

國立中興大學電機工程學系

碩士學位論文

基於 BERT 的情緒分析應用於短句自動審查

Sentiment analysis base on BERT for automatic  
censoring of short sentences

國立中興大學

NATIONAL CHUNG HSING UNIVERSITY

指導教授：林維亮      Wei-Liang Lin

研 究 生：何宗諭      Zong-Yu He

中華民國一百一十一年八月

國立中興大學電機工程學系

碩士學位論文

題目(中文)： 基於 BERT 的情緒分析應用於短句自動審查

題目(英文)： Sentiment analysis base on BERT for automatic censoring  
of short sentences

姓名： 何宗諭 學號： 7109064351

經 口 試 通 過 特 此 證 明

論文指導教授

林維亮

論文考試委員

林維亮

黃穎聰

陳封岳

中 華 民 國

111 年

07 月

26 日

## 誌謝辭

從考上研究所到現在，時間也過去了兩年，真的過得很快，原本我以為會很漫長的，現在想想這或許也是過得很充實的一種體現。

回想起來這個過程，從一開始修課，再接射月計畫，老實說對我來說蠻有挑戰的，因為我剛來的時候程度並不好，很多事情都要花比別人多時間去想辦法，或者是請教實驗室的同學，不過想想成長總是伴隨痛苦，還是有得到一些心得和方法，這段時間至少比我在大學四年還有更多意義。同時也很幸運地，有一幫很友善的同學，願意在我迷惘的時候分享自己的經驗或者交換意見，一起渡過難關。

最後特別感謝老師的指導，meeting 的時候願意聽我的報告，給我方向讓我知道該做什麼，對我幫助很大。



國立中興大學

NATIONAL CHUNG HSING UNIVERSITY

## 摘要

隨著資訊傳播的發展，越來越多的資訊會出現在生活中，如何判斷資訊是有意義或是惡意是非常重要的問題。NLP 的技術，讓機器能夠理解人類的語言，我們可以藉由其中情緒分析的應用，過濾掉不適當的言論，使以往需要透過人工審核的部分能夠由機器接手，提升審查的效率。

本文使用 BERT 架構做中文短句的情緒判斷，目標是用於審核廣告詞、廣告看板或留言板，過濾掉不適當的言論。透過蒐集網路上的各種言論，當中除了包含網購評論、餐點評論、店商評論等等，還有情緒字典等等資料去訓練演算法，讓演算法正確判別一般短句的情緒，除此之外，也設計了新的架構做判斷，解決對一些帶有諷刺性的語句的判斷問題。

關鍵字：深度學習、NLP、BERT、情緒分析、中文短句。

國立中興大學

NATIONAL CHUNG HSING UNIVERSITY

# Abstract

With the development of information dissemination, more and more information will appear in life. How to judge whether the information is meaningful or malicious is a very important issue. NLP technology enables machines to understand human language. We can filter out inappropriate speeches through the application of sentiment analysis, so that the parts that used to be reviewed by humans can be taken over by machines, improving the efficiency of review.

This paper uses the BERT architecture to make emotional judgments of Chinese short sentences. The goal is to review advertising words, advertising boards or message boards, and filter out inappropriate speeches. By collecting various comments on the Internet, including online shopping reviews, meal reviews, store reviews, etc., as well as sentiment dictionary and other data to train the algorithm, the author let the algorithm correctly identify the emotions of ordinary short sentences. In addition, a new structure is also designed to make judgments to solve some sarcastic sentences.

Keywords: Deep learning 、 NLP 、 BERT 、 Sentiment analysis 、 Chinese short sentences

# 目錄

摘要 .....	i
Abstract.....	ii
目錄 .....	iii
表目次 .....	iv
圖目次 .....	v
第一章 序論 .....	1
第二章 基礎知識及相關文獻探討 .....	2
2.1 Natural Language Processing .....	2
2.2 RNN(Recurrent Neural Network)[9] .....	4
2.3 LSTM(Long Short-Term Memory)[8] .....	4
2.4 Transformer .....	5
2.4.1 self-attention[7] .....	6
2.4.2 Multi-Head Attention .....	7
2.5 BERT .....	8
2.5.1 輸入的處理 .....	9
2.5.2 Pre-train .....	10
2.5.3 下游任務(downstream task) .....	11
2.6 中文的情緒分析 .....	14
第三章 蒐集資料 .....	17
3.1 網路留言與評論 .....	18
3.2 情緒字典 .....	18
第四章 架構和模型訓練 .....	19
4.1 整體架構 .....	19
4.2 模型訓練[13] .....	20
4.2.1 輸入前處理 .....	20
4.2.2 訓練[13] .....	21
第五章 研究結果與比較 .....	23
第六章 結論與未來展望 .....	28
6.1 結論 .....	28
6.2 未來展望 .....	28
第七章 參考文獻 .....	29

## 表目次

表 1 使用參數 .....	23
表 2 有斷詞處理混淆矩陣 .....	25
表 3 未斷詞的混淆矩陣 .....	25
表 4 卓騰語言科技的混淆矩陣 .....	25
表 5 精確度比較 .....	26
表 6 諷刺語句 .....	26



國立中興大學

NATIONAL CHUNG HSING UNIVERSITY

## 圖目次

圖 1 情緒分析[14]	1
圖 2 NLP 和核心任務	2
圖 3 NLP 處理流程	3
圖 4 RNN model	4
圖 5 LSTM 架構	5
圖 6 Transformer 架構[6]	6
圖 7 self-attention 運算過程	7
圖 8 Multi-head-attention 計算過程	8
圖 9 BERT 的階段	9
圖 10 BERT 輸入的處理	9
圖 11 Mask LM model	10
圖 12 mask 的做法	10
圖 13 Next sentence prediction	11
圖 14 單句分類[12]	12
圖 15 Q & A 問答系統[12]	12
圖 16 詞性標註[12]	13
圖 17 因果關係判斷[12]	14
圖 18 SINM 架構[19]	15
圖 19 情緒分佈示例[19]	15
圖 20 改良版的 CNN 架構[20]	16
圖 21 整理後的資料	17
圖 22 網路留言資料集	18
圖 23 所提架構	19
圖 24 訓練架構	20
圖 25 輸入的轉換過程	21
圖 26 維度變化	22
圖 27 訓練集和驗證集準確度	23
圖 28 判斷正面的結果	24
圖 29 判斷負面的結果	24
圖 30 網路語錄	27



# 第一章 序論

隨著時代的進步，越來越多廣告的手段出現在市面上，像是公車的紋身、路上的電子看板、甚至是網路留言板等等，這些都是傳達資訊的方法，差別在於平台的不同。假設今天我們是處於廣告業者的立場，接到了訂單需要幫別人刊登廣告，但是手上可能有上百上千條資料，必須判斷哪些適合刊登，那些不適合刊登，才不會出問題，像這種非常大的資料量，人工審核非常困難，如果這個時候有一個情緒分析系統可以直接幫使用者做好分類，就非常有幫助。

情緒分析，目前有很多研究人員投入時間和精力投入研發，應用的地方有很多[1][2][3]，甚至是 google、微軟都有投入研發，在台灣也有一些新創公司在研究。實際在商業上的應用也有很多，像是某些大企業會蒐集大量的產品意見回饋，然後用相關的情緒分析 API 得到一些數據，用以這些數據當作指標，評估未來是否需要做更进一步的改進。也可以使用在股市，像是風險情緒分析，用來確定投資者對某一特定股票或資產的意見，這些資訊可能會暗示未來的價格走勢。

本文希望研究一個用在中文且不限定話題的情緒分析系統，基於 BERT[12] 的架構結合情緒字典還有諷刺言論的判斷，應用在馬路上的電子看板、廣告招牌、或是網路上的留言板等等各種透過廣告詞傳播資訊的平台，將不適合的語句自動過濾，取代人工審查的負擔。

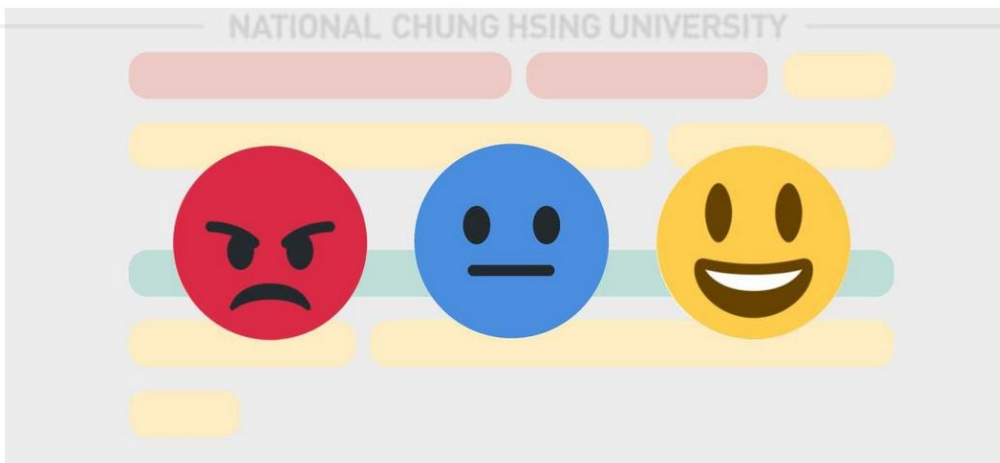


圖 1 情緒分析[14]

## 第二章 基礎知識及相關文獻探討

### 2.1 Natural Language Processing

Natural language processing 的中文是自然語言處理，一般會簡稱為 NLP[11]，主要目的是讓電腦能夠理解人類的語言表達意思。

在技術的早期是基於類似統計的概念去訓練 model，不能很好的辨識複雜的句子和理解上下文關係，但隨著深度學習的發展有很大程度改善了早期的問題，其中誕生了一些有名的網路架構，像是 RNN、LSTM、Transformer 和 BERT 等等，它們的出現讓先前的任務取得更好的成果，至今仍然在發展當中。

它主要有兩個核心的任務，自然語言理解 (Natural Language Understanding, NLU) 和自然語言生成 (Natural Language Generation, NLG)，如圖 2

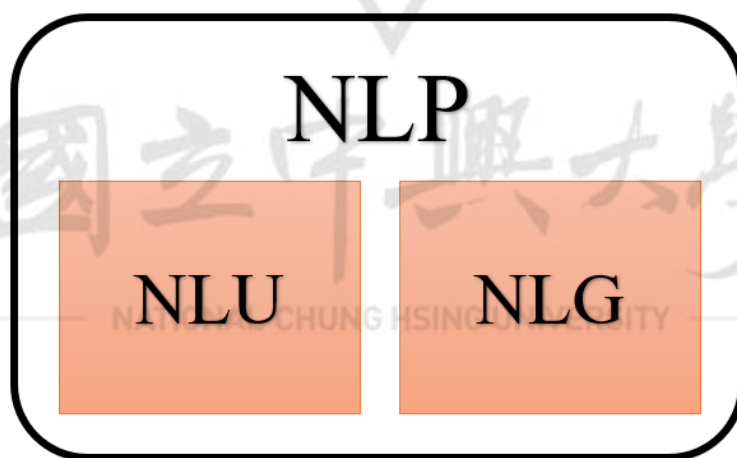


圖 2 NLP 和核心任務

NLU 是指希望機器像人一樣可以具備理解語言的能力，像是做情緒分析或翻譯。想要完成這個事情並不是很簡單，扣除有數十種語言的多樣性外，還有要知道正確的意思，往往需要先理解上下文之間的關係。

