

基於依存關係感知能力的深度學習模型進行金融推文之數值關係檢測

Numerical Relation Detection in Financial Tweets using Dependency-aware Deep Neural Network

Yu-Chi Liang¹, Min-Chen Chen², Wen-Chao Yeh³, Yung-Chun Chang^{*}
^{1,2,*}Graduate Institute of Data Science, Taipei Medical University, Taipei, Taiwan
³Institute of Information Systems and Applications, National Tsing Hua University,
Hsinchu, Taiwan

^{1,2,*}{m946108001, m946109001, changyc}@tmu.edu.tw
³wych@m109.nthu.edu.tw

摘要

近年來，許多研究是以金融文本資料進行分析，然而，我們發現在這些文本資料中的數字亦包含值得探討的豐富資訊。本論文藉由分析金融議題相關之推特文章，探討文本中目標數字與目標標籤是否具有關聯性。本研究採用基於變換器的雙向編碼器表示式作為模型架構之主要表示機制，並將依存關係作為特徵轉成依存關係矩陣後放入卷積神經網路中，使模型透過依存關係學習到文本中詞與詞間的關聯性。根據實驗結果顯示，本研究採用之方法對於此任務有良好之預測能力，其 Macro-averaging F1 Score 為 71.05%。

Abstract

Machine learning methods for financial document analysis have been focusing mainly on the textual part. However, the numerical parts of these documents are also rich in information content. In order to further analyze the financial text, we should assay the numeric information in depth. In light of this, the purpose of this research is to identify the linking between the target cashtag and the target numeral in financial tweets, which is more challenging than analyzing news and official documents. In this research, we developed a multi model fusion approach which integrates Bidirectional Encoder Representations from Transformers (BERT) and Convolutional Neural Network (CNN). We also encode dependency information behind text into the model to derive semantic latent features. The experimental

results show that our model can achieve remarkable performance and outperform comparisons.

關鍵字：金融社交媒體、基於變換器的雙向編碼器表示技術、卷積神經網路、依存語法
Keywords: Financial Social Media, BERT, CNN, Dependency Grammar

1 Introduction

隨著人工智慧的蓬勃發展，各行業正在研究如何將技術應用在產業實務上。金融服務業拓展一個新穎的研究領域，稱為金融科技 (Financial Technology, FinTech)。目前 FinTech 已融入自然語言處理 (Natural Language Processing, NLP) 技術，透過機器學習與文字探勘分析財務數據來獲取許多與財務相關的有用訊息。例如，財務情緒分析可以獲得專家對於股票市場趨勢的意見讓客戶能夠做出更好的決策 (Sohangir et al., 2018)；透過銀行員工和客戶之間聯繫過程的歷史紀錄分析產業關聯以及決定當地產業之信心水準 (Sakaji et al., 2019)。

以往分析財務文本資料時，主要是針對文字部分做探索分析；然而要瞭解金融資料的細節，不僅須要分析文字內容還須依靠數字資訊。若能取得財務資料中數字所包含的意義可以更快速的掌握關鍵內容。隨著網際網路的發展，人們透過社交媒體表達自己意見的頻率隨之增高，也直接讓社交媒體平台成為當今訊息的主要來源之一。Vilas 等人研究推特用戶對金融市場相關事件的影響，分析結果顯示儘管推特只是社交媒體而非專業金融論壇，但推文內容對於金融事件仍具影響

