

# 基於網路社群之旅遊經驗及對應情境之情感意見分析研究

## Sentiment Analysis in Scenic Spot

### Experiences from Web Community Using Semantically

### Relevant Contextual Information

甯格致 Ke-chih Ning

國立政治大學資訊科學系

Department of Computer Science

National Chengchi University

[98753502@nccu.edu.tw](mailto:98753502@nccu.edu.tw)

賴昆棋 Kun-chi Lai

國立政治大學資訊科學系

Department of Computer Science

National Chengchi University

[98753504@nccu.edu.tw](mailto:98753504@nccu.edu.tw)

### 摘要

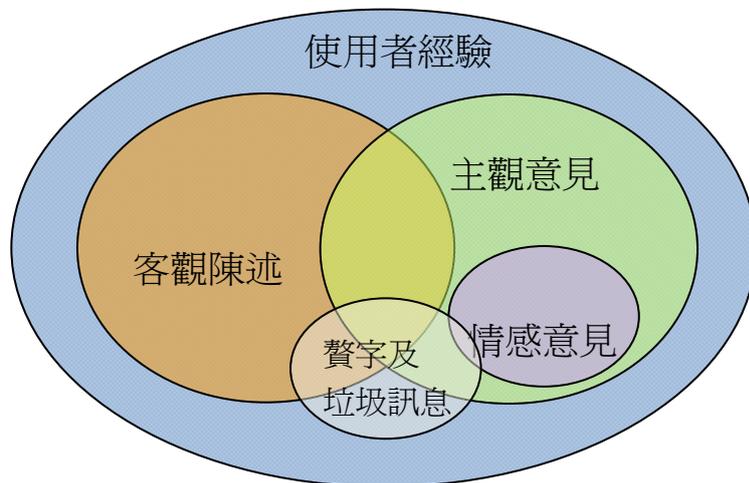
網際網路擴張迅速，Web 2.0 服務擴張造就了一具有大量使用者意見的資訊載體。而使用者意見中包含了客觀的描述(fact)及主觀的意見(opinion)，主觀的意見中又包含了情緒、情感(sentiment)的資訊。這些資訊不論對提供商品的業者、使用服務的客戶或是人際關係都具有很高的應用價值。以旅遊經驗為例，使用者的經驗回饋產生了大量龐大的意見資訊，然而目前的資訊技術尚很難藉由自動化的方式大量從中挖掘出有用的參考資訊。本研究蒐集了大量來自網路社群之旅遊景點意見(包含有八大旅遊景點分類，包含了 1,275 個店家、景點等營業場所，共達 9,771 筆使用者回饋的訊息)，並提出詞組規則探索之方法，試圖建立起一個可自動化大量進行使用者意見分析的流程，並基於當前其他學者提出的方法及研究貢獻加以改良，尤其是加入了情境相關詞彙的偵測及量化，對於情感分析能產生更好的效果。本研究提出的分析架構以上述使用者的資料作為訓練及測試用，並從中標記並建立出包含有 4,865 筆旅遊意見相關的詞彙庫，再合併以 SVM 為分類器，於單獨判斷極性的部分得到最佳 F<sub>1</sub> Score 99.8%；於綜合判斷極性及情感強度則得到最佳 F<sub>1</sub> Score 89.1%，皆呈現出此混和式的分類架構有相當理想的成效，可見適當之詞彙庫使用策略配合調校合適之分類器應能對特定領域之網路情感意見分析有相當大助益。

關鍵詞：情感分析、網路資料探勘、自然語言處理

## 一、緒論

近年網路的活躍造就了新型態資料整合與分享模式，其中最有效者為 Web2.0，其最大的特色之一為網站內容是由使用者所提供，而非管理者的蒐集，相較於傳統（早期）網站，此類網站提供了更多貼近生活化的議題，使用者很容易透過個人式部落格或討論區性質網站中，來分享自己的生活經驗、對事物的看法等。Web2.0 基本的概念即為提供多人共同創作平台環境，讓使用者可以在上面瀏覽資料、編輯內容、創建內容及分享內容，最著名成功案例如：Wikipedia、Facebook、Twitter、Plurk、Flickr...等。然而由於越來越大量的資訊藉由各種 Web 2.0 的服務匯流進網際網路中，此類訊息多半是屬於非結構化的資訊(unstructured)被儲存著，因此如何能夠有效率地從中提取有用的訊息就成了相當重要的議題。

自然語言分析是目前相當熱門之研究主題，其中相關的許多技術可以幫助我們以機器自動化的方式來大量解析網路中非結構化存取的使用者意見，取得重要訊息後可再轉由結構化、非結構化方式作為其他如市場行銷、商業調查、人際互動、社會網路分析等應用。這些意見通常包含了主觀的陳述、客觀的意見或是帶有情緒、情感的訊息，並且存在有一或多個意見的主題(target)、附屬之子題(feature)。意見分析者通常要能夠盡可能取得這些資訊才可能產生有意義的分析結果。某些情況下，意見提供者(opinion holder)相關的背景資料也是一重要的評估用訊息 [2][3]。