NLG 是指將非語言的資料數據，轉換成人類能夠理解的格式，常見應用像是生成文章或聊天機器人。會遇到的難點是需要考量語言是開放性的，有

千百萬種排列組合，有時候甚至沒有任何規律。

要讓機器認識語言前，要先將文字化成數字的向量讓機器可以計算，方法有 one-hot encoding 和 word embedding。以前常用 one-hot encoding，作法是先建立一個包含所有字的字典，如果這個字典有 3000 字，要表達某個字，會以(3000, 1)維的方式表達，其中內容都是由 2999 個 0 和 1 個 1 組成，好處是簡單直接，壞處是這種稀疏資訊的方式對深度學習效果不太好，且若是多義字可能就無法表示。目前主流是 word embedding，是將 one-hot encoding 向量映射到空間中，在這個空間內，意義越是有關連的詞語，彼此之間會越是靠近，好處是更好表達詞語之間的關係，以及計算上維度遠低於 one-hot encoding。

NLP 的流程可以參考圖 3，流行的應用有很多，例如問答系統，還有翻譯，像是中翻英或英翻中，詞性標註、語音辨認、文字的分類，也可以是情緒分析和垃圾信件分辨。除了文字外，現今的技術甚至可以應用在影像上面[4]，例如搭配 CNN[10]取出影像特徵後，能夠讓機器可以描述出這個影像的訊息。

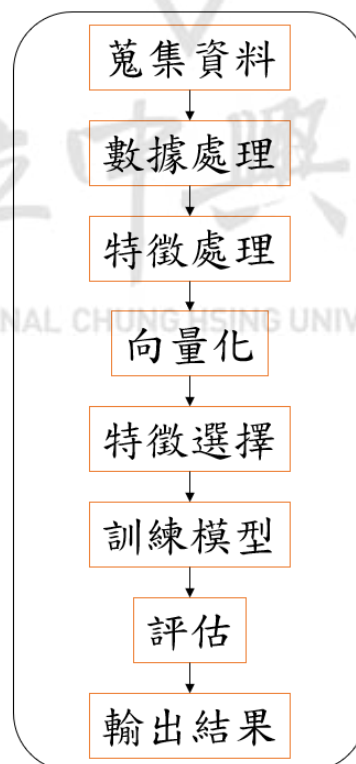


圖 3 NLP 處理流程

## 2.2 RNN(Recurrent Neural Network)[9]

RNN 中文稱為遞迴神經網路(圖 4)，是一種常出現在 NLP 領域的網路架構，是一種 sequence to sequence model，可用來實現聊天機器人或是做語言的上下文預測。

它的出現解決了以往神經網路在輸入之後，無法考慮前面輸出的影響，過程是基於在當前輸入端額外加上前一時刻的 hidden layers 的輸出，而前一時刻的輸出又考慮了前前時刻的 hidden layers 輸出，參考式 1、式 2。

$$h_t = W \cdot h_{t-1} + U \cdot x_t + b \quad \text{式 1}$$

$$y_t = g(V \cdot h_t) \quad \text{式 2}$$

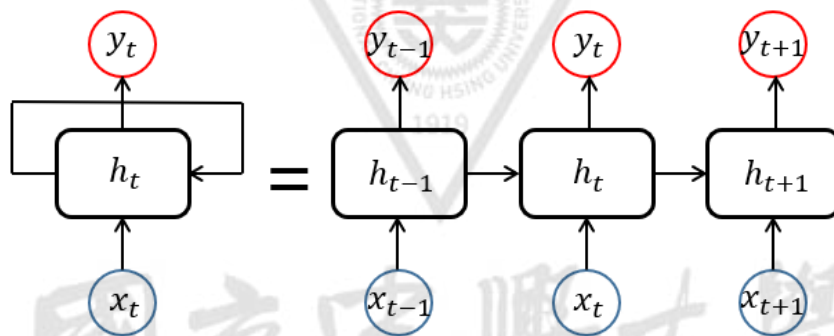


圖 4 RNN model

儘管 RNN 對於 sequence 之間的關係有更好的理解，但還是有一些問題存在，像如果不被告知要結束為止，RNN 會持續的循環，像這樣持續地疊加就會有梯度爆炸(Gradient vanishing)或梯度消失(Gradient exploded)。

## 2.3 LSTM(Long Short-Term Memory)[8]

LSTM 是變形的 RNN，如圖 5，是用來改善 RNN 的梯度爆炸和梯度消失問題，讓它能夠更長的輸入 sequence 表現，對於 RNN 只考量上一時刻的 hidden layer 數值  $h_t$ ，LSTM 還會多考量一個狀態  $c_t$ (cell state)。

它的過程主要多了控制 input gate、forget gate、output gate 的動作，還有

memory cell 暫時儲存狀態。其中 input gate 負責有多少資料該進入 memory cell 儲存，forget gate 負責是否放棄一些不重要的資訊，或是繼續保留，最後 output gate 負責有多少資料要輸出，3 個 gate 的動作可以透過後續訓練做出最佳化。

作為 RNN 的改良版，LSTM 的推出無疑是改善了梯度的消失和爆炸，也可以更好的保留更長的資訊，不過缺點是訓練時間過長。

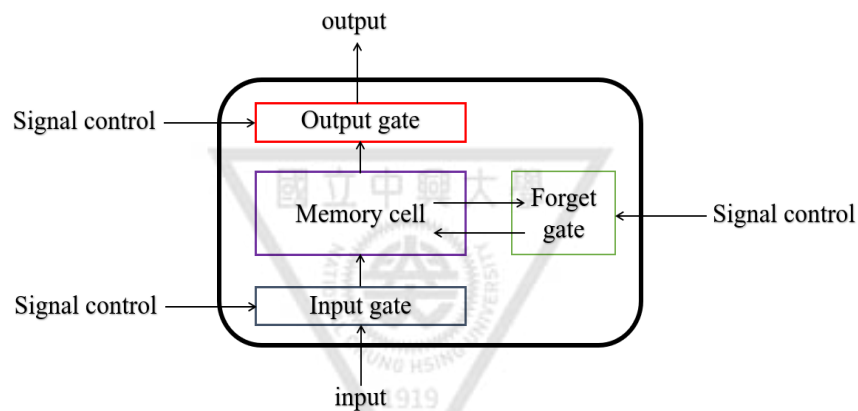


圖 5 LSTM 架構

## 2.4 Transformer

Transformer[6]是一種常被用於文字翻譯的 sequence to sequence model，主要包含了兩個內部結構，encoder 和 decoder，如圖 6。

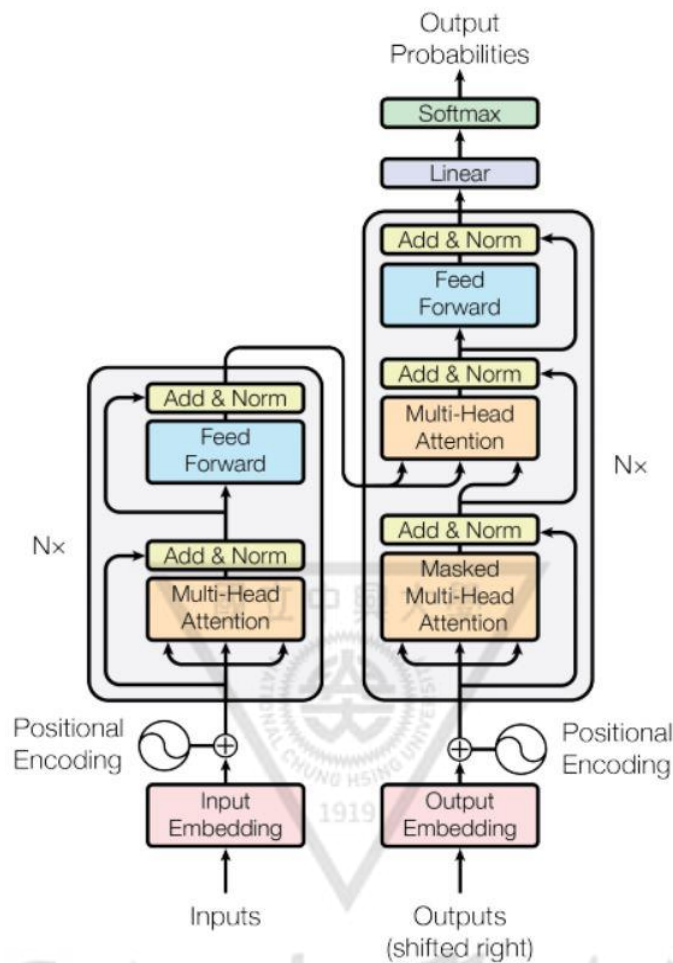


圖 6 Transformer 架構[6]

### 2.4.1 self-attention[7]

讓 Transformer 變得比其他架構更流行、泛用、甚至是取代的原因，在於 self-attention 的方法，主要是基於 attention[5] 的改良。這個方法有別於以往像是 RNN、LSTM 會根據輸入的位置，在運算時會參考前一時刻的結果，造成輸入 sequence 較多時容易發生資訊的損失和缺乏平行運算能力，self-attention 最大的好處是能夠平行運算，也改善了資訊的損失，讓機器更能理解上下文之間關係。

而 self-attention 在 Transformer 輸入端的處理，會讓 input sequence  $x_{1 \sim 4}$  做 embedding 使其向量化，再與 Positional Encoding 相加(式 1) 可以得到  $a_{1 \sim 4}$ ，接著分別乘上不同的權重( $w_q$ 、 $w_k$ 、 $w_v$ )，得到  $Q$ (Query)、 $K$ (Key)、 $V$ (Value) 矩陣，其中  $q$  是用來和  $v$  匹配， $k$  是被匹配， $v$  是要提取的資訊。