力 (Vilas et al., 2019)，若善加利用社交媒體上的文章並加以分析，相信可以獲得對於市場和客戶都有幫助的相關資訊。相較於新聞文章，經由社交媒體產生的資料並無固定格式，內容可能包含表情符號、連結網址和網路用語。因此，在使用自然語言處理技術分析社交媒體數據時需考慮到比較複雜的情況，如何掌握數據中的關鍵訊息成為一項具挑戰性的任務 (Farzindar and Inkpen, 2015)。考量文本中可能包含多個數字或是多個目標標籤，因此需決定推特文章中特定數字與特定附加目標之間是否具有關聯性。如圖 1 的例子，當目標標籤為股票「BAC」時，數字「3」表示股票「BAC」每一股的現金支出為 3。因此，數字「3」是表示股票「BAC」，而數字「2009」表示年份，它與股票「BAC」並無直接關聯性。此為二元分類任務，透過金融推文判斷文本中目標數字與目標標籤之間是否具關聯性：「1」表示目標數字與目標標籤間具有關聯性，而「0」則是目標數字與目標標籤間沒有關聯性。本研究計畫藉由依存句法分析應用在深度學習模型上來探討與金融事件相關之推特文章中目標數字與目標標籤之間是否存在相關性。

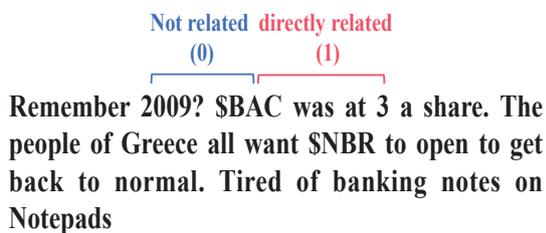


圖 1. 推文中數字與股票代碼間之關係示意圖

2 Related Work

使用自然語言處理技術進行數值分析任務之應用相當廣泛，其中包含臨床醫學領域、科學領域、新聞領域等。然而，隨著 FinTech 議題逐漸受到重視，有許多任務著重在如何將自然語言處理技術應用於金融領域，像是語義評估工作坊 (International Workshop on Semantic Evaluation, SemEval) 於 2017 年舉辦之金融微博和新聞的情緒分析 (Cortis et al., 2017)、耶拿大學語言和訊息工程實驗室舉辦之第一屆經濟學和自然語言處理研討會 (1st Workshop

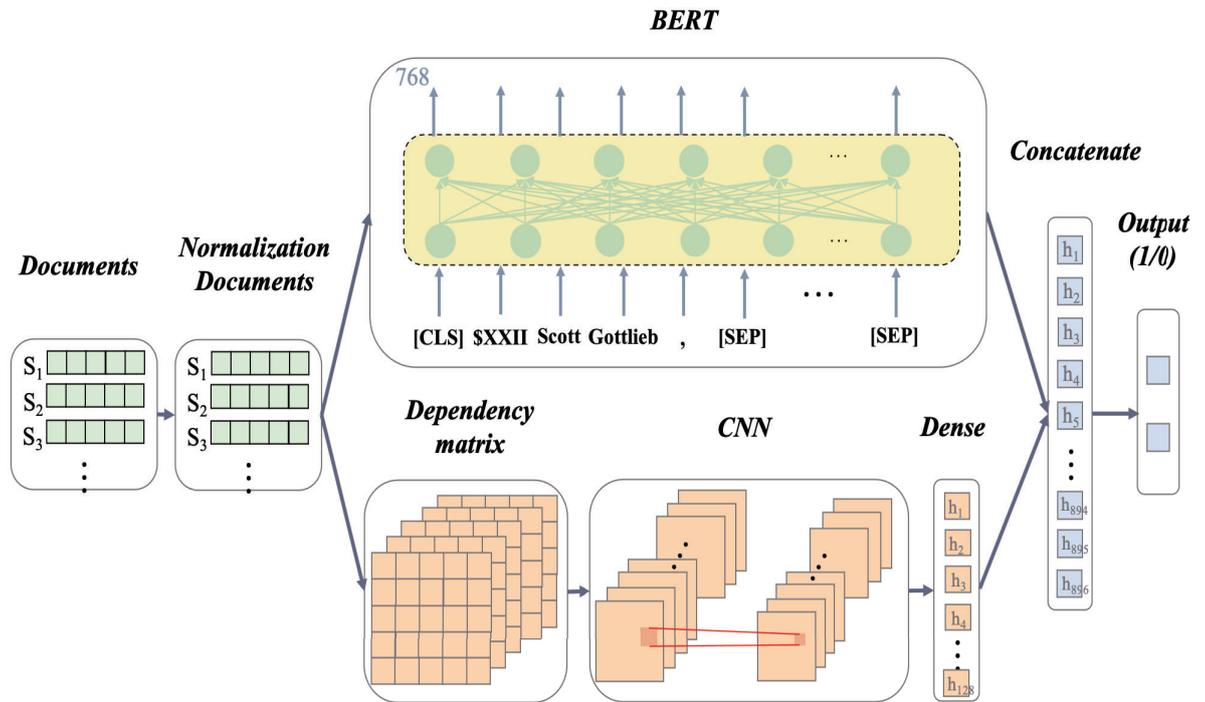
on Economic and Natural Language Processing, ECONLP 2018) (Hahn et al., 2018)、國際資訊檢索評估會議 (NII Testbeds and Community for Information access Research, NTCIR) 第十四屆的理解金融推文中的數字 (Fine-Grained Numeral Understanding in Financial Tweets, FinNum) 任務 (Chen et al., 2019) 以及第十五屆金融推文的數字附件 (Numeral Attachment in Financial Tweets, FinNum-2) 任務 (Chen et al., 2020)。

