

國立中興大學資訊管理學系

碩士學位論文

利用多任務學習以財經新聞預測美國重要經濟
領先指標

Forecasting US Economic Leading Indicators from
Financial News using Multi-Task Learning

指導教授： 英家慶 教授 Jia - Ching Ying

研 究 生： 劉家蓁 Chia - Chen Liu

中華民國 一一一年 七月

國立中興大學資訊管理學系

碩士學位論文

利用多任務學習以財經新聞預測美國重要經濟
領先指標

Forecasting US Economic Leading Indicators from
Financial News using Multi-Task Learning

指導教授： 英家慶 教授 Jia - Ching Ying

研 究 生： 劉家蓁 Chia – Chen Liu

中華民國 一一一年 七月

誌謝辭

在研究所兩年的時間裡，碩士論文能夠如期完成，需要感謝許多人，在此獻上最大的敬意。

首先想要感謝指導教授英家慶老師悉心的指導，在每週的團體會議報告，老師總能問出許多關鍵的問題，幫助我看見報告中的盲點，讓我從問題中學習掌握報告重點，以及學習用正確的角度分析出核心問題，讓我在研究所期間能夠持續進步。

我也想要感謝張雁婷學姐給予的協助，不時撥空討論研究上遇到的瓶頸，讓我可以即時改進，也在我報告時提醒我需要注意的細節，並提供我許多寶貴的建議。感謝實驗室的夥伴宛儒，在研究所兩年期間，學業與生活上的互相幫忙與支援，感謝其他實驗室的朋友，讓研究所生活變得有趣。

感謝我的家人，讓我可以專心學習，沒有經濟壓力，有良好的生活環境；感謝我的朋友，謝謝他們支持我做想做的事，時時給予我鼓勵與實質上的陪伴；感謝教會的夥伴們，你們是每週的精神糧食與快樂的泉源，感謝神帶領我順利完成學業，將一切榮耀歸與祂。

最後，我想要特別感謝我的爺爺，謝謝他相信我，在我選擇跨領域考研究所時，給我最大的支持與鼓勵，在這兩年常常關心我的研究與生活，謝謝你的教導與照顧，願能與在天國的你分享這份喜悅。

劉家蓁 謹誌

國立中興大學 資訊管理所 人工智慧與資料科學組

人工智慧與應用研究室

摘要

近年來，人工智慧的發展日漸成熟，在金融領域的應用也多元發展。由於美國經濟指標對制定政策、美國股市等多方面造成影響，因此，本研究以預測美國重要經濟領先指標為目標，利用現有的公開財經新聞資料，經語言預訓練模型進行資料預處理，並運用階層式時間序列模型對新聞資料提取有價值的資訊，搭配注意力機制計算不同新聞主題內容對預測經濟指標之重要性，將新聞主題之重要程度加入至特徵向量計算，最後以多任務學習架構預測多個經濟領先指標。經過實證研究，本研究對美國消費者物價指數和信心指數走勢預測的正確率達八成，對個人消費支出物價指數與零售銷售指數預測正確率超過七成，並且能在經濟指標公布前一周至一個月前，就能預測下一個月指標的走勢。

關鍵字：深度學習、時間序列模型、多任務學習、自然語言處理

Abstract

In recent years, the development and application of artificial intelligence has been mature, and there are a lot of successful AI applications in many aspects, also in Fintech. Economic leading indicators are key statistics based on economy, and have an impact on policies, stock market, etc. Therefore, we propose a hierarchical, multi-task learning approach that uses textual data to make predictions for four important economic leading indicators. First, we group news by topic. Then we summarize and extract information by time-series method and attention. Finally, we concatenate news information and historical data to predict economic leading indicators. This study has an 80% accuracy in predicting the trend of the U.S. Consumer Price Index (CPI) and Consumer Confidence Index(CCI), and over 70% accuracy in predicting the Personal Consumption Expenditure Price Index (PCEPI) and Retail Sales Index. The experimental results show that our approach successfully forecasts economic leading indicators. Besides, our approach is able to predict four indicators one week beforehand.

Key words: Deep Learning, Time Series Model, Multi-Task Learning, Natural Language Processing

目錄

摘要	i
Abstract.....	ii
目錄	iii
表目錄	v
圖目錄	vi
第一章 緒論	1
1.1 研究動機與背景	1
1.2 研究目的	2
1.3 論文組織與架構	3
第二章 文獻探討	4
2.1 深度學習應用於總體經濟指標預測	4
2.1.1 美國重要經濟指標	4
2.1.2 深度學習應用於總體經濟指標預測的相關文獻	6
2.2 應用深度學習模型於財經新聞資料	7
2.2.1 預訓練語言模型	7
2.2.2 應用深度學習模型於財經新聞的相關文獻	9
2.3 多任務學習(Multi-Task Learning)	10
2.3.1 多任務學習	10
2.3.2 多任務學習應用於金融領域之相關文獻	12
2.4 長短期記憶模型(Long Short Term Memory, LSTM)	12
2.5 注意力機制(Attention Mechanism).....	14
第三章 研究設計與實施	16
3.1 研究流程	16
3.2 資料集	17
3.2.1 新聞資料搜集	17

3.2.2	經濟指標資料	18
3.3	資料預處理	19
3.3.1	經濟指標標準化	19
3.3.2	文本資料預處理方法	20
3.4.	模型架構	21
3.4.1	整合新聞主題特徵	21
3.4.2	整合一段時間區間內的新聞特徵	23
3.4.3	結合歷史資料做預測	24
第四章	實驗結果與分析	25
4.1	實驗數據設置	25
4.2	實驗結果評估函數	25
4.3	分類模型評價指標	26
4.3.1	正確率(Accuracy)	26
4.3.2	精確率(Precision)和召回率(Recall).....	27
4.3.3	F-Score.....	27
4.3.4	Matthews Correlation Coefficient(MCC).....	27
4.4	實驗結果	28
4.4.1	模型實驗結果	28
4.4.2	與其他模型比較結果	36
第五章	結論與未來研究方向	39
5.1	結論	39
5.2	未來研究	39
參考文獻	41

表目錄

表 1 未標準化的經濟指標數值	20
表 2 標準化的經濟指標數值	20
表 3 混淆矩陣	26
表 4 有無使用新聞文本資料預測效果比較表	28
表 5 有無使用新聞主題對文本做初步分類預測效果比較表	29
表 6 注意力機制對經濟領先指標走勢預測效果比較表	30
表 7 歷史經濟指標資訊對經濟領先指標預測效果比較表	32
表 8 歷史經濟指標資訊對經濟領先指標預測效果比較表	33
表 9 不同新聞主題對經濟領先指標預測效果比較表	33
表 10 不同時間區間的新聞對經濟指標預測效果比較表	34
表 11 多任務學習架構對經濟指標預測效果比較表	35
表 12 其他模型預測經濟指標效果比較表	37

圖目錄

圖 1 BERT 訓練流程示意圖	8
圖 2 Hard Parameter Sharing 模型架構圖	11
圖 3 Soft Parameter Sharing 模型架構圖	12
圖 4 研究流程圖	16
圖 5 數據集標籤分佈直條圖	19
圖 6 BERT 文本預處理方法流程圖	21
圖 7 經濟指標預測模型	23
圖 8 注意力機制分析	31

第一章 緒論

本章節將以研究的動機及背景說明本研究之重要性，後以研究的主要目的與任務說明本研究之貢獻，最後針對論文的整體組織與架構作介紹。

1.1 研究動機與背景

近年來，人工智慧的發展應用都是人們所關注的議題，在不同的領域中已各自發展許多實際的應用科技，在金融科技方面，價格預測、投資組合管理、偵測詐欺嫌疑、量化交易、風險管理等多種與金融相關的深度學習應用模型。金融市場上，經濟領先指標的走向一直是政府、銀行、投資者都相當關注的數據[1]，因為經濟領先指標代表了各國總體經濟景氣狀況、反應了國家經濟將會步入衰退或是復甦的重要數據，用於判斷消費、出口貿易、投資，以及經濟成長等情況。

對政府而言，經濟領先指標走向有助於政策制定、以及擬定貨幣策略；而對於銀行中的分析師或是投資者而言，總體經濟指標對投資市場而言是揭露公開資訊的重要管道，被視為表達經濟狀況的重要指標，Kyereboah 等人研究[2]指出政府所公布的總體經濟指標新聞會影響到股票市場的報酬率和波動性，股票市場的投資參與者對於即將公布的總體經濟指標會有某種程度的預期，因此當實際公布的總體經濟指標數據與市場預期不符，投資參與者對經濟指標所表現出的資訊做不同的解讀，或是受到衝擊，導致投資參與者會用不同的回應方式，反映在市場上，假如經濟領先指標公布的結果與市場相符，有高度共識時，所引起的股價波動性不會很大；但是若發布結果出乎市場預期，交易量可能會有大幅度的增減，進而引起價格波動。經濟領先指標與金融市場有高度連動性，經濟指標數據結合股市脈動，精確掌握社會經濟環境的變化，進而研判未來。

美國自二次世界大戰後，力圖透過建立制度和規則，改變國際秩序，並透過全球性制度，拓展和維護美國的國家利益[3]。美國成功成為世界秩序中的領頭

羊，並在全球經濟上扮演主導的角色，通過其美元在貨幣市場上的地位，操控原料市場中黃金、石油與農產品的交易價格，讓美國經濟表現深深影響著全球經濟脈動與方向。因此，美國的領先經濟指標是全球金融機構、政府單位關注的重要指標，各方透過美國經濟分析局(BEA)的經貿報告，以及聯準會會議紀錄，從財經分析角度預測經濟成長的走勢。

近年來，深度學習在金融領域上有許多實際的應用，Xu 等人[4]提取新聞與社交媒體上的文字資訊的形成序列，並使用深度生成網路搭配長短期記憶模型推論未來的股價的表現，Sawhney 等人[5]透過強化學習模擬股票交易市場環境，並藉長短期記憶模型與注意力機制建立預測股票交易收益的模型，Agarwal 等人[6]則是利用圖神經網路連結企業之間的資訊，並結合新聞資料預測股價走勢；從以上金融相關預測的文獻回顧中可以發現大多數的研究都是應用深度學習於股票市場的預測任務，鮮少研究將深度學習應用於經濟指標的分析預測，並且在少數預測經濟指標之文獻中，尚未有使用新聞資料建立預測經濟指標之模型，在文獻探討的第一節將會有更進一步的說明。

因此，希望能利用深度學習中多任務學習(Multi-Task Learning)架構，透過時間序列模型對新聞文字資料進行特徵提取，了解目前市場經濟與社會情勢的狀態，建立能夠同時預測多個經濟領先指標的模型。

1.2 研究目的

此研究主要目的為利用多任務學習框架搭配深度學習方法，建置美國重要經濟領先指標預測模型，推測美國未來的經濟表現，以投資的角度而言，目的是領先預測市場價格變化，經濟指標預測值與實際值的落差對股價造成極大的影響，經濟指標在公佈之前，市場對不同經濟指標有其預測值，一旦公布真實的指標數據，產生低於預期或是高於預期的情況，而這些情況影響投資人的心理狀態以及對市場的預估，進而影響市場的變化；鑑於過去沒有針對經濟指標預測有完整的研究，因

