

元 智 大 學

資 訊 工 程 學 系

碩 士 論 文

中文 Valence-Arousal-Irony 語料庫的創建和評估  
Construction and Evaluation of a Chinese Valence-  
Arousal-Irony Affective Corpus

研 究 生： 劉 靜

指導教授： 賴國華 博士

中華民國 一〇六 年 一 月

中文 Valence-Arousal-Irony 語料庫的創建和評估  
Construction and Evaluation of a Chinese Valence-  
Arousal-Irony Affective Corpus

研 究 生：劉靜

Student : Jing Liu

指 導 教 授：賴國華 博士

Advisor : Dr. K. Robert Lai

元 智 大 學

資 訊 工 程 學 系

碩 士 論 文

A Thesis

Submitted to the Department of Computer Science and Engineering  
Yuan Ze University  
in Partial Fulfillment of the Requirements for the Degree of  
Master of Science  
in

Computer Science and Engineering

January 2017

Chungli, Taiwan, Republic of China

中華民國 一〇六年一月

# 中文 Valence-Arousal-Irony 語料庫的創建和評估

研 究 生: 劉靜

指導教授: 賴國華 博士

元智大學資訊工程學系

## 摘要

反諷研究在自然語言研究領域受到了越來越多的重視, 基於反諷表達與常用自然語言表達的不同, 針對反諷表達的情感強度做分析研究對反諷領域的研究有重要的意義。基於人類文本語言資源豐富以及深度神經網路學習方法成熟化, 我們追求自然語言的語意或者情感進行自動化識別。文本的情感表示主要有兩種方式, 類別型以及維度型。類別型將文本情感表示為離散的情感分類, 例如正面與負面的二元分類, 六大情緒分類等。維度型表達方法是用連續的數值表示情感不同的維度, 例如 Valence-Arousal, 分別表示文本正負面程度和激動程度。反諷表達的主要特徵是, 字面意思與作者想要表達的真實情感不同。已有對於反諷的研究都是二分類研究, 屬於反諷判別。反諷的情感強度研究是一個新的挑戰, 我們使用不同於分類・表達的維度型表達對反諷文本進行情感強度分析。首先, 我們通過手工標記, 創建了一個 Valence-Arousal-Irony(VAI)三維的中文反諷語料庫。通過畫散點圖的方法我們對 VAI 之間的關係進行了分析。為了評估所建語料庫的可用性, 我們採用四種不同的深度學習方法, 分別使用 5-fold 交叉驗證方法對 VAI 值進行預測。實驗結果表明, CNN\_LSTM 模型對語料庫的 VAI 維度值預測效果最好。

關鍵字: 情感分析, 反諷語料庫, 維度型表達, 情感強度預測, 機器學習

# Construction and Evaluation of a Chinese Valence-Arousal-Irony Affective Corpus

Graduate student : Jing Liu

Advisor : Dr. K. Robert Lai

Department of Computer Science and Engineering  
Yuan Ze University

## ABSTRACT

An increasing amount of research has recently focused on representing affective states as continuous numerical values on multiple dimensions, such as the valence-arousal (VA) space. In irony sentiment analysis, most of the study are focused on irony detection, and no numerical values on irony intensity have been used as representing states. Compared to the categorical approach that represents effective states as different classes (e.g., irony and not irony), the dimensional approach can provide more fine-grained sentiment analysis. But no effective resources with irony ratings has been constructed. Therefore, we constructed an effective corpus called Chinese valence-arousal-irony NTU irony text containing 1,004 sentences based on NTU Irony Corpus[4]. To improve the annotation quality, a corpus cleanup procedure is used to remove outlier ratings and improper texts. Word2vec was used to represent words as mathematical vectors. Experiment using varies deep-learning algorithms (CNN, LSTM, CNN-LSTM, LSTM-CNN) to predict the VAI ratings of the corpus.

Keyword: Sentiment Analysis, Irony corpus, Dimensional expression, Sentiment Intensity Prediction, Machine Learning

## 志 謝

感謝元智大學和雲南大學給與我們這次在元智大學資訊學院學習的機會，在這一年的學習過程中，我在自然語言勘探實驗室學到了新的知識和技能，在元智大學的學習生活是我人生中具有重要意義的學習和生活經驗。來臺灣學習的機會是難得可貴的，我一定會珍惜並且銘記這一段學習的時光。

感謝元智大學資訊學院賴國華老師和禹良治老師對我的悉心指導和教誨，讓我在研究領域有了新的認知，特別感謝禹良治老師在研究過程中，教授與我的學習方法，以及在遇到問題時，老師對我的指導和幫助，讓我受益匪淺。資訊學院的老師們嚴謹的治學態度和淵博的學識，令我非常崇敬，感謝任課老師還有書報討論課與我們分享知識的老師們。

感謝實驗室同學們的無私幫助，感謝林柏霖，方冠傑，潘桓毅，何原野和林淑鶯同學，在我的研究工作過程中，向他們請教問題的解決辦法時，都給與我很好的建議，並且在和他們的討論過程中，我們互相建立了友誼，做到了互相學習進步。

感謝我的朋友和親人對我在物質和精神上的支持，讓我能夠順利的學習畢業。

對於上述所有幫助過我的老師，同學，朋友和家人，再次表示衷心的感謝，如果沒有他們，我的研究工作和碩士畢業論文就不能夠順利的完成。他們對我的幫助和關懷將是我心中最寶貴的記憶。

# 目錄

摘要 .....	iii
ABSTRACT.....	iv
志 謝.....	v
目錄 .....	vi
圖目錄 .....	viii
表目錄 .....	ix
第一章、引言.....	1
1.1 動機.....	1
1.2 挑戰.....	4
1.3 貢獻.....	5
1.4 組織結構.....	6
第二章、背景和相關工作.....	7
2.1 維度型表達的概念.....	7
2.2 反諷文本的研究現狀.....	9
2.2.1 反諷語料庫構建的現狀.....	9
2.2.2 非英文資源的研究現狀.....	10
2.2.3 購物網站和新聞領域的研究現狀.....	11
2.2.4 Twitter 文本的反諷分析 .....	11
第三章、中文反諷 VAI 語料庫的構建 .....	11
3.1 NTU 反諷語料庫 .....	11
3.2 中文反諷 VAI 語料庫的標記規則 .....	12
3.3 維度關聯分析.....	14
第四章、基於機器學習方法的反諷文本情感維度預測.....	18
4.1 反諷文本的 VAI 預測流程 .....	18
4.2 卷積神經網路 (CNN) 預測 VAI.....	20
4.3 長短期記憶人工神經網路 (LSTM) 預測 VAI.....	22

4.4 卷積-長短期記憶人工神經網路 (CNN_LSTM) 預測 VAI.....	25
4.5 長短期記憶人工神經網路卷積神經網路 (LSTM_CNN) 預測 VAI .....	26
第五章、實驗過程和結果.....	27
5.1 實驗說明.....	27
5.2 參數調整結果.....	27
5.3 實驗結果和分析.....	30
第六章、總結.....	31
參考文獻.....	32



# 圖目錄

圖 1、情感的 Valence-Arousal 空間標記法，不同情緒的詞可以映射在 VA 坐標系中。	8
圖 2、中文反諷 VAI 語料庫的 VAI 三個維度間關係的散點圖 .....	17
圖 3、中文反諷語料庫的 valence-arousal-irony 預測模型系統流程圖 .....	18
圖 4、卷積神經網路 CNN 預測句子維度值的架構圖.....	20
圖 5、CNN 模型中卷積層窄卷積操作示例圖[32].....	21
圖 6、CNN 模型中卷積層寬卷積操作示例圖[32].....	22
圖 7、標準 RNN 架構圖[29].....	23
圖 8、LSTM 輸入門結構[29] .....	24
圖 9、LSTM 遺忘門結構[29] .....	24
圖 10、LSTM 輸出門結構[29] .....	25
圖 11、卷積神經網路&長短期記憶人工神經網路 CNN_LSTM 架構圖 .....	25
圖 12、長短期記憶人工神經網路&卷積神經網路 LSTM_CNN 架構圖 .....	26





## 表目錄

表 1、反諷相關研究中所使用反諷判別方法的統計表 .....	9
表 2、中文 VAI 維度反諷語料庫資訊說明表 .....	14
表 3、傳統詞向量和分散式詞向量特徵對比表 .....	19
表 4、模型 LSTM 預測 VAI 維度值的最佳參數表 .....	28
表 5、模型 CNN 預測 VAI 維度值的最佳參數表 .....	28
表 6、模型 CNN_LSTM 預測 VAI 維度值的最佳參數表 .....	29
表 7、模型 LSTM_CNN 預測 VAI 維度值的最佳參數表 .....	29
表 8、四種模型對中文反諷 VAI 語料庫三個維度值的預測結果表 .....	30



# 第一章、引言

## 1.1 動機

隨著社交網路的蓬勃發展以及網路中文本的大量生成，科研工作者可以從網路中獲取足夠的自然語言文本進行自然語言分析研究工作。使用電腦對自然語言文本進行自動的情感語意識別，是我們不斷追求的目標。在社交工具和大型購網站中，有大量的使用者生成文本，從這些文本中能夠挖掘出使用者的觀點/意向/情感，等有用的資訊。深度學習方法的發展和應用越來越受到人們的重視，在自然語言研究方面，深度學習方法在語意/情感上的自動預測有良好的表現。大量文本資料和深度學習方法的結合使得機器自動預測自然語言的語意和情感成為可能。在文本表達中，有一種異於一般表達的方式，就是反諷表達。反諷表達，用於表示不同於字面意思的語意和情感。本文中我們的工作就是對反諷文本的構建和情感值的預測。

文本的情感分析在自然語言分析中佔據重要地位。目前對於文本情感的表達包括兩種方法，一種是離散型的分類方法(categorical)，有二分類（正面，負面），多分類（喜，怒，哀，樂）等分類方式；另外一種是連續性的維度型表示方法，如[1]所描述詞彙的 Valence-Arousal(VA)二維表示，其中 valence 表示正負面程度，arousal 表示激動程度。維度型的表示方法相對於類別型表示方法能夠更加細微性的描述文本的情感強度，因為維度型表示方法能夠使用連續不同的數值表示文本的正面或者負面程度，在第二章中我們會講述維度型表達的概念。在文本情感分析中維度型表達的分析方式是有必要的，例如句子“我喜歡那個包包。”和“我真的好喜歡那個包包啊！”，這兩個句子都是表達喜歡一樣東西，但是情感強度是後者強於前者很多。所以維度型的表達方法有助於在情感分析中能夠更準確的預測句子的情感強度。這對於更加準確的預測文本作者的情感傾向是非常有幫助的。

