

國立高雄科技大學

資訊管理系(所)

碩士論文

基於數據增強的情感辨識技術  
Emotion recognition technology  
based on data Augmentation

研究生：黃智暘

指導教授：黃文楨 博士

中華民國 一 一 一 年 六 月

基於數據增強的情感辨識技術

Emotion recognition technology based on data Augmentation

研究生：黃智暘 HUANG, CHIH-YANG

指導教授：黃文楨 HUANG, WEN-CHEN

國立高雄科技大學

資訊管理碩士班

碩士論文

A Thesis

Presented to

Department of Information Management

National Kaohsiung University of Science and Technology

in Partial Fulfillment of the Requirements

for the Degree of

Master of Engineering

in

Information Management

June 2022

Kaohsiung, Taiwan, Republic of China

中華民國 一 一 一 年 六 月

國立高雄科技大學(第一校區)研究所學位論文考試審定書

資訊管理系碩士班[第一]碩士班

研究生 黃智暘 所提之論文

論文名稱(中文): 基於數據增強的情感辨識技術

論文名稱(英/日/德文): Emotion recognition technology based on data Augmentation

經本委員會評審，符合碩士學位論文標準。

學位考試委員會

召集人

殷堂凱

簽章

委員

黃承龍

黃文楨

殷堂凱

指導教授

黃文楨

簽章

系所主管

蔡國河

簽章

中華民國 111 年 06 月 24 日

保存期限：永久

## 國立高雄科技大學學位論文著作權歸屬協議書

論文名稱：基於數據增強的情感辨識技術

論文種類：☐博士論文 ☒碩士論文

研究生：黃智暘

系所名稱：資訊管理系

指導教授：黃文楨

共同指導教授：

茲為保障著作人著作權益，並就論文著作權之歸屬及事後權利行使方式，包括論文應如何公開發表、發表時應如何標示著作人姓名、論文事後可作何種修改以及未來應如何授權他人利用等事項，碩、博士生與指導（含共同指導）教授依下列原則達成協議：

- 一、碩、博士生所撰寫之論文，如指導（或共同指導）教授僅為觀念之指導，並未參與內容表達之撰寫，依著作權法規定，學生為該論文之著作人，並於論文完成時，即享有該論文之著作權，指導教授無法於事後主張為共同著作人，亦不得共同掛名為著作人。（著作權法第10條之1）
- 二、如指導（或共同指導）教授不僅為觀念的指導，且參與內容之表達而與學生共同完成論文，且各人之創作，不能分離利用者，則為共同著作，學生與指導教授為論文的共同著作人並共同享有著作權，此等共同著作著作權（包括著作財產權及著作人格權）的行使，即應取得碩、博士生與指導（或共同指導）教授之共同同意後，始得為之。（著作權法第8條、著作權法第40條之1第1項）

三、依上述原則，本論文之著作權歸屬：

☐ 研究生單獨擁有。

☒ 研究生與指導教授共同擁有。

☐ 研究生、指導教授及共同指導教授共同擁有。

研究生：黃智暘 日期：111年7月20日

指導教授：黃文楨 日期：111年7月20日

## 基於數據增強的情感辨識技術

### Emotion recognition technology based on data Augmentation

## 摘要

近年來，隨著科技的進步與社群媒體的普及，有更多的人在網路上發表自己對某些事情的看法與意見。若能即時分析作者在做出這個言論時的情緒狀態，則可以讓我們更加快速的瞭解大眾對於現今各項議題的情緒反應。進而將這些資訊提供給企業、政府、機構、組織等等，並做出相對應的回應。

本研究資料集(DataSet)來源是自然語言處理與中文計算國際會議(NLPCC)中的NLPCC 2018 Shared Task1。其中情緒類別有快樂、悲傷、憤怒、恐懼、驚訝、以及無任何情緒六大類別。總共有6000個句子，且各項情緒是不平衡的。所以需要先做資料前處理，接著對於較少情緒的類別做隨機抽樣。最後透過數據增強技術增加句子數量，使得各項情緒趨近平衡。目的為防止訓練時過擬合以及增加其分類準確率。

本研究採用Word2Vec模型來進行詞向量的訓練，以詞向量的結果來做同義詞替換，接著採用BERT-wwm-ext模型進行分類與評估。評估的方式主要為Accuracy、Precision、Recall、和F1-score。本研究將探討數據增強前後的差異，對同義詞替換法不斷嘗試、修正與調整，以提高準確率。在不改變原本句子結構的前提下做增強，分別比較並分析其個別差異：替換(1)單一詞彙、(2)雙詞彙以及(3)多詞彙與(4)單純只透過翻譯來增強數據集。又以(3)最佳，準確度為79%，分別高出(1) 4%、(2) 3%、(4) 16%。

關鍵字：Word2Vec、數據增強、Bert、多標籤分類

## Abstract

In recent years, with the advancement of technology and the popularity of social media, more people have expressed their views and opinions on certain things on the Internet. If we can analyze the emotional state of the author when he made this speech in real time, it will allow us to understand the public's emotional response to current issues more quickly. And then provide this information to enterprises, governments, institutions, organizations, etc., and make corresponding responses.

The source of this research data set (DataSet) is the NLPCC 2018 Shared Task 1 in the International Conference on Natural Language Processing and Chinese Computing (NLPCC). Among them, there are six categories of emotions: happiness, sadness, anger, fear, surprise, and no emotion. There are a total of 6000 sentences, and the emotions are unbalanced. So you need to do the data pre-processing first, and then do random sampling for the less emotional categories. Finally, the number of sentences is increased through data enhancement technology, so that the various emotions are closer to balance. The purpose is to prevent overfitting during training and increase its classification accuracy.

This research uses the Word2Vec model to train the word vector, uses the result of the word vector to do synonym replacement, and then uses the BERT-wm-ext model for classification and evaluation. The evaluation methods are mainly Accuracy, Precision, Recall, and F1-score. This study will explore the differences before and after data enhancement, and constantly try, modify and adjust the synonym substitution method to improve the accuracy. Make enhancements without changing the original sentence structure, compare and analyze individual differences: replace

(1) single vocabulary, (2) double vocabulary and (3) multiple vocabulary and (4) simply enhance the data set through translation. And (3) is the best, with an accuracy of 79%, which is higher than (1) 4%, (2) 3%, and (4) 16%.

Keywords : Word2Vec 、 Data Augmentation 、 Bert 、 Multi-label Classification

## 誌 謝

回憶起兩年前剛成為研究生時，總有點擔心與害怕，害怕自己沒有足夠扎實的資訊科技相關領域技術，擔憂自己會不會找不到方向，導致最後浪費兩年的時間，所幸系上有完備的相關活動與演講，甚至於必修課中邀請許多專業人士到學校演講，所涵蓋的領域相當廣泛，包含像是智慧物聯網、長照2.0應用軟體和促進人類行為的技術設計等等，非常感謝資管系學院讓我能夠涉及到諸多領域的學識，讓我順利找到感興趣的方向。

兩年的碩士生涯就這樣過去了，很慶幸自己能夠遇到指導教授黃文楨老師，總是能在我迷惘時給予我許多論文上的提點與建議，並時常撥空和我討論進度，使我能順利完成論文。也感謝老師們給我機會參與DIGI+ Telant跨域數位人才加速躍升計畫，在整個活動過程中著實收穫了不少。

感謝口試委員殷堂凱老師和黃承龍老師撥空於論文口試時給予指正與建議，使本論文能夠更加完善，也讓我知道了我還有許多的不足，也感謝老師們對學生論文的肯定，學生於此衷心感謝。

在此也感謝研究所一起打拼的宋宏鈞、曾建智、蔡宗傑、以及蔡絜婷，還有在我碩士兩年生活中曾給予陪伴、鼓勵和幫助的人，很高興能與你們相遇。

最後，我要感謝我的家人、同事、業師總是給予我精神上的支持，還有生活上無微不至的照顧，非常感謝您們，您們辛苦了。

黃智暘 謹誌於  
國立高雄科技大學 資訊管理學所  
中華民國一百一十一年六月



# 目 錄

摘要 .....	i
Abstract .....	ii
誌 謝 .....	iv
目 錄 .....	v
圖目錄 .....	vii
表目錄 .....	viii
壹、緒論 .....	1
一、研究背景 .....	1
二、研究動機與目的 .....	2
三、研究貢獻 .....	3
(一) 多元情感辨識技術 .....	3
(二) 數據增強之研究 .....	3
貳、文獻探討 .....	4
一、情感分析 .....	4
二、情緒識別 .....	4
三、機器翻譯 .....	6
四、字詞相似性 .....	6
參、研究方法 .....	8
一、研究架構 .....	8
二、同義詞替換方法 .....	13
三、訓練分類模型 .....	18
四、評估分類器 .....	19
五、各模型參數設計 .....	21
(一)、Word2Vec .....	21
(二)、BERT-wwm-ext .....	22