此本研究希望能夠透過深度學習模型從新聞資料中提取出表現社會與經濟資訊的特徵並進行經濟領先指標預測，希望能夠透過本研究所建構之預測模型，使數據量少的經濟指標資料能夠得到好的預測效果，留意經濟指標的變化，提升對市場與總體經濟走向的掌握度。利用自然語言處理的特徵提取技術為基礎，建構適合進行經濟指標的預測模型並訓練，通過深度學習搭配時間序列模型提取細部特徵，建立模型框架，以達成下列目的：

1. 提出用於美國重要經濟領先指標走勢預測模型。
2. 使用多任務學習框架預測多個經濟領先指標。
3. 應用注意力機制於深度學習中，提升對重要經濟領先指標預測的效果。

1.3 論文組織與架構

本研究利用多任務學習進行美國重要經濟領先指標走勢預測模型建置共分成五大章節，第二章節為針對本研究資料與方法相關應用之研究文獻的比較討論；第三章節詳細介紹本研究使用之新聞文本資料、經濟指標資料、資料前處理流程以及及研究方法設計；第四章節進行模型實驗評估的結果呈現，第五章節是總結本研究結論以及說明未來可研究的方向。

第二章 文獻探討

本章節將介紹研究使用的方法，鑒於本研究利用深度學習中的多任務學習作為研究方法框架，將先於章節 2.1 說明領先經濟指標的定義與類別，並闡述先前學者利用深度學習模型應用於經濟指標走勢預測；章節 2.2 是財經新聞之語意分析以及過去研究如何應用新聞資料於金融領域的研究文獻概況；章節 2.3 是多任務學習框架之介紹；本研究的研究方法設計，主要是利用深度學習的時間序列模型提取新聞內容產生的時序資料中的語意特徵以及時間特徵，並透過注意力機制對重要資訊給予更多的關注，用以提升模型的預測準確率與解釋力，因此於章節 2.4 介紹時序資料的時間序列模型及其在金融分析領域的應用；在最後的章節 2.5 則是對注意力機制的基本原理運作與實際應用說明。

2.1 深度學習應用於總體經濟指標預測

本研究因使用總體經濟指標作為研究標的並利用深度學習模型進行分析預測，因此將先對經濟指標定義與說明，後說明深度學習模型應用於預測經濟指標的相關研究。

2.1.1 美國重要經濟指標

在經濟學中，不同經濟指標分別代表的涵義有所不同，且每個都能從數據與走勢中表現出總體經濟的狀態，主要分成物價、消費、經濟、金融、循環、勞動等六大項[7]。其中，受到市場高度關注且與美國股市標準普爾 500 指數走勢有顯著正相關的經濟指標[8]有以下七個指標：

- (1) 消費者物價指數(Consumer Price Index, CPI)：反映與社會相關的商品及勞動價

格統計出來的物價變動指標，為衡量通貨膨脹的主要指標。

- (2) 非農就業人數(Non-farm payrolls)：美國就業人口中非農業就業人數的數據，由美國的勞動部於每月月初的禮拜五公佈，其數據反映美國的經濟現況。非農就業人數影響美聯儲(The Federal Reserve)如何制定美元貨幣政策，經濟發展走勢的好壞會使美聯儲會選擇降息或是升息，進而影響美元的貨幣價值。
- (3) 零售銷售指數(Retail Sales Index, RSI)：零售銷售是統計美國各州，各個型態與不同規模的零售商每月業績所得，表現出美國人民在非耐久財和耐久財的消費規模。民間消費占美國 GDP 組成比重的七成，零售銷售額佔 GDP 中消費項目的四成，因此，零售銷售額幾乎佔美國 GDP 三成的表現。因此，零售銷售指數能夠反映消費市場的興衰。
- (4) 消費者信心指數(Consumer Confidence Index, CCI)：反應消費者對經濟社會環境的信心，藉著抽樣調查，反映消費者對目前與未來半年的經濟景氣、個人財務狀況與就業市場的感受和看法。
- (5) 個人消費支出物價指數(Personal Consumption Expenditures Price Index, PCEPI)：個人消費支出是由美國經濟分析局(BEA)於每月月底公佈，用來衡量過去一個月內個人消費支出加總。而 PCEPI 是經由個人消費支出計算再將它指數化的數據。個人消費支出可以表現目前社會經濟通貨膨脹的程度，因此個人消費支出也是美聯儲在制定貨幣政策時，主要參考的通貨膨脹指標。
- (6) 耐久財新訂單 (Durable Goods Orders)：美國普查局在每月月底提供的製造業出貨、存貨與訂單報告中非常重要的數據。耐久財是指使用年限超過 3 年的商品。由於非耐久財的需求量變化小，值得注意並會影響製造業市場表現的是耐久財，而製造商必須先有訂單才能規劃工廠生產時程，因此，耐久財新訂單的數據被視為是觀察製造業景氣表現的領先指標。

- (7) 初領失業金補助人數(Unemployment Insurance : Initial Claims)：美國勞動部在每週四公佈上一週勞動市場的表現，反映美國上一週失業或無薪假的人數變化，可用於預測失業率及非農就業人數，為勞動市場的領先指標。
- (8) 續領失業金補助人數(Unemployment Insurance : Continued Claims)：美國勞動部在每週四公佈上一週持續申請失業補助金的人數，意即在領過上上週的失業補助金仍未找到工作、連續申請補助金的勞工。續領失業金補助的數據較能表現真實就業市場的情況。

2.1.2 深度學習應用於總體經濟指標預測的相關文獻

隨著人工智慧領域蓬勃發展，深度學習廣泛應用於各個領域，在金融領域中也有諸多成功的應用[9]，而在經濟指標方面，例如 Zahara[10]所提出的利用當月 28 種食物價格來預測泗水的消費者物價指數，消費者物價指數內包含食物、醫藥保健、衣著、交通及通訊、居住、教養娛樂、雜項七大類的分類指數，此篇論文使用其中一種項目的資料來做預測；在 Cheng 等人的研究中[11]，使用中國河南省經濟指標數據建立深度神經網路模型，預測河南省南陽市的年失業率。

從相關文獻的表述中可以發現，目前預測經濟指標的模型[12]，都是使用結構性數據來預測，但如前面所述，因為是領先指標，表示數據中隱含的資訊領先於其他數據，因此能夠用來預測領先指標的數據種類稀少。除此之外，指標之間互有關聯，例如失業率越低，顯示就業情況良好，同時有利消費，零售消費指數升高。倘若不考慮關聯性的後果，就如同 Cheng 等人[11]使用多個經濟指標預測失業率，模型表現非常好，但用領先指標互相預測，指標之間相關性高，且發布時間間隔長，並無法達到真正提前得知經濟領先指標的走勢。因此，我們嘗試使用其他可以預測經濟走勢的資料來源。

2.2 應用深度學習模型於財經新聞資料

2.2.1 預訓練語言模型

大規模預訓練模型 (Pre-Training Model)，例如 BERT 和 GPT[13]，是自然語言處理中被廣泛使用的模型。由於大規模預訓練模型可以從海量未標記的數據中有效獲取有用的知識。透過將知識儲存到巨量的參數並在特定任務中訓練進行模型微調(fine-tune)，預訓練模型隱含豐富的知識以使各種下游任務受益，在廣泛的實驗和應用中已得到證明。現在自然語言處理領域的共識是採用大規模預訓練模型作為下游任務的支柱，而不是針對各種自然語言任務從頭進行模型訓練。以下介紹在自然語言領域中已被廣泛使用的 BERT (Bidirectional Encoder Representations from Transformers)[14]。

BERT 是以非監督學習方式訓練，運用龐大的無標籤文本資料成功訓練成具有優秀語言表示能力的預訓練模型，BERT 的架構是 Transformer 中的編碼器 (Encoder)[15]。Transformer 主要由編碼器(Encoder)和解碼器(Decoder)組成，編碼器是把每一個單詞輸入轉成向量，再經由解碼器讀取向量產生符合需求的資訊。BERT 作為一個編碼器，把輸入句子的每個字輸出成向量，我們就可以將文字的特徵向量進一步丟到其他子模型完成任務。

BERT 分成預訓練與微調兩個訓練階段。在預訓練的階段，研究團隊使用 BookCorpus 以及維基百科等無標記的文本資料，以非監督式學習進行模型訓練。在微調階段，面對不同性質的下游任務，使用該任務資料集對模型進一步訓練，對模型內隱藏層層數、調整學習率等，幫助在任務上有最佳的表現。

在預訓練階段，BERT 利用兩種方法進行模型預訓練，分別是 Masked Language Model(MLM)以及 Next Sentence Prediction(NSP)。首先，在克漏字測驗

任務，設計一個特別的 token 稱為[Mask]。輸入的詞彙中有 15%機率會以[MASK]的 token 隨機替換，模型根據輸入的上下文含義，預測[Mask]可能代表的詞。過往使用於語言表示模型都是有順序地來預測下一個詞，沒有雙向理解上下文的能力，使用克漏字訓練可以讓模型依上下文來預測被遮蓋的字詞，使模型能夠學習如何根據上下文推測出語意。而 NSP 任務，則是希望 BERT 能夠具有理解句子之間關係的能力。任務的主要內容是給予上下相連的兩個句子，訓練模型判斷兩句相連是否符合邏輯。為了進行 NSP 任務，針對輸入的句子需經過一些預處理，在兩句話的中間加入[SEP] token，讓模型知道兩句話的交界處，以及放入[CLS]的 token 在句子起始位置，讓模型知道此位置要進行分類的動作，並將[CLS]的輸出接到線性二元分類器，由該分類器輸出這兩個句子是否相連。

BERT 的資料集包含 8 億個在 BooksCorpus 上的單詞以及 25 億個在維基百科的單詞等未經標籤的資料進行預訓練，海量的訓練資料耗費大量的運算資源。但 Google 開源釋出已訓練好的模型，因此，我們能夠下載已預訓練好的 BERT 模型，將我們蒐集的文本轉成向量。將文本特徵向量應用在我們設計的任務上，模型無需花費訓練取得文本特徵的成本。

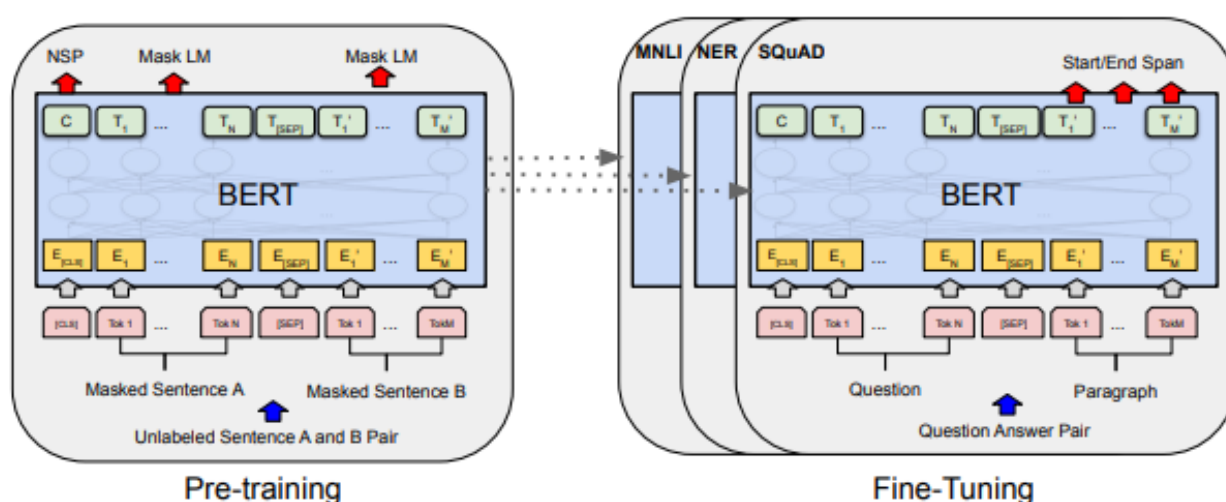


圖 1 BERT 訓練流程示意圖

從上所述可以知道 BERT 有以下優勢：(1)無監督模型的數據量大，因為不需要人工標記數據，任何文本資料都能成為用來預訓練的資料集。(2) BERT 能夠從預訓練的任務中學會語法結構、解讀上下文語意等語言學中的重要知識，透過微調參數有效訓練下游的子任務，並優化其表現。(3) 減少處理不同自然語言處理任務所需要的訓練成本。