FinNum 是一項理解金融社交媒體文本資料中數字屬性的任務。當我們實際執行 FinNum 任務時，發現一項值得探討的主題：金融社交媒體中的資料通常包含多組數字，在進行數值分類任務前應先判別數字是否與股票文字代碼有關聯性。適逢 NTCIR 第十五屆提出新的任務以改善 FinNum 任務實際應用時所遇到的問題，任務名稱為 FinNum-2。FinNum-2 是一項二元分類任務，主要是探討金融社交媒體數據中已標註完成的目標數字與目標標籤之間是否具有關聯性。主辦單位共提供兩組比較基準 (Baseline)，其中 Baseline-Majority 是將全部資料都視為數字和目標標籤之間存在關聯性之實驗結果。Baseline-Caps-m 則是合併單詞的嵌入向量 (Token embedding)、字元嵌入向量 (Character Embedding)、單詞的位置嵌入向量 (Position Embedding) 以及單詞的幅度嵌入向量 (Magnitude Embedding)，並放入膠囊網路中進行訓練 (Chen et al., 2019)。除了上述由主辦單位提供之比較基準組外，

- TLR-3：採用基於變換器的雙向編碼器表示技術 (Bidirectional Encoder Representations from Transformers, BERT) 與 Robustly optimized BERT approach (RoBERTa) 進行集成學習 (Ensemble learning) 並根據預測機率最小值選出預測結果 (Moreno et al., 2020)。
- CYUT-2：針對預訓練 XLM-RoBERTa 模型進行微調 (Jiang et al., 2020)
- MIG-2：在 BERT 預訓練模型後添加雙向長短期記憶模型 (Bi-directional Long Short-Term Memory, BiLSTM) 並

將損失函數權重設為 0.8 與 0.2 (Chen et al., 2020)。

依存關係矩陣將做為卷積神經網路 (Convolutional Neural Network, CNN) 輸入特徵，



- TMUNLP-1：取出 BERT 產生之向量表達式，將之放入 BiLSTM 模型並續接一層注意力機制 (Attention) (Liang et al., 2020)。
- WUST：透過 TF-IDF 生成文本表示法再放入支援向量機 (Support Vector Machine, SVM) 模型進行分類預測 (Xia et al., 2020)。

3 Methodology

本研究之模型架構稱為 Dependency Grammars-infused BERT-CNN Model，如圖 2。先將金融短文本經預處理步驟後放入 BERT 模型學習完整之短文本內容，並將短文本轉為向量後取出備用。同時，將金融短文本進行依存句法分析 (Dependency parsing) 並將結果轉換成依存關係矩陣 (Dependency matrix)。

並將 CNN 模型輸出結果使用全連接層 (Dense Layer) 來降低維度。最後，將上述兩個輸出結果合併後放入全連接層進行最終分類預測，判斷文本中股票代碼與數字間是否具關聯性。