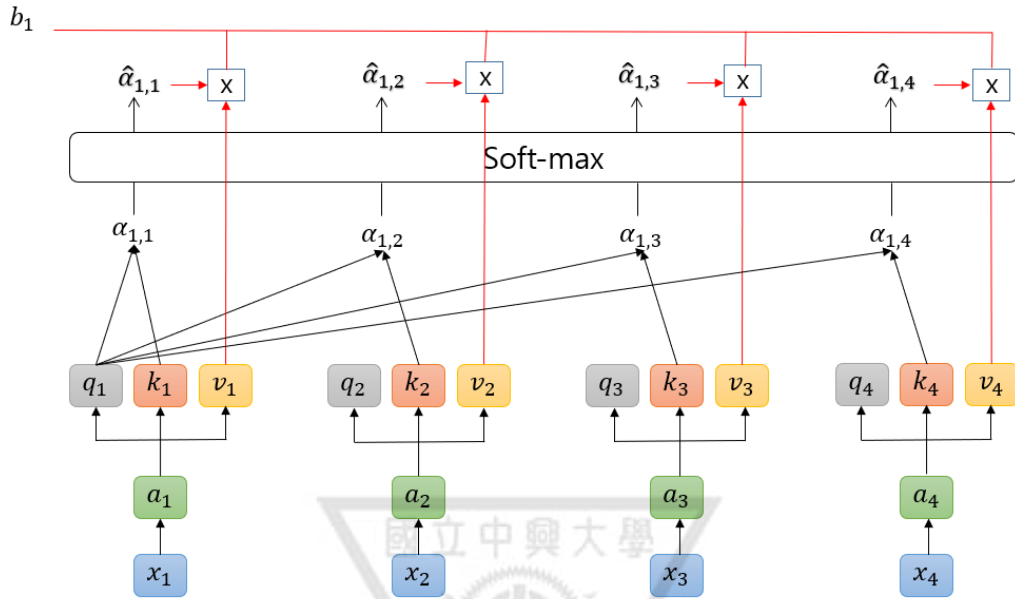


圖 7 self-attention 運算過程

之後運算過程如圖 7，用 Q 對每個 K 做內積得到  $\alpha$ ，如式 1，可以知道匹配的程度，再做 soft-max 算出  $\hat{\alpha}$ ，接著與 v 相乘可以算出輸出 b，如式 2。

$$\alpha_{1,i} = \frac{q_i \cdot k_i}{\sqrt{d}} \quad \text{式 3}$$

$$b = \text{softmax}\left(\frac{q_i \cdot k_i}{\sqrt{d}}\right) * v \quad \text{式 4}$$

## 2.4.2 Multi-Head Attention

Multi-head-attention 的運算跟 self-attention 一樣，差別在 q、k、v 有乘以維度不同的權重(w)，投射成不同維度的向量，如圖。

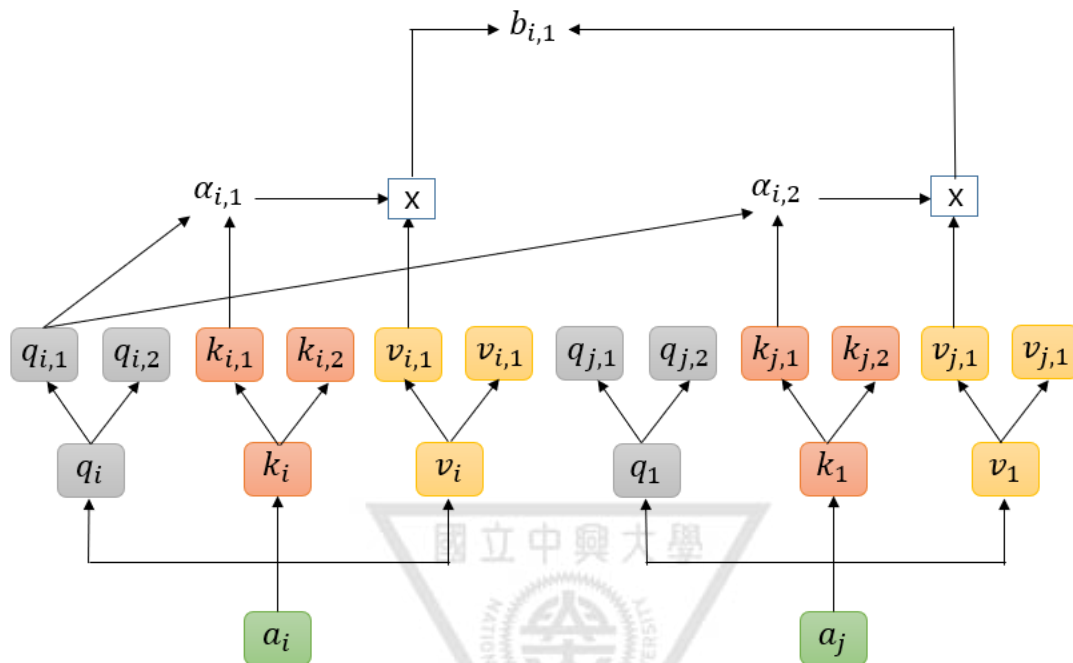


圖 8 Multi-head-attention 計算過程

## 2.5 BERT

Bert 的全名是 Bidirectional Encoder Representations from Transformers，是 2018 年 google 發表的 NLP pre-train model，它的主要構成是由 transformer 的 encoder 組成，應用可以是影像、聲音，但最廣泛是用於文字。

它主要分為兩個階段(圖 9)，pre-train 和 fine-tune，不同階段會有不一樣的任務，以下會分別介紹它的輸入處理、然後如何做 pre-train 以及這個 pre-train model 在一般使用上會搭配下游任務(downstream task)做實務上的應用。



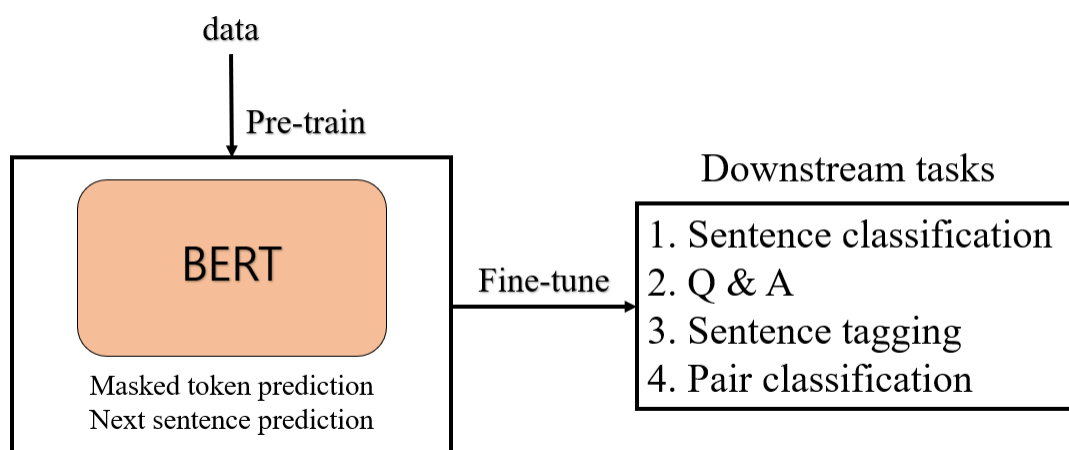


圖 9 BERT 的階段

### 2.5.1 輸入的處理

輸入與 Transformer 不同，需要做 3 種 embedding 向量化，然後將這些結果相加，如下圖 10。

Input	[CLS]	中	興	大	學	[SEP]
Token Embeddings	$E_{[CLS]}$	$E_{[中]}$	$E_{[興]}$	$E_{[大]}$	$E_{[學]}$	$E_{[SEP]}$
	+	+	+	+	+	+
Segment Embeddings	$E_{[A]}$	$E_{[A]}$	$E_{[A]}$	$E_{[A]}$	$E_{[A]}$	$E_{[A]}$
	+	+	+	+	+	+
Position Embeddings	$E_{[0]}$	$E_{[1]}$	$E_{[2]}$	$E_{[3]}$	$E_{[4]}$	$E_{[5]}$

圖 10 BERT 輸入的處理

其中[CLS]是輸入句子的開頭，[SEP]是分隔符號。Token Embedding 代表將 input 做 embedding，Segment Embedding 代表是第幾句，第一句是 0，第二句是 1。Position Embedding 代表經過學習的向量。

## 2.5.2 Pre-train

pre-train 階段會用大量沒有被標註(label)的文字，做類似人類在學習某一個語言的時候的學習歷程，包含以下兩種任務。

### 2.5.2.1 Mask LM

類似於填空題，如圖 11、12。是指輸入一句話，並且從這句話隨機抽 15%的字預測，而這被抽掉的字裡面，80%的機會用一個[mask]的符號取代，10%的機會用任意一個字取代，剩下 10%什麼事都不會做，可以參考圖 8，這樣做的好處是讓 model 有能夠讀懂上下文的能力。