六、系統環境.....	23
肆、研究結果 .....	25
一、分類結果.....	25
二、研究探討 .....	31
三、研究限制.....	32
伍、結論與未來研究 .....	33
陸、參考文獻 .....	34

## 圖目錄

圖 1：研究架構圖.....	9
圖 2：XML節點示意圖 .....	10
圖 3：去除Emoji示意圖 .....	11
圖 4：使用WikiCorpus 套件提取文章之虛擬碼.....	14
圖 5：WikiCorpus提取文章.....	14
圖 6：Jieba斷詞之虛擬碼.....	15
圖 7：Jieba斷詞結果 .....	15
圖 8：Word2vec模型訓練參數虛擬碼 .....	16
圖 9：測試模型結果之虛擬碼.....	17
圖 10：以開心一詞測試模型結果.....	17
圖 11：同義詞替換示意圖.....	17
圖 12：各方法Precision分數.....	29
圖 13：各方法Recall分數.....	29
圖 14：各方法F1-Score分數.....	30
圖 15：各方法準確度 .....	30
圖 16：節點格式之錯誤訊息.....	31

## 表目錄

表 1：翻譯比較 .....	11
表 2：數據增強前各項情緒統計 .....	12
表 3：數據增強後各項情緒統計 .....	12
表 4：各項情緒訓練集、驗證集及測試集數量（單位：句數） .....	18
表 5：混淆矩陣(Category) .....	19
表 6：Word2Vec訓練參數 .....	22
表 7：BERT-wwm-ext 訓練參數 .....	23
表 8：系統規格 .....	24
表 9：開發工具 .....	24
表 10：原始資料各情緒Precision、Recall、F1-Score分數 .....	25
表 11：翻譯增強各情緒Precision、Recall、F1-Score分數 .....	26
表 12：單詞彙同義詞替換各情緒Precision、Recall、F1-Score分數 .....	26
表 13：雙詞彙同義詞替換各情緒Precision、Recall、F1-Score分數 .....	27
表 14：多詞彙同義詞替換各情緒Precision、Recall、F1-Score分數 .....	27
表 15：準確度比較 .....	28

# 壹、緒論

## 一、研究背景

近年來，隨著科技的進步，AI人工智慧已經成為許多業界、學界的發展目標，而AI發展的項目不計其數，其中以深度學習(Deep Learning)技術出現突破性發展；以現今來說，蘋果公司的Siri(Speech Interpretation and Recognition Interface，語音解析及辨識介面)就是深度學習的一種應用，藉由大量的數據資料進行分析訓練，進而讓程式能夠辨認人類的說話內容以及文法結構。對於人工智慧中的其中一項應用-情感分析(Sentiment Analysis)而言，常被應用在細粒度情感分析、情感檢測、意圖分析。

舉例來說，現今社群媒體的普及例如：Facebook、Twitter……，有更多人會藉由網際網路分享自己對某些事情的看法，分析作者在做出這個言論時的情緒狀態稱為情感分析，因此，能夠即時且準確地分析網路上評論的情感傾向，是情感分析重要的研究方向。

情感分析是自然語言處理(Natural Language Processing, NLP)領域中的一項常見任務，該任務的目的是檢測情緒。以往的研究主要集中在二元(正向和負向)以及三元(正向、負向和中立)的情感分析。此外，僅僅分類兩類或三類情緒是不夠的，人們可能是快樂的、悲傷的、憤怒的、恐懼的或者是驚訝的，而不是只有兩種情緒，只從正負向情緒很難判斷作者當下的情緒以及使用者體驗，進而無法給予更多的協助。

## 二、研究動機與目的

在以往的研究中，主要集中在二元(正向和負向)以及三元(正向、負向和中立)的情感分析。如何處理多種情緒的情感分析，並且研究如何準確的分析情感傾向，將各種評論加以整理，並且將相同情緒的評論分為同一類，由於現今網路上評論過於龐大，要獲得大量的評論相當花費時間成本，如何在數據有限的情況下，能夠有高準確的情感辨識，是本文研究所要探討的動機。

對於數據增強的方式有很多種，例如：同義詞替換(Synonyms Replace)、隨機插入(Randomly Insert)、隨機交換(Randomly Swap)、隨機刪除(Randomly Delete)等方法，而分類的準確率也有所不同，只求效果能夠越高越好。

本文研究方式先將原始資料集做前處理，接著針對情緒較少的類別做同義詞替換增加資料量，最後在做分類訓練。此外本文採用同義詞替換的數據增強方法加以改良實驗，將原本隨機從句子中抽取N個詞彙改良成分別抽取單一詞彙、雙詞彙以及多詞彙，之後將抽取這些詞的同義詞，用同義詞將原詞替換。例如："今天是個好天氣"替換成"今日是個好天氣"。通過同義詞替換後句子大至上還是會有相同的標籤，其目的是希望能夠在實驗中找出較有效的方式以提升分類的準確率。

### 三、研究貢獻

#### (一)多元情感辨識技術

本研究使用資料集為自然語言處理 (Natural Language Processing, NLP) 和中文計算 (Chinese Computing, CC) 領域的國際會議所提供的資料集，此資料集中情緒類別分為Happiness (快樂)、Sadness (悲傷)、Anger (憤怒)、Fear (恐懼)、Surprise (驚訝)，此外，本研究將無任何情緒的句子獨立出來為第六類命名為「None」，一共六個類別，在樣本數較少的情況下保有一定的準確度。

#### (二)數據增強之研究

本研究透過簡單數據增強之中的同義詞替換加以研究，以替換單一詞彙、雙詞彙、以及多詞彙的同義詞，來增強數據，針對資料筆數較少的情緒類別進行增強。最終分別以75%、76%、79%準確度優於純透過翻譯做數據增強的方法。

## 貳、文獻探討

### 一、情感分析

情感分析(sentiment analysis)，也稱為意見挖掘(Opinion mining)，涉及檢查針對特定實體的情緒。文本中表達的情緒或觀點進一步分為正面、負面、中性或更細粒度的分類（最正面、最不正面、最負面和最不負面）。情感分析對許多領域非常有用，例如股市預測[1]、文字情緒檢測[2]、產品意見調查[3]。應用於醫學領域，可評估患者的心理健康狀況。此外，電影的情感和類型也可以通過觀看預告片來預測。如今，人們更多地關注用於分析情緒的多模態數據[4]，[5]。因此，除了以文本形式結合星級和用戶評分外，還可以藉助多模態數據選擇詳盡的產品評分，其中可以使用語音、圖像或表情符號等各種形式對產品的評論進行合併[6]。中文情感分析研究主要有兩種方法：單語(monolingual)方法和雙語(bilingual)方法。在前者中，極性檢測(polarity detection)等典型的情感分析任務直接基於中文進行。後者利用現有的英語語言資源和機器翻譯技術來翻譯中文自然語言文本[7]。本研究主要是使用雙語的方法來進行中文的情緒分析處理。

### 二、情緒識別

情緒識別(emotion identification/recognition)已經應用於安全駕駛[8]、社會保障[9]等多個領域。識別情緒的方法主要有三類。第一種是基於人的身體信號，如語音[10]、身體語言[11] - [13]、面部表情[14]，[15]等，具有易於收集的優點。儘管如此，仍無法保證準確性。因為人們相對容易操縱面部表情或語



音等身體信號來隱藏他們的真實情緒，尤其是在社交互動中。人們在正式的社交場合中可能會微笑，即使他們正在經歷負面情緒。第二類使用生理信號，例如腦電圖（EEG）、溫度測量（T）、心電圖（ECG）、肌電圖（EMG）、皮膚電反應（GSR）、呼吸信號（RSP）等。Shu等人[16]對基於生理信號的情緒識別進行了廣泛的回顧，包括情緒模型、啟發方法、已發布的生理情緒數據集、特徵、分類器以及基於生理信號的情緒識別。第三類是用來辨識文字所表達的情緒。隱式情感識別是最具挑戰性的問題，因為這種情感通常隱藏在文本中，因此，其解決方案需要了解上下文。文本中隱性情感識別的主要方法有四種：基於規則的方法、基於經典學習的方法、深度學習方法和混合方法[17]。Batbaatar等人[18]提出了一種新的神經網絡架構，稱為 SENN（語義情感神經網絡），它可以通過採用預先訓練的詞表示來利用語義/句法和情感信息。SENN 模型主要有兩個子網絡，第一個子網絡使用雙向長短期記憶（BiLSTM）捕獲上下文信息並關注語義關係，第二個子網絡使用卷積神經網絡（CNN）提取情感特徵和側重於文本中單詞之間的情感關係。他們採用了埃克曼六種基本情緒的概念來加以辨別。本研究也是使用這六類的情緒來加以識別。Erenel等人[19]提出一種新的從文本中進行情感檢測領域的詞選擇方案。建議使用中等頻率的術語作為選擇術語的基礎。通過考慮它們的相關性分數來評估所有術語，因為中等頻率的術語也被認為對區分有價值。多項研究表明，他們所提出的特徵選擇方案優於傳統的基於過濾器的特徵選擇措施，如卡方和基尼文本。對於文檔表示，採用詞袋方法(bag-of-words approach)，其中每個選定的單詞如果出現在文檔中，則被賦予 1 的權重，如果沒有出現，則賦予 0 權重。提議的方案結合了 Chi-Square 或 Gini-Text 未選擇的術語。使用基準數據集，可以觀察到由於適度頻繁的術語提高了術語子集的表示能力，準確度有了顯著提高。