BERT 成功應用在自然語言處理中的任務，例如 Yang 等人[16]利用預訓練好的 BERT 模型結合 Anserini 資訊檢索工具，訓練問答系統並成功在 SQuAD 閱讀理解與問答的任務上超越 SOTA(State-of-the-art)模型；Hoang 等人[17]應用 BERT 預訓練模型以及帶有附加生成文本的方法，解決 SemEval2015 與 SemEval2016 情緒分析數據集的子任務，相較之前的方法效果有顯著地提升。從過去研究可以知道 BERT 在自然語言處理的領域上應用已經相當成熟，在各項自然語言處理的分支任務都有極其優越的表現。

2.2.2 應用深度學習模型於財經新聞的相關文獻

過去的研究[18][19]使用新聞來做特徵擷取的資料，是因為新聞記事講求真實、公開、準確以及即時性，因此新聞能夠表現出經濟、社會、政治、文化所發生的事件與狀態。隨著網路科技的發展越來越先進，數位新聞的閱讀比例逐年增加[20]，因此網路新聞平台提供的新聞內容越來越完整且專業，從網路上能收集到的新聞資料也越加豐富。

在金融領域方面，先前學者的研究有許多利用財經新聞資料預測匯率與股價走勢[21]，例如 Chen 等人[22]提出一個基於 BERT 模型應用，從不同主題的新聞中提取出關鍵的新聞，並利用時間序列模型總結大量的財經新聞預測外匯走勢，成功預測世界流通的六大幣別匯率；Ding 等人[23]將企業與企業之間的關係轉換成知識圖譜，將圖譜中的額外訊息結合新聞資料中的事件，預測股票市場走勢。從上述

提及的研究中，可以表現出財經新聞能夠有效地提取金融市場、經濟事件的特徵，並成功應用於預測重要的投資標的。因此，本研究將透過深度學習方法利用財經新聞預測美國重要的經濟指標。

2.3 多任務學習(Multi-Task Learning)

2.3.1 多任務學習

傳統機器學習方法主要是單任務的模式進行學習，對於複雜的任務則透過多個獨立的任務進行學習，並學習將不同任務結果進行整合得到最終的結果。在1993年 Caruana[24]提出多任務學習的概念，多任務學習的關鍵在於尋找任務之間的關係，如果任務之間的關係衡量恰當，那不同任務之間就能相互提供額外的有用訊息，利用這些額外的訊息，可以訓練出表現更好的模型。多任務學習的優勢是在模型效果不下降的情況下提升多任務預測效率。由於可以共享部分神經網路結構，多任務學習的模型的參數量相較於多個單任務的模型參數量總和低，意旨在於實時預測多任務場景下模型的效率會更高。除此之外，針對標籤樣本少的數據集，單任務學習往往不能獲得足夠的訊息，多任務學習能補足樣本少的缺點，從其他任務中獲得有用訊息，提升任務學習的效果與模型泛化能力。

多任務學習是遷移學習的分支，因為多任務學習的核心概念在於共享參數，與遷移學習的概念相同。根據 Ruder[25]多任務學習的概述，多任務學習參數共享的形式主要分成兩種，以下將針對此兩種方法介紹。

2.3.1.1 Hard Parameter Sharing

Hard Parameter Sharing 形式如下圖所示，在多個任務學習過程中直接共享部分神經網路的參數，不同任務的模型共同使用底層的共享隱藏層(Shared Layers)，用來提取共通特徵，再根據不同任務設計符合其任務的輸出層(Task-specific layers)，降低模型過擬合(Overfitting)風險，同時學習的任務越多，模型越難找到

一個能表示所有任務的特徵向量，而過擬合於特定任務的可能性就越低，Hard Parameter Sharing 是多任務學習的參數共享中較為簡單且常見的方法。

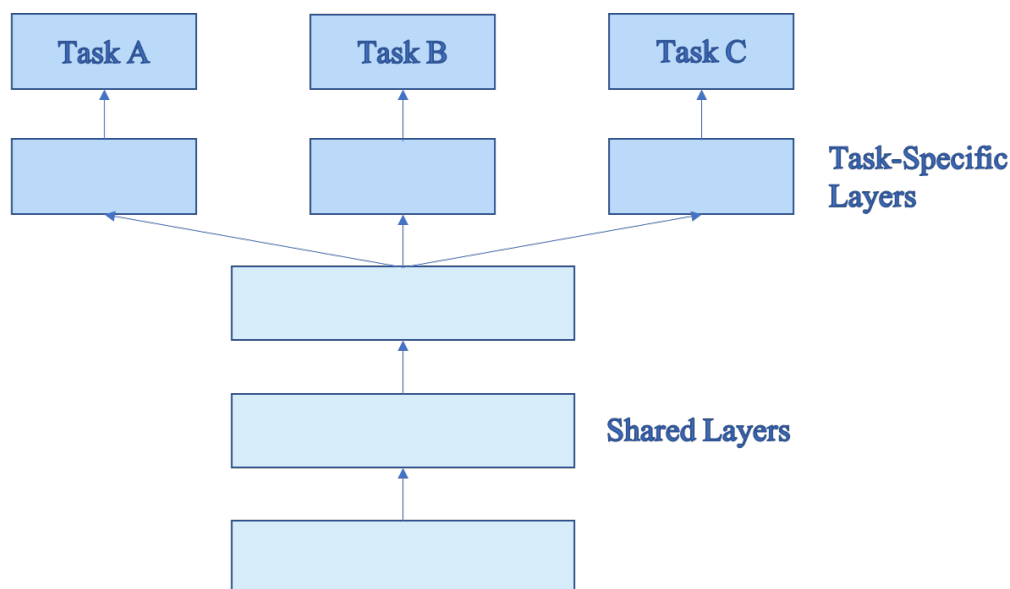


圖 2 Hard Parameter Sharing 模型架構圖

2.3.1.2 Soft Parameter Sharing

Soft Parameter Sharing 形式如下圖示。每個任務分別有自己的神經網路參數，但是會對多任務的網絡參數之間加約束，使參數彼此相似，約束指標如 L2 Norm[26]，或是利用複製其他任務的隱藏狀態獲得訊息，以及使用注意力機制主動選取有用訊息[27]。Soft Parameter Sharing 在實際場景中較不常見，用於任務之間關聯性較弱的場景。

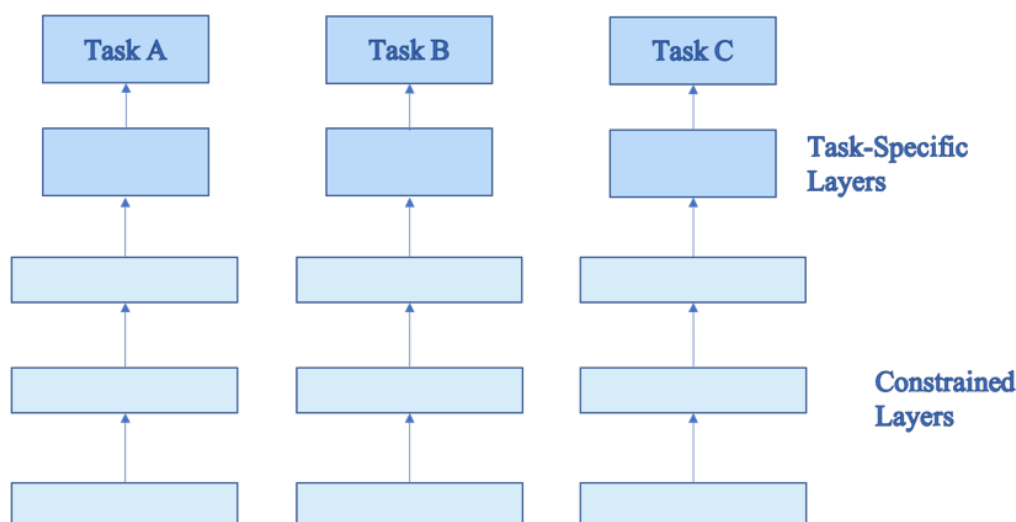


圖 3 Soft Parameter Sharing 模型架構圖

2.3.2 多任務學習應用於金融領域之相關文獻

多任務學習在股票價格、收益、波動率議題上皆有成功的應用，例如在多任務學習概念提出三年後，Ghosn[28]利用多任務學習架構搭配人工神經網路模型預測股票的收益，在時間序列模型被提出後，結合時間序列資料與多任務學習架構應用於股票交易[30]，以及利用公司法說會文字紀錄與語音建立財務風險與股價的多任務預測模型[31]。從文獻回顧中可以發現，過去沒有應用多任務學習於經濟指標相關的研究，而經濟指標之間的相關性符合多任務學習架構建立之要點。因此，本研究將使用多任務學習架構，建立預測經濟領先指標模型。

2.4 長短期記憶模型(Long Short Term Memory, LSTM)

長短期記憶模型是一種遞迴神經網路(Recurrent Neural Network)，遞迴神經網路是指一種有記憶的神經網路，用於處理時間序列型資料，遞迴神經網路特點是想理解一個時間段內的資料時，能夠將一時間段的資料根據時間點的不同，分段輸入至神經網路中，神經網路被循環使用，並在循環過程中，記憶分段輸入的資訊，最後整合記憶產生能代表整段輸入的資訊。在 RNN 的模型訓練情況中，如

果輸入的資料長度太長，模型會記不住前面的資訊，進而出現梯度消失的問題，造成後續預測錯誤，因此發展出有特別記憶機制的長短期記憶模型(Long Short Term Memory, LSTM)[32]，能夠比基礎的遞迴神經網路記住更多先前的資訊，進而得到較好的預測準確率。

如上所述，為了解決模型訓練時梯度消失的問題，LSTM 利用兩種傳輸狀態紀錄資訊，以及三種閘門(Gates)決定是否儲存輸入資料的特徵或是刪除先前留下來的特徵資訊。兩種傳輸狀態分別為單元狀態(Cell State)與隱藏狀態(Hidden State)，前者存儲的是長期記憶，後者則是當前時間點的短期記憶。而 LSTM 的運作流程是先利用 Forget Gate 選擇性遺忘部分的長期記憶，再用 Input Gate 對當前時間點的輸入選擇要新增至長期記憶的內容，有了要遺忘的部分以及新增的部分後就可更新長期記憶，也就是單元狀態；在短期記憶的更新，先將當前時間點的輸入與上一時間點的隱藏狀態結合後，利用 Output Gate 選擇多少內容作為當前時間點的短期記憶，再與長期記憶合併成為當前時間點的輸出。