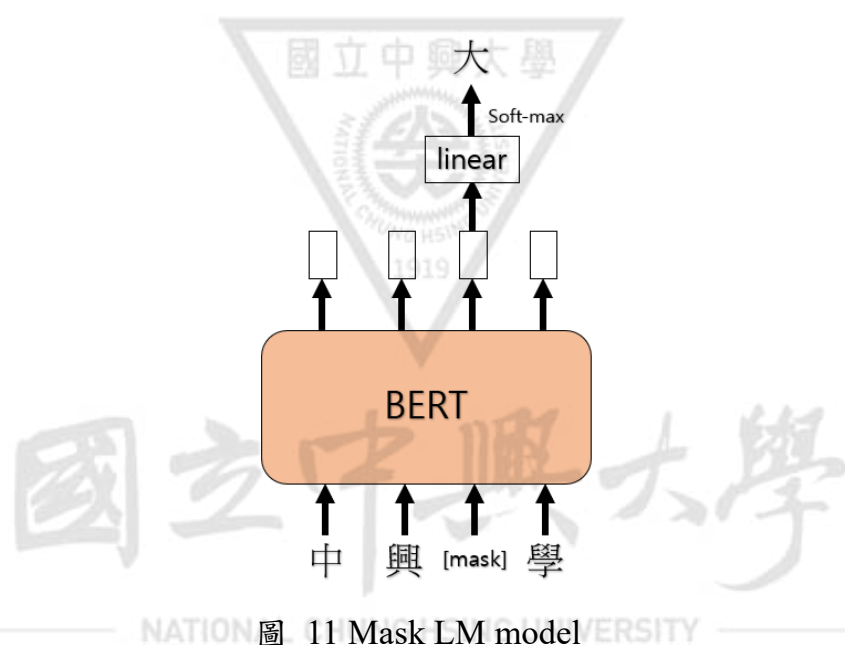


圖 11 Mask LM model

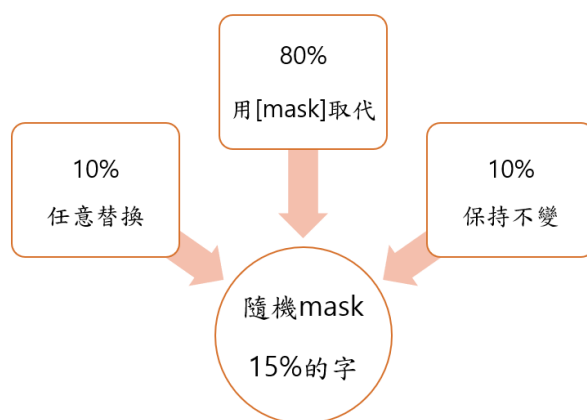


圖 12 mask 的做法

### 2.5.2.2 Next sentence prediction

輸入 2 句話，判斷兩者間是否有上下文關係，如圖 13。

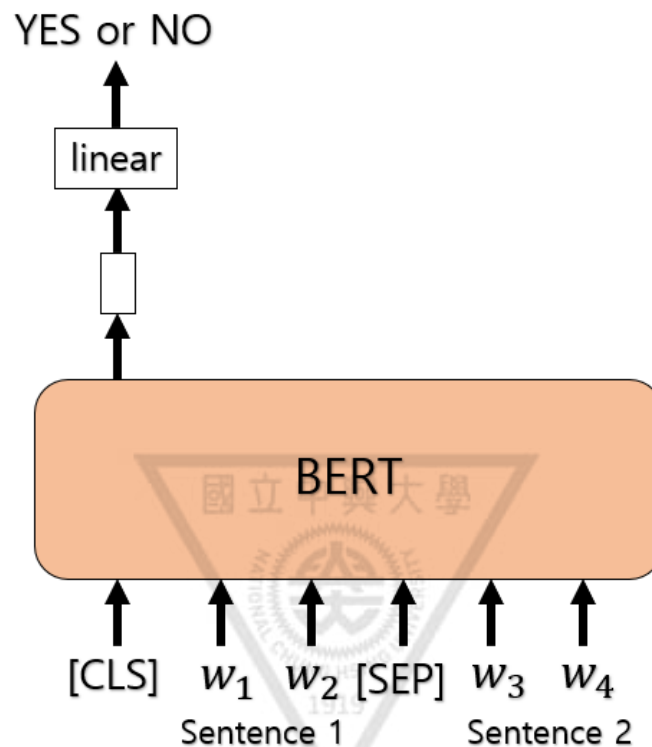


圖 13 Next sentence prediction

### 2.5.3 下游任務(downstream task)

另一個訓練的階段，會用有標註的資料進行訓練。作者指出在應用上，並不會直接拿 BERT 使用，而是會用訓練好的 pre-train model，並在之後搭配一個簡單的神經網路做分類，是一種遷移式學習，這個組合才是真正的使用方式。這部分主要是根據下游任務訓練，回頭對 pre-train model 做參數上的調整(fine-tune)，達到更好的效果。

其中下游任務有 4 種，以下做簡單介紹。

#### 2.5.3.1 Sentence classification，單句分類(圖 14)

常見的應用是情感分析，輸入一個 sequence，然後在 BERT 的第一個輸出後面接一個分類器，用 softmax 輸出分類。

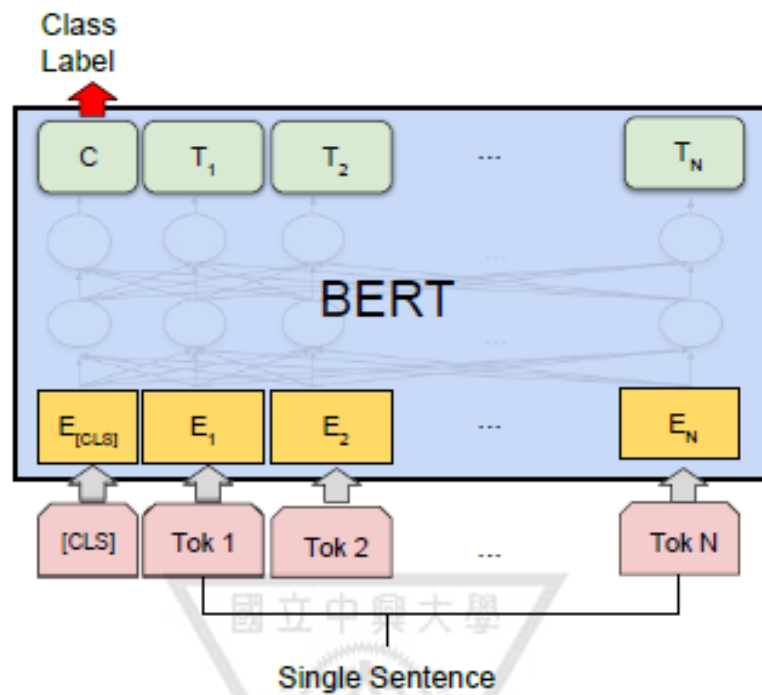


圖 14 單句分類[12]

### 2.5.3.2 Q & A(問答系統)(圖 15)

能夠對一個人給出的文章做出回答，然後回答的內容一定是文章裡面的字。輸入一篇文章和一個想問的問題，會回傳 2 個正整數  $s$  和  $e$ ，這兩個正整數代表在文章中答案開始的位置和答案結束的位置，將這兩個字之間所有包含的文字連接起來，就是答案。

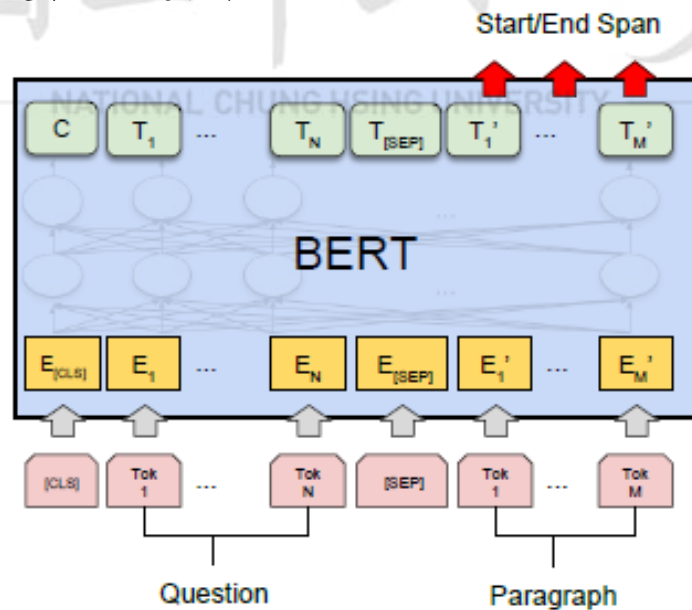


圖 15 Q & A 問答系統[12]

### 2.5.3.3 Sentence tagging(詞性標註)(圖 16)

讓機器詞性的標註，像是動詞形容詞名詞等等。輸入一段 sequence，在 BERT 除了第一個輸出外的向量接分類器，以 softmax 輸出結果。

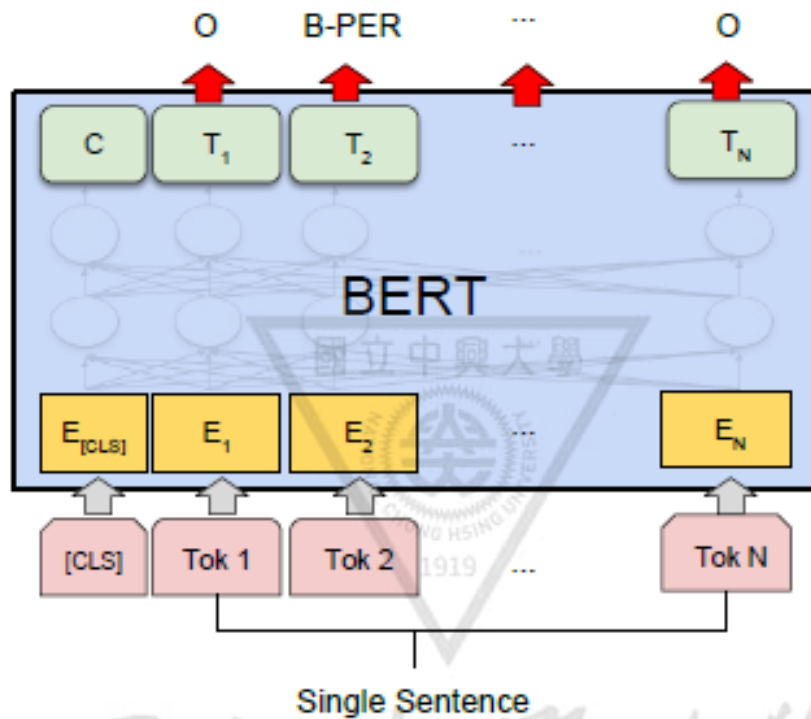


圖 16 詞性標註[12]

### 2.5.3.4 Pair classification(圖 17)

判斷兩個句子的相似度，可以用於評價和留言系統，判斷兩者之間有無矛盾。輸入 2 個 sequence，中間以 [SEP] 符號做分隔，然後取 BERT 第一個輸出的向量，接一個分類器，以 softmax 輸出。

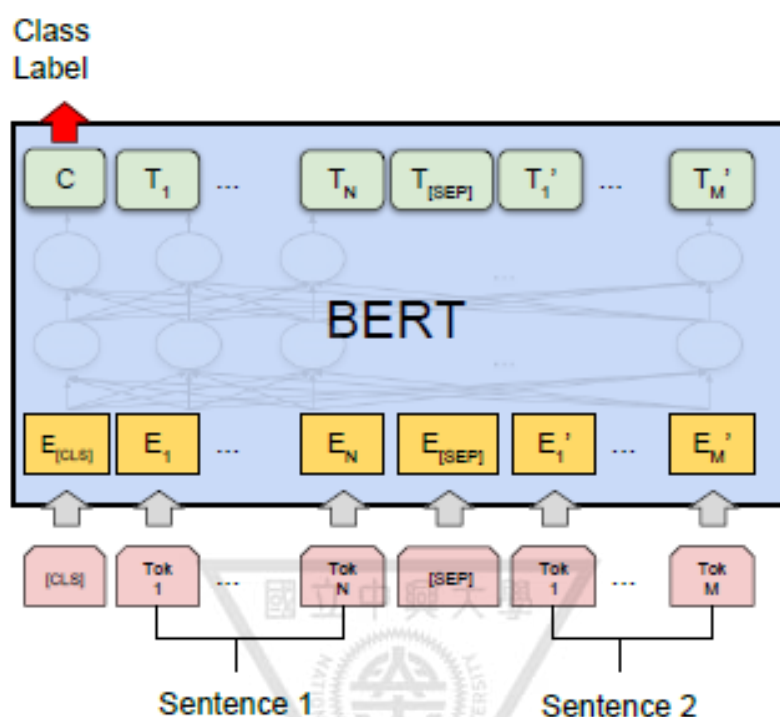


圖 17 因果關係判斷[12]