圖一、使用者經驗之意見包含範圍

然而此類意見分析的工作並不容易達成(圖一)，因為 1. 語言的彈性非常大，較難以精簡的規則去表達或是了解分析；2. 再者語言所存在的時空背景、情境狀況等也會影響相同語言下所表現出來的意義；3. 用以表達客觀陳述或是主觀意見的語言並沒有明確的界線，常常情境變化了，意義就會有很大的不同；4. 網

路中尚存在有許多的垃圾訊息、廣告訊息、甚至連言談之間也有可能出現冗贅的語句。以上諸點都有可能影響意見分析的成效，因此採用多種方法整合起來以利有效的分析工作進行，就目前而言應是較為可行的方向 [5][6]。

本研究以旅遊景點為意見範圍蒐集了相當大量的網路回饋意見，接著利用詞頻、詞組頻率從中過濾出含有情感、主觀意見的詞組規則，並藉著此已發掘之規則以人工標記的方式建立出一具大量詞彙之詞彙庫，以輔助單純使用分類器上的不足。本研究更延伸探索詞組規則至語句中可能之意見主題(情境)，試圖建立出意見方向與情境之間的關聯性。不同情境條件下，相同的描述用語有可能表達了不同極性的意見，例如：房間尺寸是”越大”代表”越好”；手機尺寸卻是”越小”代表”越好”。不同的情境有可能會讓後述之意見產生”逆轉”的作用，同樣道理句子中的形容詞(ADV)也有類似的”增強”、”減弱”及”逆轉”作用存在。適當地發掘這些規則並加以利用即有助於辨識時正確率的提升 [7]。

本研究分別以意見極性(正、負向)、意見極性及情感強度(意見有情感、情緒或是不具情感、情緒)，以及情境資訊加入與否以及是否使用詞彙庫來建立主要的實驗項目。訓練及測試用之語料樣本包含 1,275 個不同的景點、店家或遊憩相關營業場所，使用者意見數量共 9,771 筆 (提取主要目標句型後消去相同者後訓練樣本總數 4,865 筆)，在使用  $F_1$  Score 為評估指標下得到最佳結果：99.8% 為單獨判斷意見極性時，並同時考慮情境因素，以及使用詞彙庫輔助分類 (不使用詞彙庫時最佳結果為 75.42%)；而納入情感意見的綜合辨識部分最佳結果為：89.1% 並同時考慮情境因素及使用詞彙庫輔助 (不使用詞彙庫時僅有 37.68%)。以情境輔助的成效上同樣以  $F_1$  Score 為評估指標，若在各組使用詞彙庫的情況下則績效不顯著，不使用詞彙庫條件下則極性判斷項目有較好的成效 (未考慮情境因素：63.65%；有考慮情境因素：75.42%)。

本研究採用分類器加上情境資訊及自建詞彙庫之輔助分類方法，經由多組實驗證實可得到較佳之分類績效。雖其中加入情境資訊後改善不甚顯著的情況則有可能是因本實驗樣本限定於幾乎相同意見領域(domain)的緣故，意即以旅遊經驗此單一意見領域來看，意見描述用語的語意極性方向大致上是相似的，較無涉及不同領域容易發生之意見極性逆轉現象，但也因為這種單一意見領域的專屬性而造就了詞彙庫相當顯著的績效增長。

## 二、 文獻探討

### 2.1 CKIP 中文斷詞系統

CKIP 中文斷詞系統的技術是由中研院資訊科學所詞庫小組所開發，將自然語言中最小的語言單位 ”字詞” 抽離，廣泛的被運用於機器翻譯、語言分析、資料挖掘等領域中，其關鍵的技術在於利用其所收錄語料（語詞）與分析的本文進行比對，自動的找出文句中的斷句及詞類標記。CKIP 中文斷詞系統除了提供線上展示與短句的示範性操作外，也提供了一開放 API 供使用者大量批次轉換，透過 XML 格式交換簡化線上操作。

中文斷詞的準確性依據所中文詞庫蒐集各領域的特殊詞彙與專有名詞，因此廣泛的蒐錄各領域詞典，為目前最有效之方式，對我們的專案而言，因旅遊議題所延伸評論，出現專業領域的用詞其相對機會相當低，斷詞的精確度可達預估的效果。但近年許多的網路用語，則幾乎未收錄於中文詞庫中，此屬本研究及其他網路意見分析研究者極可能會遇到的問題，以本研究而言，此部分的資料雖不多，但是因 CKIP 段詞系統無法辨識這類詞語，因此有仍很可能會影響到情感分析的成效。

常見的網路用語包括：

1. 同音錯字：北七(白痴)。
2. 注音文：如我ㄊ歡去天母。
3. 縮寫：如 ”醬啊” 表示 ”這樣子啊” 的感嘆意思。
4. 表情符號：如 Orz 表示 “無可奈何” 或沮喪失敗的意思。
5. 特殊符號：(\*\_\*) 表不妙，^^ 表開心，亦有使用表情圖片者。
6. 數字符號：如 0837 表示你別生氣。