另外，長短期記憶模型讀取資料的方向也可以是反向的，例如在看一段句子時，不只可以從句首讀到句尾，也可以從句尾到句首的方向閱讀內容，這樣雙向的理解較全面，此模型稱為雙向長短期記憶模型(Bi-directional Long Short-Term Memory, Bi-LSTM)，由兩個 LSTM 疊加在一起組成，差別在於輸入資料的順序，一者是由前往後，一者由後往前輸入，再將兩個 LSTM 的輸出結果拼接成為最終的輸出，整合讓模型對上下文含義有更完整的理解。

透過以上的說明，可以發現 LSTM 能有效地將不需要的資訊遺忘，留下有用的資訊進行訓練。因此，在各領域被廣泛地使用且得到很好的效果，例如在金融領域方面，利用 LSTM 建立時間序列模型預測股市指數的走勢[33]；利用 LSTM 加上注意力機制，幫助模型有時序性地整合一個時間段的新聞、或是社群媒體上的文本資料，成為可代表一時間段訊息的特徵向量，在模型預測上有很好的表現

[34][35][36]；加上時間變量的 LSTM[37]，在 LSTM 的計算中加入資料之間的時間差，讓 LSTM 輸出的特徵向量可以更完整表現出文字資料中隱含的資訊。

2.5 注意力機制(Attention Mechanism)

注意力機制想要解決的問題是源自於 Neural Machine Translation(NMT)任務，而解開 NMT 任務的模型架構是 Sequence to Sequence(Seq2Seq)，Seq2Seq 是一種用序列預測序列的 Encoder-Decoder 架構[38]。Seq2Seq 中的 Encoder 實作方式是藉由 LSTM 將序列型的文本資料轉換成以特徵向量表示，而這特徵向量是取最後一個時間點的輸入產生的隱藏狀態(Hidden State)作為整串輸入的表示向量(Context Vector)。然而，問題在於 Encoder 與 Decoder 之間，只有一個表示向量串接，意即不論給予 Encoder 輸入長度的長短，都會被固定維度的表示向量限制。除此之外，LSTM 在輸入的長度變長，越早輸入的資料仍避免不了被遺忘或是被新資訊取代。

注意力機制解決了這個問題，注意力機制的核心思想是計算屬於每個時間點的表示向量，透過此機制可以提升處理輸入的準確性，並且提供對資訊的可解釋性。注意力機制主要分為兩種類型，分別是 Global Attention 與 Local Attention，Global Attention 是將注意力放在全部的資料，如此有一個缺點是，每一個時間點產生的隱藏狀態都需要參與注意力計算，當輸入長度較長時，計算的成本較大，效率低；而 Local Attention 只將注意力放在少部分的輸入，減少計算，和 Global Attention 各有優缺，而在實際應用中，Global Attention 應用更為普遍。

Global Attention 計算分成三個步驟，第一個步驟是為每一個時間點的隱藏狀態計算對目標表示向量的重要分數，此步驟的輸出一個與輸入長度一樣長的向量，計算的方式有以下幾種：

$$a(s_{i-1}, h_j) = \begin{cases} s_{i-1}^T h_j & \text{dot} \\ s_{i-1}^T W_\alpha h_j & \text{general} \\ v_\alpha^T \tanh(W_\alpha [s_{i-1}; h_j]) & \text{concat} \end{cases} \quad (2-1)$$

第二步是計算注意力的權重，分數的向量經由 Softmax 函數歸一化，凸顯其中的最大值並抑制低於最大值的其他分數。

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_z} \exp(e_{ik})} \quad (2-2)$$

第三步是計算目標表示向量，將各時間點的隱藏狀態乘上其權重加總求和後得到輸出的目標。

$$c_i = \sum_{j=1}^{T_x} \alpha_{ij} h_j \quad (2-3)$$

除了 Bahdanau 等人提出的基礎注意力機制[38]與 Luong 等人在注意力機制基礎上進行權重計算的變化[39]，還有隨著任務不同進而發展出不同的注意力機制，運用在不同領域的任務中，如多變量時序預測[40]，Raehyun Kim 等人[41]應用階層式圖注意力機制(Graph Attention Networks)，搭配公司之間的關係結構圖，預測個股股價走勢，Jintao Liu 等人[42]提出利用多個自注意力機制組成的 Transformer 的膠囊網絡提取社交媒體文本中的語意結構，預測個股股價走勢。由此可見，注意力機制可以靈活地變化且使用在不同類型的任務上。

由上述可知注意力機制可以靈活地變化且使用在不同類型的任務上，因此，本研究將利用多任務學習搭配注意力機制，使模型能專注在重要資訊上學習且使預測結果更富有可解釋力。

第三章 研究設計與實施

在預測美國重要經濟指標的分析中，需要解決長時序、高維度特徵與多個預測目標的問題。因此本研究針對美國財經相關新聞，透過預訓練語言模型將文字轉換成向量，成為適合進入分析模型之資料格式，使用多任務學習搭配時間序列模型有效地提取特徵並解決多個預測目標的問題，最終得到重要經濟指標預測之結果；以上所述，即為利用多任務學習進行預測美國重要經濟領先指標方法之框架。本章節將說明研究步驟的設計與實作，並介紹模型系統架構。

3.1 研究流程

圖 4 為本研究提出的一個基於財經新聞標題為主要資料的研究流程，本研究想透過分析一段時間內的財經相關新聞來了解目前金融市場狀態並預測與其有相關的指標之未來表現。第一步是為了將非結構化的資料轉換成結構化資訊而進行的資料預處理，將新聞文字資料經過預訓練的語言模型進行向量轉換並使用金融情緒分類模型過濾資料，保留有正向與負向金融情緒的新聞標題內容，作為適合預測模型分析之新聞資料集，第二步是運用自然語言處理的預訓練模型加上時間序列模型建立領先經濟指標預測方法，先將新聞向量依固定時間間隔整合成表示矩陣，且使用時間序列模型加上注意力機制，針對多個經濟指標進行預測模型訓練，最後輸出經濟指標走勢的預測結果。將於下一節介紹本研究所使用的資料集，於 3.3 節說明資料預處理的過程，並於 3.4 節詳細介紹預測經濟指標的模型架構。

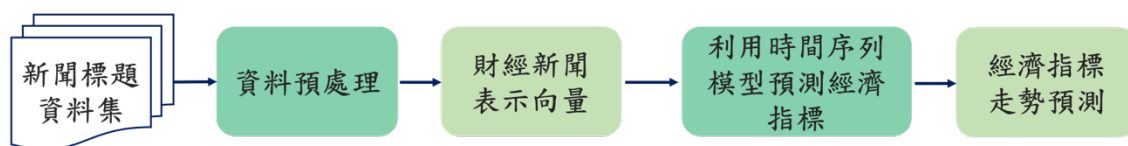


圖 4 研究流程圖

3.2 資料集

為預測美國重要經濟指標，主要搜集美國財經新聞網站 Forbes 和 CNBC 的歷史新聞資料，由於新聞網站的歷史資料最早能追溯到 1997 年 6 月，因此搜集的資料集從 1997 年 6 月到 2021 年 7 月，共 24 年的資料。而經濟指標資料集主要從 FRED(Federal Reserve Economic Data)線上數據庫[43]取得。此章節分為新聞資料與經濟指標資料進行說明。

3.2.1 新聞資料搜集

本研究的資料是透過網路爬蟲 (Web Scraping) 的方式將有關財經的新聞擷取下來，由於需要透過關鍵字搜尋才能夠瀏覽歷史新聞，而 Forbes 和 CNBC 的新聞類別名稱不同，若用個別的類別名稱做關鍵字搜尋，新聞類別將不一致，因此，參照 Deli Chen 等人[44]的研究中所參考的路透社新聞網中的新聞類別作為關鍵字，分別是 Fed, Currencies, Consumer, Finance, China, Asia, Autos, Commodity, Energy, Environment, Europe, Defense, Media, Sustainable, Middle East, Telecommunications, Politics, United Kindom, Retail, Transportation, United States, Aerospace 共 22 個關鍵字，在財經新聞網站中擷取與關鍵字相關的新聞。

Scrapy 是一個網頁爬蟲框架，針對網頁爬蟲具有完善的專案架構及執行流程，在維護上也較容易，另外，在執行效率方面，Scrapy 框架的表現也非常好，適用於大型的爬蟲專案。由於 Forbes 和 CNBC 網站都有反爬蟲機制，為克服反爬蟲機制，採用代理伺服器處理。我們先從透過爬蟲抓取國外的第三方伺服器列表[45]，再使用第三方主機代為發送請求，因此對於網站方而言，他收到的請求是來自於第三方的，避免網站認為是同一請求方，進而免除了封鎖爬蟲的運作。

從 1997 年 6 月至 2021 年 7 月，刪除重複的新聞標題，兩個網站一共爬取了 1,321,777 篇新聞。

3.2.2 經濟指標資料

從 FRED 數據庫下載指標歷史資料，並根據不同指標的特性與代表的意義，設定與其相符的標籤。在與美國股市標準普爾 500 指數走勢有顯著相關的七個經濟領先指標中，我們將預測其中四個重要經濟領先指標，分別是消費者物價指數(CPI)、消費者信心指數(CCI)、零售銷售指數(Retails)與個人消費支出物價指數(PCEPI)，因為這四個指標的更新頻率相同，皆為每月更新一次，從 1997 年 6 月至 2021 年 7 月，一共有 290 筆樣本數。

從聯準會會議紀錄[46]可以發現，聯準會對通膨的控制就是以消費者物價指數年增率在兩個百分比內為目標時，以及會將個人消費支出物價指數年增率是否大於兩個百分比作為升息的參考基準點。因此，我們將消費者物價指數與個人消費支出物價指數年增率大於兩個百分比的樣本給予標籤 1，否則為 0。

在消費者信心指數方面，由於在經濟學上消費者信心指數常以月增比(相對上個月增加比率)來比較，因此我們將以消費者信心指數的月增率為預測目標，若月增率大於 0，則標籤為 1，否則為 0。零售銷售指數的表現一般會比較去年同期時的增加比率(相對於去年同個月份的增加率)，由於零售銷售指數沒有像消費者物價指數以及個人消費支出物價指數一樣有經濟意義上的基準點，因此我們以零售銷售指數年增率的中位數作為分割資料的標準，年增率大於中位數則標籤為 1，否則為 0。以經濟學角度給予指標適當的標籤後，資料集整體分佈如下圖所示：

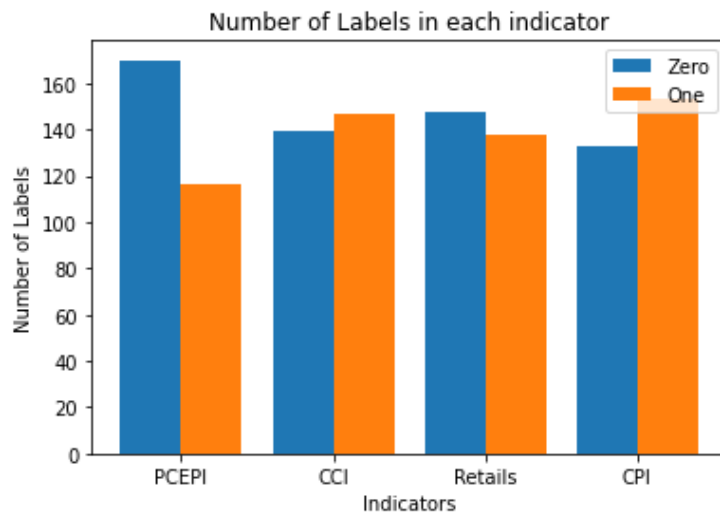


圖 5 數據集標籤分佈直條圖

3.3 資料預處理

由於我們期待使用財經新聞預測經濟指標，而新聞文本資料為非結構化資料，並非能直接投入深度模型計算的資料格式，因此，需要透過本章節的資料預處理方法，利用語言模型進行文本資料轉換與計算，以及將經濟指標原始數據進行標準化，提高數值分布之範圍，再進入分析模型中。