## 2.6 中文的情緒分析

Gen Li 團隊[19]在 2020 年發表了一篇關於中文情緒分析的 paper，提到他們設計了新的架構，結合 Transformer 的 encoder 和加入情緒資訊分辨器輔助 Bi-LSTM 訓練，並以 F1 score 做為比較基準，相較於其他架構有比較好的成果。架構如圖 18，輸入先經過斷詞的處理後，會再經過 Embedding layer 和 Transformer Encoders layers，接著會分成兩種任務，分別是情緒資訊分辨器和 Bi-LSTM 的訓練。

情緒資訊分辨器的目的是能夠判斷詞的情緒特徵，之後會根據情感的強度分成三個類別，以”餐廳的早餐不錯”為例，”不錯”會被分類成有情緒的特徵類別，然後”早餐”則是會分類成某個情緒特徵的目標類別，兩者之間的關係類似形容詞和名詞，最後會把除了以上分類之外的全部分類為較低情緒的特徵的類別。Bi-LSTM，作者在這有引入 attention 機制，透過訓練情緒資訊分辨器在 transformer encoders 可以得到有情緒資訊分布的特徵向量，再以這些向量當作輸入再進行訓練，最後會經過 softmax 輸出正面或是負面情緒兩類。

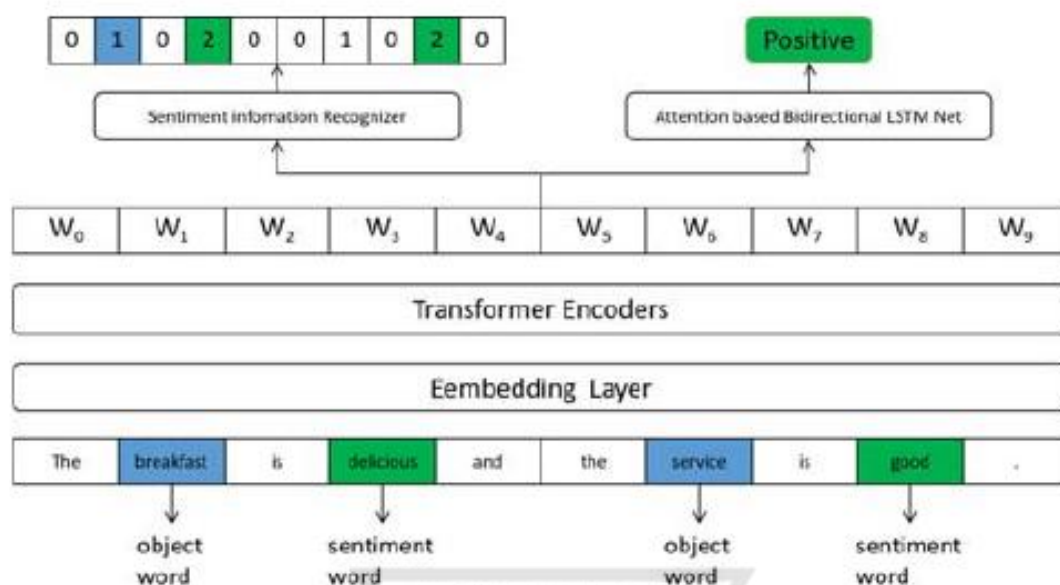


圖 18 SINM 架構[19]

圖 19 是他們提供的範例，藍色代表情緒特徵的目標，紅色和綠色分別代表負面 and 正面的情緒詞。

Sentiment information distribution of text	Sentiment polarity
Business big bed rooms, the room is (very big), the bed is 2 meters wide, the overall feeling: affordable and good!	Positive
The breakfast is (just so so), the room is clean, but the noise is (too loud), and the price is (relatively expensive)!	Negative
Good, delicious, and inexpensive, fast delivery!	Positive
Delivery time is too long, noodles are soft, and (not chewy)!	Negative

圖 19 情緒分佈示例[19]

Xing Lin 和 Chunyan Han[20]也提出一種改良版的 CNN 網路架構 PCCNN 做中文的情緒分析，基於 Ping Li 團隊的 CNN 架構[21]，如圖 20。考慮到大多數人使用拼音輸入法，會出現諧音輸入錯誤和網絡流行語的諧音問題，加入利用漢字拼音的特點的第三種通道，一樣有很好的成果。在模型的訓練上，輸入使用 Word2vec 分別作向量化，分成傳統的 word、char 和新增的 Phoneticize char 三種，多通道的好處是避免參數過多和時間過長，之後再分別經過卷積層和 maxpooling 層提取特徵，最後經過 dropout 和 softmax 可以得到預測的結果。

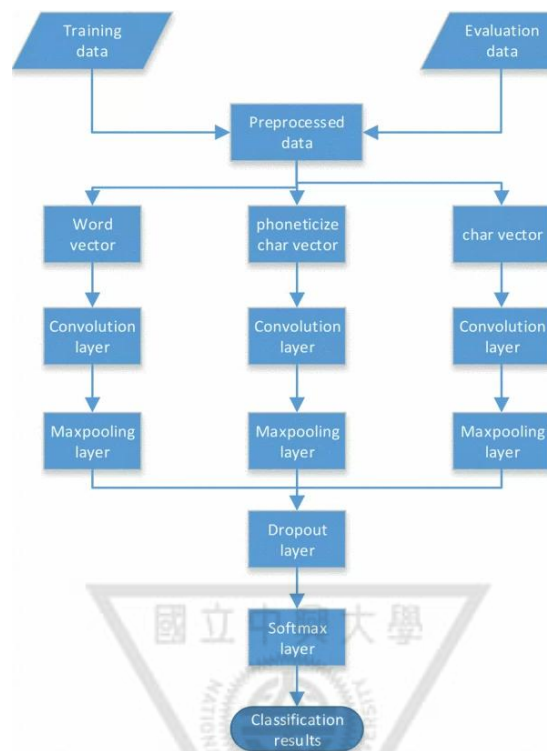


圖 20 改良版的 CNN 架構[20]

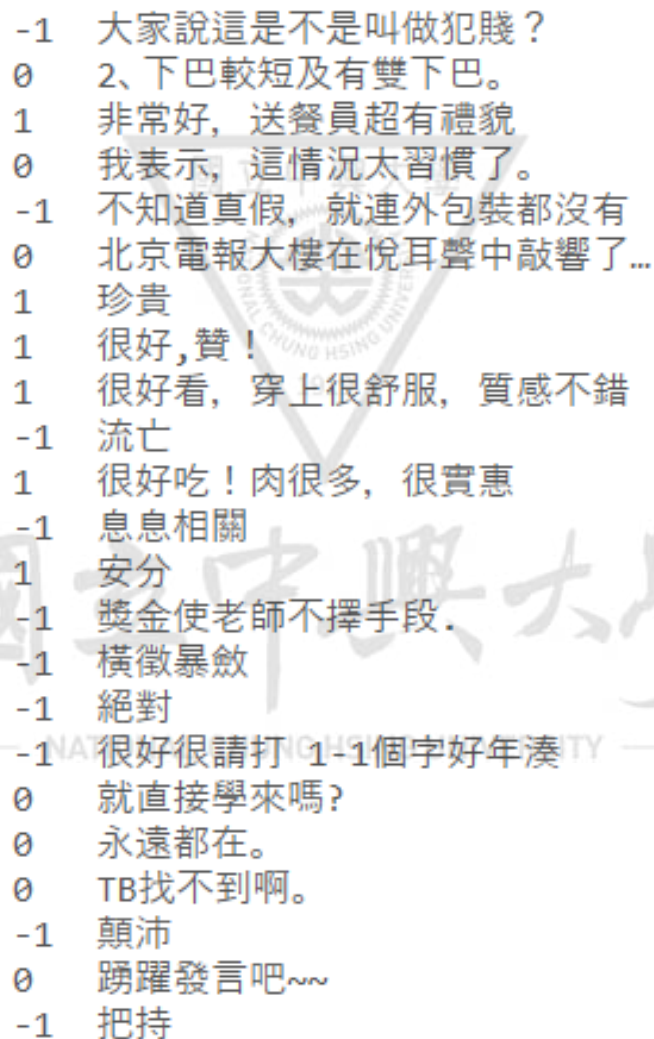
國立中興大學

NATIONAL CHUNG HSING UNIVERSITY



### 第三章 蒐集資料

我們希望能夠將短句做正確的分類，在給資料的限制是 15 個字元，包含標點符號，再將結果分為三類還有相應的標註(label)，分別是正面(1)、中立(0)、負面(-1)，如圖 21，當中包含網路上的開源資料集和各種留言或是評價，還有情緒字典。



-1 大家說這是不是叫做犯賤？  
0 2、下巴較短及有雙下巴。  
1 非常好，送餐員超有禮貌  
0 我表示，這情況太習慣了。  
-1 不知道真假，就連外包裝都沒有  
0 北京電報大樓在悅耳聲中敲響了...  
1 珍貴  
1 很好,贊!  
1 很好看，穿上很舒服，質感不錯  
-1 流亡  
1 很好吃！肉很多，很實惠  
-1 息息相關  
1 安分  
-1 獎金使老師不擇手段。  
-1 橫徵暴斂  
-1 絕對  
-1 很好很請打 1-1個字好年湊  
0 就直接學來嗎？  
0 永遠都在。  
0 TB找不到啊。  
-1 顛沛  
0 踴躍發言吧~~  
-1 把持