預處理步驟旨在標註目標標籤與目標數字，作法為在目標標籤與目標數字的前後皆分別使用「£」與「§」符號標註單詞。透過此法標註所有文本讓模型於訓練時能學習到文本中的目標標籤與目標數字，進而獲取更多目標標籤與目標數字之訊息。BERT 模型是由 Transformer 中的編碼器訓練而成 (Devlin et al., 2018)，可由 BERT 官方網站下載預訓練模型。本研究採用 BERT-Base Uncased 模型，包含 12 層 Transformer 區塊、768 維度之隱藏層尺寸。使用 Uncased 模型能一併將文本中大寫英文字母轉為小寫以及移除文本中之重音符號。將金融短文本放入此模型訓練後取出第一個 [CLS] 產出之 768 維度向量，與依存關係矩陣向量結合並進行最終預測。

圖 2. Dependency Grammars-infused BERT-CNN Model 架構圖

圖 3 為文本輸入 BERT 模型至產出向量之流程。首先將文本加入特殊標籤後放入 BERT 模型，即為變換器 (Transformer) 模型之編碼器 (Encoder) 架構。BERT 模型會根據選用的預訓練模型將文本中每個文字皆輸出成大小與隱藏層尺寸相同之向量，在此步驟中選用第一個單詞 [CLS] 產出之向量，續接依存句法向量進行最終模型訓練。

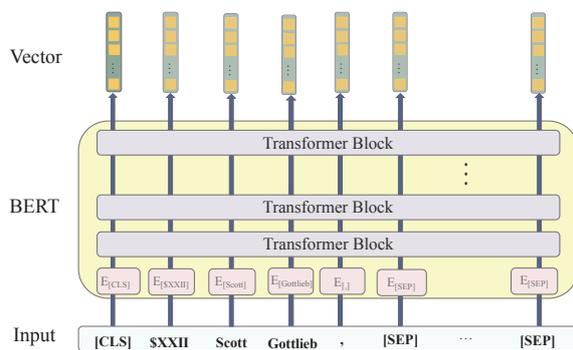


圖 3. 文本表示法生成圖

依存句法分析 (Dependency Parsing) 為 NLP 一項關鍵技術，本研究使用 StanfordCoreNLP 套件進行依存句法分析。Stanford CoreNLP 是由史丹佛 (Stanford) 大學自然語言處理實驗室提出之自然語言分析套件，除了依存句法分析外亦包含其他多種功能；例如，分詞 (Tokenization)、詞形還原 (Lemmatization)、命名實體識別技術 (Named Entity Recognition, NER) 與詞性標註 (Part-Of-Speech, POS) 等。從 3.5.2 版本開始，Stanford CoreNLP 預設參照 Universal Dependencies v1 (Nivre et al., 2016) 輸出語法關係 (Schuster and Manning, 2016)，此文檔一共包含 40 種不同之依存關係。圖 4 上半部為使用 StanfordCoreNLP 生成依存句法之範例，箭頭連接之兩分詞表示含有依存關係，箭頭旁之文字顯示兩者依存關係。統整文本放入依存句法分析產生之依存關係種類並將依存關係進行編號，以供後續產生依存關係矩陣使用。

圖 4 下半部為依存關係矩陣範例，其縱軸與橫軸皆以文本表示，在含有依存關係兩個分詞之對應位置放入相對應的依存關係編號，

若是分詞之對應位置並無依存關係則在對應位置中填入「0」。舉例如下，在分詞「SDPW」與「Glad」之間具有依存關係「acl:relcl」且此關係編號為「12」，在進行依存關係矩陣轉換時便會在縱軸為「SDPW」且橫軸為「Glad」以及縱軸為「Glad」且橫軸為「SDPW」兩處皆放上「12」。另考量每一個文本之長度可能不同且為因應後續接 CNN 模型時輸入特徵圖需為固定尺寸之向量，故需先判斷全部資料中各單一文本的單詞數量。以各文本中單詞數量最大者為依存關係矩陣之固定長度 N 。若是尺寸小於 N 的文本則以 0 填補長度不足處，於是可將依存關係矩陣尺寸皆處理成大小為 $N \times N$ 矩陣。