反諷表達是自然語言中一種較為特殊的表達方式，特點是句子中具有正面詞彙，但上下文中含有負面語意。根據上下文會發現句子中的字面意義與作者想要表達的情感意向相反。以下是取自於 NTU Irony Corpus[4]的範例。

(1a) 真是太帥了。

(1b) 好事情在我身上不會超過一天，真是太帥了。

(2a) 很好~

(2b) 很好~ 昨天忘了吃藥~ 今天耳朵又淹水了。

(3a) 真是太"幸運"了。

(3b) 今天去打球打到眼鏡鏡片掉下來，回家途中車子還拋錨壞在路邊...真是太"幸運"了。

上面的例子中，(1a)、(2a)和(3a)是不考慮上下文的例句，可以看出這些都是正面的句子，但是讀取上下文之後會發現，作者想要表達的意義是有不好的事情發生，情感偏向於負面消極，所以這些例句都是反諷表達。除了 VA 二維表示之外，反諷表達還存在第三維表示，也就是反諷 (Irony)，因為除了正負面(V)程度和激動(A)程度之外，反諷表達還存在反諷程度，所以 Irony(I)是對反諷強度的維度型描述。對反諷表達從三個維度(VAI)同時進行預測分析，就可以分別從正負面程度，激動程度和反諷程度對反諷文本進行情感分析，這是一種對於反諷表達更加細細微性和全方位的分析方式。

至今，對反諷文本分析領域的研究中，我們發現在反諷研究領域中反諷語料庫非常稀缺，連續性的維度型反諷研究也還沒有人涉足，大多研究者的工作都集中於對反諷表達的判定，已有的反諷語料庫有 (Amazon 產品評論語料庫) [6] (reddit irony corpus)新聞評論語料庫與[12]，這兩個反諷語料庫中的文本標記都是類別型標記，也就是被標記為屬於反諷或者不屬於反諷。然而連續型的維度型標記的中文反諷語料庫沒有被創建過，因此，在本文中我們的主要任務就是通過人工標記的眾包方法建立一個中文反諷 Valence-Arousal-Irony 三維的語料庫，然後對該語料庫使用深度學習方法進行實用性評估。

在對反諷文本的分析中，神經網路深度學習方法被用來對反諷文本進行分類預測，[27]使用 CNN (卷積神經網路)，LSTM (長短期記憶神經網路) 和 DNN (深度學習網路) 以及他們的組合方式 (CNN-CNN)，(LSTM-LSTM)，(CNN-LSTM-DNN) 對反諷文本進行分類預測。[26]創建了 Content and User Embedding Convolution Neural Network (CUE-CNN)方法對反諷文本進行分類。使用深度學習方法對反諷文本進行分析獲得了較好的評估結果，因此我們計畫對本文中創建的

維度型反諷語料庫使用深度學習方法進行維度預測。

分散式詞嵌入技術 word2vec[7]在自然語言研究領域的貢獻非常顯著，[27][28]都是使用 word2vec 作為訓練詞向量的方法，在此基礎之上對反諷文本進行分類型預測，並取得好的結果（F-score of .92）。因此我們考慮以 word2vec 技術為基礎訓練我們所創建的語料庫中的文本資料，在此之上，使用深度機器學習方法對反諷文本進行維度預測分析。



## 1.2 挑戰

目前反諷語料庫比較稀缺，純反諷文本預料庫更加稀缺。因為，反諷表達並不是人們常用的情感表達方式，所以資源相對於一般正常表達較少，獲取過程非常複雜。同時基於反諷表達的特殊性，我們認為反諷文本語料庫的構建和評估存在挑戰性。具體挑戰性描述如下：

反諷表達的特殊性主要在於文本結構和用詞的特殊。反諷文本結構與常用文本結構不同，通常具有上下文語意相悖的文本環境，這種結構會增加對反諷文本情感的自動預測難度。例如：“中午開三小時會議，晚上又加開三小時.. 真是太精實太有效率了..”，如果不結合上下文，“精彩”，“有效率”這些正面詞彙極易導致這條句子被預測為正面積極的表達。但實際上，結合上文就能知道這條句子的語意是抱怨會議時間太長，效率太低，是負面消極的情感表達。所以如何有效的結合上下文來準確的預測反諷文本是具有挑戰性的工作。

反諷文本的獲取工作比較複雜。網路上的自然語言文本資源大多是常用表達，相對而言，反諷文本的量比較少。另外，由於反諷文本結構特殊，所以並沒有關於反諷文本結構的規範定義，如何確認一個文本是否是反諷句需要花費很多的人力去判斷。基於上述原因，在大量的網路資源中確定一個文本是否是反諷文本，並將它歸於反諷文本語料庫是一個具有挑戰性的工作過程。

維度值的預測比較困難。維度型表示方法能夠細細微性的表示情感強度，因此有助於準確地預測文本的情感傾向，但是如何能夠更加準確地預測維度值是重中之重，也是比較複雜的任務。維度型表示方法相對於類別型方法更具有難度，因為維度型是對於情感表達使用連續性的數值來描述的，比如“很好，感冒了”和“很好，又感冒了”兩個句子都是反諷表達，但是在強度上後者的負面強度、激動程度和反諷強度都相對於前者更加強烈，所以準確的預測出反諷句子的強度是有難度的。綜上，我們得出結論，探索能夠在反諷文本 VAI 值的自動預測上表現良好的自動預測方法是我們在本文中要做的工作。

### 1.3 貢獻

在本文中，我們首先建構了一個中文反諷 VAI 語料庫，最大的貢獻就是為以後的反諷研究提供了資料基礎和研究基線。創建過程是基於臺灣大學創建的 NTU 反諷語料庫[4]，對 1005 條反諷文本進行 VAI 三個維度的標記。標記工作使用 google cloud platform 作為平臺，以(SAM)[5]模型作為標記原則。標記工作完成之後，通過標記資料處理過程之後，獲得了每個句子的 VAI 標記值，對每個句子的標記值進行異常處理和求平均，用最後得到的平均值作為語料庫中句子的真實標記值。我們對 VAI 三個維度之間的關係通過畫散點圖的方法，在第三章的第三節進行了關係分析和總結。最後，對語料庫進行實用性評估，評估方法是使用不同的機器學習方法對語料庫進行學習和自動預測 VAI 值，找出預測結果最好的方法。為了表示詞向量特徵，我們使用 word2vec[7]模型訓練的維琪百科詞向量作為我們語料庫中出現的詞彙的向量表示來源。實驗過程是使用三個機器學習方法(CNN, LSTM, CNN-LSTM)分別對該語料庫的 VAI 進行 5-fold 交叉驗證維度值預測。



## 1.4 組織結構

第一章是引言，主要介紹了自然語言分析和反諷文本情感分析的工作動機，還有挑戰性和本文工作的貢獻。第二章主要介紹情感分析的研究背景知識和反諷文本研究的相關工作。第三章是中文反諷 VAI 語料庫的建構過程。第四章是對語料庫的評估工作，內容包括 VAI 維度關聯分析和基於機器學習方法的反諷文本情感強度預測。第五章是總結。



## 第二章、背景和相關工作

本章介紹情感分析的背景知識和反諷文本的研究現狀。維度型表達的概念。反諷研究中類別型的研究現狀。

### 2.1 維度型表達的概念

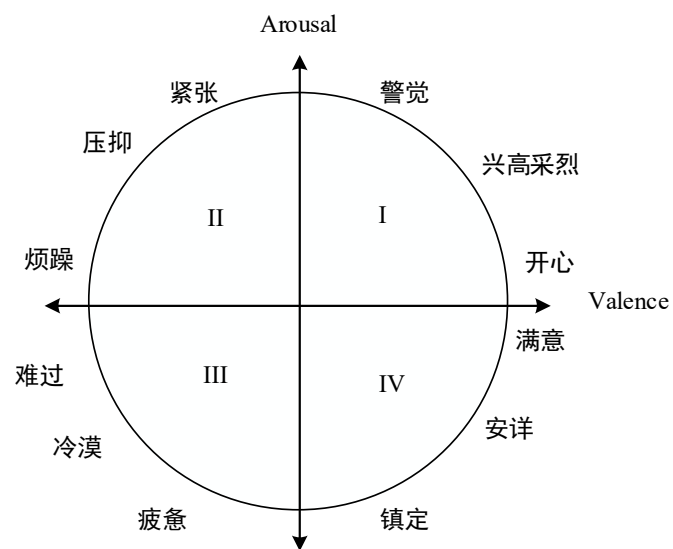
維度型表達，就是使用連續的數值對詞彙或者文本某種情感的連續性表達。目前常用的對於詞、句子，或者長文本進行維度型描述的維有 Valence 和 Arousal。如圖 1 所示，橫軸表示 valence，縱軸表示 arousal，他們的數值分別代表正負面程度和激動程度，區域 I 表示的情感為正面並且激動，比如詞彙（警覺、興高采烈、開心），區域 II 表示的情感為負面並且激動（緊張、壓抑、煩躁），區域 III 表示的情感為負面並且冷靜（難過、冷漠、疲憊），區域 IV 表示的情感為正面並且冷靜（鎮定、安詳、滿意）。該圖的繪製引用了雲南大學資訊學院何雲超同學的研究生畢業論文《使用聚類及卷積神經網路於類別型及維度型情感分析之研究》圖 2。其中 VA 的數值範圍可以是[1,9]，對於 V 值（橫坐標），1 表示非常負面，5 表示中性既不正面也不負面，9 表示非常正面。對於 A 值（縱坐標），1 表示非常平靜，5 表示一般激動，9 表示非常激動，從 1 至 9，表示激動程度不斷增強。在這裡我們用詞作為實例物件對這兩個維度進行描述，例如，“興高采烈”這個詞位於座標空間的右上角位置，表示這個詞所表達出的二維情感是正面程度和激動程度都非常高。詞“滿意”，位於橫坐標的最右端的位置，位於縱坐標中間位置，情感傾向為非常正面並且一般激動。