圖 21 整理後的資料

### 3.1 網路留言與評論

包含nlpcc2014.以及各種網路電商和外賣平台網路留言,其中nlpcc2014,如圖 22,將原本標記是厭惡(disgust)、傷心(sadness)、生氣(anger)的歸類成負面,快樂(happy)、驚喜(surprise)、喜歡(like)歸類成正面,無情緒(none)則歸類成中立。

none,“这一做法属全国首例,抽空专程去学习和了解一下。”  
none,男人,女人其实一样,除了一个你,一定还有一个不一样的他(她)。  
none,这个人可以是朋友,也可以是前女友。  
none,无论他们承不承认,在他心里她都是特别的存在。  
none,你无法替代,也改变不了。  
none,宽容些吧,人都有过往,都有无法忘记的人,无法释怀的感情。  
sadness,兴奋滴计划了两天的踏青活动宣告终结,更改为在偶家里滴成人益智类游戏,偶们灰常悲催的慨叹着岁月是这般的催人老哟!!!  
none,就是这个家伙让我果断摒弃了ip4。  
like,它真的太完美了`莫非...  
none,莫非我真是它手最小的机主??  
none,[白眼]  
none,今晚和一个九零后论战一晚,想起鲁迅的话:我一向是相信进化论的,总以为将来必胜于过去,青年必胜于老人,.....  
none,后来我明白我倒是错了。  
surprise,我在广东,就目睹了同是青年,而分成两大阵营,或则投书告密,或则助官捕人的事实!  
disgust,我的思想因此衰退,后来便时常用了怀疑的眼光去看青年,不再无条件的敬爱了。  
disgust,你说说挺好一辆大奔,非得在四个轮毂上弄四条红布条子,在后挡风上贴个对联,居然还有横批。  
disgust,这也算中国特色吧。  
surprise,震惊全国的---婚礼:法拉利黄色F360-345万(停产) 破黑F430-360万 黑色F430 spider-404万  
红色575M-320万(停产), 4台劳斯莱斯(金标)幻影(均价710万左右), 4台宾利飞驰(均价300万以上), 20台奥迪A8-W12  
quattro还有两台S8(均价240万左右), 打杂的40多辆卡宴,路虎,X5,Q7都不在婚车队伍中  
none,今天得闻一位曾经的旅游媒体兄长吴永平先生于前段时间突发脑溢血去世,享年50岁。  
sadness,吴兄入旅游较早,于教育电视台旅游栏目,吃苦耐劳,辛勤耕耘,后离开旅游界,攻硃萼病重之时,其也多有关照,热心人啊。  
sadness,人生短暂,各位当把握之。  
sadness,并顺祝吴兄家人以及各位端午安好。  
surprise,(居然无合影照片可供留念)  
like,阿森纳球员从不使诈'教授从不教坏小盆友'所以阿尔沙文被推倒还把球射出去错失点球'枪手踢的是世上最华丽而又最老实的球风'我因此热爱他们  
none,//@麦蒂无罪:世界是四维的,排!  
none,[不要] 今年以来两市IPO新股已多达156只,合计募资高达1654亿元,占2009年IPO重来近20%!  
happiness,按目前速度以及全球IPO总体放缓的情况,上半年中国A股的IPO融资额很可能再度冠绝全球!  
none,在西方,现代民族国家的建立大概首先从13世纪的英国、法国开始,其典型特征是王权与教权、封建主权力的斗争,而在此过程中,国家相对于市民社会的独立性也日益凸显。  
sadness,心情不好的时候特别想找个个人撒娇或者哭一哭。  
sadness,身边的人都很热闹自己却不敢上前去的时候也很希望有人过来陪我。  
sadness,但是这样的人从未出现过。  
none,品牌是消费者与产品的关系,消费者才是品牌的最后拥有者,品牌是消费者经验的总和。  
none,一大卫-奥格威(David Ogilvy)  
disgust,工商局你们太有才了,临时可以为我开创新的业务受理流程?  
disgust,凸凸 我没什么好说的了, 出来做生意除了求财就是求和气,这也太雷人了。  
disgust,早说啊!我搬办公室也不受这个气了。

圖 22 網路留言資料集

### 3.2 情緒字典

由於本文是專注於短句判斷,考慮到情緒的判斷有時甚至可以因為某幾個字就判斷出正、負面的關係,以及為了避免某些判斷情緒的關鍵字會有重複性過高的問題,額外加上了台大情緒字典做為新增的訓練集,它包含大量正面和負面的各種詞語。

## 第四章 架構和模型訓練

本文的目標是希望知道中文短句的情緒對應關係，除了對一般日常對話或新聞標題之類的话题之外，也能夠對某些包含諷刺性的語句做出判斷，目標是提升判斷成正面的準確度。

### 4.1 整體架構

圖 23 是本文所提出的架構。首先會輸入一句中文短句，然後經過訓練過的 BERT model 做第一次情緒判斷，這一階段若是被判斷成負面或是中立會直接輸出，若輸入是長度超過 7 個字元且被判斷成正面也會直接輸出，但若是輸入小於 7 個字元，會進入斷詞處理的階段然後做額外判斷。斷詞的處理是我們為了改善 BERT 對於同時包含好或壞字眼的諷刺短句判斷不是很好的問題，例如愚人快樂。

斷詞處理是基於 jieba[18] 的 python 套件，它可以對中文的句子做裁切，例如自信過了頭，經過斷詞的處理後，可以分成，自信/過了頭，一共 2 段，這裡的處理我們會把被斷詞後的詞語，每一個詞都用 BERT 重新判斷，若這些被斷詞後的所有詞語在這判斷後仍然都是正面，就會把這個輸入判斷成正面，而如果這些斷詞的判斷中有出現負面詞語，就把輸入判斷成中立。

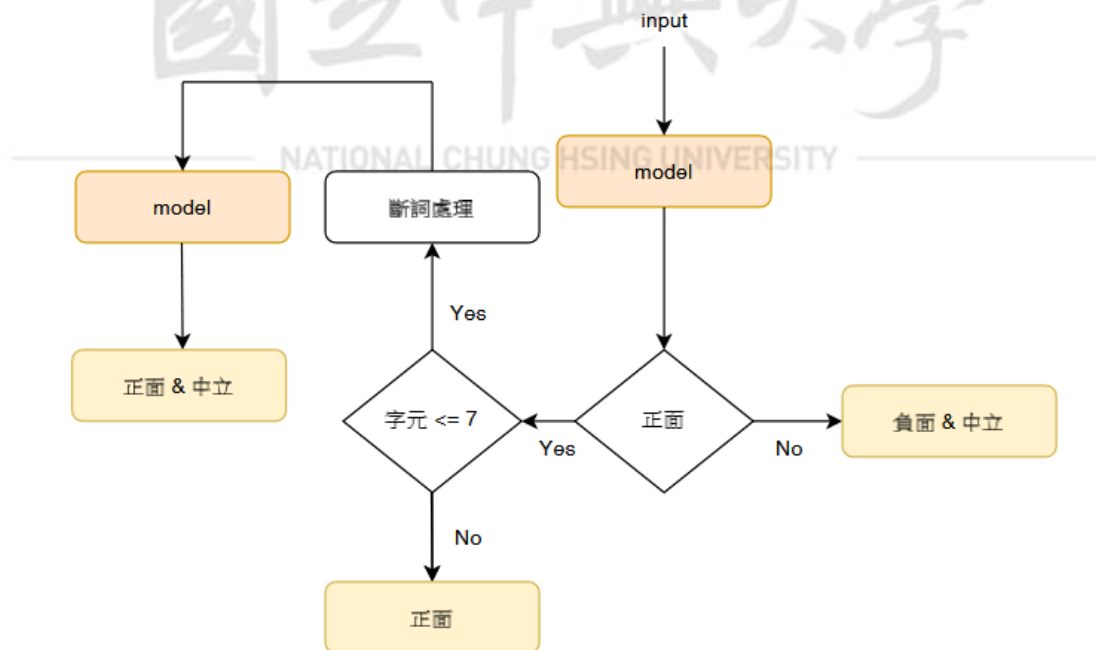


圖 23 所提架構

## 4.2 模型訓練[13]

輸入一句話，經過訓練過的 BERT model 和分類器，最後輸出正面、中立、負面，如圖 24。

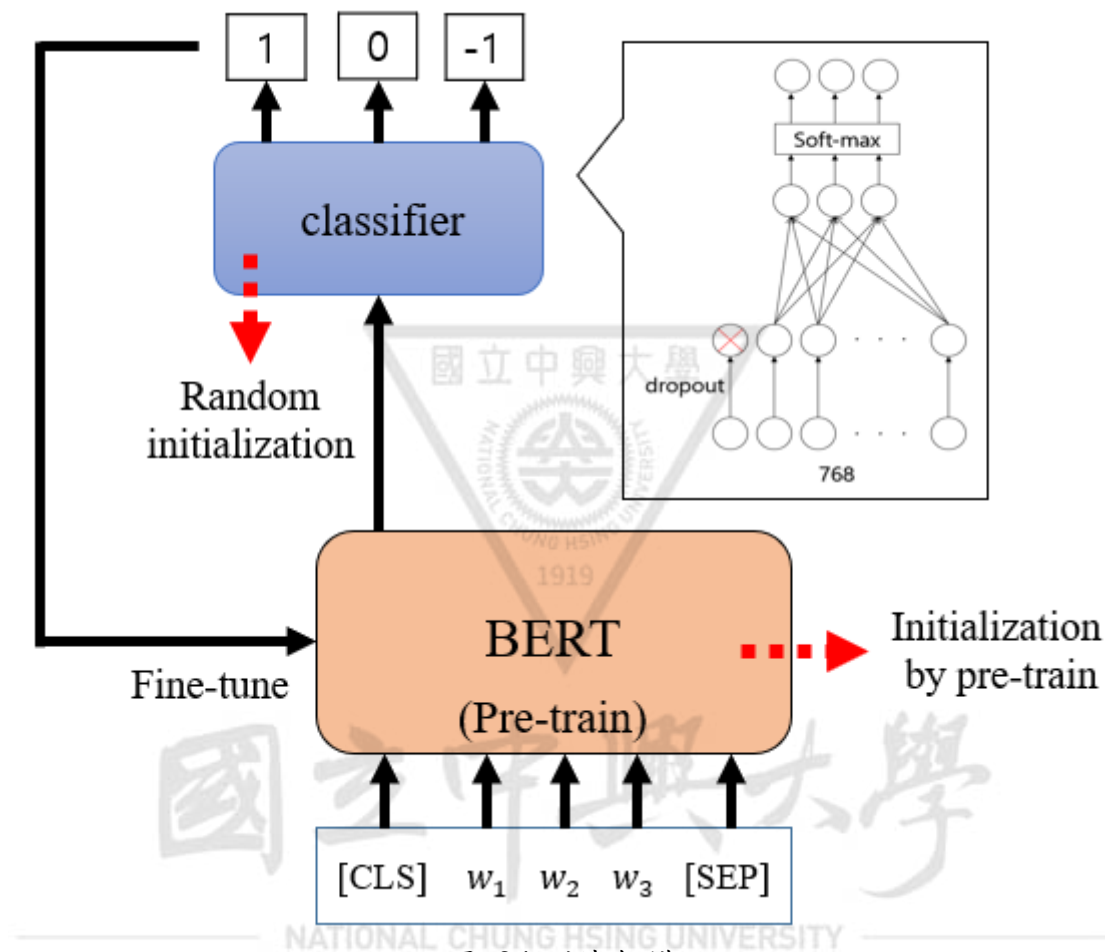


圖 24 訓練架構

### 4.2.1 輸入前處理

首先，中文短句化成一串輸入 sequence，然後在開頭和結尾加上  $[CLS]$  和  $[SEP]$  符號，分別代表句子的開始和結束，接著會對這些序列不足 128 個的部分加上  $[PAD]$  記號做擴充，因為有預設的 128 的限制，處理好之後再對每一個 sequence 和符號做轉換，如下圖 25。

圖 25 輸入的轉換過程

Attention mask 是用來界定 self-attention 範圍，1 是輸入 sequence，後續會有運算的環節，0 是[PAD]，不進入運算。

### 4.2.2 訓練[13]

接在 pre-train model 之後的分類器內部是一個簡單的類神經網路，需要重新訓練。除了有 dropout layer 之外，這個分類器在最後會通過 soft-max 得