3.3.1 經濟指標標準化

雖然經濟指標的指數升降幅度看似微小，實際上影響非常大，例如經濟成長比率下降 1%，台灣 GDP 約 19 兆台幣，差 1 個百分比就差了 1900 億元。然而，在深度學習模型中，數值間極小的差距讓模型誤以為沒有太多差別，造成模型無法理解經濟數值的小差距代表的意義，進而無法成功訓練模型。因此，將經濟指標數據標準化，提升模型收斂度與精準度。

DATE	PCEPI	CCI	Retails	CPI
2001/4/1	0.023735	-0.0612	0.036334	0.032691
2001/5/1	0.025567	-0.0528	0.036625	0.036152
2001/6/1	0.023917	-0.0596	0.023622	0.032483
2001/7/1	0.021019	-0.1262	0.02355	0.027199

2001/8/1	0.020957	-0.234	0.02821	0.027199
----------	----------	--------	---------	----------

表 1 未標準化的經濟指標數值

DATE	PCEPI	CCI	Retails	CPI
2001/4/1	0.574225	-0.362043	-0.011372	0.874547
2001/5/1	0.753671	-0.308491	-0.004370	1.144666
2001/6/1	0.592040	-0.351843	-0.317650	0.858265
2001/7/1	0.308198	-0.776438	-0.319380	0.445835
2001/8/1	0.302068	-1.4637	-0.207117	0.445835

表 2 標準化的經濟指標數值

3.3.2 文本資料預處理方法

欲使用 BERT 預訓練模型，需先將文字轉換成符合 BERT 輸入的特殊格式，以下為文字預處理步驟：首先，先去除新聞標題中的標點符號，接下來將文字分詞(tokenization)，分詞是將句子分解為單詞，通過分析單詞的順序來幫助解釋文本的含義，這些標記有助於理解上下文或開發自然語言模型。分詞目前有許多不同的方法可以應用，我們使用 BERT 預訓練模型中訓練好的套件，將新聞標題分詞，並在句頭與句尾加上特殊[CLS]和 [SEP] token，以符合 BERT 模型輸入的格式，BERT 預訓練模型在被訓練完後，每個分詞都有一個屬於自己的 ID，所以我們能透過 BERT 套件將新聞標題切割後的分詞轉換成相對應的 ID。由 ID 組成的句子成為進入 BERT 模型的輸入，輸出代表每一個分詞含義的向量，而在 Choi 等人[47]的研究中，發現平均分詞向量在文本相似的任務表現不錯，以及在 Sawhney[5]等人研究裡也是使用平均分詞向量作為新聞表示向量。因此，我們使用此方法輸出新聞表示向量。

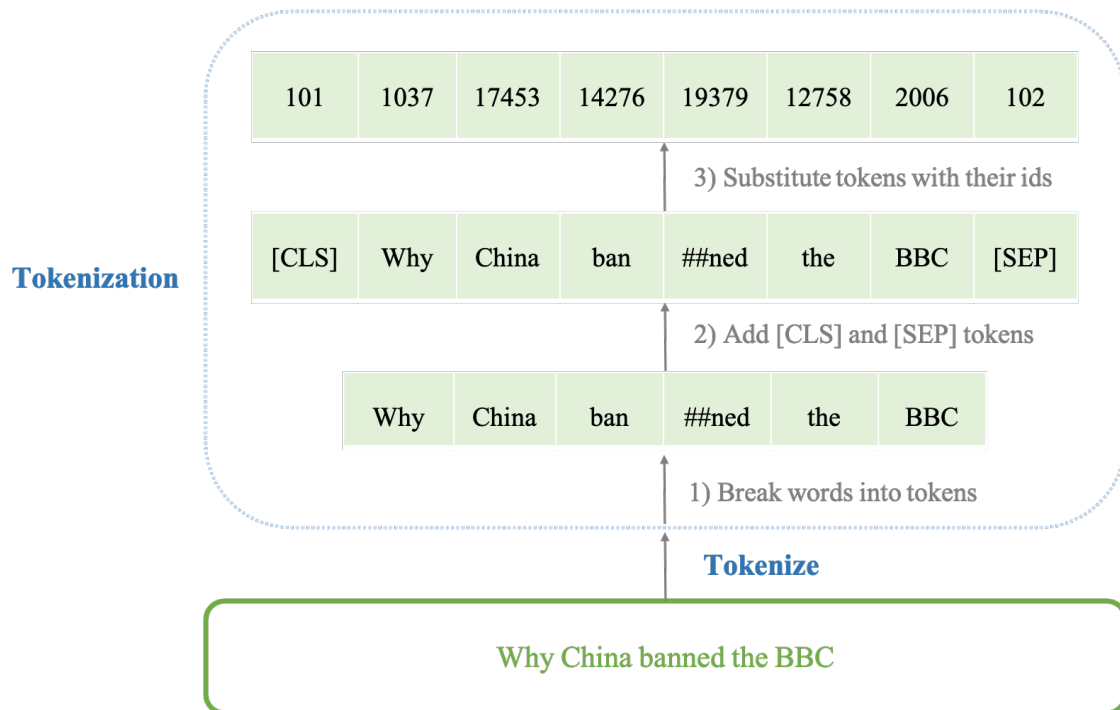


圖 6 BERT 文本預處理方法流程圖

3.4. 模型架構

本研究使用時間序列模型作特徵擷取，模型架構如圖，包含 LSTM 專門處理含有時序特徵資料的時間序列模型，再加入注意力機制的計算不同新聞主題代表的重要性，最後加上前一個月的經濟指標表現特徵，輸出下個月對經濟指標指數的預測。以下針對模型架構與流程進行講解。

3.4.1 整合新聞主題特徵

本研究使用時間序列模型擷取並整合新聞標題的特徵資訊。以預測日往回推 T 個天數，每一天都有 n 個主題的新聞，每個新聞主題內有 m 則新聞標題，因此每一天會有 $n * m$ 則新聞。經過資料預處理的 BERT 將新聞文字內容轉成向量後，從整合每一天的新聞開始。將一天內各個主題的新聞表示向量輸入至時間序列模型，如圖 7 中下方多個新聞標題表示矩陣(News Embedding)作為輸入資料，而其中一個二維的矩陣代表一個新聞類別所有新聞標題的表示向量，維度是 $m *$

768，經過第一個藍色代表的時間序列編碼器，輸出 n 個分別代表各個新聞主題內容的特徵表示向量(Topic Embedding)，因新聞主題數量多，藉由第二個時間序列編碼器整合一天內多個新聞主題的內容，並結合 Topic Attention 捕捉不同新聞主題的重要程度，輸出成代表一天的特徵表示向量。經由神經網路進行新聞標題間時間關係以及文字訊息的特徵捕捉，為避免造成過擬合(Overfitting)的問題，本研究設定兩層的時間循環網路及 128 個隱藏層神經元進行時間序列模型的建置。

本研究在注意力機制的設計是使用 Global Attention 來關注 T 個新聞主題對預測經濟指標數值的重要程度，注意力機制是透過時序模型 T 個新聞主題的隱藏層輸出與最後一個隱藏狀態計算出各新聞主題的資訊對預測經濟指標數值的重要程度，權重計算方式為在 2.5 節中提到的 general 算法，圖中 β_T 代表經過計算後的注意權重，計算如公式(3-1)，

$$e_{ij} = a(\text{Output}, h_T) \quad (3-1)$$

公式(3-2)將各時間點的注意權重進行標準化，使其相加起來等於 1，

$$\alpha_{ij} = \frac{\exp(e_{ij})}{\sum_{k=1}^{T_z} \exp(e_{ik})} \quad (3-2)$$

公式(3-3)再將時序模型的輸出與權重相乘得出加權過後的輸出向量 c_i 。

$$c_i = \sum_{j=1}^{T_x} \text{Output}^\top \alpha_{ij} \quad (3-3)$$

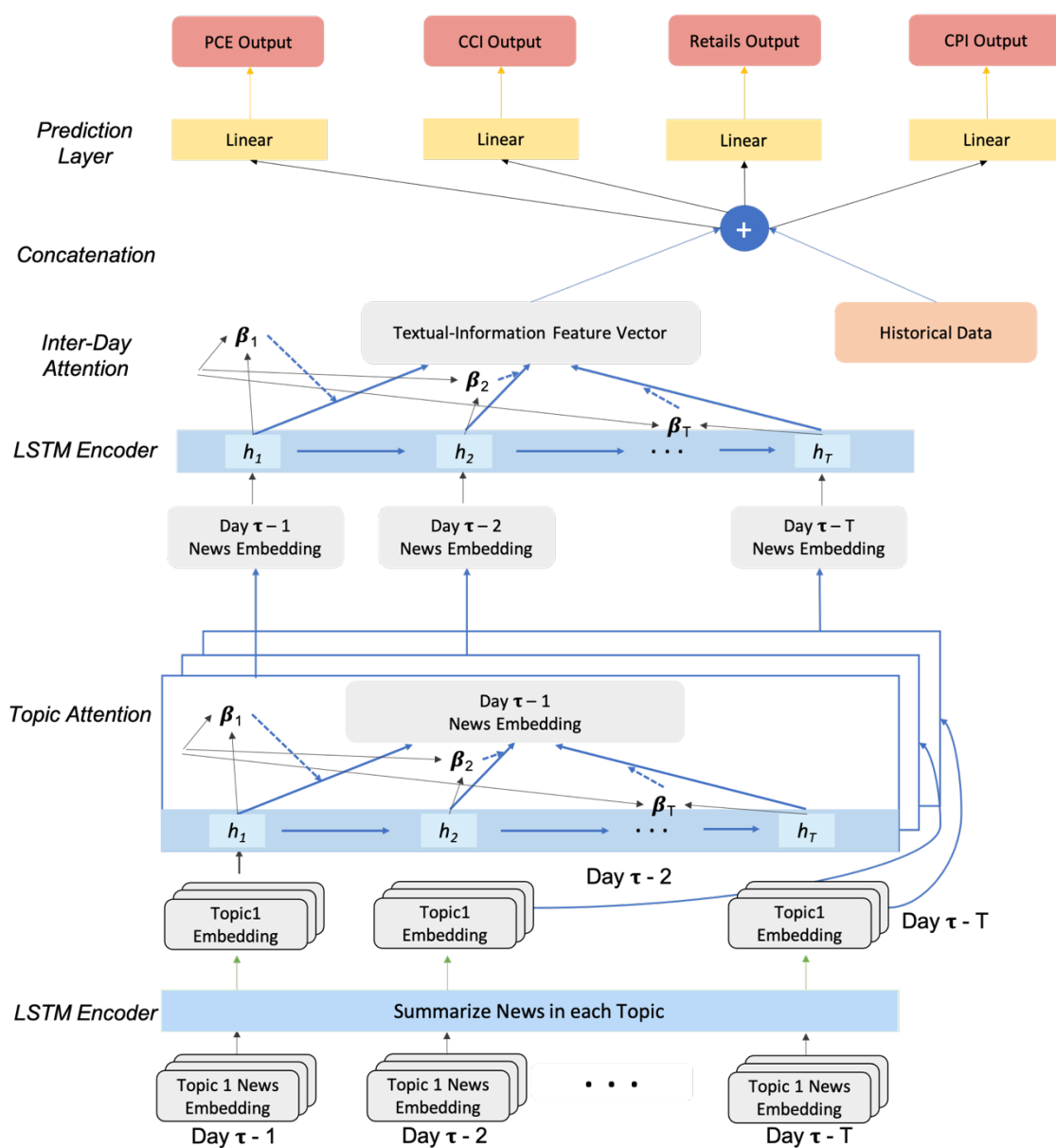


圖 7 經濟指標預測模型

3.4.2 整合一段時間區間內的新聞特徵

在前一小節說明時間序列編碼器輸出各個新聞主題特徵向量後，接著進入第三個時間序列編碼器，為的是整合不同日期的特徵向量，並結合注意力機制計算，即為圖示中的 Inter-Day Attention，此處使用的計算方法與前一小節相同，讓