諸如此類的問題，傳統的解決方法是先將這些網路用語儘可能收集，再利用統計方式篩選出現頻率較高出現者，利用一個對照表還原這些網路用語的意思。就本研究而言，次部份的問題解決方式大致上也是沿用上述做法的概念，並人工建立出所需的情感詞彙庫以利後續的分類器使用。

### 2.2 特徵擷取

經由斷詞後結果進行特徵擷取，希藉由各方式取出有意義或代表性的詞彙，常見的三種方法詞庫比對法、統計法及文法剖析。

表一、特徵擷取方式比較表

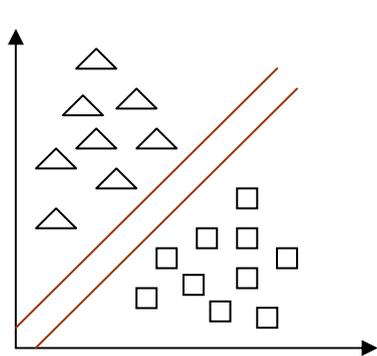
特徵擷取方法	基本作法概念	優點	缺點
詞彙庫比對法	事先建立詞庫，對文件中詞彙進行比對	實做容易、精確度高 (precision)	詞彙庫建置不易、比對速度慢、無法辨識未建置詞庫
統計法	對文件中的詞頻進行分析，將詞頻出	不需要大量詞庫 新生的詞可以被	出現頻率少的詞容易被忽略

	現在某一個區間次數，取出當關鍵字	找到	
文法剖析法	透過自然語言的文法剖析，再運用特定準則過濾不適合的詞彙	可有效避開無意義詞彙	需要詞庫為基礎 需完整的句子進行剖析(缺乏完整句型或是口語句型相對較無法辨識)

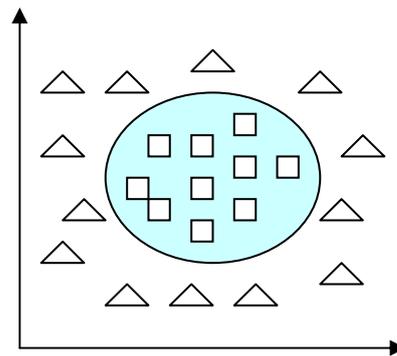
### 2.3 決策支援向量機 (Support Vector Machine, SVM)

1990 年代後因自然語言處理開始採用機器學習(Machine Learning)技術逐漸萌芽，一般而言，機器學習可以兩大面向，非監督式(Unsupervised)摘要模型與監督式(Supervised)摘要模型的使用（林士翔，2010），非監督式摘要模型基本上在建立摘要模型時並不需參照或使用任何人工事先標記過的語料庫，相對而言監督式模型則通常需要有人工事先標記過的語料再提供模型訓練使用，概念上以一組訓練文件集以訓練機器辨讀每篇訓練文件的每一語句是否屬於摘要語句(Summary Sentence)，利用類似的概念作為訓練之基礎與修正。著名的相關研究包括簡單貝氏分類器(Naïve-Bayes Classifier)、高斯混合模型(Gaussian Mixture Model, GMM)、隱藏式馬可夫模型(Hidden Markov Model, HMM)、支援向量機(Support Vector Machine, SVM)、條件隨機域(Conditional Random Fields, CRF)等。

近年因決策支援向量機容易使用、高正確率、相對快速的特性，因此被用於自然語言處理中或需要自動辨別分類研究上，SVM 主要的理念是將向量對應到高維度的空間上，並於空間上找到最大邊界將其區分成二元類別，核心技術上主要分線性 (linear) 與非線性 (non-linear) 兩種。



圖二、線性分割



圖三、非線性分割

使用支援向量機進行分類之研究相當廣泛，如利用嘉義地區福衛二號(Formosat-2)影像為基礎，利用支援向量機進行水稻辨識，經實驗結果 SVM 與高斯最大似然分類法相比，具有較高的穩定性（陳承昌，2006）；利用部落格中代表表情符號最為語料，利用 SVM 訓練機器進行情緒分類，藉以判斷文章中的情緒傾向（楊昌樺等，2007）[1]。

## 2.4 條件隨機域 (Conditional Random Field, CRF)

條件隨機域 (Conditional Random Fields, CRF) 常用於建構機率模型並對序列資料作分段，並計算出最大條件的機率值，其為無方向圖形模型 (Undirected Graphical Model) 的格式，給定一段觀測序列，無方向圖形模型可以定義一個單一對數線性分配。目前亦有許多工具 CRF，如 CRF++、LingPipe CRF 及 Flex CRFs 等。文獻上有以條件隨機域 (CRF) 為理論基礎之自動書目文獻擷取系統，將書目中姓名、期刊、年份等資訊擷取出，最後產生著作分類統計 (邱欣萍，2006)；在機器學習分類演算法中，分別利用 SVM 與 CRF 進行部落格情緒分類實驗，而其結果採用 CRF 之結果通常較 SVM 佳 [1]。

然而使用條件隨機域為分類器將會涉及得重新利用諸如 IOB1 (其他類似標準還有像如 IOB2、IOE1、IOE2 等) 的標記規範來標記資料，若面臨到的處理資料相當龐大，則此過程乃需要極大量之人工作業才可完成。