CNN 模型是以深度神經網路為基礎架構再增添兩種新的神經元，分別為卷積層 (Convolutional Layer) 與池化層 (Pooling Layer) (Murugan, 2017)。CNN 不只被廣泛應用於影像辨識與聲音辨識等任務上，也被發現可在自然語言處理上有良好效能。我們將文本轉為固定尺寸之依存關係矩陣後便放入 CNN 模型中訓練並將經由平坦層輸出之向量透過全連接層轉為 128 維度之向量。

將上述由依存句法分析產生之特徵向量與 BERT 模型產生之文本表示法進行串連 (Concatenate) 後放入全連接層進行最終預測。

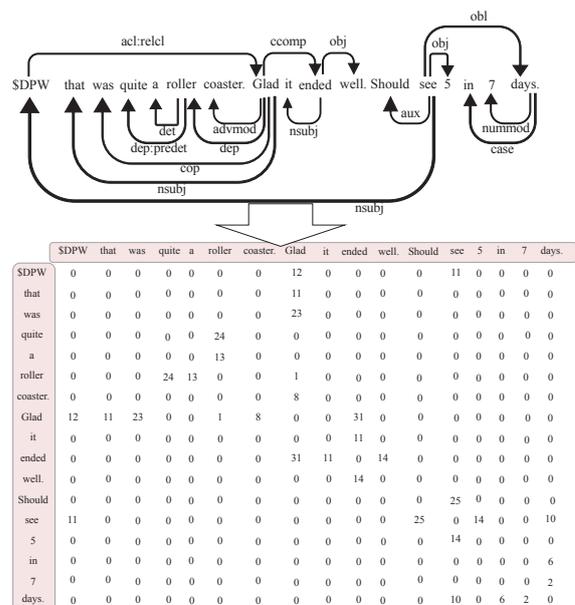


圖 4. 生成依存句法與依存關係矩陣範例圖

4 Result and Discussion

本研究使用 NTCIR 第十五屆會議的金融推文數字探討 (Numeral Attachment in Financial Tweets, FinNum-2) 任務之 NumAttach 資料集，探討與金融事件相關之推特文章 (Tweets) 中數字和目標標籤間是否存在相關性。此資料集包含推特文章 (tweet)、目標數字 (target number)、目標標籤 (target cashtag)、關係 (relation) 以及目標數字位置 (offset) 共五個欄位。參考表 1，整體資料集包含 10,340 句短文本，切分 70% (7,187 句) 做為訓練資料集、10% (1,044 句) 為開發資料集以及 20% (2,109 句) 為測試資料集。本研究採用 macro F1 分數作為實驗評估效能標準。

表 1. NumAttach 資料集分佈

資料集	具關聯性	不具關聯性	總和
訓練集	5,827	1,360	7,187
開發集	850	194	1,044
測試集	1,721	388	2,109
總和	8,398	1,942	10,340

表 2 為採用三種不同方式表達依存句法分析特徵並比較三者實驗結果與未添加依存句法分析特徵之實驗結果。

- BERT：指不使用任何依存句法分析特徵僅將原句經預處理後放入 BERT 模型中。
- BERT+Dep (SDP)、BERT+Dep (OHM) 以及 Our Method 分別表示將上述三種依存句法分析特徵分別放入 CNN 模型中並與 BERT 模型之結果合併，其中 Our Method 即為本論文所提出之 Dependency Grammars-infused BERT-CNN Model。

「Dep (SDP)」為透過目標標籤與目標數字間最短依存關係路徑 (Shortest Dependency Path, SDP) 並整理出依存關係進行編號，而後將 SDP 透過編號轉為 10 維向量，若 SDP 之長度大於 10 則取出前 5 個以及後 5 個依存關係，若 SDP 之長度小於 10 則是使用 0 將向量擴充至長度為 10。