模型訓練計算不同時間的重要性，並將時間重要性程度的特徵融入最終新聞資訊表示向量(Textual Information Feature Vector)。

3.4.3 結合歷史資料做預測

在取得整合新聞主題後的特徵資訊後，經過一層全連接層，再結合前一個時間點的歷史資料，也就是前一個月的領先經濟指標，其標準化後的數值，再將合併後的特徵向量放入預測各領先經濟指標的全連接層，輸出經濟指標數值的預測結果。

第四章 實驗結果與分析

本章節將探討應用財經新聞預測領先經濟指標的效能，將先描述訓練數據設置與實驗結果評估定義，後針對效能評估進行說明。

4.1 實驗數據設置

重要經濟領先指標預測模型是使用 1997 年 6 月至 2021 年 7 月共 290 筆經濟指標數據進行訓練，20%樣本數作為測試資料，其餘 80%的訓練資料在每個 epoch 訓練隨機抽出 10%作為驗證資料，用以評估模型的效能，由於經濟指標數據資料量稀少，模型訓練的 batch size 設定為 1，epoch 設定為 50，並設定當驗證集損失函數小於訓練集損失函數兩次時，模型訓練提早結束(Early stopping)，學習率為 $5e-3$ ，優化器使用 Adam。

4.2 實驗結果評估函數

深度學習的損失函數(Loss Function)是用來評估模型的預測值與真實值之間的差異程度，因此，亦成為深度學習模型訓練過程要優化的目標函數，深度學習模型訓練的過程就是使損失函數最小化的過程，當損失函數越小代表預測值與真實值的差異越小，意即模型的預測效果越好。

損失函數可以大致分為分類與回歸兩個面向，本研究的模型主要任務是預測經濟指標數據的數值，所以是屬於回歸面向。本研究使用在回歸預測中最常被使用的均方誤差(Mean square error, MSE)作為訓練模型的損失函數，計算如公式(4-1)，

$$MSE = \frac{1}{mn} \sum_{j=1}^m \sum_{k=1}^n (Y_{j,k} - \hat{Y}_{j,k})^2 \quad (4-1)$$

其計算方法是求預測值(Y_i)與真實值(\hat{Y}_i)之間差異的均方值，公式中的平方是為了避免相減後的數值有負號而影響累加後的損失值， m 代表預測的經濟指標數， n 代表要預測的目標數目，均方的計算是因為模型訓練時會因在訓練時每一次訓練多個樣本，因此必須計算所有樣本的平均值。透過公式(4-1)可以知道計算出來的 MSE 越趨近 0 為最好的預測效果。而本研究在結果的評估函數亦直接使用模型訓練時的損失函數值進行預測效果的評估。

4.3 分類模型評價指標

評價指標用來判斷分類系統的優劣，是設計深度學習系統一個重要的環節。分類器的分類情形可以用表 3 混淆矩陣(Confusion Matrix)表示，而透過混淆矩陣設計出多種評價分類器表現的指標如正確率(Accuracy)指標、精確率(Precision)、召回率(Recall)、F1-Score，以及專門發展用於二分類問題的 Matthews Correlation Coefficient(MCC)指標。

	實際為真	實際為假
預測為真	TP(True Positive)	FP(False Positive)
預測為假	FN(False Negative)	TN(True Negative)

表 3 混淆矩陣

4.3.1 正確率(Accuracy)

正確率表示模型預測正確數量占所有樣本的比例，正確率考量了所有情況，但是當樣本數據有數量不平衡的情況時，此項指標就失去其表示能力，不具參考價值。

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \quad (4-2)$$

4.3.2 精確率(Precision)和召回率(Recall)

精確率和召回率都是專注在被預測為真的資料，但注重的點有所不同。從公式可以發現，精確率在意的的是在預測為真的結果中，真正為真的比例；而召回率則是注重在實際為真的資料中被預測出來的比例。精確率重視有沒有把實際為假的樣本誤判為真，召回率可以看出模型預測為真的結果有沒有盡可能觸及到實際為真的樣本。

$$\text{Precision} = \frac{TP}{TP+FP} \quad (4-3)$$

$$\text{Recall} = \frac{TP}{TP+FN} \quad (4-4)$$

4.3.3 F-Score

F-Score，亦稱為 F-Measure，結合精確率和召回率計算的指標，透過加權平均同時考慮兩個指標表現，平衡反應模型的精確度。

$$F - \text{Score} = \frac{(1+\beta^2)\text{Precision} \times \text{recall}}{\beta^2 \text{Precision} + \text{recall}} \quad (4-5)$$

β 是自行定義的參數，當 β 為 1 時，表示精確率和召回率權重相同，亦是所謂的 F1-Score，表示同樣重視精確率和召回率，F1-Score 可說是兩指標的平均，公式如下：

$$F - \text{Score} = 2 \frac{\text{Precision} \times \text{recall}}{\text{Precision} + \text{recall}} \quad (4-6)$$

F1-Score 理想數值是趨近於 1，表示精確率和召回率都有好的表現。

4.3.4 Matthews Correlation Coefficient(MCC)

MCC 即為馬修斯相關係數，用以測量二分類的分類性能指標，它綜合地考量了混淆矩陣中的四個基礎評價指標。馬修斯相關係數描述的是實際樣本與預測樣本之間的相關係數，範圍在-1 到 1 之間，即使在兩類別的樣本數量不平衡時，也可以使用 MCC 計算[48]。

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (4-7)$$

4.4 實驗結果

本節將針對研究所提出的方法對於經濟領先指標預測效果進行呈現與討論，將根據不同新聞主題進行經濟領先指標預測比較，以了解新聞內容對預測的影響；以及取不同時間段的新聞內容作為輸入，比較新聞時間點對預測經濟領先指標的影響；比較多任務學習架構是否發揮其作用；以及加入前一個月時間點的經濟指標數據對實驗結果的影響，及與其它處理財經相關資訊預測經濟數據模型的外部實驗比較。

4.4.1 模型實驗結果

此小節將探討本研究方法的內部比較實驗結果，將先討論模型有無加入新聞文本、有無使用新聞主題對文本做初步分類、注意力機制與歷史資訊對模型預測的有效性；接下來比較使用不同新聞主題與不同時間區間的資料對模型表現影響，最後比較多任務架構對預測的效果。

4.4.1.1 探討有無使用新聞文本資料對經濟領先指標預測的影響

本小節針對有無新聞文本資料對經濟領先指標預測是否影響預測效果，比較如下表。沒有使用新聞文本資料，僅使用經濟指標前四個月的數值資料，建立時間序列模型預測，從結果可以發現，僅用經濟指標的歷史數值資料無法有好的預測結果，相較有加入文字資料的模型，效果差異大，由此可以看出，有新聞文本資料加入，能整合出有價值的特徵資訊，對預測經濟指標有實質的效益。

表 4 有無使用新聞文本資料預測效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
With text as input	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957
W/o text as input	PCE	1.1964	0.6552	-0.168	0.0	0.0	0.0
	CCI	1.0945	0.6379	0.1148	0.66	0.8919	0.7586
	Retails	1.1182	0.5345	0.0357	0.5641	0.6875	0.6197
	CPI	1.2002	0.5	0.0	0.0	0.0	0.0

4.4.1.2 探討有無使用新聞主題對文本做初步分類對經濟領先指標預測的影響

本小節針對新聞文本資料透過新聞主題做初步分類，是否影響預測效果，比較如下表。新聞主題是根據從財經新聞網爬取新聞資料時所下的關鍵字整理而成，當模型輸入的文本資料沒有新聞主題的初步分類，而是全部新聞以日期排序，從結果可以發現，使用沒有新聞主題分類的資料進行預測時，整體表現相較差，由此可以看出，新聞透過關鍵字分類能夠幫助模型提取有價值的資訊，進一步提升模型預測的能力。

表 5 有無使用新聞主題對文本做初步分類預測效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
With topics	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957
W/o topics	PCE	0.1524	0.8596	0.6861	0.7	0.875	0.7778
	CCI	0.3660	0.7719	0.5054	0.8108	0.8333	0.8219

Retails	0.4868	0.7018	0.5282	1.0	0.4688	0.6383
CPI	0.1669	0.7368	0.5357	0.9375	0.5172	0.6667

4.4.1.3 探討有無加入注意力機制對經濟領先指標預測的影響

本小節針對加入注意力機制的預測效果比較在下表，表 6 呈現了注意力機制對預測任務的預測效果，可以發現針對加入注意力機制在不同新聞主題以及時間的特徵向量上，使模型在整合不同新聞主題的資訊時可以根據權重整合，預測效果有明顯的提升。將兩個注意力分別拆開來看，可以看到將注意力機制放在新聞主題上時，提升消費者信心指數(CCI)與零售銷售指數(Retails)表現；而將注意力機制放在時間時，對耐久財新訂單指數(PCE)以及消費者物價指數(CPI)走勢預測較有幫助，兩個注意力機制都有發揮其效用，在不同指標上優化模型預測表現。

表 6 注意力機制對經濟領先指標走勢預測效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
With topic and inter-day attention	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957
W/o inter-day attention	PCE	0.1486	0.8448	0.6377	0.6842	0.8125	0.7429
	CCI	0.3716	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4735	0.7241	0.5563	1.0	0.5	0.6667
	CPI	0.1721	0.7241	0.4924	0.8824	0.5172	0.6522
W/o topic attention	PCE	0.1491	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3701	0.7759	0.5101	0.8158	0.8378	0.8267
	Retails	0.4793	0.7069	0.5324	1.0	0.4688	0.6383
	CPI	0.1664	0.7586	0.5682	0.9412	0.5517	0.6957

w/o topic and inter-day attention	PCE	0.1514	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3899	0.7586	0.4893	0.8286	0.7838	0.8056
	Retails	0.4783	0.7069	0.5324	1.0	0.4688	0.6383
	CPI	0.166	0.7414	0.5401	0.9375	0.5172	0.6667

在下圖可以看到在訓練時注意力機制的權重分布，以特定主題分類中的財經主題為例，如圖所示，橫軸代表新聞主題名稱，縱軸代表預測的時間點，圖中的權重表示在預測時間點往前七天提取的新聞內容中，不同新聞主題的內容對預測領先經濟指標的影響程度。首先，可以看到在 2013 年 6 月 1 日進行預測時，新聞主題的注意力權重在 Finance 較於其他主題高，表示與 Finance 相關的新聞，對預測經濟指標較有幫助；而在 2013 年 9 月 1 日進行預測時，與 Fed 相關新聞對模型預測較有幫助；在 2014 年 1 月 1 日預測時，則是 Technology 與 Fed 主題注意力權重較高。從結果可以發現，在不同時間點，權重分佈不同，專注的新聞內容亦不相同，表示注意力機制成功發揮其作用。