到 3 個輸出，分別代表短句正面、中立、負面的機率，我們會從機率取其中最高者，給予代號 1、0 或-1 當成最後的結果。

維度的變化可以參考下圖 26。

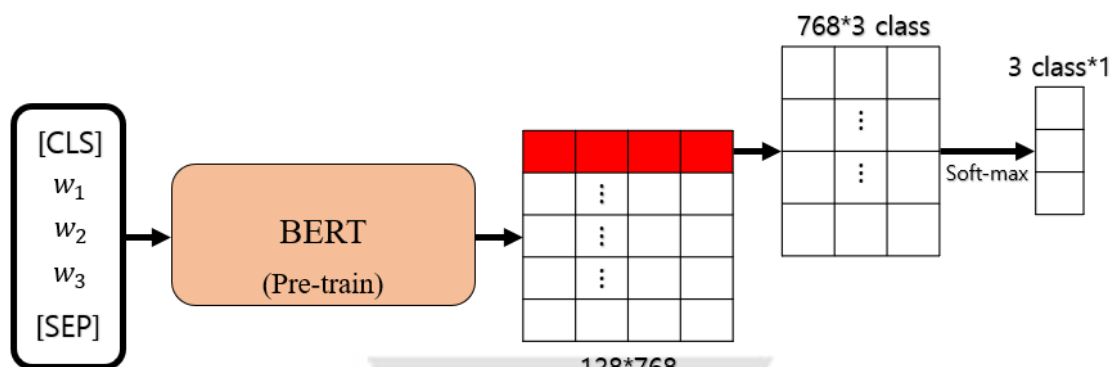


圖 26 維度變化

## 第五章 研究結果與比較

訓練有使用預訓練模型，整體過程大約花了2至3小時，一共8個epoch，batch size=16，主要參數可以參考下表，而 Loss function 是用 sparse categorical cross-entropy，測試集的部分我們用了新聞標題和 google 餐廳留言等等一般向的短句做實際預測，一共 204 筆資料，59 筆正面，61 筆中立，84 筆負面。而我們關注的部分是被預測正確的有多少是真的正確，不能接受將來在實際使用會顯示出一些不適當的句子，所以結果與比較會由混淆矩陣和精確度來做成效的展示。

表 1 使用參數

參數名稱	數值
Pre-process model	bert_zh_preprocess[16]
Pre-train model	bert_zh_L-12_H-768_A-12[17]
優化器	adamw
Batch size	16
Epochs	8
Learning rate	2e-5
訓練集數量	28410
驗證集數量	7131
測試集數量	204

圖 27 是訓練的準確度曲線，紅色是訓練集，藍色是驗證集，可以看到訓練集在第 8 個 epochs 開始收斂，而驗證集一直保持在大約 0.86 左右的準確度。

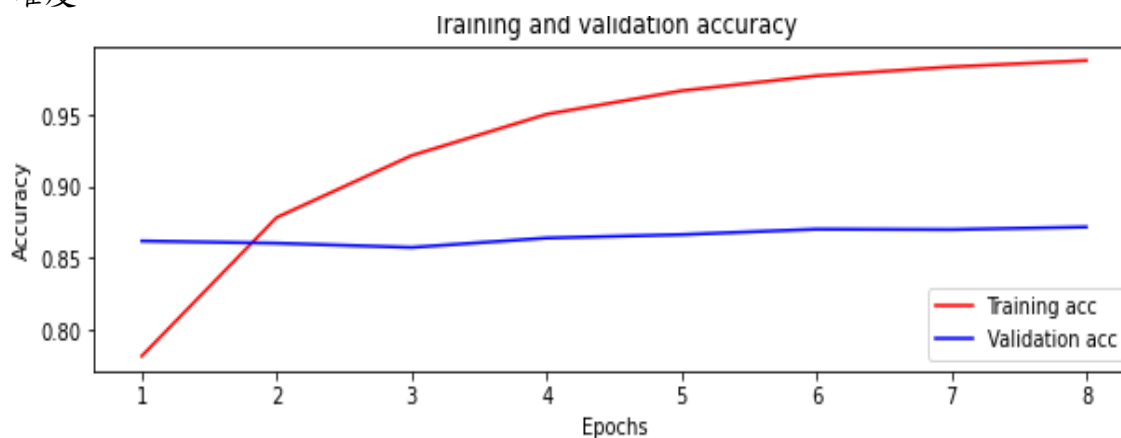
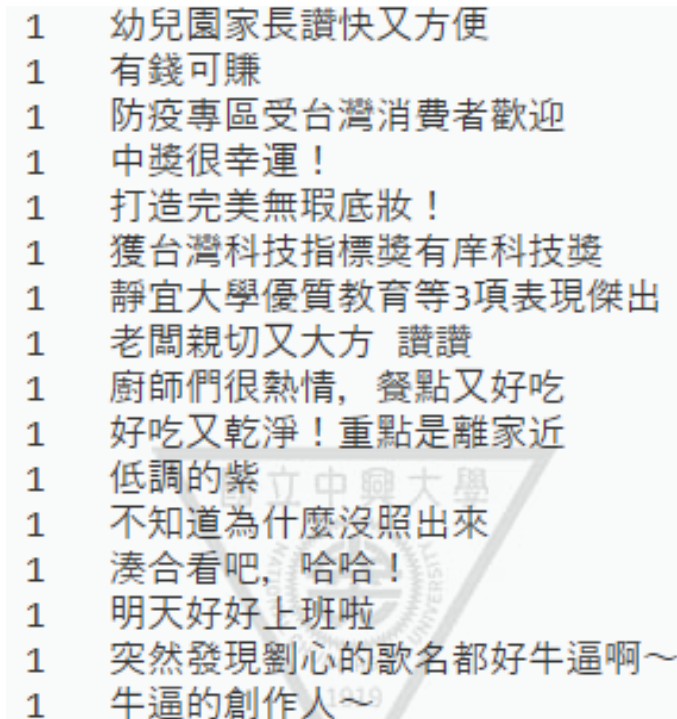


圖 27 訓練集和驗證集準確度

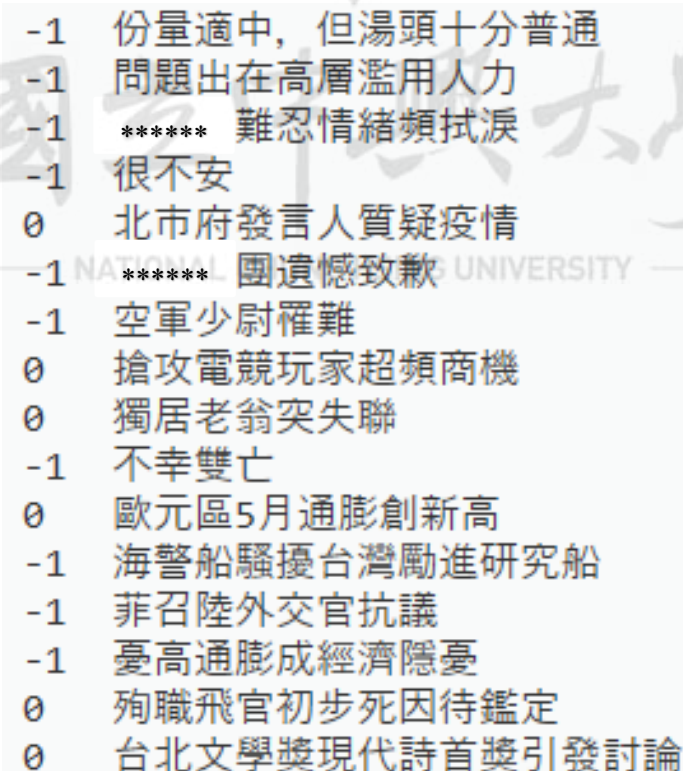


圖 28 和圖 29 是測試集的預測結果，當中不包含諷刺性的短句，可以發現對於明顯的讚美詞或是批評的詞語都能被判斷出來。



1 幼兒園家長讚快又方便  
1 有錢可賺  
1 防疫專區受台灣消費者歡迎  
1 中獎很幸運!  
1 打造完美無瑕底妝!  
1 獲台灣科技指標獎有庠科技獎  
1 靜宜大學優質教育等3項表現傑出  
1 老闆親切又大方 讚讚  
1 廚師們很熱情, 餐點又好吃  
1 好吃又乾淨! 重點是離家近  
1 低調的紫  
1 不知道為什麼沒照出來  
1 湊合看吧, 哈哈!  
1 明天好好上班啦  
1 突然發現劉心的歌名都好牛逼啊~  
1 牛逼的創作人~

圖 28 判斷正面的結果



-1 份量適中, 但湯頭十分普通  
-1 問題出在高層濫用人力  
-1 \*\*\*\*\* 難忍情緒頻拭淚  
-1 很不安  
0 北市府發言人質疑疫情  
-1 \*\*\*\*\* 團遺憾致歉  
-1 空軍少尉罹難  
0 搶攻電競玩家超頻商機  
0 獨居老翁突失聯  
-1 不幸雙亡  
0 歐元區5月通膨創新高  
-1 海警船騷擾台灣勵進研究船  
-1 菲召陸外交官抗議  
-1 憂高通膨成經濟隱憂  
0 殉職飛官初步死因待鑑定  
0 台北文學獎現代詩首獎引發討論

圖 29 判斷負面的結果



表 2 有斷詞處理混淆矩陣

混淆矩陣 (單位:數量)		預測		
		正面	中立	負面
實際	正面	46	13	0
	中立	14	42	5
	負面	1	17	66

表 2 是有加入斷詞處理的混淆矩陣，數據左下角可以發現有一筆資料實際是負面，但卻被預測成正面，這筆資料的內容是:感謝師傅，讚嘆師傅。原本我們有打算過濾一些宗教性的言論，但這筆資料正面詞語過多，包含感謝和讚嘆，且銜接的詞語是師傅，也並無特別之處，導致很容易就被判斷成正面。

表 3 未斷詞的混淆矩陣

混淆矩陣 (單位:數量)		預測		
		正面	中立	負面
實際	正面	48	11	0
	中立	15	41	5
	負面	2	16	66

表 3 是未斷詞的混淆矩陣，數據左下角可以發現有一筆資料實際是負面，但卻被預測成正面，有兩筆資料，其中一筆是:感謝師傅，讚嘆師傅，會出現的原因就如同上述。另一筆是:自信過了頭，這筆資料是由一個正面詞語自信，和一個負面詞語過了頭組成，像這類短句同時包含負面詞語和正面詞語是演算法比較弱的地方，所以我們才加入斷詞處理這個問題。

以上表 2 和表 3 是加入斷詞與未加入斷詞的混淆矩陣，除此之外我們也有使用卓騰語言科技公司的免費 API[15]測試不包含諷刺性的語句的測試集，可以參考表 4。