「Dep (OHM)」是將依存句法分析特徵轉為 One-Hot 矩陣 (One-Hot Matrix, OHM) 後放入 CNN 模型，One-Hot 矩陣是先將矩陣之橫軸與縱軸皆以原句表示並逐一比對兩單詞間是否具有依存關係，若兩單詞間具有依存關係則以「1」表示，反之，「0」表示單詞間不具有依存關係。

表 2. 金融推文探討數字之依存句法分析實驗比較

Method	Macro F ₁ score (%)
BERT	70.20%
BERT + Dep (SDP)	58.74%
BERT + Dep (OHM)	64.25%
Our Method	71.05%

比較 BERT + Dep (SDP)、BERT + Dep (OHM) 以及 Our Method 三組實驗結果可看出，使用最短依存關係路徑產生 10 維向量以及 One-Hot 矩陣皆比僅採用 BERT 模型之效能差，前者之效能減少 11.46%，後者則是減少 5.95%。BERT+Dep (SDP) 在找尋最短路徑時是透過計算空白字元得到句子的最短依存關係路徑，然而目標數字可能包含著其他符號使得句子無法找出最短依存關係路徑僅能使用全句替代，此特徵設計方式未能良好表達出目標數字與目標標籤之間依存關係。

將依存關係特徵設計成 One-Hot 矩陣時，僅能表達出單詞間有無依存關係但無法學習到依存關係種類，使得模型未能透過學習 Dependency parsing 特徵來增加識別推特文章中的目標標籤與數字間關聯性的能力。

本論文所提出之 Dependency Grammars-infused BERT-CNN Model 效能為 71.05%，不僅相較 BERT 模型效能略高 0.62%。且為全部實驗中最具辨識能力之模型。此結果顯示使用編號之方式表達依存句法分析特徵能夠完整表達出句子中所有依存關係及種類，此特徵設計方式有助於模型學習到每一句話中單詞之間不同種類之依存關係，加入此特徵亦能幫助模型對於推特文章中的目標標籤與數字間關聯性有更好之辨識能力。

表 3 為本研究提出之模型架構 (Our Method) 與兩組 Baseline、五組參賽模型之實驗結果比較。Our Method 於 NTCIR FinNum-2 任務中排名第三，且明顯高於第四名。雖非最優異之效能，但在預測與金融事件相關之推特文章

中數字和目標標籤之間是否存在相關性方面已表現出眾。Baseline-Majority 之方式未經模型訓練僅將全部資料皆視為有關聯性，故此 Baseline 於此比賽中效能排最後。WUST 使用 TF-IDF 生成文本表示法並搭配 SVM 模型訓練，文本經過模型訓練有助於辨識沒有關聯性之文本，因此，此模型之效能較 Baseline-Majority 提升約 10%。Baseline-Caps-m 考量四種詞嵌入並放入膠囊網路中，不僅學習到常見之單詞的嵌入向量，同時學習了字元嵌入向量 (Character Embedding)、單詞的位置嵌入向量 (Position Embedding) 以及單詞的幅度嵌入向量 (Magnitude Embedding)，因此相較於 WUST 採用 TF-IDF 生成文本表示法，此方法學習到較多資訊且在放入膠囊網路中可以更有效地學習資訊中的特徵。因此，此模型相較於 WUST 之模型效能提升約 10%。