## 三、 研究方法

### 3.1 資料蒐集及標記

本研究設計了符合特定網站格式之網路機器人 (Web Crawler)，以自動化的方式大量擷取網路上旅遊相關之意見、心得、回覆、短文等資訊，並加以分類及歸納整理，資料的概況大致上如(表二)所示：

表二、景點資料包含類型

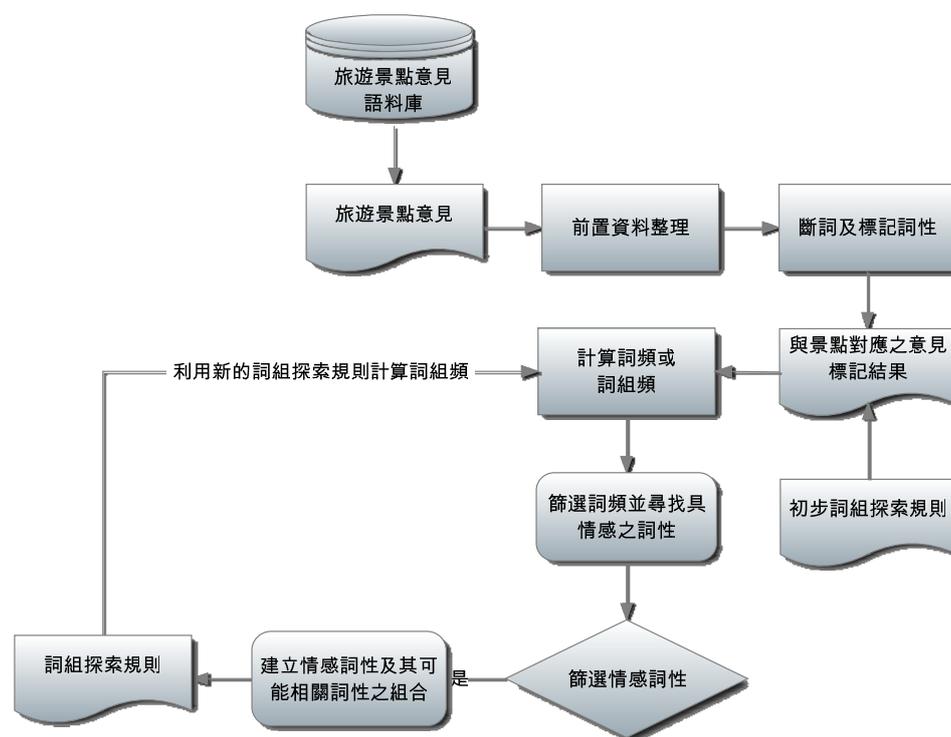
景點類型	商家數	意見數
公園	97	546
其他遊樂場所	380	1623
夜市	6	20
花園	30	117
紀念堂	13	67
動物園	8	71
渡假中心	192	2162
樂園、遊樂場	51	967
觀光農場、果園、茶園	498	4198
	1275	9771

其中商家數為 1,275 家，意見總數為 9,771 筆，然而每則意見又依不同的詞組探索規則(template, 見 3.2.2 節)，會將之拆解產生出更多的詞組(instance)，因此這些詞組的數量會遠高於 9,771 筆，實際初步測試我們以 16 種簡單的詞組探索規則進行探索，即從中產生了約 77,756 筆詞組，我們可再進一步依據詞彙相同的詞組予以收合(collapse) 即可歸納出此探索規則所產生之詞組特徵，以前例而言則約產生出 1,619 筆詞組特徵。然而詞組探索所採用的規則乃是最主要的影響原因。

## 3.2 尋找較具情感意見之詞性或詞性組合

### 3.2.1 斷詞處理

本研究將上節所示大量的意見資料送往斷詞處理單元 (即中研院 CKIP 斷詞處理系統) 進行詞性標記 (Part-of-speech, POS) 作業, 完成後即可獲得已標記詞性之原資料。然而極大量的意見及詞性資料仍不容易處理, 因此我們得建立一半自動化之方式來進入資料中探索, 即詞組探索規則。初步詞組探索規則可以先嘗試以簡單的詞性開始探索, 掌握了目標資訊可能的詞性所屬後, 可再建立精細一些的探索規則。此部分工作尚須人工之半自動化建立探索規則及產生出來之結果篩選。以 (圖四) 為例, 即為本研究使用旅遊景點意見為基礎, 以具情感詞性為目標之流程圖。



圖四、斷詞處理及篩選情感詞性、建立探索規則

### 3.2.2 詞組探索規則 (template) 之建立

詞組探索規則是本研究自行開發讓我們以指定的詞組規則探索大量來自 CKIP 斷詞系統已標記資料的方式。(表三)為簡化之 CKIP 斷詞系統詞性表, 初步探索規則可先由其中較簡單且常見的詞性開始, 例如 N (名詞)、ADV(副詞, 修飾動詞用)、ADJ(形容詞、修飾名詞用)等等。