表 4 卓騰語言科技的混淆矩陣

混淆矩陣 (單位:數量)		預測		
		正面	中立	負面
實際	正面	38	21	0
	中立	13	40	8
	負面	3	38	43

表 5 精確度比較

	正面	中立	負面
斷詞	<b>0.754</b>	0.583	0.929
無斷詞	<b>0.738</b>	0.603	0.929
卓騰語言科技	<b>0.703</b>	0.404	0.843

精確度計算公式如式 5， $y_t$ 代表預測和實際相同， $y_0$ 代表預測和實際不同

$$\frac{y_t}{y_t + y_0} \quad \text{式 5}$$

觀察表 5 可以發現，以一般的測試集測試，對於沒有斷詞處理的精確度高達 0.738，而加入斷詞處理後，對於正面的精確度僅有不到 2% 的提升，這是因為經過斷詞後會把短句切割成一個個短詞語做單獨判斷，這個做法除了會考驗訓練集的資料豐富度之外，也包含演算法本身對過短語句判斷並不是特別強，導致斷詞後的詞語比較容易有誤判的情況。

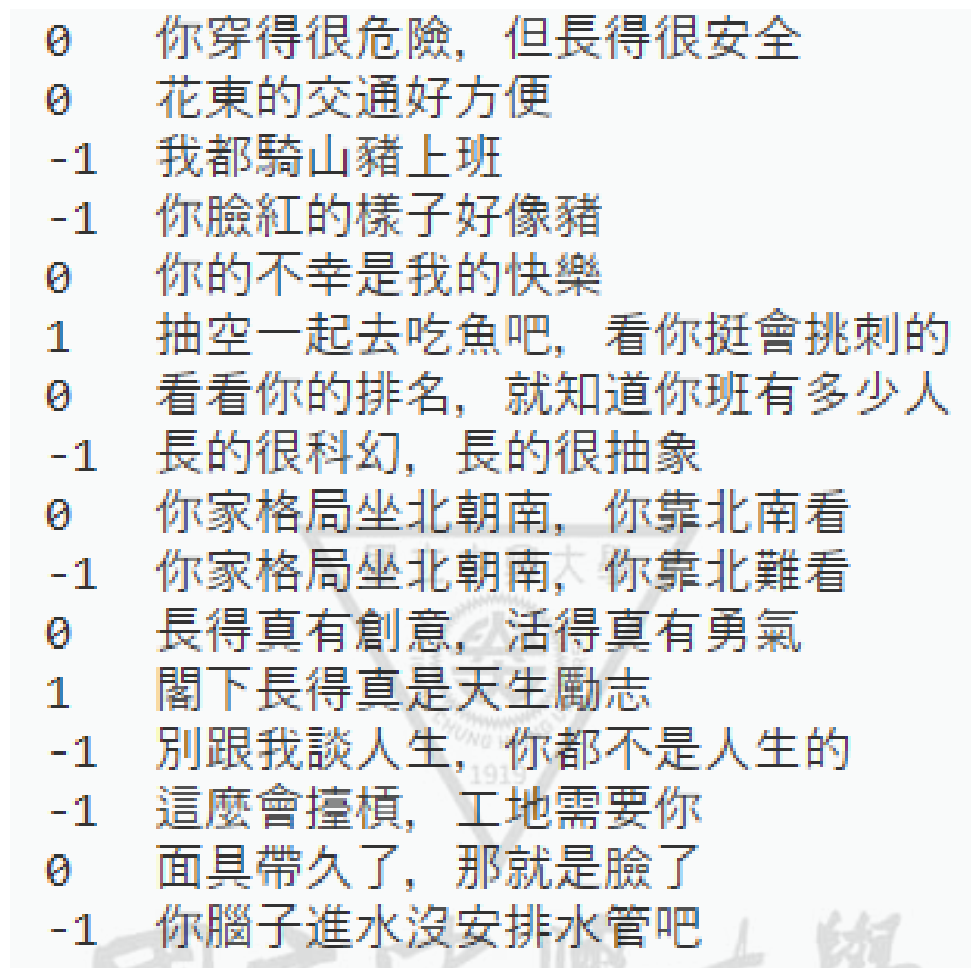
與卓騰語言科技公司相比，可以發現我們對於短句的判斷精確度在 3 種類別都有比較高的表現，這個原因是因為我們是針對短句去研究，而他們則是沒有對範圍做設限，用途不同，實際使用 API 也可以發現他們做到能夠應用在更長的語句或是文章，也有不錯的成果。

諷刺性的語句因為是屬於非主流，所以判斷會單獨展示，如表 6，可以發現加入斷詞的處理後可以將某些同時包含好或壞的詞語判斷成中立，相較於沒有斷詞處理的方式有改善，也可以看出卓騰語言科技公司對於這類短句的判斷不夠精準。

表 6 諷刺語句

	斷詞	無斷詞	卓騰語言科技
阿不就好厲害	中立	正面	負面
有錢人終成眷屬	中立	正面	正面
祝智障生日快樂	中立	正面	正面
愚人快樂	中立	正面	正面

我們也有準備一些非主流的網路和語錄測試我們的判斷能力，如圖 30。



0	你穿得很危險，但長得很安全
0	花東的交通好方便
-1	我都騎山豬上班
-1	你臉紅的樣子好像豬
0	你的不幸是我的快樂
1	抽空一起去吃魚吧，看你挺會挑刺的
0	看看你的排名，就知道你班有多少人
-1	長的很科幻，長的很抽象
0	你家格局坐北朝南，你靠北南看
-1	你家格局坐北朝南，你靠北難看
0	長得真有創意，活得真有勇氣
1	閣下長得真是天生勵志
-1	別跟我談人生，你都不是人生的
-1	這麼會擡槓，工地需要你
0	面具帶久了，那就是臉了
-1	你腦子進水沒安排水管吧

圖 30 網路語錄

沒有判斷好的地方，像是閣下長得真是天生勵志，字面上理解是正面，但用勵志去形容長相應該是諷刺人，這種例子對於現在的階段還是屬於比較難解決的問題。

## 第六章 結論與未來展望

### 6.1 結論

本文的目的是希望讓合適的言論能正確地被刊登，這考驗我們對於正面的言論判斷，是否足夠準確。

觀察前面的結果可以發現，本文除了能夠對某些諷刺性的言論做出判斷之外，對於負面的短句也有很高的精確度，儘管如此還是有一些地方需要克服，例如某些同時包含好或壞的言論可能還是會有問題，或是一些比較高端的諷刺性言論，像是工業區的空氣聞起來特別清新，有時候連人類都還要稍微思考，要讓機器做出正確的判斷還是非常困難的。

另外，文字的排列組合可以說是無限，有時候要判斷情緒除了需要考量上下文之間的關係，還有字面意義也要列入考量，並不能單純以字面意義去判斷，對現在的我們來說，我們期待之後能做更多的上下文考量。

### 6.2 未來展望

未來我們希望能夠把這項研究用在更多地方，除了要提升整體的判斷精確度之外，也能夠對更長的文字去進行分析，以及對更多諷刺性的言論能有更好的判斷能力。同時也要對斷詞的效果做出改進，包含將詞跟詞之間的切割得更有意義和提升短詞語的準確度。

## 第七章 參考文獻

- [1] S. Tiwari, A. Verma, P. Garg and D. Bansal, "Social Media Sentiment Analysis On Twitter Datasets," 2020 6th International Conference on Advanced Computing and Communication Systems (ICACCS), 2020, pp. 925-927, doi: 10.1109/ICACCS48705.2020.9074208.
- [2] G. Li, Q. Zheng, L. Zhang, S. Guo and L. Niu, "Sentiment Information based Model For Chinese text Sentiment Analysis," 2020 IEEE 3rd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), 2020, pp. 366-371, doi: 10.1109/AUTEEE50969.2020.9315668.
- [3] L. -C. Cheng and S. -L. Tsai, "Deep Learning for Automated Sentiment Analysis of Social Media," 2019 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2019, pp. 1001-1004, doi: 10.1145/3341161.3344821.
- [4] Varsha Kesavan, Vaidehi Muley, Megha Kolhekar, "Deep Learning based Automatic Image Caption Generation", In 2019 Global Conference for Advancement in Technology (GCAT).
- [5] Dzmitry Bahdanau, Kyunghyun Cho, Yoshua Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate", In ICLR, 2015
- [6] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin, "Attention Is All You Need.", In NIPS'17: Proceedings of the 31st International Conference on Neural Information Processing.
- [7] Hengshuang Zhao; Jiaya Jia; Vladlen Koltun, "Exploring Self-Attention for Image Recognition", In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
- [8] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," in *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, 15 Nov. 1997, doi: 10.1162/neco.1997.9.8.1735.

- [9] Jeffrey L. Elman, "Finding Structure in Time". Cognitive Science. Volume 14, Issue 2, Pages 179-211 (April–June 1990)
- [10] Alex Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton "ImageNet Classification with Deep Convolutional Neural Networks". In NIPS, 2012.
- [11] Elizabeth D. Liddy, "Natural language processing", 2nd edn. Encyclopedia of Library and Information Science, Marcel Decker, 2001.
- [12] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova, "BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding", arXiv:1810.04805 [cs], May 2019, Accessed: Aug. 09, 2020. [Online]. Available: <http://arxiv.org/abs/1810.04805>.
- [13] Classify text with BERT,  
[https://www.tensorflow.org/text/tutorials/classify\\_text\\_with\\_bert](https://www.tensorflow.org/text/tutorials/classify_text_with_bert)
- [14] <https://www.freecodecamp.org/news/what-is-sentiment-analysis-a-complete-guide-to-for-beginners/>
- [15] <https://api.droidtown.co/#keymoji>
- [16] [https://tfhub.dev/tensorflow/bert\\_zh\\_preprocess/3](https://tfhub.dev/tensorflow/bert_zh_preprocess/3)
- [17] [https://tfhub.dev/tensorflow/bert\\_zh\\_L-12\\_H-768\\_A-12/4](https://tfhub.dev/tensorflow/bert_zh_L-12_H-768_A-12/4)
- [18] <https://github.com/fxsjy/jieba>
- [19] G. Li, Q. Zheng, L. Zhang, S. Guo and L. Niu, "Sentiment Information based Model For Chinese text Sentiment Analysis," 2020 IEEE 3rd International Conference on Automation, Electronics and Electrical Engineering (AUTEEE), 2020, pp. 366-371, doi: 10.1109/AUTEEE50969.2020.9315668.
- [20] X. Lin and C. Han, "Chinese Text Sentiment Analysis Based on Improved Convolutional Neural Networks," 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), 2018, pp. 296-300, doi: 10.1109/ICSESS.2018.8663888.

- [21] Ping Li, Y M Dai and D H. Wu, Application of dual-channel convolutional neural network in sentiment analysis[J]. Journal of Computer Applications, vol. 38, no. 334, pp. 22-26, 2018.



國立中興大學

NATIONAL CHUNG HSING UNIVERSITY