對於句子級別和長句字級別的文本，VA 維度型情感表達方式同樣適用。反諷句除了 VA 維度之外還有第三維度（反諷強度），不同於一般表達方式，所以反諷強度是反諷表達特有的一個維。因此本文中，我們對反諷文本進行三個維度的預測，即 Valence-Arousal-Irony，Irony 就是指反諷強度，在本文中我們定義反諷強度範圍為[1,9]，其中 1 表示反諷強度很弱，5 表示反諷強度一般，9 表示反



諷強度很強。

圖 1、情感的 Valence-Arousal 空間標記法，不同情緒的詞可以映射在 VA 坐標系中。



## 2.2 反諷文本的研究現狀

表 1、反諷相關研究中所使用反諷判別方法的統計表

Method	Paper_ID ( 參考文獻編號 )
SVM	[8],[9],[11],[12],[13],[14],[16],[18],[24],[25],[27],[28],[30]
Decision Tree	[2],[8],[9],[15],[30]
Random Forest	[2],[8],[15]
Naïve Bayes	[8],[9],[30],[31]
Logistic Regression	[8],[20]
CNN	[26],[27]
LSTM	[27]
DNN	[27]
other	[3],[10],[21],[17],[26],[23]

至今，對於反諷研究的工作都是集中於對反諷文本的分類性預測，也就是，使用分類型(categorical)表達描述反諷文本，研究目的是判斷一個文本是不是反諷表達。下面，我們基於目前反諷的研究工作，著重對反諷語料庫和預測方法進行了統計分析。其中，語料庫的構建都是來源於網路資源 (Twitter、微博、Amazon、新聞網站)，分類方法有決策樹(DT)、隨機樹(RF)、支援向量機(SVM)、朴素貝葉斯(NB)和邏輯回歸(LR)，如表 1 所示，是反諷相關研究工作使用的預測方法的分類表，其中使用 SVM 和 DT 比較多，而使用 CNN 和 LSTM 比較少。因此我們在本文中會採用 CNN 和 LSTM 方法對反諷文本進行預測。

### 2.2.1 反諷語料庫構建的現狀

目前，以不同的網路應用領域作為資源來源，已經建構了一些有效的反諷語料庫。例如，[6]以亞馬遜網站的產品評論為文本建立了一個反諷語料庫，語料庫中的所有評論文本針對一個特定的產品，同時包含普通評論和反諷評論，研究人員在此語料庫的兩個級別（文檔和文本表達）上進行反諷判別的研究。(Warriner et al. 2013)將(Bradley and Lang (1999))創建的英文語料庫擴展至 14,000 個詞彙。這個語料庫是在三維標準(valence, arousal,和 dominance)之上建立的，並且包括幾乎所有分類標準的激勵因數，例如疾病種類，職業種類，和禁忌詞彙。

在反諷研究領域中，使用用維度型表達（也就是連續型標記）來表示反諷強度的語料庫還沒有被創建過。第一個反諷中文語料庫[4]，語料庫中的句子被標記為反諷或者非反諷，是分類型表達而不是連續型表達。因此，至今為止，連續型標記的中文反諷語料庫也未被創建過。我們在本文的第三章將會描述維度型中文反諷語料庫的創建過程和結果。

反諷研究在情感分析中有重要的作用，很多的研究者在不同語言不同領域做了很多自動識別反諷表達的研究工作。一些研究者[9 10 12 15 17 18 20 23 28 2 19]基於已經存在的一些分類器做反諷自動判別研究，例如，SVM（支持向量機），DT（決策樹）。還有一些研究者[3 11 21 24 26 27 29]基於自己創建的分類系統模型做反諷的自動判別研究。

反諷語料資源大都是來自于網路應用，例如，博客，微博，購物網站，新聞網站等。絕大多數的反諷研究者[12 17 18 20 23 28 2 19 3 21 26 27 29]都是從 Twitter, Amazon 獲取他們的語料資源，少部分研究者[9 10 15 24 11]是從博客和新聞網站獲取預料資源。這些研究工作都集中于判別文本是不是反諷表達，也就是，研究者們將文本標記為兩類（反諷和非反諷），在此基礎之上對文本進行預測分類。在這些研究中並沒有維度型的表達方式來表示反諷情感強度。因此本文中，我們對於反諷表達維度型的研究是具有創新性的。

### 2.2.2 非英文資源的研究現狀

在對反諷研究的語言種類進行統計時，我們發現除了英文之外還有對其他語言的研究，其中[17 20 23 28 2 19]是從 tweeter 中獲取的英文語料資源。還有一些研究 [12 18 3 21]是其他種類的語言。[3]使用推特 API 收集了法語推特文本，他們提出了一個創新性的三步模型：1)僅基於文本內部資訊，使用監督學習方法判別推文是否是反諷的。2)測試內部上下文而不是外部表達的上下文。3)當字面意義不能被理解時，就意味著推文很可能傳達了非字面的意義。該研究得到的結論是，當資料包含所有否定和非否定的推文時，準確性比較好，並且內部上下文特徵有助於改善分類問題。2010 年荷蘭科技期刊提供的荷蘭語推文[17]，該研究使用的分類演算法是平衡篩選器（Balanced Winnow）(Littlestone, 1988)。他們發現大多數的推文包含字面意義是正面的文本資訊和三種類型的詞彙（明確的標記，

加強詞和感歎詞)，所研究推文的主題目標是普通的青少年主題（學校，家庭作業和家庭生活）。[11]收集了兩種主要的特徵，域依賴特徵和域非依賴特徵（[15]提議），該研究的語言是義大利推文，本文被標記為四種規模（主觀性/客觀性，正面性/負面性，反諷/非反諷和政治的/非政治的），屬於類別型標記。[14]創建了一個大型的捷克語推文語料庫，包括 7000 條手工標記的推文，其中反諷文本的指示器是反諷標籤（#sarcasm hashtag），他們的實驗結果表明，使用 SVM 分類器和特徵資料集 “B1+POS characteristics+ Emoticons+ Word-case” 的效果優於 n-gram 基線，並且還得出結論在反諷研究中，捷克語的研究比較有挑戰性。

### 2.2.3 購物網站和新聞領域的研究現狀

購物網站和新聞網站的評論文本經常被用於作為語料資源的來源，例如，reddit 子網反諷語料被[10]和[12]用來創建了來源於六個 reddit 子網的帶評論注解的語料庫。通過分析語料庫，他們提供了一個經驗主義的結論，在反諷判別問題中語境是必須考慮的因素。

[21]基於 Amazon 創建了一個包含 66000 條產品評論的語料庫，分類實驗過程使用的是半監督式反諷識別方法(SASI)[33]。[8]使用[6]基於 Amazon 創建的語料庫，六個不同的分類器被用來判別反諷表達，其中，基於星號評級，bag-of-words 和特定特徵的邏輯回歸方法表現最優。[9]通過收集一系列的區分性特徵，使用 Amazon 和 Slashdot 的正面資料和負面資料來重現反諷，所有的文檔通過代表性比率(representativeness ratio)被描述為頻率加權向量。

### 2.2.4 Twitter 文本的反諷分析

Twitter 文本是反諷研究者最常用語料來源。[22]提供了預測反諷的一種特殊方法，通過使用作者的歷史推文，來預測現在發佈的文本是否為反諷表達，整個方法包括兩個組成部分：1) 一個基於對比的預測器；2) 一個基於歷史推文的預測器。其中，測試資料集，來自於[18]，基於歷史推文預測器的條件下，他們獲得的評估結果與(Diloff et al.2013)相當。對比之下，當資料集是一組來源於情感詞字典和含蓄情感短語的兩極詞彙時，使用基於對比的預測器獲得了較好的預測結果。[23]使用自助(Bootstrapping method)方法來訓練分類器，為了識別網路對話

中特殊類型的主觀表達，他們使用眾包技術創建了一系列反諷和不潔詞的語言指示器（與[6]中使用的方法相似），並且創建了分類器，這個分類器的創建借鑒了前人創建的自助主觀性的分類器(bootstrapping subjectivity classifiers )工作[19][20]，當他們使用句法模式來創建更多的一般反諷指示器時，就獲得了更高的精確度和召回率。[25]收集了一組或者普通或者具有反諷意義的目標詞彙，提出了一個兩步方法：1) 一個眾包任務，收集一系列平行的資料集，它們是作者的反諷表達和能夠傳達作者情感意圖的改述；2) 能夠判別語義對立的詞/短語的無監督佇列技術。他們使用了三種不同的方法獲取詞嵌入，分別是全文本矩陣因數(WTMF) (Guo and Diab, 2012b)；word2vec. (2013a)，和 GloVe (Pennington et al., 2014)。實驗結果表明，使用詞嵌入和改進的 SVM 核方法得到了最優的實驗結果。

[13]使用上下文的不一致來判別推文是不是反諷表達，在使用不同的特徵組合訓練了分類器 SVM 之後，他們的系統表現優於[18]；Maynard and Greenwood, 2014)。[16]創建了他們自己的反諷推文語料庫，語料庫中文本的反諷標記是由原作者來決定的，最終此語料庫包括 900 條推文，隸屬於三種分類：反諷(S)，正面(P)，負面(N)。他們得出的結論是：不管是人類判斷還是機器學習，反諷判別都是比較困難的。[24]的研究工作集中於四種方法：1) 情感分析；2) 概念級別和普遍的常識知識；3) 一致性和 4) 機器學習分類方法。SVM 方法基於兩種不同的特徵集相比於兩種基線方法得到了較高的精確度。這兩種基線方法分別是：反諷定義和 SVM 方法基於 N-gram 特徵。有些研究者將推文分類為幾個不同的主題，比如，[2]將推文分為四個不同的主題：反諷，教育，幽默和政治。最後三個主題被認為是非反諷的。[15]將推文分為六個不同主題：反諷，教育，幽默，諷刺，政治和新聞。[2]與[15]的實驗結果類似，所使用的分類方法中 RF 和 DT 表現比較好。

綜上，目前對於反諷的研究都是分類研究，使用的判別方法各不相同。在本文中，我們對反諷的研究擴展至維度型的預測，使用的預測方法是深度學習網路，深度學習網路具有更加智慧化的優勢，例如 CNN 和 LSTM，深度神經網路能夠針對一條句子的不同區域進行不同權重的分析，能夠有效地整合跨句資訊，從而對句子進行維度型預測。