以 N(名詞)而言，初步探索規則可建立如下：

1. 單一詞 (僅尋找語料中有出現的名詞)：

N

2. 雙詞 (容許名詞的前後各一任意詞的探索範圍)：

\* N ; N \*

3. 多詞 (容許名詞的前後各任意數量任意詞的探索範圍)：

\* \* N ; \* N \* ; N \* \*

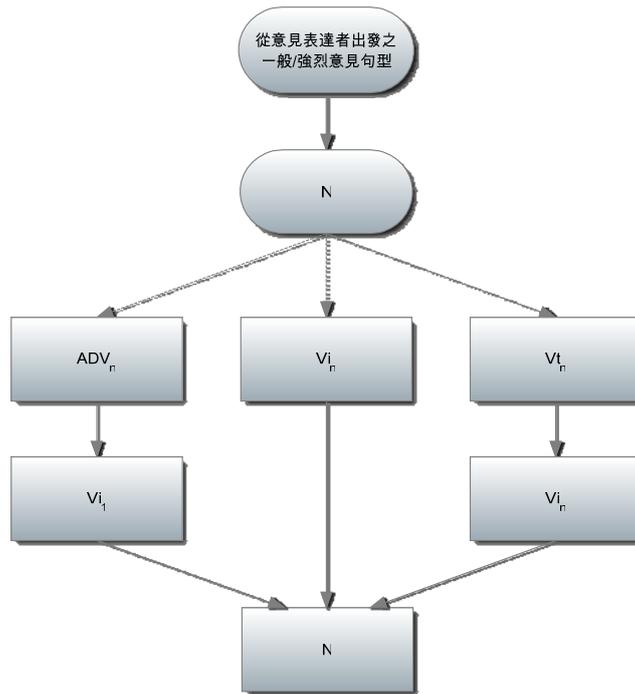
4. 特定多詞 (容許名詞及形容詞的前後各任意數量任意詞的探索範圍)：

\* N ADJ \* ; N ADV \*

\*(星號)即代表任意詞，意即該位置可容許任意詞的組合之探索，然而過多的任意詞會導致呈現出來的資料語意不通順，尚需要妥善的篩選及設計探索規則。以本研究來說最主要的目標為具有情感、情緒類的用語，因此經過初步詞組探索規則及結果的詞頻統計篩選後(過濾掉不正確、不常用的語句)，即可發現 Vi (狀態類使動動詞) 為具備此要求之詞性，因此我們可再進一步以 Vi 為中心建立探索規則，逐步擴展後可歸納所求之詞性組合如(圖五)所示，其中置詞組前後之 N(名詞) 常被視為意見或情緒句型描述的目標，一些歸納出來的詞組呈現概況如(表三)：

表三、情感、情緒詞性組合例子

語句	詞組
我有點失望、環境很舒服、床單很恐怖 十分沮喪挫折、很適合老人家、最愛玩金礦山 不爽、很有趣、非常噁心、超刺激	(N ADV Vi) (ADV Vi N) (ADV Vi)
印象深刻、態度爛、通體舒暢 好地方、做人阿莎力、適合情侶、爛老闆 舒服、熱情、有趣、刺激、倒楣、傻眼	(N Vi) (Vi N) (Vi)
態度惡劣、民宿佈置溫馨、老闆娘好客 感到窩心、覺得怕怕、獨享浪漫 缺乏歡樂氣氛、享受悠閒時光、想要清靜一下	(N Vt Vi) (Vt Vi) (Vt Vi N)



圖五、情感、情緒意見之詞性組合

### 3.3 可能詞性組合篩選及詞彙庫建立

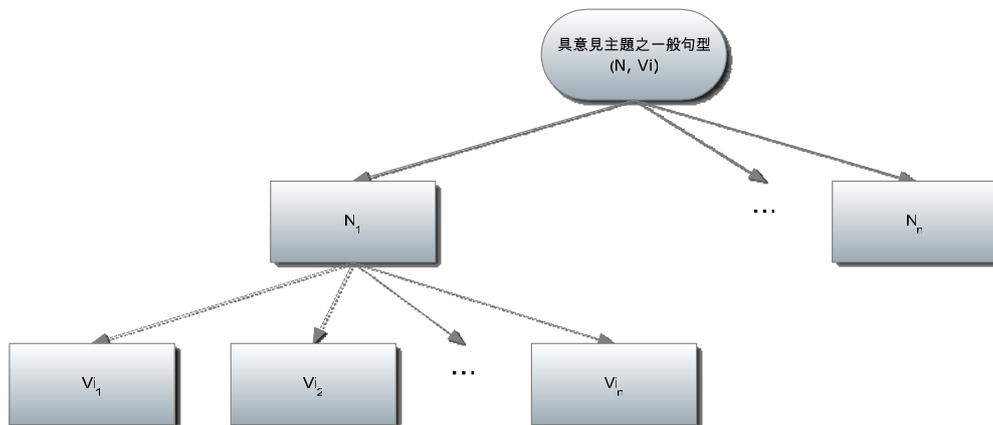
依據 3.1~3.2 節之半自動方法，我們不但可以依此建立起專屬意見領域之詞彙庫，甚至也可以應用此方法於其他領域，建立不同的需求目標出對應之詞彙庫。(圖六) 中“標記情感強度及意見極性”可能是一較不具穩定性的程序，因標記涉及個人主觀的語言認知及知識背景，且片面的詞性不見得就可反映出確切之真實語意。例如：