### 三、機器翻譯

在機器翻譯 (MT) 輸出質量有了大幅度的改進以及為該輸出發現的新的使用案例的背景下，Moorkens等人 [20]報告了使用統計和神經機器翻譯系統翻譯文獻的實驗。六位具有文學翻譯經驗的專業翻譯人員在三種情況下完成了英語到加泰羅尼亞語的翻譯：從頭翻譯、神經機器翻譯後編輯和統計機器翻譯後編輯。他們通過問卷和訪談的形式在翻譯前後提供反饋。雖然所有參與者都喜歡從頭開始翻譯，主要是因為可以自由發揮創造力而不受段級分割的限制，但那些經驗較少的人發現機器翻譯的建議很有用。Google 的神經機器翻譯 (GNMT) [21]系統使用基於 8 個編碼器和 8 個解碼器層的注意力和殘差連接 LSTM 網絡。注意機制將解碼器的底層與編碼器的頂層集成在一起，從而提高了並行性並減少了訓練時間。在推理計算過程中，使用低精度算法來加快最終的翻譯速度。在輸入和輸出中，將單詞劃分為一組有限的公共子單詞單元 (wordpieces)，以改進對稀有單詞的處理。因此，它在字符定界模型的靈活性和詞定界模型的效率之間提供了良好的平衡，有效地處理了罕見詞的翻譯，並提高了系統的整體準確性。通過波束搜索，將長度歸一化過程和覆蓋懲罰相結合，生成一個最有可能覆蓋源語句中所有單詞的輸出語句。GNMT 在 WMT' 14 英法和英德基準測試中取得了有競爭力的結果。與Google基於短語的生產系統相比，對一組孤立的簡單句子進行人工評估平均減少了 60% 的翻譯錯誤。本研究是使用Google的翻譯軟體來進行中文和英文之間的翻譯。

### 四、字詞相似性

在字詞的相似性比較下，word2vec[23]包含了兩種新穎的模型架構，用於表示從非常大的數據集中計算單詞的連續向量。這些表示的質量在單詞相似性任務中進行測量，並將結果與之前基於不同類型神經網絡的最佳性能技術進行比較。

在計算成本低得多的情況下，準確度有了很大的提高，即從 16 億個單詞的數據集中學習高質量的詞向量需要不到一天的時間。此外，這些向量在測試集上提供了最先進的性能，用於測量句法和語義詞的相似性。針對中文 BERT 的全字屏蔽 (Whole Word Masking, WWM) 策略，Cui 等人[24]開發了一個新模型 MacBERT 並使用一系列中文預訓練語言模型作為基線，包括BERT、RoBERTa、ELECTRA、RBT等。作為一種語言校正方法，以掩碼語言模型 (Masked Language Model, MLM) 方式進行了修改，從而減少了預訓練和微調階段之間的差異。MacBERT 在幾個中文 NLP 數據集上進行了測試，結果表明，提出的 MacBERT 在大多數任務中都取得了顯著的進步。此外，詳細分析的結果表明，MLM 任務應優先於 NSP 或其變體，因為NSP 沒有表現出優於任何其他任務的優勢。本研究所使用的資料集為NLPPC 2018 Shared Task1 [25]。在數據增強技術中，Wei和Zou[22]介紹了 EDA(Easy Data Augmentation)：用於提高文本分類性能的簡單數據增強技術。EDA 中有四種簡單但功能強大的操作：同義詞替換、隨機插入、隨機交換和隨機刪除。當應用於較小的數據集時，EDA 會產生特別令人印象深刻的結果。對五個數據集的研究發現，使用 EDA 訓練使用 50% 的可用訓練集產生的準確度與使用所有數據的正常訓練相同。本研究是使用EDA數據增強技術來強化訓練的結果。

## 參、研究方法

### 一、研究架構

本研究使用的資料集為自然語言處理 (Natural Language Processing, NLP) 和中文計算 (Chinese Computing, CC) 領域的國際會議所提供的資料集，以 NLPC C 2018 Shared Task1 資料集作為本研究的實驗對象，資料集內容為多種語言混雜而成，內有6000筆資料，每筆資料可能包含零、一個或多個情緒，這些情緒數量是不平均的，本研究的處理流程大致上分四個步驟，架構圖如圖1所示。

步驟一、解析資料集：其內容主要解析資料，包含資料格式轉換，以利後續研究。

步驟二、資料前處理及數據增強前置作業：將資料集內容翻譯成繁體中文，移除表情符號，Word2Vec詞向量模型訓練。

步驟三、針對較少類別數據增強：將情緒較少的類別做數據增強。

步驟四、訓練分類模型並分類結果：BERT模型訓練，情緒分類並分析其分類結果。

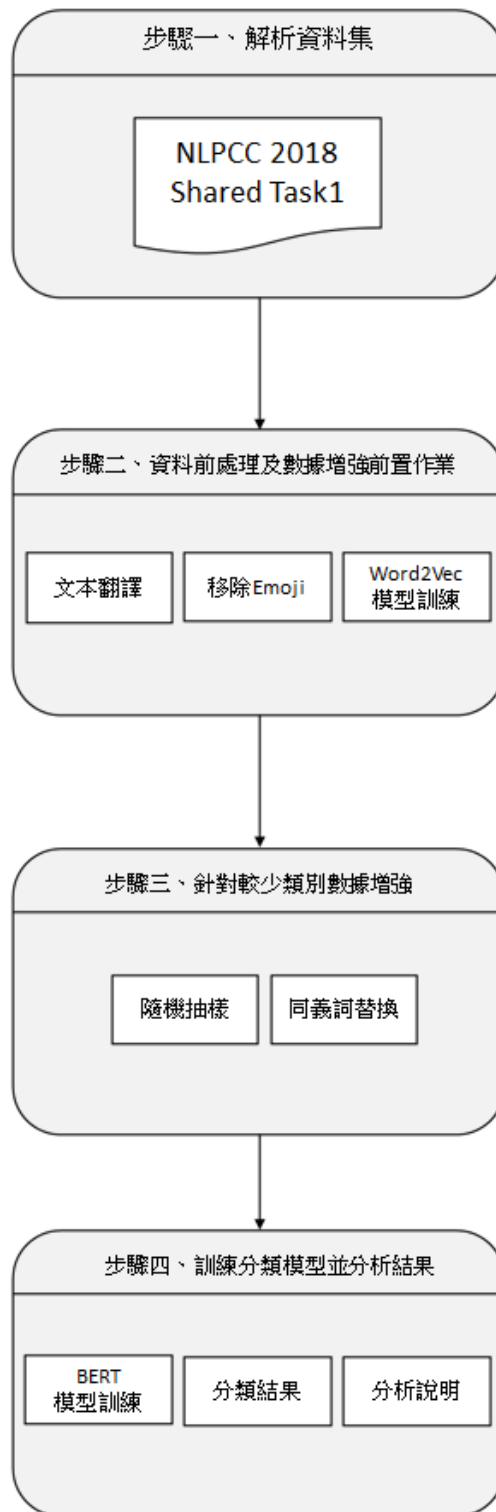
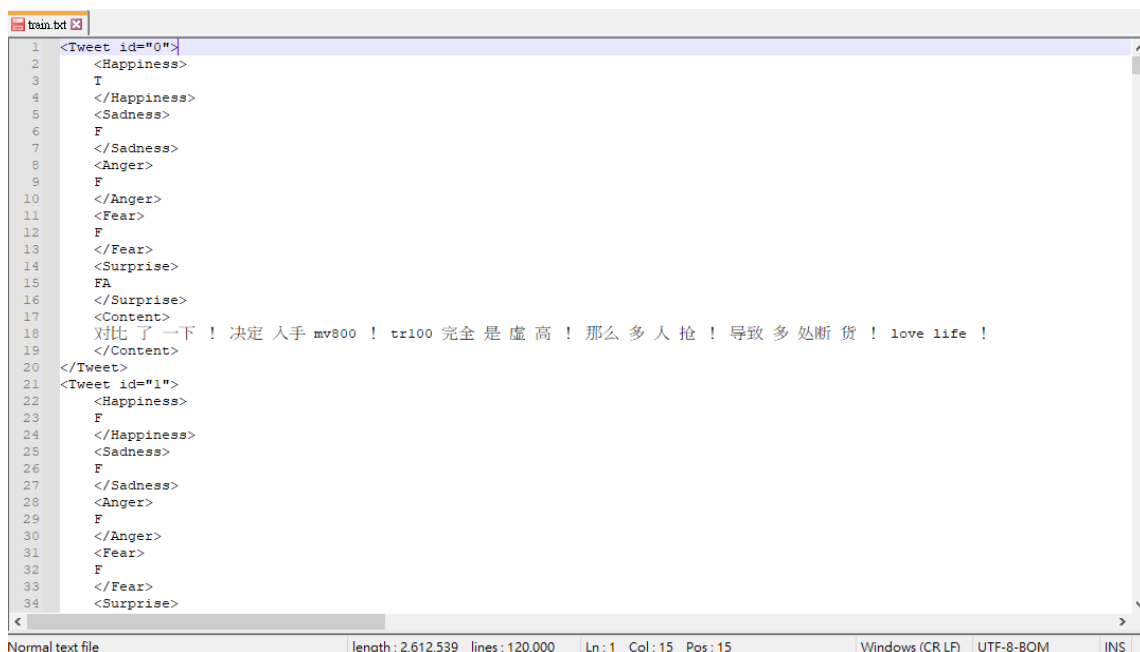