## 第三章、中文反諷 VAI 語料庫的構建

### 3.1 NTU 反諷語料庫

中文反諷 VAI 語料庫是基於臺灣大學的 NTU 反諷語料庫所創建的。下面部分對於 NTU 反諷語料庫的介紹引用[4]中的內容

首先介紹 NTU 反諷語料庫的創建過程。在 NTU 反諷語料庫的創建過程中有三點重要的貢獻：1) 反諷語料庫構建。2) 反諷模式挖掘。3) 反諷結構分析。其中，突出貢獻是，NTU 反諷語料庫是第一個可用的中文反諷語料庫。語料庫的創建第一步，是通過使用自展方法（bootstrapping methodology）從微博中選出 3,178,372 條既包含消極情感又包含積極詞彙的文本，第一步的完成就構建了反諷語料庫的候選資料集。第二步，過濾掉含有“but”類連接詞的文本，這些連接詞包括：“但”，“但是”，“可是”，“只是”，“不過”，“然而”，“卻”，“可惜”，“偏偏”，“反而”，“倒是”，去掉含有此類連接詞的文本之後，文本集剩餘 254,836 條文本。第三步，是通過一個自展處理過程，去發現更多的反諷文字模式，每一次迴圈過程，都根據最新發現的反諷模式去尋找符合新模式的句子，得到了最終的 1,004 條被確認為反諷文本的語料庫，這些文本最終構成 NTU 反諷語料庫。

第三步的自展處理過程如下：

- 1) 首先使用已被確認的反諷模式[degree adverb + positive adjective]。
- 2) 文本中屬於 1) 中所提到模式的文本被取出。
- 3) 讓標記者來決定 2) 中取出的文本是否是真正的反諷文本。
- 4) 如果標記者在被重新查看的文本中發現了新的反諷模式，則使用新的反諷模式從 1) 開始重複整個過程。

得到的反諷模式如下[4]：

- 1) Degree Adverbs + Positive Adjective

此模式中，必須包含如下兩個組成部分：

- (a) Degree adverb phrase + positive adjective phrase (b) Negative context

(s1) 點餐都要等半小時，服務還真是好阿

## 2) The Use of Positive Adjective with High Intensity

此模式中，必須包含如下兩個組成部分：

Positive adjective with high intensity (b) Negative context

(s2) 我的 plurk「又」發生不明錯誤了...這真是這世紀最偉大的發明啊

## 3) The Use of Positive Noun with High Intensity

此模式中，必須包含如下兩個組成部分：

(a) Positive noun with high intensity (b) Negative context

(s3) 中秋節收到的大禮是.....長了一堆肉

## 4) The Use of “很好” (very good)

此模式中，必須包含如下兩個組成部分：

(a) Sentence boundary + 很好 + punctuation (b) Negative context

(s4) 感冒... 很好!! 我的假期飛了

## 5) “可以再...一點”

此模式中，必須包含如下表達：

可以再 + negative adjective + 一點

(s5) 零下十一度...你可以再冷一點

這五種反諷模式所占總的句子總數的比例以及每種模式的句子數目，在 NTU 反諷語料庫中符合反諷模式 3 和模式 4 的句子占總體的 90%，表明在微博中，使用“很好”，“可以再...一點”這些表達來表示反諷情感是一種較普遍的現象。其他模式的反諷表達數量較少，因為其他模式在句子結構和表達方式上比模式 3 和模式 4 都更加有難度。

最終得到的語料庫 NTU 有 1005 條句子，其中，最長句子中詞的個數為 87。最短句子中詞個數為 4。另外句子中的詞被標記了 3 類標籤：(1) ironic word/phrase, (2) context, and (3) rhetoric element。例子：<context sentiment="pos">才剛買的書，竟然掉頁了，</context>這品質<rhetoric>也太</rhetoric><ironic sentiment="neg">好</ ironic>了<rhetoric>吧</rhetoric>。

## 3.2 中文反諷 VAI 語料庫的標記規則

語料庫的標記工作是通過使用 google cloud platform 平臺將標記應用發佈至伺服器，使用眾包（crowdsourcing）的方式對 NTU 反諷語料庫進行標記，標記者通過訪問網址並通過郵箱帳號登陸就可以開始進行標記工作。

標記工作要求每個句子至少要被標記 10 次，也就是，每個句子的 V、A、I 值三個維度分別至少都要有 10 條標記，並且，標記系統能夠保證每條句子最多只能被標記十次。三個維度的數值範圍是 (1-9)，每條句子的三個維度標記由同一標記者同時完成，標記者通過選擇 9 個數值中一個數值的方式來進行點選標記。標記頁面採用 Self Assessment Manikin (SAM) [5]規則，特點是，頁面中提供的圖像會給予標記者提示作用，可以說明標記者做出準確快速的標記判斷。

V 值表示句子的正負面程度，以 5 為分界值表示中性，小於 5 都表示負面，不同數值分別表示不同程度的負面，1 表示非常負面。大於 5 表示正面，不同值分別表示不同正面，9 表示非常正面。A 值表示句子激動程度，從 1 至 9 表示激動程度越來越大，1 表示非常平靜，5 表示一般激動，9 表示非常激動。I 值表示反諷程度，從 1 至 9 表示反諷程度越來越大，1 表示反諷程度很弱，5 表示反諷程度一般，9 表示反諷程度非常大。

標記工作完成之後，使用一個資料處理程式對標記資料進行 outlier 處理，也就是，對每個句子的三個維度分別進行合理範圍外的標記資料清理工作。以  $[\text{mean}-1.5*\text{SD}, \text{mean}+1.5*\text{SD}]$  為合理標記範圍。將不符合規範的標記值去除之後，再對每個句子的三個維度標記值進行求平均，最後得到的平均值將作為最終的真實標記值。最終得到每個句子的三維標記值，就得到了我們的中文反諷 valence-arousal-irony 語料庫，語料庫的屬性如表 2 示。



### 3.3 維度關聯分析

表 2、中文 VAI 維度反諷語料庫資訊說明表

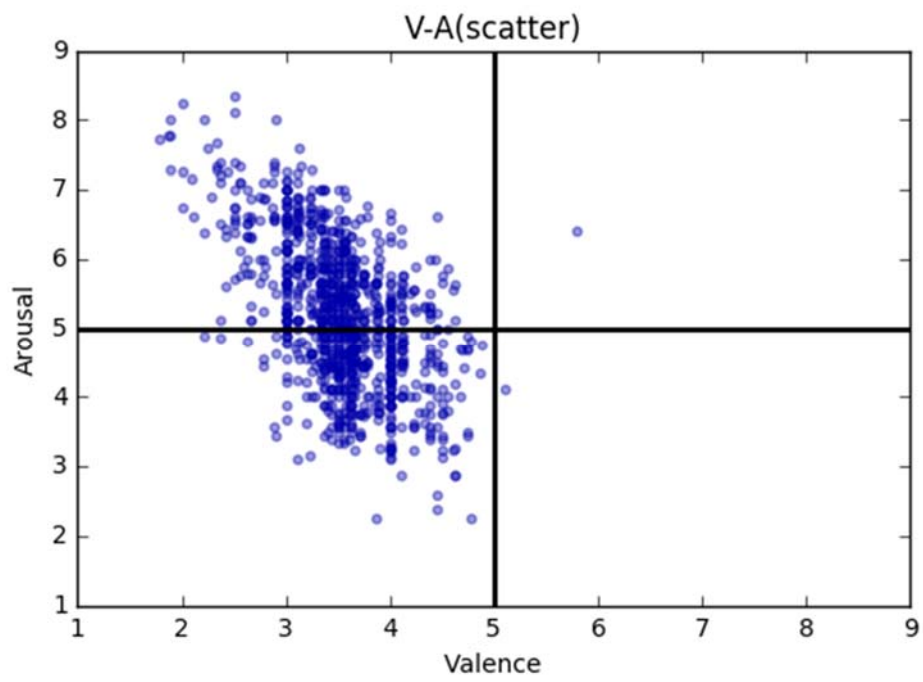
語料庫名稱	中文反諷 VAI 語料庫
語料庫類型	反諷
語料庫來源	微博留言板 from Plurk website
語言	中文繁體
句子個數	1005
最長句子詞數	87
最短句子詞數	4
標記維度	Valence、Arousal、Irony

標記工作結束後，我們通過畫散點圖的方式對 VAI 三個維度之間的關係進行分析。如下圖 2 中的(a)，(b)，(c)三個分圖，分別表示，VA，VI，AI 之間的散點圖分佈關係，圖中的散點表示語料庫中所有的句子分佈，橫坐標和縱坐標分別表示該句子的 V，A 或者 I 的標記值。

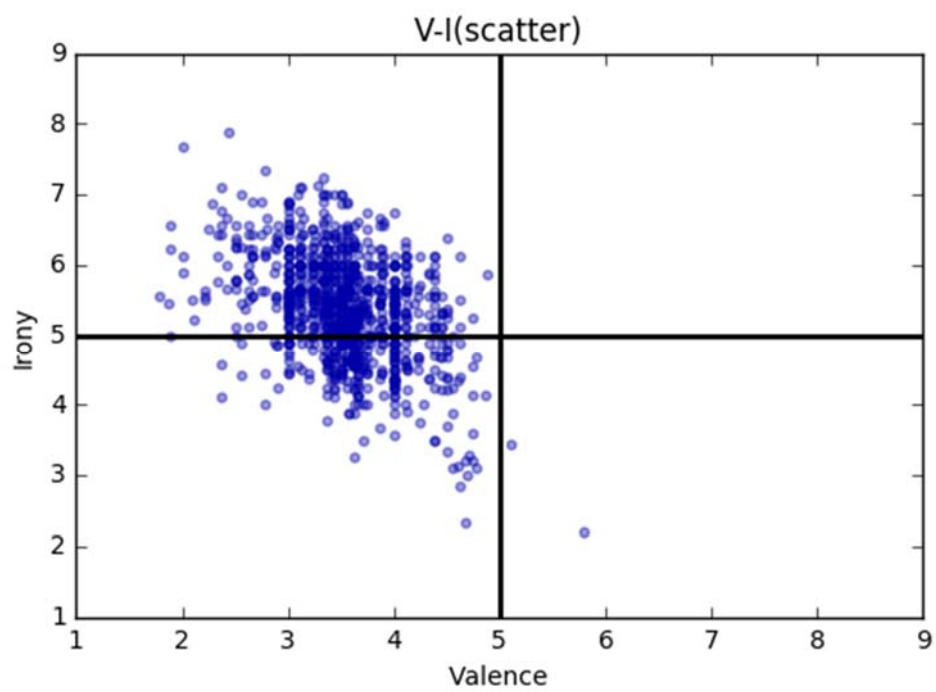
圖(a)中，V 和 A 兩個維度呈負相關線性關係，絕大多數的 V 值小於 5，說明基本上所有句子所表達的感情傾向都是負面的，也說明了，我們所創建的語料庫中的句子都是負面反諷文本。A 值分佈相對於 V 值，佔據的數值範圍較廣，即語料庫中的反諷句情感表達中既有激動的情感也有平靜的情感，從線性關係中，我們可以看出句子的負面程度越大，句子情感的激動程度也越大，因此可以看出，反諷表達越負面越激動。圖(b)中，VI 之間也呈現負相關線性關係，隨著句子的負面程度越來越大，反諷程度也越來越強，說明在反諷表達中，人們習慣於使用更強烈的反諷強度來表示更加負面的情感。圖(c)中，AI 之間的分佈呈現正相關線性關係，句子的激動強度越來越強，反諷的強度也越來越強。綜上可以看出，反諷表達中，當人們使用反諷強度非常強的表達時，情緒上傾向於激動程度比較大，並且所表達的語義傾向於更加負面的情感。同時，我們還能夠發現，網路中