“很大”、“很小”(ADV Vi)：尚未呈現極性之表達用語句；

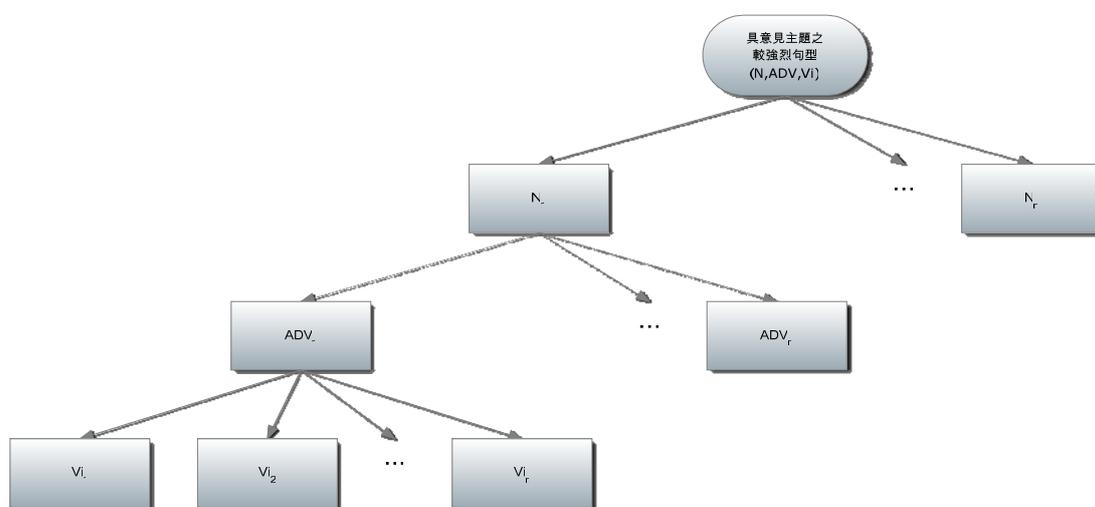
“壓力很大”、“選擇很少”(N ADV Vi)：呈現負面情緒、負面評價之語句；

“房間很大”、“煩惱很少”(N ADV Vi)：呈現正面評價之語句、正面情緒之語句；

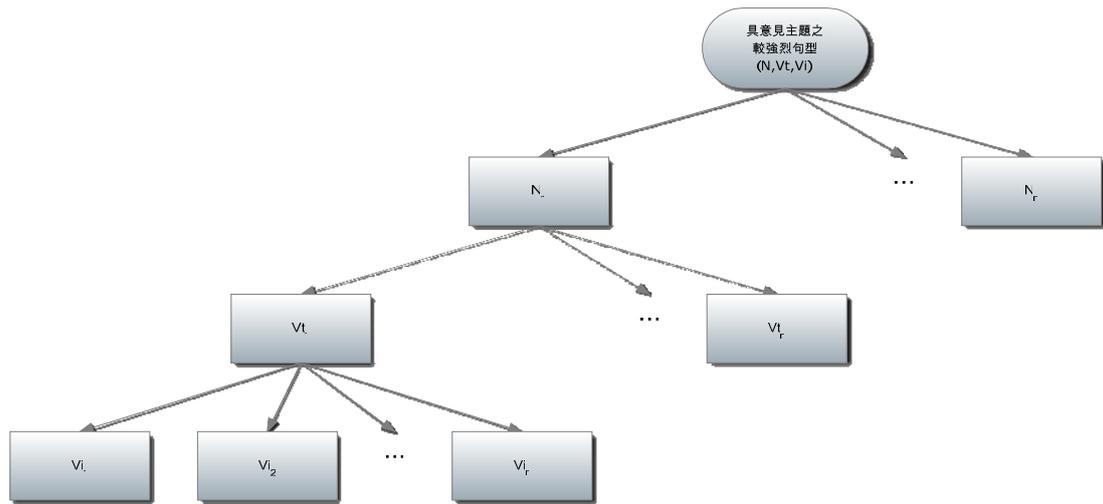
此例子即是說明當我們尚未確定片面的詞組前後是否還有要搜尋的目標詞性前，貿然認定該詞的極性或是是否有情緒可能會造成後續訓練測試時大量的錯誤。Vi 具有一預先給定的“極性”及“情緒強度”；ADV 對於 Vi 可具有“增強”、“減弱”、“逆轉”的作用；N、Nd、Ng (屬性不同之名詞) 對於 Vi 或 (ADV Vi) 可具有“維持”、“逆轉”的作用 [7]。也因此，充分地對於意見資料應是必要的過程，我們可歸納探索到的規則及對應之語句(instances)成爲一表狀或樹狀架構如(圖六)、(圖七)及(圖八)所示：



圖六、具意見主題之一般句型 (N Vi)