圖 1：研究架構圖

步驟一、解析資料集：意指擷取文件中所需要的段落，由於NLPCC所提供的資料集為XML格式，如圖1所示，文件節點依序為Tweet id（代號）、Happiness（快樂）、Sadness（悲傷）、Anger（憤怒）、Fear（恐懼）、Surprise（驚訝）、Content（主內容）。本研究使用BeautifulSoup這個程式庫方法對XML檔案做資料的提取，BeautifulSoup是一個可以從HTML或XML文件中提取數據的Python庫，將提取出的內容存成csv格式檔案，作為往後的研究資料。



```
1 <Tweet id="0">
2   <Happiness>
3     T
4   </Happiness>
5   <Sadness>
6     F
7   </Sadness>
8   <Anger>
9     F
10  </Anger>
11  <Fear>
12    F
13  </Fear>
14  <Surprise>
15    FA
16  </Surprise>
17  <Content>
18    对比了一下！决定入手 mv800！tr100 完全是虛高！那么多人抢！导致多处断货！love life！
19  </Content>
20 </Tweet>
21 <Tweet id="1">
22   <Happiness>
23     F
24   </Happiness>
25   <Sadness>
26     F
27   </Sadness>
28   <Anger>
29     F
30   </Anger>
31   <Fear>
32     F
33   </Fear>
34   <Surprise>
```

圖 2：XML節點示意圖

步驟二、資料前處理及數據增強前置作業：文本翻譯，意指將資料集的主內容藉由翻譯轉換語言，將主內容中英文、簡體中文、日文統一翻譯成繁體中文。現今普遍的翻譯有Google的Google翻譯、微軟的Microsoft Translator翻譯以及德國科隆GmbH開發的DeepL翻譯。Google翻譯是一項由Google於2006年開始提供的翻譯文段及網頁的服務，在2016年2016年，Google翻譯正式於英語-拉丁語翻譯中引入Google神經機器翻譯系統，並已於2021年3月前成功將其拓展至所有語言上。Microsoft Translator翻譯是一個免費的個人翻譯應用程序，支持70多種語言，可以翻譯文本、語言、對話、相機照片。DeepL翻譯是一項神經機器翻譯

服務，目前有付費版本以及免費版本，於2017年8月28日推出，此處以免費版本作比較。各項翻譯比較如表1，由於Google支援整份檔案翻譯，速度優於其他翻譯工具，故本研究使用Google翻譯作為本研究的翻譯工具。

表 1：翻譯比較

	Google	Microsoft Translator	DeepL (Free)
優點	1. 翻譯109種語言 2. 免費翻譯 3. 線上翻譯方便，速度快 4. 可以離線翻譯 (app)	1. 翻譯自然度和準確度高 2. 可以離線翻譯 (app) 3. 簡單易用	1. 翻譯自然度和準確度高 2. 免費翻譯 3. 可以離線翻譯 (電腦版) 4. 簡單易用
缺點	1. 翻譯不自然 2. 翻譯文法參差不齊	1. 60種語言	1. 翻譯語言種類較少 (24種) 2. 不支援語音或圖像翻譯

步驟二、資料前處理及數據增強前置作業：刪去多餘，意指刪去csv中的表情符號，表情符號來自日本，日文是"繪文字"，而其日文羅馬字拼法就是"Emoji"。第一個表情符號是由日本的栗田穰崇在1999年左右創造出來，到現在隨著社群平台的發展，Emoji越來越普及，已經成了新世代的共通語言。本研究使用Python的emoji庫來過濾多餘的表情符號，如圖2所示。接著訓練Word2Vec詞向量模型，於後續章節詳細介紹。

```

import emoji
import re
test_str = "下雪天❄️❄️, hotpot最适合不过, 两盘毛肚进肚, 整个人都舒服咻啦! 暖暖哒🔥"
test_str = emoji.demojize(test_str)
test_str = re.sub(':\S+?:', '', test_str)
print(test_str)

```

❏ 下雪天, hotpot最适合不过, 两盘毛肚进肚, 整个人都舒服咻啦! 暖暖哒

圖 3：去除Emoji示意圖

步驟三、針對較少類別數據增強：將情緒較少的類別做數據增強：增強數據集，意指針對數量較少的類別做數據增強，本研究使用NLPPC 2018 Shared Task1 資料集作為研究對象，此資料集共有6000筆資料，大小約為2.5MB，每筆資料可能包含零、一個或多個情緒，這些情緒數量不平均，此外，本研究將無任何情續獨立出來取名為「None」，如表2所示。本研究參考了Jason Wei等人（2019）提出的簡單數據增強技術（EDA: Easy Data Augmentation），使用同義詞替換方法增強數據集。針對數量較少的情緒做增強，透過隨機抽樣的方式將各項情緒平均，如表3所示，接著使用Word2Vec訓練中文詞向量將重複的句子做同義詞的替換。

表 2：數據增強前各項情緒統計

	Happiness	Sadness	Anger	Fear	Surprise	None
資料筆數	1,824	1,086	571	648	651	1,896
百分比	30.4%	18.1%	9.5%	10.8%	10.9%	31.6%

表 3：數據增強後各項情緒統計

	Happiness	Sadness	Anger	Fear	Surprise	None
資料筆數	1,824	1,629	1,713	1,944	1,953	1,896
百分比	17.7%	15.8%	16.6%	18.9%	19.0%	18.4%

步驟四、訓練分類模型並分類結果：建立分類器，訓練模型建立文本分類器，本研究使用Yiming Cui等人（2019）提出的預訓練模型BERT-wwm-ext（Whole Word Masking for Chinese BERT）訓練情緒分類器，wwm（Whole Word Masking）稱全詞Mask或整詞Mask，是Google在2019年發布的一項BERT的升級版本，主要更改了原預訓練階段的訓練樣本生成策略。原始的BERT-base, Chinese中，沒有考慮



傳統NLP中的中文分詞（CWS）。BERT-wwm-ext使用中文維基百科（包括簡體和繁體）進行訓練，在分類任務中有較高的準確度。故本研究使用BERT-wwm-ext作為預訓練模型。

步驟四、訓練分類模型並分類結果：評估分類器，評估分類器分數，本研究主要用準確率（Accuracy）、精確率（Precision）、召回率（Recall）、以及F1-Score來評估模型結果，分別對使用同義詞增強方法替換單一詞彙、雙詞彙以及多詞彙還有與單純透過翻譯來增強數據集做比較其差異。

## 二、同義詞替換方法

本研究使用Gensim套件以及Word2Vec訓練將文字轉成向量，以Wiki上的中文資料為訓練語料，選擇2021年11月20日時間點的資料，資料檔名為zhwiki-20211120-pages-articles-multistream.xml.bz2，檔案大小約2.3GB，接著，使用Gensim裡面的一個函式WikiCorpus從Wiki的資料中提取乾淨的文章出來，再將標點符號的地方以空白隔開，如圖5所示。使用Jieba斷詞，如圖7所示。透過Gensim訓練 Word2Vec 模型，以開心一詞測試模型結果，如圖10所示。最後透過隨機抽樣方法，做同義詞替換增強資料集。為了保留原本句子結構與隨機替換和多次翻譯不同的是針對特定字詞依照Word2Vec模型相似度分別做單、雙以及多詞彙的替換，如圖11所示。

```
#import WikiCorpus套件
from gensim.corpora import WikiCorpus

#匯入下載下來的維基百科資料
wiki_corpus = WikiCorpus('zhwiki-20211120-pages-articles-multistream.xml.bz2', dictionary={})
#初始化text_num
text_num = 0

#開始做資料前處理，將處理完的資料存成wiki_text.txt
with open('wiki_text.txt', 'w', encoding='utf-8') as f:
    for text in wiki_corpus.get_texts():
        f.write(' '.join(text)+'\n')
        #每處理完一筆text_num+1
        text_num += 1
        #每當text_num是1萬的倍數就印出來，方便查看進度
        if text_num % 10000 == 0:
            print('{} articles processed.'.format(text_num))
#全部執行完印出來，可看出資料量有多少
print('{} articles processed.'.format(text_num))
```

圖 4：使用WikiCorpus 套件提取文章之虛擬碼

歐幾里得 西元前三世紀的古希臘數學家 現在被認為是幾何之父 此畫為拉斐爾的作品 雅典學院 数学 是利用符号语言研究數量 结构 变化以及空间等概念的一門学科 ...