的反諷表達絕大多數是負面的情感表達，人們習慣於使用反諷來表示對人或者事的不滿和嘲諷，而反諷語氣的強度和情感表示的強度差距很大，有輕度平靜的反諷也有激動的強烈的反諷。

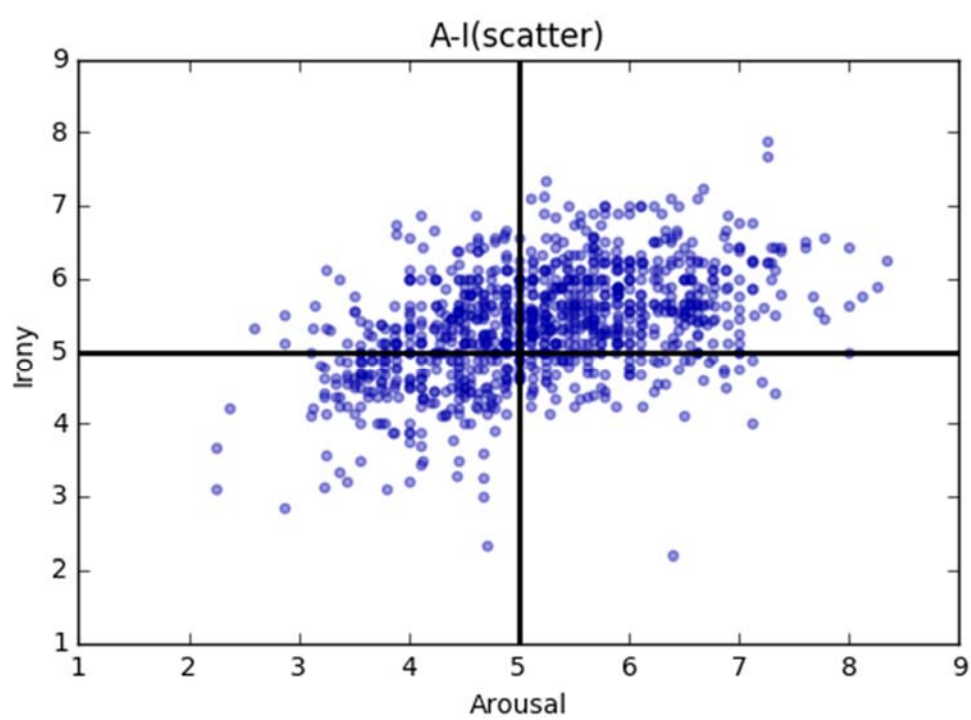




(a)



(b)



(c)

圖 2、中文反諷 VAI 語料庫的 VAI 三個維度間關係的散點圖



## 第四章、基於機器學習方法的反諷文本情感維度預測

在本章中，我們將介紹對於維度型反諷語料庫進行預測的整個過程，以及如何使用詞向量和三類機器學習方法對反諷文本的情感強度進行預測。實驗過程中，我們對語料庫中句子的 VAI 三個維度進行預測分析，採用 5-fold 交叉驗證方式。將五次實驗結果的平均值作為最終評估結果。這三類機器學習方法分別是：卷積神經網路（CNN），長短期記憶人工神經網路（LSTM），前兩種神經網路方法的組合方法：（CNN\_LSTM）。

### 4.1 反諷文本的 VAI 預測流程

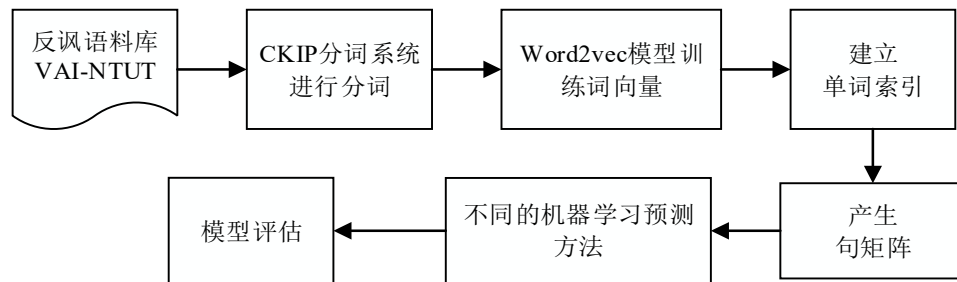


圖 3、中文反諷語料庫的 valence-arousal-irony 預測模型系統流程圖

如圖 3 所示，我們通過系統架構圖對語料庫的預測系統進行描 CKIP 分詞系統進行分詞：使用 CKIP 分詞系統對語料庫中的反諷文本進行分詞，CKIP 是中文繁體斷詞系統。為了獲取詞向量，必須對語料庫中的句子進行斷詞處理。

**Word2vec 模型訓練詞向量：**word2vec[7]訓練詞向量的過程是使用一個雙層的神經網路對輸入的文本進行處理後，會輸出文本中詞彙的特徵向量。本文中使用的公開的 word2vec 詞向量，所訓練的文本來源是中文維琪百科。詞向量為 300 維，具體使用方法為 skip-gram 方法，這種方法處理大規模資料集的結果比

表 3、傳統詞向量和分散式詞向量特徵對比表

詞向量方式	Distributed Representation	One-hot Representation
實例	Word2vec	傳統詞向量表示
向量表示	[0.788, 0.654, -0.177, -0.542, -0.109 ...]	[0 0 0 0 0 0 1 0 0 0 0 0 ...]
缺點	計算複雜	高維度、稀疏問題、詞彙鴻溝
優點	低維、相似詞距離相近	計算簡單

較準確。如表 3 示，以往 NLP 研究中，常用的詞向量表示方式方法是 One-hot Representation，其中詞向量的維度表示詞表大小，例如一個詞向量表示為[0 0 0 0 0 0 1 0 0 0 0 0 ...]絕大多數是 0，只有一維是 1 來表示當前詞。當詞表比較大時，這種稀疏陣列的表示方式會導致每個詞的詞向量維度都特別長，造成維度災難。而 word2vec 是一種 Distributed Representation 稠密低維的實數向量表示，向量表示形如[0.788, 0.654, -0.177, -0.542, -0.109 ...]，詞向量的每一維都表示了詞的一個潛在特徵，這些特徵能夠總體捕獲句子的語法語義特徵。

**建立單詞索引：**通過對所有句子中出現的詞彙進行統計，對屬性不同的單詞賦予不同的整數值，作為單詞的索引。

**產生句矩陣：**對於給定的一個句子，通過句子中出現的單詞找到單詞索引，用代表詞彙的不同整數數值序列表示句子。句矩陣中，每一行表示一個詞的詞向量，行數表示句子中包含的詞彙個數。

**不同的機器學習預測方法：**本文中，我們使用三類機器學習方法對反諷語料庫進行 VAI 的預測。這三類方法分別是 CNN, LSTM, CNN-LSTM。後文中我們會分別對這三類方法進行分析。

**模型評估：**我們使用平均絕對誤差評估指標對模型的而結果進行評估，是

MAE (平均絕對)。MAE 是對預測值與真實值之間差距的評估。

## 4.2 卷積神經網路 (CNN) 預測 VAI

我們使用一維的卷積神經網路對語料庫中的反諷句子進行 VAI 預測，一維是指卷積核的寬度與詞向量的維度相同，卷積核只在一個方向上移動，對句矩陣進行卷積操作。這裡我們用一個簡單的模型對一維卷積神經網路進行描述，如圖 4 所示，句子“很好一顆饅頭從早吃到晚”一共有 8 個詞彙。假設詞向量維度為 10，則用句矩陣表示是 8 行 10 列的矩陣。

圖中句矩陣的每一行表示一個詞的詞向量，詞向量的維度為 10，在實際操

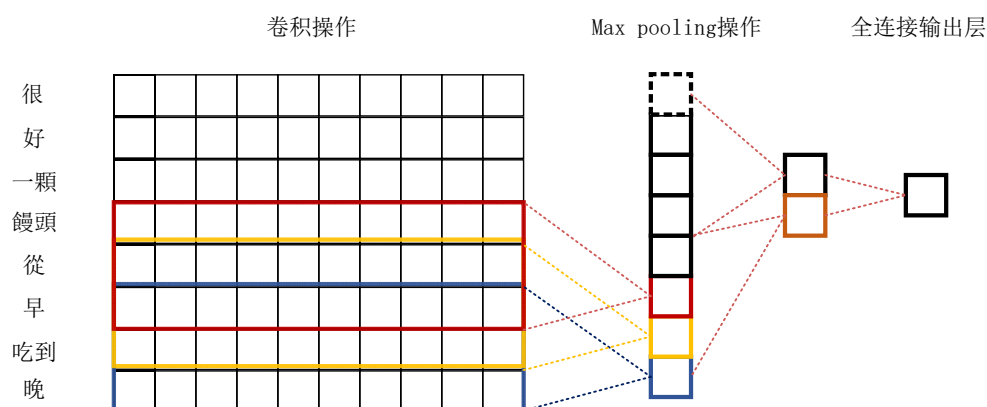


圖 4、卷積神經網路 CNN 預測句子維度值的架構圖

作中詞向量的維度為 300。圖示卷積操作中卷積核的數目 nb\_filter=7，卷積核的長度為 10，卷積核的尺寸 (kernel\_size) 為 filter\_length=3。