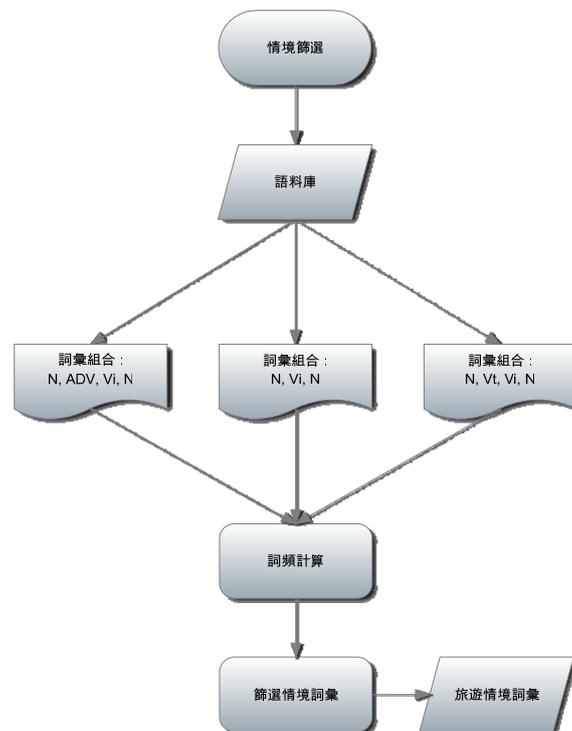
圖七、具意見主題之較強烈句型 I (N ADV Vi)



圖八、具意見主題之較強烈句型 II (N Vt Vi)

### 3.4 情境用詞與詞性組合

情境 (context) 通常是代表了意見的背景，它可以是一個季節、時間、地點、代稱或是甚至有時候口語上是被忽略不提的了。但在 3.3 及緒論中我們有提到情境的確有可能會讓意見的極性逆轉，因此若能建立情境用詞的詞彙庫，並配合意見、情感、情緒的詞彙庫一同進行標記，便能較準確地判斷實際的語意。在 CKIP 斷詞系統的規格中我們可以發現許多不同的名詞的種類，甚至有符合上述情境的詞性，然而經過實際測試，回傳的資料並沒有標記得這麼細節，否則我們也可以直接利用詞性作為情境的區分 [1]。



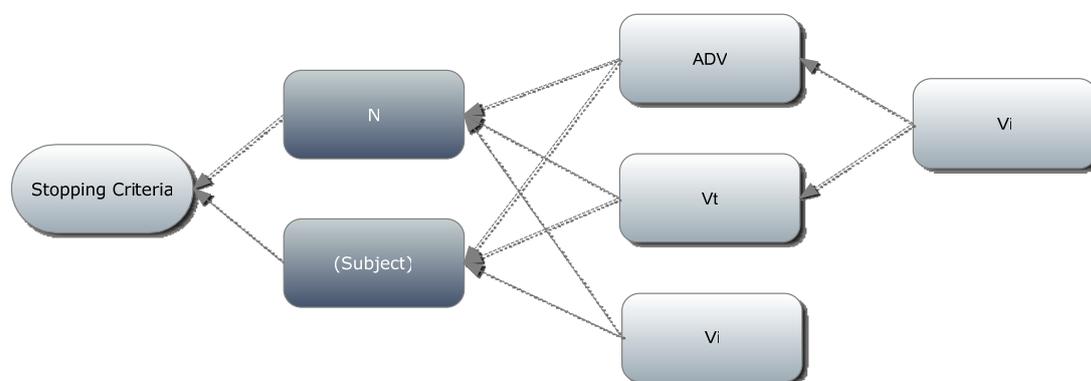
圖九、篩選情境詞彙

類似於 3.2 節的作法，我們可經由類似的步驟來列舉出所以情境的詞彙，示意圖如(圖九)，當中的三個詞彙組合即為經由 3.2.2 節之詞組探索規則建立得來。

語句中屬於情境的用語通常會在句子的前段，例如(圖六)、(圖七)及(圖八)所示的關係，而一則意見通常會有不只一個情境，端看意見提供者的意見以及每個意見前面是否有所屬的情境詞彙出現。而當情境無法於句子中被探索到時，通常意見的情境或是目標即為旅遊景點、店家整體 (圖十)。經由這樣子的探索我們可以逐步歸納出有意義的情境用語，茲依詞頻、所占比重摘錄部份展示如(表四)。

表四、情境詞組概率統計表(節錄)

