ' <i>claim</i> ', ' <i>worker</i> ', ' civilian ', ' popul ', ' <i>monthli</i> ', ' <i>lowest</i> ', ' <i>payrol</i> ', ' <i>war</i> ', ' <i>ii</i> ']
(3) 降息文檔	unemploy : [' popul ', ' african ', ' <i>age</i> ', ' civilian ', ' workweek ', ' <i>forc</i> ', ' <i>claim</i> ', ' <i>hour</i> ', ' <i>group</i> ', ' <i>steadi</i> ', ' hispan ', ' white ', ' american ', ' <i>job</i> ', ' <i>averag</i> ', ' <i>insur</i> ', ' <i>hurrican</i> ', ' <i>payrol</i> ', ' <i>employ</i> ', ' <i>roughli</i> ']
(4) 利率不變文檔	unemploy : [' <i>forc</i> ', ' popul ', ' <i>age</i> ', ' workweek ', ' <i>hire</i> ', ' <i>jobless</i> ', ' <i>delinqu</i> ', ' <i>work</i> ', ' <i>addendum</i> ', ' <i>save</i> ', ' <i>lengthi</i> ', ' <i>weekli</i> ', ' <i>worker</i> ', ' <i>job</i> ', ' <i>reason</i> ', ' <i>suspend</i> ', ' <i>manifest</i> ', ' <i>interest</i> ', ' <i>payrol</i> ', ' <i>durat</i> ']

附表 1 列出四類文檔中，各自與 *unemploy* 的前 20 大近義詞，並發現：

(一) 出現在全部四類文檔的字詞：['forc', 'workweek', 'popul']

(二) 只出現在升息、降息的字詞：['african', 'hispan', 'white', 'american', 'civilian']