具體的操作過程是，長為 10 寬為 3 的卷積核對句矩陣進行卷積運算之後，輸出 7 維的輸出向量。卷積操作分為廣卷積 (wide convolution) 和窄卷積 (narrow convolution)，其中窄卷積對處於邊緣的輸入資料所做的卷積運算次數要少於處於中心位置的輸入資料，而廣卷積操作過程中，對於超出卷積核的部分用 0 補充，可以做到對所有輸入資料均量次數的特徵提取。卷積操作之後，進入池化 (max pooling) 操作，池化操作分為最大池化操作 (max pooling) 和平均池化操作 (mean pooling)。一般採用最大池化操作，池化操作之後能夠輸出固定大小的矩陣，也就是說無論輸入的句子長度為多少，經過池化操作之後都能夠得到一個

相同維度大小的輸出結果。如圖 4 所示使用了 2 個濾波器，就會得到一個 2 維的輸出。總之，池化操作降低了輸出結果的維度，並且保留了顯著特徵。池化層輸出的向量將會被輸入到線性回歸層，通過特徵的線性組合來預測句子的 VAI 值。

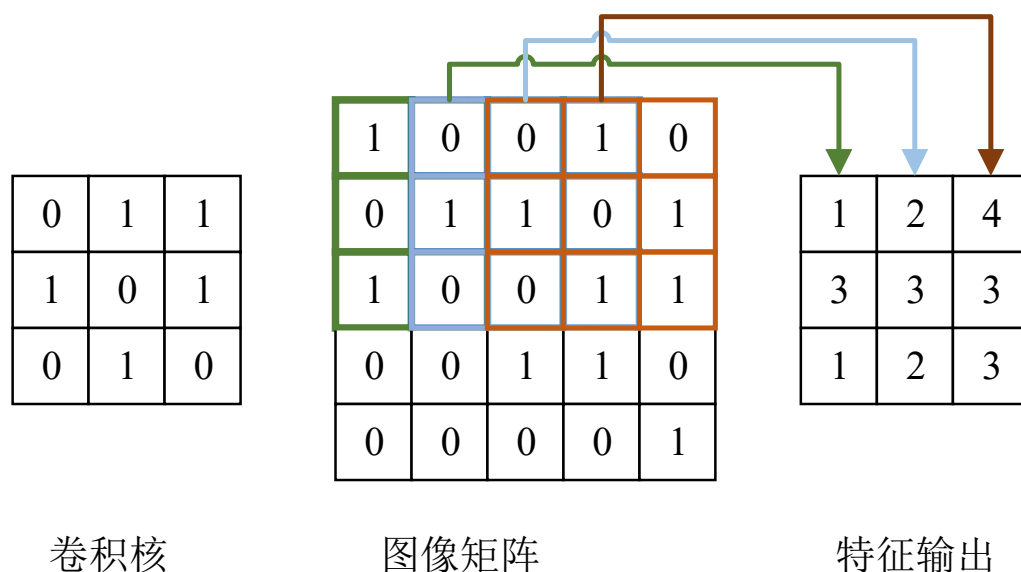


圖 5、CNN 模型中卷積層窄卷積操作示例圖 [32]

關於卷積操作，在影像處理中，卷積核在圖像矩陣中的卷積過程是二維的，如圖 5 所示

是一個  $3 \times 3$  大小的卷積核對  $5 \times 5$  大小的圖像矩陣進行卷積操作，卷積核(filter or kernel)的滑動步長為 1，也就是每卷積一個點之後滑動一格的距離繼續進行卷積操作。行方向卷積完成之後，卷積核就往下移動一格，重複第二行的卷積。每次的卷積操作都是卷積核中的數值與對應的矩陣點進行點積加和運算。這個過程演示的是對圖像矩陣的窄卷積過程，最後的輸出是一個  $3 \times 3$  大小的特徵輸出。

上述過程是二維卷積，而我們所研究的句子卷積過程是一個一維卷積，也就是卷積核只需要一個方向的移動。如圖 6 所示，卷積核尺 (filter\_length=3)，卷積核的維度與句矩陣中的詞向量維度相同(dims=5)，在實際實驗中，我們將詞向量維度設置為 300。句矩陣中每一行表示一個詞，行數表示詞的個數。實際的詞向量是 float 類型的資料，在該圖中我們用數字 0,1 來做簡化說明。與上圖不同，這



是一個寬卷積過程，在句矩陣的上下邊緣做補 0 操作，實現寬卷積。這樣縱向做了 7 次卷積操作後得到了卷積層的輸出資料。

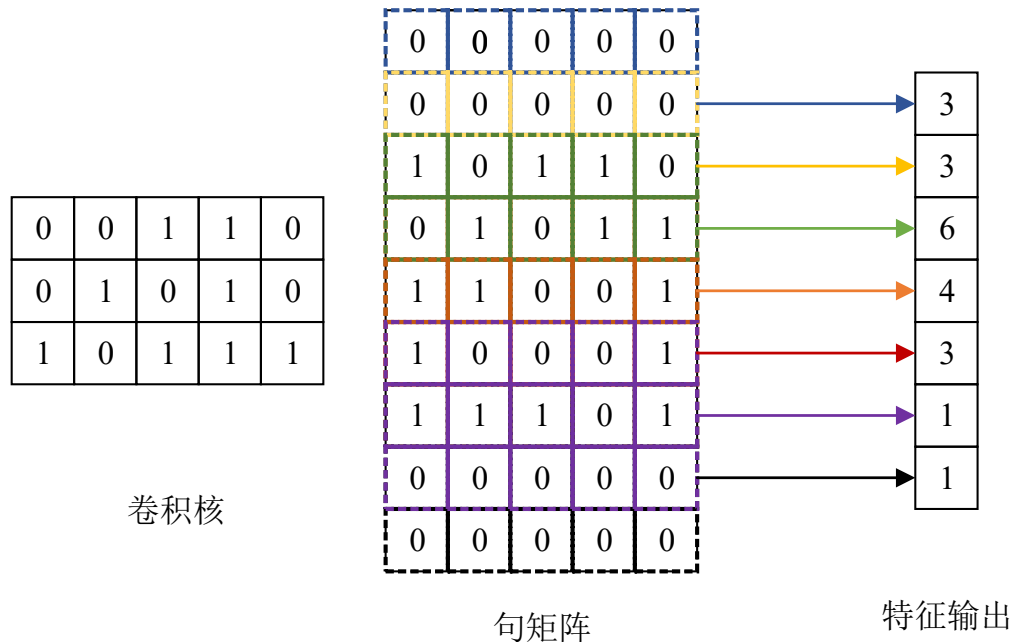


圖 6、CNN 模型中卷積層寬卷積操作示例圖 [32]

#### 4.3 長短期記憶人工神經網路（LSTM）預測 VAI

長短期記憶人工神經網路是一種特殊的時間遞迴神經網路(RNN),LSTM 與一般 RNN 的外部系統結構相同，如下圖 7 所示。RNN 適合解決時間序列輸入輸出的問題，對於一個句子來說，其中的詞彙也是序列關係。比如說“很好，一顆饅頭從早吃到晚”，我們的目的是預測這個句子的 VAI 值，需要考慮到這個句子的先前的相關資訊“很好”，才能對這個句子的三個情感維度做更準確地預測。標準 RNN 的關鍵特點就是能夠連接先前的資訊到當前的任務，然而標準 RNN 的缺陷是，句子比較長的情況下，相關資訊和當前預測位置的距離比較大，先前的相關資訊會丟失。LSTM 解決了這種問題，可以做到長期依賴資訊的學習。LSTM 與 RNN 的結構區別是，在重複的神經網路模組中，標準 RNN 結構相對簡單，只有一個 tanh 層。而 LSTM 中有多層結構以特殊的方式進行交互，我們將這個結構整體稱為記憶單元，通過特殊設計的“門”的結構能夠增加或者刪除

資訊來改變記憶單元的狀態。

門是一種資訊選擇式通過的方式，LSTM 中共有三個門，輸入門，遺忘門和輸出門。

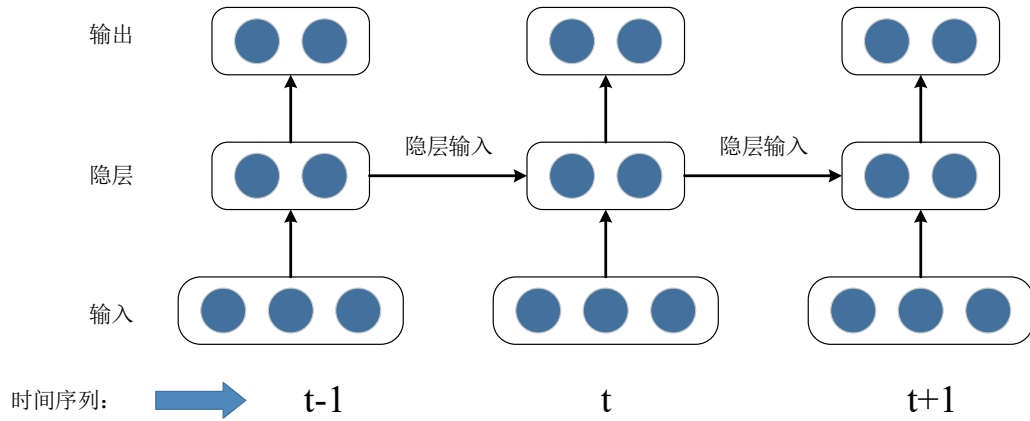


圖 7、標準 RNN 架構圖[29]

資料在記憶單元中的流動模型如圖 7 所示，其中  $W_{in}$ ， $W_{\phi}$ ， $W_{out}$  分別是輸入層，隱藏層和輸出層自有的權重集，這些權重的作用是用來對資訊進行篩選，然後決定資訊的保留和刪除。記憶單元在反覆運算過程中，使用誤差反響傳播和梯度下降調整權重，從而學習對資料的處理。

這個記憶體的輸入資料是當前的輸入資訊  $y^{in}$  和過去的記憶體狀態，經過三個門的交互處理後得到當前的記憶狀態  $s_c$ 。在資料處理過程中，三個門的組合狀態決定對資料做何種處理在每個時間步，每個門都可以打開或者關閉，每個時間步三個門的組合狀態都會更新。