圖 5：WikiCorpus提取文章

上述圖4為使用WikiCorpus套件來對維基百科的資料集做數據的提取，提取出來乾淨的文本儲存成 wiki\_text.txt 這份檔案，另外將標點符號的地方用空白隔開，畢竟可能會影響到段詞的結果，提取出的檔案內容如上圖5。

```

#import Jieba斷詞工具，以及OpenCC套件
import jieba
from opencc import OpenCC

# OpenCC參數s2t為簡體轉換繁體
cc = OpenCC('s2t')

# 開始做簡體轉換繁體的工作，將完成的檔案存成wiki_text_seg.txt
with open('wiki_text_seg.txt', 'w', encoding='utf-8') as new_f:
    #打開上一步驟處理完的維基百科檔案wiki_text.txt
    with open('wiki_text.txt', 'r', encoding='utf-8') as f:
        for times, data in enumerate(f, 1):
            print('data num:', times)
            #簡體、繁體轉換
            data = cc.convert(data)
            #Jieba斷詞
            data = jieba.cut(data)
            #為了明顯看出斷詞，這邊使用換行來方便查看
            data = [word for word in data if word != ' ']
            data = ' '.join(data)

            new_f.write(data)

```

圖 6：Jieba斷詞之虛擬碼

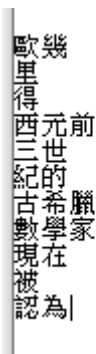


圖 7：Jieba斷詞結果

上述圖6為Jieba斷詞之程式碼，透過OpenCC這項工具將簡體中文轉為繁體中文，再進行Jieba斷詞。將處理好的文本儲存為wiki\_text\_seg.txt，圖7為處理好的文本結果

```

#import Word2Vec套件
from gensim.models import word2vec

#Word2Vec套件參數設定
#亂數種子
seed = 666
#Skip-gram，選擇了訓練速度較快的CBOW，所以設定為0
sg = 0
#周圍詞彙要看多少範圍，取前10的範圍即可，再往後差異太大就不考慮
window_size = 10
#轉成向量的維度
vector_size = 100
#詞頻少於 min_count 之詞彙不會參與訓練，設定1意味著只要有出現字就會參與訓練
min_count = 1
#訓練的並行數量
workers = 8
#訓練的迭代次數
epochs = 5
#每次給予多少詞彙量訓練
batch_words = 10000

#開始做訓練，匯入上一步驟所斷詞完的檔案
train_data = word2vec.LineSentence('wiki_text_seg.txt')
#模型調整上述設定完的參數
model = word2vec.Word2Vec(
    train_data,
    min_count=min_count,
    vector_size=vector_size,
    workers=workers,
    epochs=epochs,
    window=window_size,
    sg=sg,
    seed=seed,
    batch_words=batch_words
)
#訓練完的模型儲存下來
model.save('word2vec.model')

```

圖 8：Word2vec模型訓練參數虛擬碼

上述圖8為Word2vec模型訓練的參數調整，seed為亂樹種子；sg為Skip-gram，本研究選擇了訓練較快的CBOW故設定為0；window\_size為詞彙的範圍；vector\_size為向量的維度；min\_count代表詞頻少於 min\_count 之詞彙不會參與訓練；workers為訓練的並行數量；epochs為訓練的迭代次數；batch\_words則是每次給予多少詞彙量做訓練。最後將訓練完的模型儲存為word2vec.model。

```
#import Word2Vec
from gensim.models import word2vec
#匯入上一步驟所儲存的模型
model = word2vec.Word2Vec.load('word2vec.model')
#用開心一詞來做模型測試，most_similar為抓出較相似的詞彙
for item in model.wv.most_similar('開心'):
    print(item)
```

圖 9：測試模型結果之虛擬碼

```
('歡笑', 0.7036511301994324)
('高興', 0.6754605770111084)
('歡樂', 0.673014760017395)
('歡喜', 0.6690046191215515)
('感動', 0.6654621362686157)
('興奮', 0.6616406440734863)
('驚喜', 0.658953070640564)
('樂意', 0.6584629416465759)
('驕傲', 0.6558607816696167)
('關心', 0.6552302241325378)
```

圖 10：以開心一詞測試模型結果

上述圖9為測試模型結果的程式碼，首先，先將訓練好的模型讀入，接著以『開心』這個詞彙來做測試，使用most\_similar找出向量最接近的詞彙由於先前訓練參數設定只抓前10筆的資料，故結果顯示10筆，再往後差異太大就不考慮。

虽则一面已很足夠很開心真的你是特別的存在

->虽则一面已很足夠很歡笑真的你是特別的存在

今天嗨了壹下午雖然有點累不過很開心漸漸找到了壹些感覺相信會越來越好

->今天嗨了壹下午雖然有點累不過很高興漸漸找到了壹些感覺相信會越來越好

圖 11：同義詞替換示意圖

### 三、訓練分類模型

本研究使用BERT-wwm-ext預訓練模型進行訓練，將資料集依序完成上述步驟一到步驟三的步驟後分為訓練集、驗證集。訓練集與驗證集比例為90：10，即90%的資料用於訓練。資料集統計結果如下表4所示。

表 4：各項情緒訓練集、驗證集數量（單位：句數）

	Total	Training(90%)	Validation(10%)
Happiness	1,824	1,642	182
Sadness	1,629	1,466	163
Anger	1,713	1,542	171
Fear	1,944	1,750	194
Surprise	1,953	1,758	195
None	1,896	1,706	190
總樣本數	10,959	9,864	1,095

## 四、評估分類器

本論文主要用藉由Pytorch來計算準確率（Accuracy）、精確率（Precision）、召回率（Recall）、以及F1-Score來評估模型結果，以下將依序介紹。一個完整的預測結果會有不同的分類，會有 True Positive(TP)、True Negative(TN)、False Positive(FP) 以及 False Negative(FN) 這四個分類，而這四個分類詳細解說如表5：

- **TP** (True Positives) 正確預測成功的正樣本。
- **TN** (True Negative) 正確預測成功的負樣本。
- **FP** (False Positive) 錯誤預測成正樣本，實際上為負樣本。
- **FN** (False Negative) 錯誤預測成負樣本。

表 5：混淆矩陣(Category)

	實際結果為「是」	實際結果為「否」
預測結果為「是」	<b>TP</b> (True Positives)	<b>FP</b> (False Positives)
預測結果為「否」	<b>TN</b> (True Negative)	<b>FN</b> (False Negative)

- **準確率 (Accuracy)**

在所有情況下，都可以正確判斷真假的比例。計算公式(1)如下式所示：

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

- 精確率 (Precision)

判斷為真的情況下，有多少是真的真。計算公式(2)如下式所示：

$$Precision = \frac{TP}{TP+FP} \quad (2)$$

- 召回率 (Recall)

在真實值是正的所有樣本中，模型預測對的比重，即：模型檢測認為是正類並且實際是正類的部分佔所有確實是正類的比例。計算公式(3)如下式所示：

$$Recall = \frac{TP}{TP+FN} \quad (3)$$

- F1-Score

指標綜合了Precision與Recall的產出的結果。F1-Score的取值範圍從0到1的，1代表模型的輸出最好，0代表模型的輸出結果最差，計算公式(4)如下式所示：

$$F1\ Score = \frac{2*Precision*Recall}{Precision+Recall} \quad (4)$$



## 五、各模型參數設計

本研究採用Word2Vec訓練詞向量模型以及BERT-wwm-ext模型進行情緒分類，本節將詳細講述個模型各參數所使用之數據。

### (一)、Word2Vec

本論文訓練詞向量採用Word2Vec模型，檢測模型時，所使用的相關參數如下表所示；訓練時Word2Vec 有兩種算法，CBOW 以及 Skip-gram，這裡選擇了訓練比較快的 CBOW，故sg為0；window\_size設定為10只需取相似度前10個的詞彙，再往後差異太大；min\_count為1指只要有出現過的詞彙都會參與訓練；workers為了節省訓練的時間，故設定為8；由於維基百科資料集內容龐大，故batch\_words設定為10000，詳細參數整理如表6。