TMUNLP-1 使用 BiLSTM 模型架構，在文本訓練上，BiLSTM 模型可以學習到句子雙向的語義關係，對於內容主要為辨識句子中兩單詞間是否具關聯性之任務來說，BiLSTM 模型可以獲取良好之效能，此模型較 Baseline-Majority 提升約 1.4%。與 TMUNLP-1 不同的是 MIG-2 之模型架構是經過完整之 BERT 預訓練模型再訓練而非僅採用 BERT 產生之文本表示法，因此在模型 Macro F₁ Score 效能上 MIG-2 比 TMUNLP-1 提升約 4%。Our Method 考量依存關係編號矩陣特徵且使用符號分別標注目標數字和目標標籤，上述之方式皆有助於模型學習辨識文本中的目標數字與目標標籤，在模型效能上，此方式較 MIG-2 提升 2.33%。CYUT-2 雖未考慮使用特徵添加至模型中但他們所使用之方式是針對 XLM-RoBERTa 模型中之學習率進行微調，XLM-RoBERTa 模型中每一層皆使用不同且與該層最合適之學習率，此舉可以提升 XLM-RoBERTa 模型整體之效能，比本論文所提出之模型架構略為提升 0.85%。TLR-3 分別同時訓練 BERT 模型與 Robustly optimized BERT approach (RoBERTa) 模型，而後根據預測出之機率選出結果，此方式可以集合兩個模型之結果，此外，因資料分布不平衡關係，透過選擇預測出機率的最小值可以更好地辨識目標數字和目標標籤之間不具有關聯性之文本，所以此方法較 CYUT-2 提升

約 2%，同時，此模型架構亦為比賽中辨識任務最優異之模型。

本研究採用 BERT 模型添加依存關係編號矩陣特徵之方式，與第一名 (TLR-3) 需訓練兩個模型並採用集成學習之方法以及第二名 (CYUT-2) 針對 XLM-RoBERTa 模型進行微調之方式相比；本論文之模型訓練時間相對較短且對於電腦硬體設備要求相對較低。本研究提出的方法可以在相對較低的能耗時間情況下獲取與前二名近似之預測效能，顯示出此模型對於辨識任務內容具有貢獻度。

表 3. 嵌 FinNum-2 任務模型效能比較

Method	Macro F ₁ score (%)
TLR-3	73.95%
CYUT-2	71.90%
Our Method	71.05%
MIG-2	68.72%
TMUNLP-1	64.76%
Baseline-Caps-m	63.37%
WUST	54.43%
Baseline-Majority	44.93%

5 Conclusion

近年來隨著網際網路的蓬勃發展，人們逐漸改由在社交媒體上表達自己對於某事件的看法，因此針對從社交媒體上獲取之文本使用自然語言處理技術之研究也逐年上升。然而，過去在分析文本時大多都是著重於分析文字部分，往往容易忽略文本中的數字其實包含相當重要之資訊；尤其針對金融方面之資料，瞭解文本中數字之含義對於決策者如何根據數字結果做出適當決策更是有深遠的影響。

本論文嘗試由推特上所獲取與金融議題相關之短文資料集進行探討，計畫透過演算法辨別文章中的目標數字與目標標籤之間是否具有關聯性以助於日後瞭解文本中數字所具有之含義。本研究是以 BERT 模型為主要基礎架構並將依存句法分析所得結果轉為依存關係矩陣作為特徵放入 CNN 後，再與 BERT 串連進行最終預測。與常見進行相關任務使用之方法不同之處在於以往在設計特徵時會選用一維之方式呈現，本研究採用二維之方式呈現依存關係，透過二維矩陣搭配依存關係編號方式能夠完整表達出每一個詞與詞間之依存關係，使整體模型對於辨識社交媒體短

文本中的特定兩單詞之間是否具關聯性之效能往上提升。

考量取得完整依存關係之重要性，未來或可嘗試使用 Tweeparser 等以推特文章訓練而成的套件 (Kong et al., 2014)，期待能對寫作格式不受拘束之社交媒體短文本提取出更完整的依存關係。希望透過微調技術提升 BERT 模型架構之辨識能力，進而讓整體模型對於社交媒體短文本中的特定兩單詞之間是否具關聯性之辨識能力獲得更進一步提升之可能。