下面分別來對這三個門的功能進行描述。

輸入門：圖 8，輸入門可以決定是否允許資料登錄，要確定那些資訊要被保留在細胞的狀態中，這個操作需要兩步來完成，首先經過一個 sigmoid 層決定更新的資訊，然後經過一個 tanh 層來創建新的候選向量。

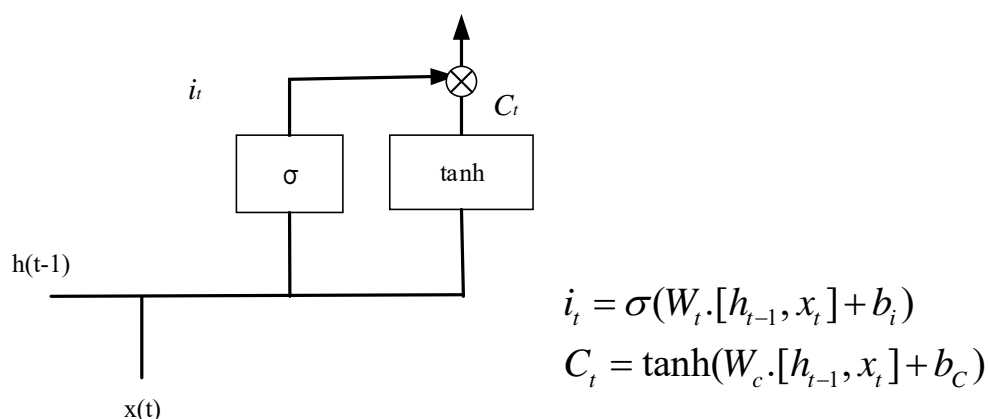


圖 8、LSTM 輸入門結構[29]

遺忘門：LSTM 中的第一步處理是，通過遺忘門來決定要丟棄的資訊。如圖 9 所示，該門的輸入是上個時間序列(t-1)的輸出  $h(t-1)$  和時間序列(t)的輸入  $x(t)$ ，經過一個 sigmoid 層的處理輸出值範圍是[0,1]，值的大小表示能夠通過的信息量。其中 0 表示全部丟棄，1 表示全部保留。

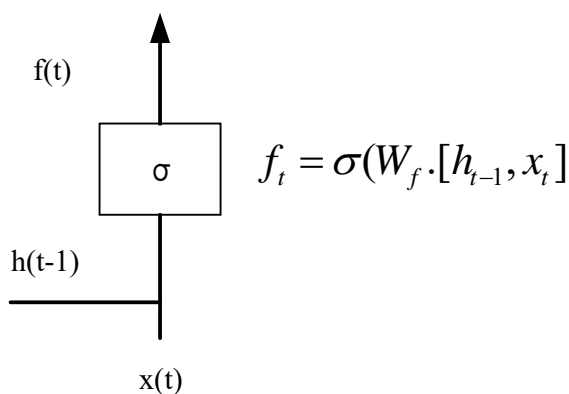


圖 9、LSTM 遺忘門結構[29]

輸出門：圖 10，輸出門的作用是決定資料的哪些部分允許被輸出，輸出門是一個 sigmoid 層和一個 tanh 層的交互處理過程。Tanh 處理之後會得到一個[-1,1]的範圍，與 sigmoid 處理之後的結果交互之後得出最終的輸出。

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

$\sigma$                        $\tanh$

$h(t-1)$

$x(t)$

圖 10、LSTM 輸出門結構[29]

#### 4.4 卷積-長短期記憶人工神經網路（CNN\_LSTM）預測 VAI

這種方法先使用卷積神經網路的卷基層，然後使用長短期記憶人工神經網路，集成了兩種神經網路的優勢。CNN\_LSTM 架構圖如圖 11 所示 CNN 的卷積操作中，卷積核對臨近詞的詞向量進行卷積，達到了降維並捕獲臨近詞之間關係特徵的目的。Pooling 層對資料特徵繼續做降採樣處理。LSTM 層有能夠存儲記憶上文資訊的結構，其中的記憶體結構能夠選擇性的保留或者遺忘資訊，尤其是改進了一般 RNN 不能保留遠距離資訊的缺點，LSTM 能夠將前文中的詞彙資訊傳遞到後文的處理中，這樣就達到了遠距離詞彙之間關係特徵的獲取。比如說句子“很好，一個饅頭從早吃到晚”，前文“很好”與後文“從早吃到晚”的特徵結合是預測 I 值的關鍵所在。最後通過 linear 啟動函數得到預測的 VAI 值。

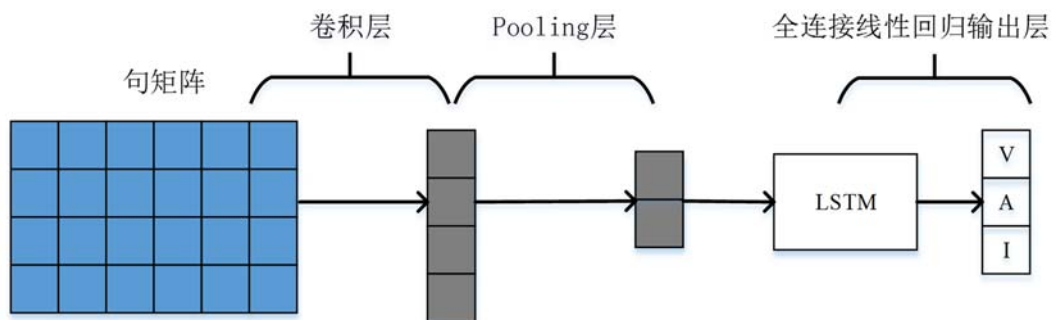


圖 11、卷積神經網路&長短期記憶人工神經網路 CNN\_LSTM 架構圖

#### 4.5 長短期記憶人工神經網路卷積神經網路 (LSTM\_CNN) 預測

##### VAI

LSTM\_CNN 與 CNN\_LSTM 順序相反，先使用長短期記憶人工神經網路再使用卷積神經網路，經過 LSTM 處理之後得到的輸出特徵將作為 CNN 的輸入資料。LSTM\_CNN 系統架構圖如圖 12 所示：

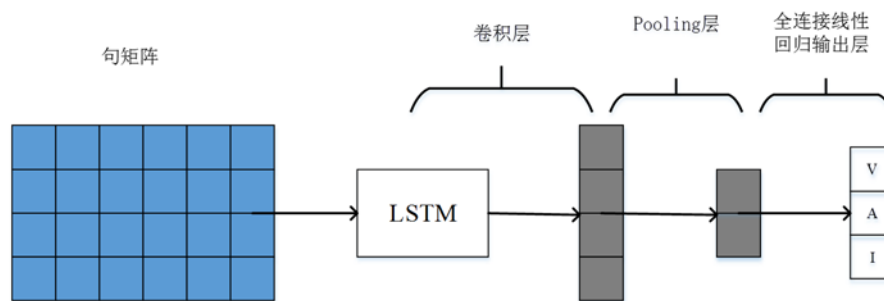


圖 12、長短期記憶人工神經網路&卷積神經網路 LSTM\_CNN 架構圖

## 第五章、實驗過程和結果

### 5.1 實驗說明

(1) 實驗方法：分別使用四種深度學習方法對語料庫進行三個維度的訓練和預測。三個維度獨立預測。採用 5-fold 交叉驗證方法，評估資料取五次結果的平均值。每次訓練隨機取訓練樣本 803 條文本，剩餘的 201 條文本作為測試資料。

(2) 評估指標：平均絕對誤差(MAE)

平均絕對誤差(MAE):

$$MAE = \frac{1}{n} \sum_{i=1}^n |t_i - p_i|$$

其中， $t_i$  表示句子  $i$  三個維度 V, A 或者 I 的標記值，VAI 三個維度獨立評估。相對應， $p_i$  表示預測的 VAI 維度值。 $n$  表示語料庫中被預測的句子個數。MAE 值越小越好，說明預測值與真實標記值更接近。

### 5.2 參數調整結果

基於龐大的資料基礎，深度學習網路的學習能力能夠得到很好的展現。在維度型情感分析中，針對一個特定的深度學習方法，通過調整參數後，預測出的維度值就能更加接近真實值。但是參數數目較多，參數組合也非常多，因此調參工作比較耗費時間。調參規則是，調整第一個參數時，其他參數保持不變，調整後面的參數時，以被調參數的最佳值作為基礎，最後得到所有參數的最佳值。對於所有的深度學習方法，採用以下兩個超參數取值分別為 `batch_size=8`，`nb_epoch=10`，因為此時的實驗結果較好。

**卷積神經網路(CNN)的最佳參數組合：**如表 4 示，V 值的卷積核個數為 10 (nb\_filter=10)；池化操作中下採樣因數 (pool\_length=1)；卷積核尺寸 (kernel\_size=10)，在模型中使用了三次 dropout 參數用來防止模型過擬合，應用 dropout 的層會斷開相應量的連接預防過擬合。在 CNN 模型中我們使用了三次 dropout，調整參數之後發現在嵌入層 (Embedding) 層 dropout 為 0 時，預測結果會比較好。卷積層和全連接層的 dropout 值分別取 0.4 和 0.2。

**長短期記憶神經網路(LSTM)的最佳參數組合：**如表 5 示，V 值的 lstm 層的輸出維度(output\_dim=50)；lstm 層的 dropout 取值為 0.5。

表 4、模型 LSTM 預測 VAI 維度值的最佳參數表

Model	参数取值		
LSTM (参数)	V	A	I
output_dim	50	60	80
droupout	0.5	0.5	0.5

表 5、模型 CNN 預測 VAI 維度值的最佳參數表

Model	參數取值		
CNN ( 參數 )	V	A	I
nb_filter	10	20	30
pool_length	1	1	1
kernel_size	10	5	10
droupout1	0.4	0.4	0.4
Droupout2	0.2	0.2	0.2

**CNN\_LSTM 的最佳參數組合：**如表 6 示，此模型對於 VAI 三個維度的預測中，卷積核的個數分別為(20,20,10)。

表 6、模型 CNN\_LSTM 預測 VAI 維度值的最佳參數表

Model	参数取值		
CNN_LSTM (参数)	V	A	I
nb_filter	20	20	10
filter_length	3	3	50
pool_length	6	6	6
dropout	0.5	0.5	0.5