表 6：Word2Vec訓練參數

參數名稱	參數
seed	666
sg	0
window_size	10
vector_size	100
min_count	1
workers	8
epochs	5
batch_words	10000

## (二)、BERT-wwm-ext

本論文訓練BERT模型時，使用的相關參數如下表所示；本節將以BERT模型之基本參數進行說明，以BERT來說，使用BERT-wwm-ext作為一開始預訓練之網路，為配合電腦規格，本實驗將MAX\_LEN設定為200，由於資料集內的句子皆為200字元以內，故作此設定，TRAIN\_BATCH\_SIZE以及VALID\_BATCH\_SIZE接調整為16，用以降低對GPU的負擔，同時EPOCHS為4，LEARNING\_RATE 調整為 1e-05，詳細參數整理如表7。

表 7：BERT-wwm-ext 訓練參數

參數名稱	參數
MAX_LEN	200
TRAIN_BATCH_SIZE	16
VALID_BATCH_SIZE	16
EPOCHS	4
LEARNING_RATE	1e-05

## 六、系統環境

本研究開發環境的作業系統為Windows 10 專業版，處理器規格為Intel(R) Xeon(R) CPU E5-1650 v4 @ 3.60GHz 3.60 GHz，記憶體16GB，硬碟空間約1TB，程式語言的開發工具使用Python3.6，在Google Colab上編譯，以表8系統規格以及表9呈現開發工具。

表 8：系統規格

項目	系統規格
OS(作業系統)	Windows 10 專業版
Processor(處理器)	Intel(R) Xeon(R) CPU E5-1650 v4 @ 3.60GHz 3.60 GHz
Memory(記憶體)	16.0 GB
Disk(硬碟)	1 TB

表 9：開發工具

開發工具	備註
Python	版本3.6
Anacoda3	Jupyter notebook
Google Colab	

## 肆、研究結果

### 一、分類結果

本實驗針對同義詞替換的數據增強方法，做了三種的實驗，再與原始資料以及純透過翻譯增強數據集的方式做比較，得知實驗結果，第一種為單詞彙同義詞替換方法，第二種為雙詞彙同義詞替換方法，第三種為多詞彙同義詞替換方法。以下表10~表14為原始資料、透過翻譯增強以及同義詞替換各情緒的Precision、Recall、F1-Score分數。

表 10：原始資料各情緒Precision、Recall、F1-Score分數

原始資料			
	Precision	Recall	F1-Score
Happiness	0.64	0.74	0.68
Sadness	0.56	0.53	0.55
Anger	0.61	0.51	0.56
Fear	0.50	0.12	0.20
Surprise	0.51	0.18	0.27
None	0.53	0.45	0.50

表 11：翻譯增強各情緒Precision、Recall、F1-Score分數

翻譯增強			
	Precision	Recall	F1-Score
Happiness	0.76	0.74	0.75
Sadness	0.62	0.64	0.63
Anger	0.66	0.64	0.66
Fear	0.53	0.56	0.48
Surprise	0.57	0.58	0.52
None	0.61	0.62	0.60

表 12：單詞彙同義詞替換各情緒Precision、Recall、F1-Score分數

單詞彙替換			
	Precision	Recall	F1-Score
Happiness	0.80	0.81	0.80
Sadness	0.74	0.78	0.76
Anger	0.70	0.68	0.66
Fear	0.78	0.74	0.74
Surprise	0.78	0.70	0.72
None	0.76	0.71	0.70

表 13：雙詞彙同義詞替換各情緒Precision、Recall、F1-Score分數

雙詞彙替換			
	Precision	Recall	F1-Score
Happiness	0.81	0.82	0.79
Sadness	0.74	0.77	0.78
Anger	0.72	0.68	0.69
Fear	0.79	0.72	0.75
Surprise	0.77	0.74	0.70
None	0.78	0.73	0.76

表 14：多詞彙同義詞替換各情緒Precision、Recall、F1-Score分數

多詞彙同義詞替換			
	Precision	Recall	F1-Score
Happiness	0.83	0.85	0.82
Sadness	0.76	0.80	0.75
Anger	0.71	0.73	0.74
Fear	0.80	0.74	0.76
Surprise	0.78	0.76	0.79
None	0.79	0.73	0.76

根據上方實驗結果的表格中，可以看出原始資料由於數據量過少導致結果不理想，透過翻譯來增強數據集的效果有限本研究認為翻譯過程中有可能會造成句

子的結構改變進而影響他的分類結果，另外本研究實驗利用同義詞替換方法來做數據增強能夠有效的提升Precision、Recall、F1-Score分數。明顯可以看出不論是單詞彙、雙詞彙或者是多詞彙的同義詞替換都明顯的比翻譯增強的效果好。

表 15：準確度比較

	原始資料	翻譯增強	單詞彙同義 詞替換	雙詞彙同義 詞替換	多詞彙同義 詞替換
Accuracy	0.51	0.63	0.75	0.76	0.79

從表15可以看出在準確度方面原始資料的準確度大約為51%，透過翻譯增強數據集的準確度大約為63%，同義詞替換方法的準確度大約介於75%~79%左右，多詞彙同義詞替換方法分別高出另外兩個同義詞替換方法3%以及4%。