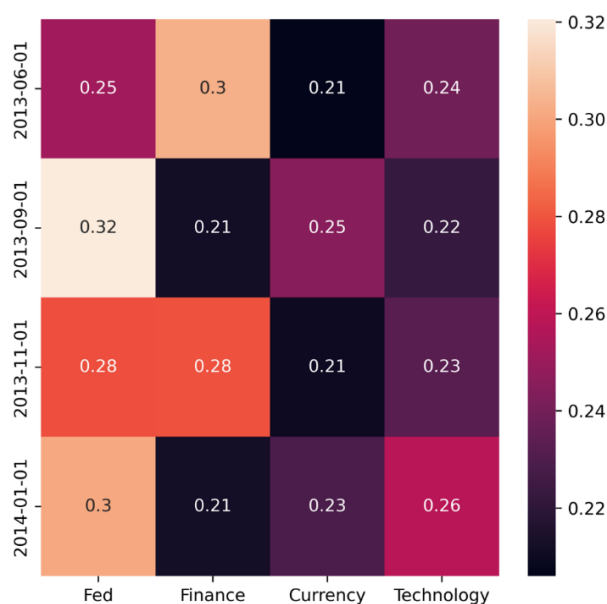


圖 8 注意力機制分析

4.4.1.4 探討有無加入前一時間點的經濟指標數據對經濟領先指標預測的影響

本小節針對加入前一時間點的經濟指標數據，也就是前一個月發布的經濟指標數據，對模型的預測效果比較在下表，表呈現了過去歷史資訊對預測任務的預測效果，可以發現加入歷史資訊後，除了對消費者物價指數預測表現的正確率較低之外，對其他經濟指標的預測效果都有明顯的提升。在消費者物價指數預測任務上，也可以發現在有加入歷史資訊後，除了正確率較低之外，其餘評量指標表現都較好。

表 7 歷史經濟指標資訊對經濟領先指標預測效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
With historical data	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957
W/o historical data	PCE	0.9694	0.7069	0.1617	0.4444	0.25	0.32
	CCI	0.819	0.5517	0.0687	0.6667	0.5946	0.6286
	Retails	0.9929	0.4828	-0.0099	0.5455	0.375	0.4444
	CPI	0.9617	0.8103	0.2155	0.1818	0.5	0.2667

4.4.1.5 探討有無經濟指標標準化對經濟領先指標預測的影響

本小節針對經濟指標數據進行標準化是否有提升模型訓練收斂與精準度做比較，比較結果如下表。從表中可以看到，經過標準化後的經濟指標，對模型訓練有效提升模型預測經濟走勢的精準度，由此可知，原始經濟指標數據範圍過小，損失函數的範圍也縮小，會導致模型無法有效學習與收斂，進而影響模型走勢的分類預測指標表現。

表 8 歷史經濟指標資訊對經濟領先指標預測效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
With Norma- lization	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957
W/o Norma- lization	PCE	0.00011	0.7241	0.0	0.0	0.0	0.0
	CCI	0.0027	0.7586	0.4775	0.8108	0.8108	0.8108
	Retails	0.00029	0.569	0.1026	0.5636	0.9688	0.7126
	CPI	0.00003	0.5	0.0	0.0	0.0	0.0

4.4.1.6 探討不同新聞主題對經濟領先指標預測的影響

本小節針對使用不同新聞主題作為模型輸入，預測的效果比較如下表。從表中可以清楚看到使用所有主題的新聞預測效果仍相對於特定主題分類有較好的表現，在特定主題的分類中，與國防政治相關的主題組合是 Defense, Middle East, Aerospace, Politics 的新聞內容作為輸入，預測耐久財新訂單指數(PCE)與消費者信心指數(CCI)效果最好，勝過使用所有主題的預測效果。

表 9 不同新聞主題對經濟領先指標預測效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
Fed Finance Currencies Technology	PCE	0.1508	0.8596	0.6861	0.7	0.875	0.7778
	CCI	0.3658	0.807	0.5816	0.8378	0.8611	0.8493
	Retails	0.4846	0.6842	0.5043	1.0	0.4375	0.6087
	CPI	0.1646	0.7544	0.5639	0.9412	0.5517	0.6957
Defense Middle East	PCE	0.1360	0.9375	0.8466	0.9231	0.8571	0.8889
	CCI	0.3701	0.8125	0.6111	0.8065	0.8929	0.8475

Aerospace Politics	Retails	0.4828	0.7083	0.5423	1.0	0.5	0.6667
	CPI	0.1546	0.75	0.55	0.9286	0.5417	0.6842
Commodity Consumer Retail Transportat- ion	PCE	0.1535	0.8545	0.6624	0.7222	0.8125	0.7647
	CCI	0.3613	0.7636	0.4951	0.8	0.8235	0.8116
	Retails	0.4801	0.6909	0.5192	1.0	0.4688	0.6383
	CPI	0.1659	0.7273	0.5197	0.9333	0.5	0.6512
All topics	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957

4.4.1.7 探討不同時間區間的新聞對經濟指標預測的影響

本小節針對使用不同時間區間的新聞作為輸入，預測效果的比較如下表，其中的時間區間是從預測日往前回推天數，從結果來看，使用距離預測日前一週的新聞預測表現較好，使用到距離預測日前兩週至前一個月的新聞，在多個指標表現上都較差一些，但是效果沒有顯著的落差，表示使用前一周至前一個月的資料都可以有不錯的預測表現。

表 10 不同時間區間的新聞對經濟指標預測效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
7 days	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957
14 days	PCE	0.1507	0.8596	0.6885	0.7	0.875	0.7778
	CCI	0.3632	0.7759	0.5101	0.8158	0.8378	0.8267
	Retails	0.4820	0.7069	0.5324	1.0	0.4688	0.6383

21 days	CPI	0.1631	0.7586	0.5682	0.9412	0.5517	0.6957
	PCE	0.1525	0.8596	0.6861	0.7368	0.875	0.8
	CCI	0.361	0.7895	0.5579	0.8529	0.8056	0.8286
	Retails	0.4919	0.7018	0.5282	1.0	0.4688	0.6383
	CPI	0.1674	0.7368	0.5357	0.9375	0.5172	0.6667
28 days	PCE	0.1524	0.8596	0.6861	0.7	0.875	0.7778
	CCI	0.3696	0.7895	0.5401	0.8158	0.8611	0.8056
	Retails	0.4842	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1669	0.7368	0.5357	0.9375	0.5172	0.6667

4.4.1.8 多任務學習架構對經濟領先指標預測的影響

本小節針對使用單任務學習與多任務學習架構的比較，預測效果比較如下表，可以看到雖然使用單任務學習架構在耐久財消費指數(PCE)與消費者物價指數(CPI)學習能力極佳，但是在消費者信心指數(CCI)以及零售銷售指數(Retails)的表現，雖然模型一樣盡力縮小預測值與真實值的誤差，所以在 loss 表現上沒有太大差異，但從走勢的預測可以看到，對於指標漲跌的判斷效果較差，單任務學習的架構在走勢預測上沒有好的表現。需要透過多任務學習整體的損失函數來自於不同任務之和，與其他指標一同學習，幫助訓練 CCI 與 Retails 指標的預測能有效預測未來走勢。

表 11 多任務學習架構對經濟指標預測效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
Multi-Task Learning	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957
Single-Task Learning	PCE	0.1419	0.9	0.7043	0.7647	0.8125	0.7879
	CCI	0.3673	0.5	0.4384	0.88	0.5946	0.7097
	Retails	0.5659	0.8	0.2769	1.0	0.1562	0.2703
	CPI	0.1651	0.9	0.5682	0.9412	0.5517	0.6957

4.4.2 與其他模型比較結果

鑒於過去研究沒有適合預測多個經濟領先指標的模型，本研究整理多個應用於金融領域如股價、匯率走勢之預測，能夠處理新聞文字資料以及時序資料之模型，對經濟領先指標進行走勢之預測，來與本研究方法做預測表現的比較。

4.4.2.1 效能評估

此小節針對本研究方法預測表現最好的預測結果與 FAST[35]、BHAM[44]、Adv-ALSTM[49]、StockNet[4]進行比較，比較如表 12。FAST 是使用新聞文本資料與歷史股價資訊，透過時間序列模型與注意力機制，加上擷取新聞發布之間的時間差作為時間序列的間隔，預測股票預期利潤並排序，使交易收益最大化；BHAM 是新聞文本資料經過時間序列模型以及學習排序新聞重要性的神經網絡來篩選新聞特徵，保留其重要的資訊，再結合歷史匯率價格資訊，預測匯率價格走勢；Adv-ALSTM 是透過對抗式學習生成具有可以表現出股票市場隨機性的假樣本，與原始資料一起訓練，預測股價走勢，由於此方法僅使用與股價相關的資訊，未使用文本資料，因此，本研究實作此方法時，使用與領先經濟指標有高度

相關的標普 500 指數歷史資訊作為模型輸入，預測領先經濟指標數值與走勢；StockNet 是以深度生成網絡為架構，透過推特(Twitter)社群媒體上與股票相關的文本資料與過去股價資訊來預測未來股價走勢。由表可以發現使用 FAST、BHAM、Adv-ALSTM、StockNet 之預測效果都沒有比使用本研究提出的方法來得好，代表專門發展來捕捉股價與匯率等時序關係的複雜模型方法不適合應用於經濟指標走勢預測；在三種不同新聞主題的資料上，本研究方法的準確率 (Accuracy)、精確率(Precision)、MCC 以及精確率與召回率的調和平均數 F1 score 表現皆較其他方法高，代表其能有效地利用新聞資訊與歷史經濟指標資料預測美國重要經濟領先指標。

表 12 其他模型預測經濟指標效果比較表

實驗結果		Loss	Acc	Mcc	Precision	Recall	F1-score
Model	PCE	0.1498	0.8621	0.6885	0.7	0.875	0.7778
	CCI	0.3687	0.8103	0.5856	0.8421	0.8649	0.8533
	Retails	0.4799	0.7193	0.5522	1.0	0.5	0.6667
	CPI	0.1655	0.7586	0.5682	0.9412	0.5517	0.6957
FAST (2021)	PCE	1.1888	0.7241	0.073	0.3333	0.1429	0.2
	CCI	1.045	0.4828	0.0	0.0	0.0	0.0
	Retails	1.1040	0.569	0.137	0.6667	0.1481	0.2424
	CPI	1.1977	0.6034	0.1186	0.5714	0.1667	0.2581
BHAM (2019)	PCE	1.0415	0.7241	0.0	0.0	0.0	0.0
	CCI	1.6529	0.6379	0.0	0.6379	1.0	0.7789
	Retails	1.003	0.5517	0.0	0.5517	1.0	0.7111
	CPI	1.060	0.4828	-0.1325	0.0	0.0	0.0
Adv- ALSTM (2019)	PCE	0.5045	0.4669	-0.0666	0.4380	0.46	0.4487
	CCI	0.455	0.5091	0.0	0.0	0.0	0.0
	Retails	0.5740	0.3349	-0.3544	0.2428	0.1619	0.1942
	CPI	0.3791	0.4575	-0.038	0.5540	0.3333	0.4162