Acknowledgments

本研究承蒙科技部計畫 MOST 109-2410-H-038-012-MY2 之補助。

References

- Sohangir, S., Wang, D., Pomeranets, A., and Khoshgoftaar, T. M. 2018. Big Data: Deep Learning for financial sentiment analysis. *Journal of Big Data*, 5(1): 1-25.
- Sakaji, H., Kuramoto, R., Matsushima, H., Izumi, K., Shimada, T., and Sunakawa, K. 2019. Financial text data analytics framework for business confidence indices and inter-industry relations. In *Proceedings of the First Workshop on Financial Technology and Natural Language Processing*, pages 40-46.
- Vilas, A. F., Redondo, R. P. D., Crockett, K., Owda, M., and Evans, L. 2019. Twitter permeability to financial events: an experiment towards a model for sensing irregularities. *Multimedia Tools and Applications*, 78(7): 9217-9245
- Farzindar, A. and Inkpen, D. 2015. Natural language processing for social media. *Synthesis Lectures on Human Language Technologies*, 8(2): 1-166.
- Cortis, K., Freitas, A., Daudert, T., Huerlimann, M., Zarrouk, M., Handschuh, S., and Davis, B. 2017. Semeval-2017 task 5: Fine-grained sentiment analysis on financial microblogs and news. *Association for Computational Linguistics (ACL)*
- Hahn, U., Hoste, V., and Tsai, M. F., 2018. Proceedings of the First Workshop on Economics and Natural Language Processing. In *Proceedings of the First Workshop on Economics and Natural Language Processing*.
- Chen, C. C., Huang, H. H., Takamura, H., and Chen, H. H., 2019. Overview of the ntcir-14 finnum task: Fine-grained numeral understanding in financial social media data. In *Proceedings of the 14th NTCIR Conference on Evaluation of Information Access Technologies*, pages 19-27.
- Chen, C. C., Huang, H. H., Takamura, H., and Chen, H. H., 2020. Overview of the NTCIR-15 FinNum-2 Task: Numeral Attachment in Financial Tweets. *Development*, 850(194): 1-44.
- Chen, C. C., Huang, H. H., and Chen, H. H., 2019. Numeral attachment with auxiliary tasks. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pages 1161-1164.
- Moreno, J. G., Boros, E., and Doucet, A., 2020. TLR at the NTCIR-15 FinNum-2 Task: Improving Text Classifiers for Numeral Attachment in Financial Social Data. In *Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies*, pages 8-11.
- Jiang, M. T. J., Chen, Y. K., and Wu, S. H., 2020. CYUT at the NTCIR-15 FinNum-2 Task: Tokenization and Fine-tuning Techniques for Numeral Attachment in Financial Tweets. In *Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies*, pages 92-96.
- Chen, Y. Y. and Liu, C. L., 2020. MIG at the NTCIR-15 FinNum-2 Task: Use the transfer learning and feature engineering for numeral attachment task. In *Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies*.
- Liang, Y. C., Huang, Y. H., Cheng, Y. Y., and Chang, Y. C., 2020. TMUNLP at the NTCIR-15 FinNum-2. In *Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies*.
- Xia, X., Wang, W., and Liu, M., 2020. WUST at NTCIR-15 FinNum-2 Task. In *Proceedings of the 15th NTCIR Conference on Evaluation of Information Access Technologies*.
- Devlin, J., Chang, M. W., Lee, K., and Toutanova, K., 2018. Bert: Pre-training of deep bidirectional transformers for language understanding. *arXiv preprint arXiv:1810.04805*.
- Nivre, J., De Marneffe, M. C., Ginter, F., Goldberg, Y., Hajic, J., Manning, C. D., ... and Zeman, D., 2016. Universal dependencies v1: A multilingual treebank collection. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 1659-1666.
- Schuster, S. and Manning, C. D. 2016. Enhanced english universal dependencies: An improved representation for natural language understanding tasks. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*, pages 2371-2378.
- Murugan, P. 2017. Feed forward and backward run in deep convolution neural network. *arXiv preprint arXiv:1711.03278*.

Kong, L., Schneider, N., Swayamdipta, S., Bhatia, A., Dyer, C., and Smith, N. A., 2014. A dependency parser for tweets. *In Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1001-1012.