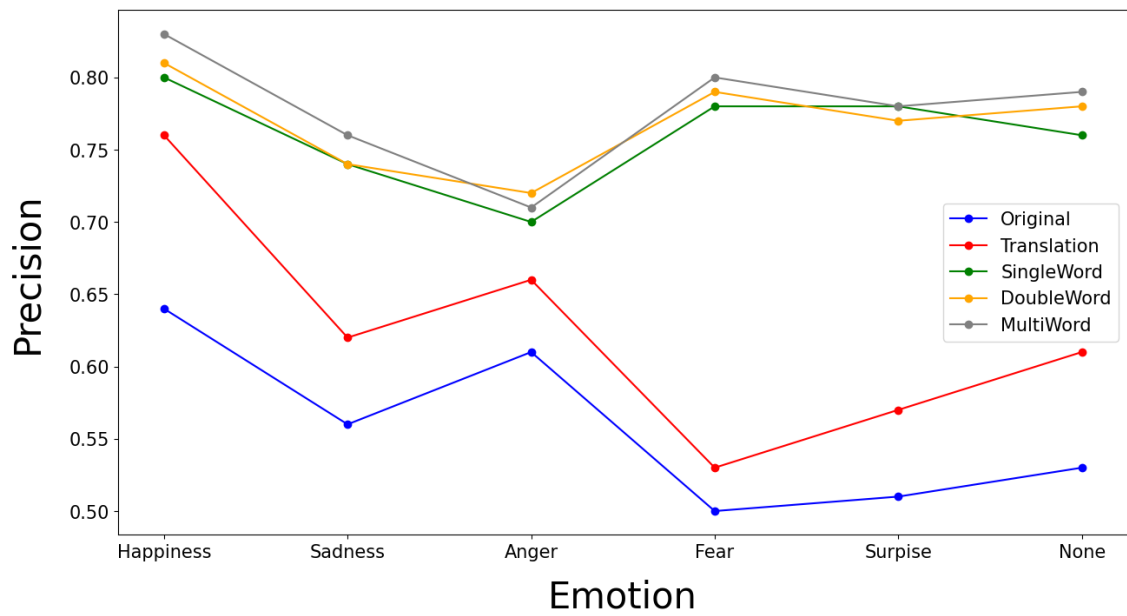


圖 12：各方法Precision分數

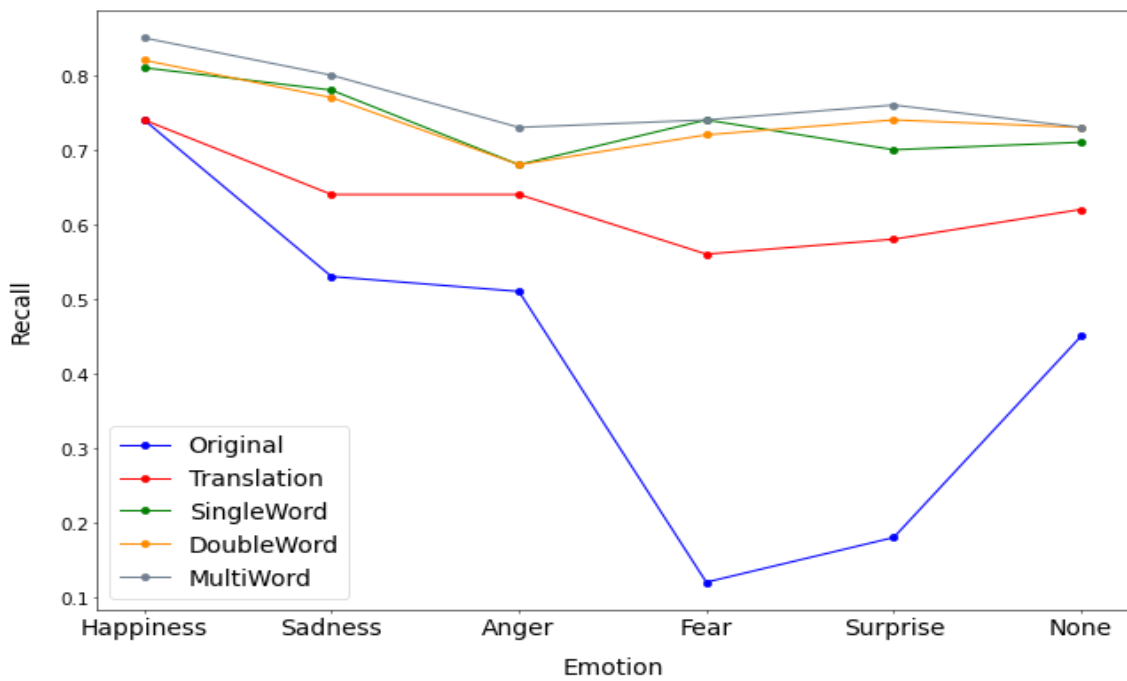


圖 13：各方法Recall分數

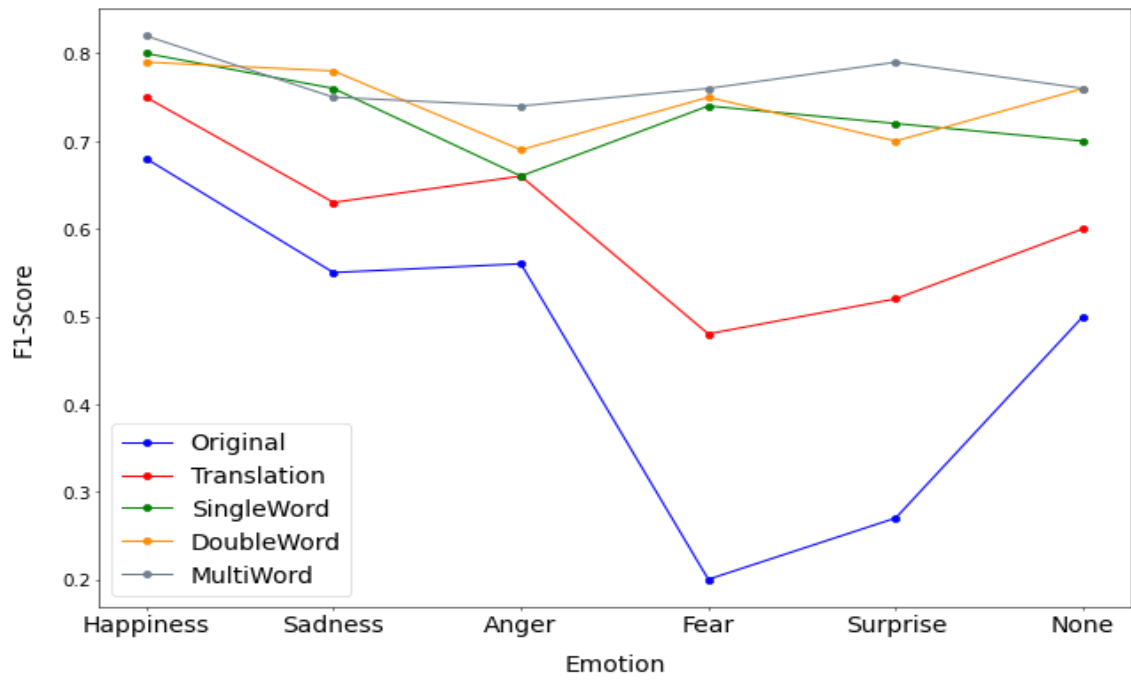


圖 14：各方法F1-Score分數

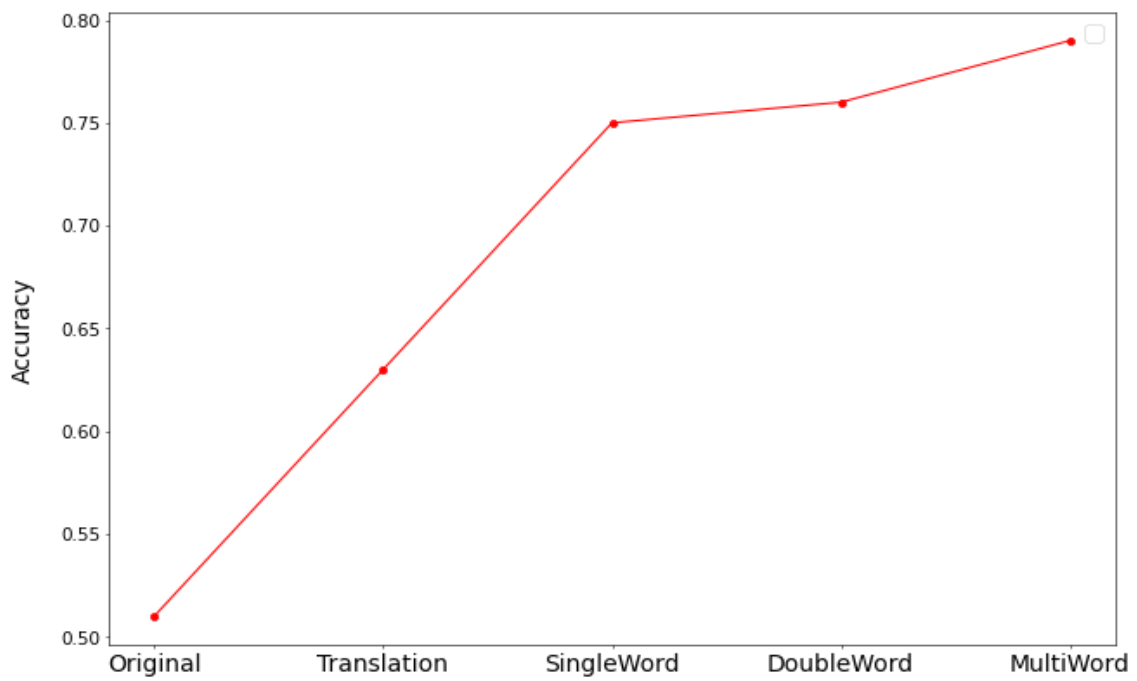


圖 15：各方法準確度

上方圖12~圖14為使用原始資料，翻譯增強以及三種同義詞替換方法增強的Precision、Recall、F1-Score分數，從圖中可以發現，原始資料的Fear、Surprise分數都偏低，由於樣本各項情緒不平均，Fear以及Surprise的樣本數過少，所以導致結果不佳。圖15為各項方法分類的準確度，表現最好的是多詞彙的同義詞替換，準確度高達79%，其次為雙詞彙的同義詞替換為76%。

## 二、研究探討

本研究在實驗的過程中遇到某些問題，並且針對這些問題，透過閱讀相關文獻後，找出其解決方式：

自然語言處理與中文計算國際會議(NLPCC)提供的NLPCC 2018 Shared Task1資料集，資料不完整，文件內缺少父節點導致讀取資料時發生錯誤訊息，如圖16為XML節點之錯誤訊息，解決的方法為以手動的方式自行新增父節點。解決問題後，可接續本研究的資料前處理，得以讓研究順利進行下去。

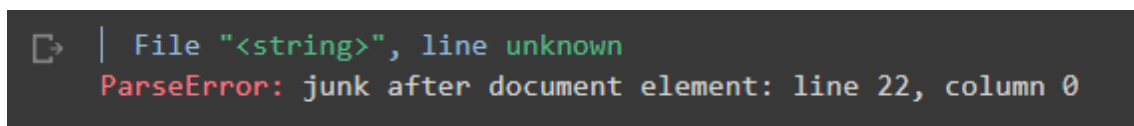


圖 16：節點格式之錯誤訊息

### 三、研究限制

本研究使用了由自然語言處理與中文計算國際會議(NLPCC)提供的NLPCC 2018 Shared Task1資料集，資料集內容為多國語言混雜而成，此資料集內容為一個句子對應一個或多個情緒，而本研究為了增加其分類準確度，故將文本統一，其中以中文佔的比例最高，所以本實驗將資料集只探討中文，以中文的資料集作為實驗對象。