StockNet (2018)	PCE	0.6988	0.5260	0.0241	0.7273	0.1778	0.2857
	CCI	0.69	0.5677	0.093	0.4259	0.8846	0.575
	Retails	0.6995	0.4843	-0.198	0.4571	0.4012	0.2546
	CPI	0.6705	0.6354	0.0912	1.0	0.0238	0.0465

第五章 結論與未來研究方向

5.1 結論

本研究的主要目的是運用文本資料進行美國重要經濟領先指標預測模型建置，以財經新聞相關資料實現，目的是為能善用公開資料在經濟指標數據發布前成功提前預測。根據實驗結果能夠看出，本研究所提出之方法擁有優異的預測效果，多任務學習架構不僅有效減少模型參數，不同任務之間相互提供額外的有用訊息，提升任務學習的效果與模型泛化能力，且本研究提出之方法取用不同時間區間的新聞資料皆能達到好的預測效果，在實驗結果比較中有使用到距離預測日前一個月的新聞，雖然考慮更多新聞，但效果沒有顯著落差，同時會增加模型計算時間，因此，本方法以預測日往前推七日的新聞資料最能有效應用於本方法進行美國重要經濟領先指標預測。而本研究方法加入注意力機制亦有效提升預測任務的效果，並且增加模型訓練與結果的解釋力。此外，本研究利用指標之間的相關性，融入歷史指標資訊，作為預測指標之特徵，提升模型預測結果，成功建立美國重要經濟指標預測的模型。

5.2 未來研究

本研究所提出利用新聞文本資料進行美國重要經濟領先指標預測的模型建立，現階段已能夠應用有效地應用於有關之經濟指標，但還有許多需要優化的部分，針對未來研究列出以下兩個方向：

1. 本研究利用搜尋關鍵字作為搜集新聞的主題，並加入注意力機制在新聞特徵的計算，然而，從實驗結果可以看到，不同主題在實驗結果上有些差異，新聞的選擇對模型訓練表現來說非常重要，本研究所提出之方法提取眾多新聞數做特徵資訊，沒有有效刪除無價值的新聞資訊，因此，

未來在新聞主題、內容的選擇上若能也以神經網絡學習決定，對於建立經濟指標預測的系統會更加完整。

2. 本研究提出之方法應用於以月為更新單位的重要經濟指標，如第二章第一節所介紹，重要經濟指標中還有以週為單位更新的指標，以及其他並非以指數方式衡量的指標，未來能以增加預測指標的數量為目標，提升模型有用性。

參考文獻

- [1] Fichtner, F., Rüffer, R., & Schnatz, B. (2009). Leading indicators in a globalised world.
- [2] Kyereboah-Coleman, A., & Agyire-Tettey, K. F. (2008). Impact of macroeconomic indicators on stock market performance: The case of the Ghana Stock Exchange. *The Journal of Risk Finance*.
- [3] G. John Ikenberry, *After Victory: Institutions, Strategic Restraint, and the Rebuilding of Order after Major War* (Princeton: Princeton University Press, 2001), p. 4
- [4] Xu, Y., & Cohen, S. B. (2018, July). Stock movement prediction from tweets and historical prices. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1970-1979).
- [5] Sawhney, R., Wadhwa, A., Agarwal, S., & Shah, R. (2021, June). Quantitative day trading from natural language using reinforcement learning. In *Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies* (pp. 4018-4030).
- [6] Sawhney, R., Agarwal, S., Wadhwa, A., & Shah, R. (2020, November). Deep attentive learning for stock movement prediction from social media text and company correlations. In *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (pp. 8415-8426).
- [7] *The Secrets of Economic Indicators: Hidden Clues to Future Economic Trends, Third Edition*
- [8] Quandt, A. K. (2020). Stock Market Correlations to Economic Indicators.
- [9] Huang, J., Chai, J., & Cho, S. (2020). Deep learning in finance and banking: A literature review and classification. *Frontiers of Business Research in China*, 14(1),

1-24.

- [10] Zahara, S. (2020, December). Multivariate Time Series Forecasting Based Cloud Computing For Consumer Price Index Using Deep Learning Algorithms. In 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) (pp. 338-343). IEEE.
- [11] Cheng, Y., Hai, T., Zheng, Y., & Li, B. (2017, July). Prediction model of the unemployment rate for nanyang in henan province based on BP neural network. In 2017 13th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD) (pp. 1023-1027). IEEE.
- [12] Ramírez, K. M., Hormaza, J. M., & Soto, S. V. (2020, September). Artificial intelligence and its impact on the prediction of economic indicators. In Proceedings of the 6th International Conference on Engineering & MIS 2020 (pp. 1-8).
- [13] Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.
- [14] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2018). Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- [15] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., ... & Polosukhin, I. (2017). Attention is all you need. *Advances in neural information processing systems*, 30.
- [16] Yang, W., Xie, Y., Lin, A., Li, X., Tan, L., Xiong, K., ... & Lin, J. (2019). End-to-end open-domain question answering with bertserini. *arXiv preprint arXiv:1902.01718*.
- [17] Hoang, M., Bihorac, O. A., & Rouces, J. (2019). Aspect-based sentiment analysis

- using BERT. In *Proceedings of the 22nd nordic conference on computational linguistics* (pp. 187-196).
- [18] Chen, Q. (2021). Stock movement prediction with financial news using contextualized embedding from bert. *arXiv preprint arXiv:2107.08721*.
- [19] Nam, K., & Seong, N. (2019). Financial news-based stock movement prediction using causality analysis of influence in the Korean stock market. *Decision Support Systems*, 117, 100-112.
- [20] Pew Research Center, January, 2021, “News Use Across Social Media Platforms in 2020”
- [21] Hu, Z., Zhao, Y., & Khushi, M. (2021). A survey of forex and stock price prediction using deep learning. *Applied System Innovation*, 4(1), 9.
- [22] Chen, D., Harimoto, K., Bao, R., Su, Q., & Sun, X. (2019). Group, extract and aggregate: Summarizing a large amount of finance news for forex movement prediction. *arXiv preprint arXiv:1910.05032*.
- [23] Ding, X., Zhang, Y., Liu, T., & Duan, J. (2016, December). Knowledge-driven event embedding for stock prediction. In *Proceedings of coling 2016, the 26th international conference on computational linguistics: Technical papers* (pp. 2133-2142).
- [24] Caruana, R. (1993). “Multitask Learning: A Knowledge-Based Source of Inductive Bias,” *Proceedings of the 10th International Conference on Machine Learning*, ML-93, University of Massachusetts, Amherst, pp. 41-48
- [25] Ruder, S. (2017). An overview of multi-task learning in deep neural networks. *arXiv preprint arXiv:1706.05098*.
- [26] Duong, L., Cohn, T., Bird, S., & Cook, P. (2015, July). Low resource dependency parsing: Cross-lingual parameter sharing in a neural network parser. In *Proceedings*

of the 53rd annual meeting of the Association for Computational Linguistics and the 7th international joint conference on natural language processing (volume 2: short papers) (pp. 845-850).

- [27] Ma, J., Zhao, Z., Yi, X., Chen, J., Hong, L., & Chi, E. H. (2018, July). Modeling task relationships in multi-task learning with multi-gate mixture-of-experts. In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (pp. 1930-1939).
- [28] Ghosn, J., & Bengio, Y. (1996). Multi-task learning for stock selection. *Advances in neural information processing systems*, 9
- [29] Ma, T., & Tan, Y. (2022). Stock Ranking with Multi-Task Learning. *Expert Systems with Applications*, 199, 116886.
- [30] Bitvai, Z., & Cohn, T. (2015). Day trading profit maximization with multi-task learning and technical analysis. *Machine Learning*, 101(1), 187-209.
- [31] Sawhney, R., Mathur, P., Mangal, A., Khanna, P., Shah, R. R., & Zimmermann, R. (2020, October). Multimodal multi-task financial risk forecasting. In Proceedings of the 28th ACM international conference on multimedia (pp. 456-465).
- [32] HOCHREITER, Sepp; SCHMIDHUBER, Jürgen. Long short-term memory. *Neural computation*, 9.8: 1735-1780, 1997
- [33] Kim, S., & Kang, M. (2019). Financial series prediction using Attention LSTM. *arXiv preprint arXiv:1902.10877*.
- [34] Hu, Z., Liu, W., Bian, J., Liu, X., & Liu, T. Y. (2018, February). Listening to chaotic whispers: A deep learning framework for news-oriented stock trend prediction. In *Proceedings of the eleventh ACM international conference on web search and data mining* (pp. 261-269).
- [35] Sawhney, R., Wadhwa, A., Agarwal, S., & Shah, R. (2021, April). FAST: Financial

- News and Tweet Based Time Aware Network for Stock Trading. In *Proceedings of the 16th Conference of the European Chapter of the Association for Computational Linguistics: Main Volume* (pp. 2164-2175).
- [36] D. M. Q. Nelson, A. C. M. Pereira and R. A. de Oliveira, "Stock market's price movement prediction with LSTM neural networks," 2017 International Joint Conference on Neural Networks (IJCNN), 2017, pp. 1419-1426, doi: 10.1109/IJCNN.2017.7966019.
- [37] Nguyen, A., Chatterjee, S., Weinzierl, S., Schwinn, L., Matzner, M., & Eskofier, B. (2020, October). Time matters: time-aware lstms for predictive business process monitoring. In *International Conference on Process Mining* (pp. 112-123). Springer, Cham.
- [38] Bahdanau, D., & Cho, K. (2014). Neural machine translation by jointly learning to align and translate. arXiv preprint arXiv: 1409.0473.
- [39] Luong, M. T., Pham, H., & Manning, C. D. Effective approaches to attention-based neural machine translation. *arXiv preprint arXiv:1508.04025*, 2015.
- [40] Li, H., Cui, Y., Wang, S., Liu, J., Qin, J., & Yang, Y. (2020). Multivariate financial time-series prediction with certified robustness. *IEEE Access*, 8, 109133-109143.
- [41] Kim, R., So, C. H., Jeong, M., Lee, S., Kim, J., & Kang, J. (2019). Hats: A hierarchical graph attention network for stock movement prediction. *arXiv preprint arXiv:1908.07999*.
- [42] Liu, J., Lin, H., Liu, X., Xu, B., Ren, Y., Diao, Y., & Yang, L. (2019, August). Transformer-based capsule network for stock movement prediction. In *Proceedings of the First Workshop on Financial Technology and Natural Language Processing* (pp. 66-73).
- [43] <https://fred.stlouisfed.org/>

- [44] Chen, D., Harimoto, K., Bao, R., Su, Q., & Sun, X. (2019). Group, extract and aggregate: Summarizing a large amount of finance news for forex movement prediction. *arXiv preprint arXiv:1910.05032*.
- [45] <https://www.us-proxy.org/>
- [46] https://www.federalreserve.gov/faqs/economy_14400.htm
- [47] Choi, H., Kim, J., Joe, S., & Gwon, Y. (2021, January). Evaluation of BERT and ALBERT Sentence Embedding Performance on Downstream NLP Tasks. In 2020 25th International Conference on Pattern Recognition (ICPR) (pp. 5482-5487). IEEE.
- [48] Chicco, D., & Jurman, G. (2020). The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC genomics*, 21(1), 1-13.
- [49] Feng, F., Chen, H., He, X., Ding, J., Sun, M., & Chua, T. S. (2018). Enhancing stock movement prediction with adversarial training. *arXiv preprint arXiv:1810.09936*.