**LSTM\_CNN 的最佳參數組合：**如表 7 示：

從上述四個參數表中可以看出，對於 VAI 的預測，同一預測模型，同一參

表 7、模型 LSTM\_CNN 預測 VAI 維度值的最佳參數表

Model	参数取值		
LSTM_CNN (参数)	V	A	I
nb_filter	10	10	10
filter_length	2	2	2
pool_length	1	1	1

數，最佳參數值比較接近，變化範圍不大，表 5 中，CNN 對 VAI 進行預測時，卷積核個數 nb\_filter 分別取值為(10,20,30),池化操作 pool\_length 取值相同，都為 1；6，模型 CNN\_LSTM 中對 VAI 進行預測，卷積核個數分別取值為(20,20,10)。參數的調整主要是與模型的結構關係比較緊密。



### 5.3 實驗結果和分析

表 8 實驗結果表, Valence, Arousal, Irony 分別是哪三個維度。從表中, 可以看出對於 VAI 的預測, LSTM\_CNN 的預測效果最好, MAE 最小值分別是 (0.409,0.769,0.583),都分別小於 CNN, LSTM, LSTM\_CNN 模型得到的預測結果。

V 值的預測中, MAE 範圍是(0.4~0.454),

A 值的預測中, MAE 範圍是(0.762~0.812),

I 值的預測中, MAE 範圍是(0.583~0.645)。

表 8、四種模型對中文反諷 VAI 語料庫三個維度值的預測結果表

	<b>Valence</b>	<b>Arousal</b>	<b>Irony</b>
Method	<i>MAE</i>	<i>MAE</i>	<i>MAE</i>
<b>CNN</b>	0.454	0.805	0.639
<b>LSTM</b>	0.412	0.784	0.593
<b>CNN_LSTM</b>	0.428	0.812	0.645
<b>LSTM_CNN</b>	0.4	0.762	0.583



## 第六章、總結

本文描述了我們工作的主要內容，維度型反諷語料庫的創建和反諷文本 VAI 值的預測。目前，自然語言領域的研究中，對於反諷表達的研究集中於類別型的研究，還沒有研究是針對於維度型的反諷表達，特別是中文反諷維度型研究。同時，因為反諷表達不是常用表達方式，所以反諷語料的獲取比較困難。因此，我們開創性的創建了維度型的中文反諷語料庫，該語料庫具有三個維度的標記值，分別是 V(Valence), A(Arousal), I(Irony)。該語料庫為以後關於反諷維度型的情感分析研究提供了資料基礎。

語料庫構建完成之後，針對於該語料庫，我們使用機器學習方法對語料庫中的樣本進行維度值的預測。使用的機器學習方法有 CNN、LSTM、CNN\_LSTM、LSTM\_CNN，我們分別使用這四種方法對 VAI 值進行預測。實驗結果表明，LSTM\_CNN 預測模型對於我們的語料庫的維度值 VAI 的預測效果最好。V 值的 MAE 最小，A 值的 MAE 最大。

在以後的研究中，為了探索能夠更準確的預測維度值的方法，我們計畫將維度之間的關聯關係加入預測方法。在預測某一個維度時，考慮其他兩個維度的詞向量權重，將三個維度的向量特徵都整合到預測模型中，希望能通過這種方式得到更加準確的預測模型。

## 參考文獻

- [1] Fontaine, J. R., Scherer, K. R., Roesch, E. B., & Ellsworth, P. C. (2007). The world of emotions is not two-dimensional. *Psychological science*, 18(12), 1050-1057.
- [2] Barbieri, F., & Saggion, H. (2014, April). Modelling Irony in Twitter. In *EACL*(pp. 56-64).
- [3] Karoui, J., Benamara, F., Moriceau, V., Aussenac-Gilles, N., & Hadrich Belguith, L. (2015). Towards a contextual pragmatic model to detect irony in tweets. *Association for Computational Linguistics (ACL)*.
- [4] Tang, Y. J., & Chen, H. H. (2014). Chinese Irony Corpus Construction and Ironic Structure Analysis. In *COLING* (pp. 1269-1278).
- [5] Peter J. Lang. 1980. Behavioral treatment and bio-behavioral assessment: Computer applications. *Technology in Mental Health Care Delivery Systems*, pp. 119-137, Ablex Publishing, Norwood.
- [6] Filatova, E. (2012, May). Irony and Sarcasm: Corpus Generation and Analysis Using Crowdsourcing. In *LREC* (pp. 392-398).
- [7] Mikolov, T., Sutskever, I., Chen, K., Corrado, G. S., & Dean, J. (2013). Distributed representations of words and phrases and their compositionality. In *Advances in neural information processing systems* (pp. 3111-3119).
- [8] Buschmeier, K., Cimiano, P., & Klinger, R. (2014, June). An impact analysis of features in a classification approach to irony detection in product reviews. In *Proceedings of the 5th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis* (pp. 42-49).
- [9] Reyes, A., & Rosso, P. (2011, June). Mining subjective knowledge from customer reviews: A specific case of irony detection. In *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis* (pp. 118-124). *Association for Computational Linguistics*.
- [10] Wallace, B. C. (2015). Sparse, contextually informed models for irony detection: Exploiting user communities, entities and sentiment. *ACL*.

- [11] Barbieri, F., Ronzano, F., & Saggion, H. How Topic Biases Your Results? A Case Study of Sentiment Analysis and Irony Detection in Italian. *RECENT ADVANCES IN*, 41.
- [12] Wallace, B. C., Do Kook Choe, L. K., Kertz, L., & Charniak, E. (2014). Humans Require Context to Infer Ironic Intent (so Computers Probably do, too). In *ACL* (2) (pp. 512-516).
- [13] Joshi, A., Sharma, V., & Bhattacharyya, P. (2015). Harnessing context incongruity for sarcasm detection. In *Proceedings of the 53rd Annual Meeting of the Association for Computational Linguistics and the 7th International Joint Conference on Natural Language Processing* (Vol. 2, pp. 757-762).
- [14] Ptáček, T., Habernal, I., & Hong, J. (2014). Sarcasm Detection on Czech and English Twitter. In *COLING* (pp. 213-223).
- [15] Barbieri, F., Saggion, H., & Ronzano, F. (2014). Modelling sarcasm in twitter, a novel approach. *ACL* 2014, 50.
- [16] González-Ibáñez, R., Muresan, S., & Wacholder, N. (2011, June). Identifying sarcasm in Twitter: a closer look. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies: short papers-Volume 2* (pp. 581-586). *Association for Computational Linguistics*.
- [17] Liebrecht, C. C., Kunneman, F. A., & van den Bosch, A. P. J. (2013). The perfect solution for detecting sarcasm in tweets# not.
- [18] Riloff, E., Qadir, A., Surve, P., De Silva, L., Gilbert, N., & Huang, R. (2013, October). Sarcasm as Contrast between a Positive Sentiment and Negative Situation. In *EMNLP* (Vol. 13, pp. 704-714).
- [19] Riloff and J. Wiebe. 2003. Learning extraction patterns for subjective expressions. In *Proc. of the 2003 conference on Empirical methods in Natural Language Processing-V*. 10, p. 105–112. *ACL*.
- [20] Thelen and E. Riloff. 2002. A bootstrapping method for learning semantic lexicons using extraction pattern contexts. In *Proc. of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, p. 214–221. *ACL*.
- [21] Davidov, D., Tsur, O., & Rappoport, A. (2010, July). Semi-supervised recognition of sarcastic sentences in twitter and amazon. In *Proceedings of the fourteenth*

*conference on computational natural language learning* (pp. 107-116).  
*Association for Computational Linguistics.*

- [22] Khattri, A., Joshi, A., Bhattacharyya, P., & Carman, M. J. (2015, September). Your Sentiment Precedes You: Using an author's historical tweets to predict sarcasm. In *6TH WORKSHOP ON COMPUTATIONAL APPROACHES TO SUBJECTIVITY, SENTIMENT AND SOCIAL MEDIA ANALYSIS* WASSA 2015 (p. 25).
- [23] Lukin, S., & Walker, M. (2013, June). Really? well. apparently bootstrapping improves the performance of sarcasm and nastiness classifiers for online dialogue. In *Proceedings of the Workshop on Language Analysis in Social Media* (pp. 30-40).
- [24] Tungthamthiti, P., Kiyooki, S., & Mohd, M. (2014). Recognition of sarcasms in tweets based on concept level sentiment analysis and supervised learning approaches. In *Proceedings of Pacific Asia Conference on Language, Information and Computing, Phuket, Thailand.*
- [25] Ghosh, D., Guo, W., & Muresan, S. Sarcastic or Not: Word Embeddings to Predict the Literal or Sarcastic Meaning of Words.
- [26] Amir, S., Wallace, B. C., Lyu, H., & Silva, P. C. M. J. (2016). Modelling context with user embeddings for sarcasm detection in social media. *arXiv preprint arXiv:1607.00976*.
- [27] Ghosh, A., & Veale, T. (2016). Fracking Sarcasm using Neural Network. In *ACL*
- [28] Van Hee, C., Lefever, E., & Hoste, V. (2016). Monday mornings are my fave:)# not Exploring the Automatic Recognition of Irony in English tweets. In *Proceedings of the 26th International Conference on Computational Linguistics* (pp. 2730-2739).
- [29] Gers, F. (2001). Long short-term memory in recurrent neural networks (Doctoral dissertation, Universität Hannover).
- [30] Joshi, A., Tripathi, V., Bhattacharyya, P., & Carman, M. (2016). Harnessing sequence labeling for sarcasm detection in dialogue from tv series 'friends'. *CoNLL* 2016, 146.

- [31] Barbieri, F., & Saggion, H. (2014). Modelling Irony in Twitter: Feature Analysis and Evaluation. In *LREC* (pp. 4258-4264).
- [32] Masci, J., Giusti, A., Ciresan, D., Fricout, G., & Schmidhuber, J. (2013, September). A fast learning algorithm for image segmentation with max-pooling convolutional networks. In 2013 IEEE International Conference on Image Processing (pp. 2713-2717). *IEEE*.
- [33] Davidov, D., Tsur, O., & Rappoport, A. (2010, August). Enhanced sentiment learning using twitter hashtags and smileys. In *Proceedings of the 23rd international conference on computational linguistics: posters* (pp. 241-249). *Association for Computational Linguistics*.