在NLPCC 2018 Shared Task1資料集中，資料筆數為6000筆，即使透過數據增強增加資料量，效果也相當有限。使用自行建立之資料集也有一定的難度，一個句子所對應的情緒沒有足夠的判斷標準，且可能容易加入主觀認知，導致判斷結果易有偏差，舉例來說，像是僅能分辨是正向或負向。本論文採用之NLPCC 2018 Shared Task1資料集中，其情緒類別分為五類，分別為Happiness(快樂)、Sadness(悲傷)、Anger(憤怒)、Fear(恐懼)、Surprise(驚訝)，所以只針對這五個情緒類別進行探討。

## 伍、結論與未來研究

由於現今網路上評論過於龐大，要獲得大量的評論相當花費時間成本，如何在數據有限的情況下，還能夠有高準確率的情感辨識分類，為此，本研究提出三種方式，利用同義詞替換方法進行單詞彙、雙詞彙以及多詞彙的同義詞替換來增強數據集，並使用Accuracy、Precision、Recall、F1-Score等值，作為實驗數據。以同義詞替換的三種方式與翻譯增強以及原始資料來做分類比較，其中，可以看出三種同義詞替換的各項實驗數據都優於翻譯增強。Accuracy分別為75%、76%以及79%又以多詞彙的Accuracy最高，因此本研究認為多詞彙同義詞替換方法最佳。

由於本實驗只針對中文的資料集作為實驗對象，並未使用其他語言進行情感辨識分類，因此認為未來可以利用此方法或提出其他不同的數據增強方法在不同語言的各式資料集做嘗試。在情緒類別方面，認為未來也可以加入更多的情緒，不僅只有5類。

## 陸、參考文獻

- [1] A. Bhardwaj, Y. Narayan, Vanraj, Pawan&M. Dutta, “Sentiment Analysis for Indian Stock Market Prediction Using Sensex and Nifty” , *Procedia Computer Science*, vol. 70, pp. 85 – 91, Jan. 2015, doi: 10.1016/j.procs.2015.10.043.
- [2] T. Chen, R. Xu, Y. He&X. Wang, “Improving sentiment analysis via sentence type classification using BiLSTM-CRF and CNN” , *Expert Systems with Applications*, vol. 72, pp. 221 – 230, Apr. 2017, doi: 10.1016/j.eswa.2016.10.065.
- [3] T. U. Haque, N. N. Saber&F. M. Shah, “Sentiment analysis on large scale Amazon product reviews” , 2018 IEEE International Conference on Innovative Research and Development (ICIRD), May. 2018, pp. 1 – 6. doi: 10.1109/ICIRD.2018.8376299.
- [4] A. Agarwal, A. Yadav&D. Vishwakarma, “Multimodal Sentiment Analysis via RNN variants” , 2019 IEEE International Conference on Big Data, Cloud Computing, Data Science & Engineering (BCD), 2019, doi: 10.1109/BCD.2019.8885108.
- [5] F. Chen, R. Ji, J. Su, D. Cao&Y. Gao, “Predicting Microblog Sentiments via Weakly Supervised Multimodal Deep Learning” , *Trans. Multi.*, vol. 20, no. 4, pp. 997 – 1007, Apr. 2018, doi: 10.1109/TMM.2017.2757769.
- [6] A. Yadav&D. K. Vishwakarma, “Sentiment analysis using deep learning architectures: a review” , *Artif Intell Rev*, vol. 53, no. 6, pp. 4335 – 4385, Aug. 2020, doi: 10.1007/s10462-019-09794-5.
- [7] H. Peng, E. Cambria&A. Hussain, “A Review of Sentiment Analysis Research in Chinese Language” , *Cogn Comput*, vol. 9, no. 4, pp. 423 – 435, Aug. 2017, doi: 10.1007/s12559-017-9470-8.

- [8] S. De Nadai et al., “Enhancing safety of transport by road by on-line monitoring of driver emotions” , 2016 11th System of Systems Engineering Conference (SoSE), Jun. 2016, pp. 1 – 4. doi: 10.1109/SYSOSE.2016.7542941.
- [9] B. Verschuere, G. Crombez, E. H. W. Koster,&K. Uzieblo, “Psychopathy and Physiological Detection of Concealed Information: A review” , PSYCHOL BELG, vol. 46, no. 1 – 2, pp. 99, Mar. 2006, doi: 10.5334/pb-46-1-2-99.
- [10] R. A. Khalil, E. Jones, M. I. Babar, T. Jan, M. H. Zafar&T. Alhusain, “Speech emotion recognition using deep learning techniques: A review” , IEEE Access, vol. 7, pp. 117327 – 117345, 2019.
- [11] F. Noroozi, C. A. Corneanu, D. Kamińska, T. Sapiński, S. Escalera&G. Anbarjafari, “Survey on Emotional Body Gesture Recognition” , IEEE Transactions on Affective Computing, vol. 12, no. 2, pp. 505 – 523, Apr. 2021, doi: 10.1109/TAFFC.2018.2874986.
- [12] Z. Shen, J. Cheng, X. Hu&Q. Dong, “Emotion Recognition Based on Multi-View Body Gestures” , 2019 IEEE International Conference on Image Processing (ICIP), Sep. 2019, pp. 3317 – 3321. doi: 10.1109/ICIP.2019.8803460.
- [13] S. Singh, V. Sharma, K. Jain&R. Bhall, “EDBL – algorithm for detection and analysis of emotion using body language” , 2015 1st International Conference on Next Generation Computing Technologies (NGCT), Sep. 2015, pp. 820 – 823. doi: 10.1109/NGCT.2015.7375234.
- [14] B. C. Ko, “A brief review of facial emotion recognition based on visual information” , sensors, vol. 18, no. 2, pp. 401, 2018.
- [15] S.-H. Wang, P. Phillips, Z.-C. Dong&Y.-D. Zhang, “Intelligent facial emotion recognition based on stationary wavelet entropy and Jaya a

lgorithm” , Neurocomputing, vol. 272, pp. 668 – 676, Jan. 2018, doi: 10.1016/j.neucom.2017.08.015.

[16] L. Shu et al., “A review of emotion recognition using physiological signals” , Sensors, vol. 18, no. 7, pp. 2074, 2018.

[17] N. Alswaidan&M. E. B. Menai, “A survey of state-of-the-art approaches for emotion recognition in text” , Knowl. Inf. Syst., vol. 62, no. 8, pp. 2937 – 2987, Aug. 2020, doi: 10.1007/s10115-020-01449-0.

[18] E. Batbaatar, M. Li&K. H. Ryu, “Semantic-emotion neural network for emotion recognition from text” , IEEE Access, vol. 7, pp. 111866 – 111878, 2019.

[19] Z. Erenel, O. R. Adegboye&H. Kusetogullari, “A New Feature Selection Scheme for Emotion Recognition from Text” , Appl. Sci.-Basel, vol. 10, no. 15, pp. 5351, Aug. 2020, doi: 10.3390/app10155351.

[20] J. Moorkens, A. Toral, S. Castilho&A. Way, “Translators’ perceptions of literary post-editing using statistical and neural machine translation” , Transl. Spaces, vol. 7, no. 2, pp. 240 – 262, Nov. 2018, doi: 10.1075/ts.18014.moo.

[21] Y. Wu et al., “Google’ s Neural Machine Translation System: Bridging the Gap between Human and Machine Translation” , arXiv:1609.08144 [cs], Oct. 2016, Accessed: Dec. 25, 2021. [Online]. Available: <http://arxiv.org/abs/1609.08144>

[22] J. Wei&K. Zou, “EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks” , Jan. 2019, Accessed: Dec. 26, 2021. [Online]. Available: <https://arxiv.org/abs/1901.11196v2>

[23] T. Mikolov, K. Chen, G. Corrado&J. Dean, “Efficient Estimation of Word Representations in Vector Space” , arXiv:1301.3781 [cs], Sep. 2013, Accessed: Dec. 25, 2021. [Online]. Available: <http://arxiv.org/abs/1301.3781>



- [24] Y. Cui, W. Che, T. Liu, B. Qin&Z. Yang, “Pre-Training with Whole Word Masking for Chinese BERT” , IEEE/ACM Trans. Audio Speech Lang. Process., vol. 29, pp. 3504 – 3514, 2021, doi: 10.1109/TASLP.2021.3124365.
- [25] Y. Zhao, N. Jiang, W. Sun&X. Wan, “Overview of the NLPCC 2018 Shared Task: Grammatical Error Correction” , Natural Language Processing and Chinese Computing, vol. 11109, M. Zhang, V. Ng, D. Zhao, S. Li&H. Zan, Edit: Cham: Springer International Publishing, 2018, pp. 439 – 445. doi: 10.1007/978-3-319-99501-4\_41.