N Vi - I		N ADV Vi - I		N Vt Vi - I		N Vi - II		N ADV Vi - II		N Vt Vi - II	
態度	1206 3.60%	態度	118 3.98%	活動	28 4.79%	菜色	119 0.36%	味道	18 0.61%	朋友	3 0.51%
價格	650 1.94%	感覺	69 2.33%	態度	16 2.74%	早餐	118 0.35%	交通	16 0.61%	木屋	3 0.51%
東西	621 1.85%	環境	67 2.26%	老闆	10 1.71%	飲料	113 0.34%	景觀	15 0.54%	大家	3 0.51%
老闆	428 1.28%	風景	63 2.13%	風景	7 1.20%	食材	113 0.34%	心情	14 0.51%	水果	3 0.51%
感覺	301 0.90%	價格	63 2.13%	地方	6 1.03%	餐點	111 0.33%	生意	13 0.47%	地圖	3 0.51%
老闆娘	254 0.76%	老闆	58 1.96%	人員	5 0.85%	蛋糕	110 0.33%	農場	13 0.44%	浴室	3 0.51%
個人	242 0.72%	房間	49 1.96%	房間	5 0.85%	湯頭	109 0.33%	停車場	13 0.44%	小朋友	3 0.51%
房間	241 0.72%	地方	36 1.65%	感覺	5 0.85%	設備	108 0.32%	時間	12 0.44%	氣氛	2 0.34%
客人	241 0.72%	老闆娘	34 1.21%	客人	5 0.85%	麵包	106 0.32%	菜色	12 0.40%	蝦量	2 0.34%
朋友	227 0.68%	空氣	32 1.15%	老闆娘	4 0.68%	咖啡	106 0.32%	床	12 0.40%	花	2 0.34%
價錢	225 0.67%	老闆人	29 1.08%	場地	4 0.68%	生活	105 0.31%	早餐	11 0.40%	印象	2 0.34%
人員	186 0.56%	價錢	26 0.98%	價格	4 0.68%	店家	105 0.31%	印象	11 0.37%	電話	2 0.34%
生意	177 0.53%	餐點	25 0.88%	餐點	4 0.68%	錢	100 0.30%	晚上	11 0.37%	水	2 0.34%
環境	159 0.47%	人員	22 0.84%	環境	3 0.51%	品質	95 0.28%	小朋友	11 0.51%	溫泉水	2 0.34%
小姐	146 0.44%	設施	22 0.74%	遊客	3 0.51%	餐廳	89 0.27%	這裡	10 0.47%	地板	2 0.34%
味道	136 0.41%	設備	22 0.74%	小孩	3 0.51%	地方	87 0.26%	木屋	10 0.44%	溫泉	2 0.34%
服務生	131 0.39%	空間	21 0.74%	民宿	3 0.51%			視野	10 0.44%		
速度	131 0.39%	品質	19 0.71%	山路	3 0.51%			山路	9 0.44%		
口味	129 0.39%	路	19 0.64%	價位	3 0.51%			景色	9 0.40%		
價位	129 0.39%	園區	19 0.64%	早餐	3 0.51%			門票	9 0.40%		
裝潢	128 0.38%	場地	18 0.64%	價錢	3 0.51%						



圖十、情境在句中位置及詞性組合的關聯

### 3.5 分類器訓練及測試

本研究使用 SVM 為分類器，然而在初步測試中非線性之 LIBSVM 訓練時間過長而無法採用，因此採用正確率稍低一些但是時間非常快速之線性 LIBLinear 進行訓練及測試的分類器。

資料集已刪除未知資料(unknown data)、亂碼及廣告資料。資料集之類別是屬於不平衡的狀態(unbalanced data)，極性資料集狀況較好：majority(正向): 3026，minority(負向): 1839；具情緒之綜合資料集則較偏頗些(biased)：majority(正向): 3766，minority(負向): 1099。也因此對於具情緒之綜合資料集在訓練時我們會對 minority class 進行加權微調 (weight 參數加權)。全部資料集皆以 5-folds cross validation (以亂數方式抽樣製作樣本)來進行測試。

實驗設定分為兩大類，依詞彙庫、情境資訊的加入與否共八種組合進行之(表五、表六)：

表五、實驗組別 – 極性分類

極性分類		詞彙庫輔助	
		有	無
情境資訊輔助	有	✓	✓
	無	✓	✓

表六、實驗組別 – 極性及情感分類

極性及情感分類		詞彙庫輔助	
		有	無
情境資訊輔助	有	✓	✓
	無	✓	✓

#### 四、 實驗結果

表六、極性分類實驗結果

極性分類		詞彙庫輔助	
		有	無
情境資訊輔助	有	Accuracy : 99.75%	Accuracy : 62.67%
		Precision : 99.61%	Precision : 63.87%
	Recall : 100.00%	Recall : 92.07%	
	<b>F<sub>1</sub> Score : 99.80%</b>	<b>F<sub>1</sub> Score : 75.42%</b>	
無	Accuracy : 99.77%	Accuracy : 63.41%	
	Precision : 99.64%	Precision : 63.88%	
	Recall : 100.00%	Recall : 94.75%	
	<b>F<sub>1</sub> Score : 99.71%</b>	<b>F<sub>1</sub> Score : 63.65%</b>	

表七、極性及情感分類實驗結果

極性及情感分類		詞彙庫輔助	
		有	無
情境資訊輔助	有	Accuracy : 95.56%	Accuracy : 55.87%
		Precision : 100.00%	Precision : 27.66%
	Recall : 80.35%	Recall : 59.05%	
	<b>F<sub>1</sub> Score : 89.10%</b>	<b>F<sub>1</sub> Score : 37.68%</b>	
無	Accuracy : 95.31%	Accuracy : 57.90%	
	Precision : 100.00%	Precision : 27.50%	
	Recall : 79.25%	Recall : 52.78%	
	<b>F<sub>1</sub> Score : 88.43%</b>	<b>F<sub>1</sub> Score : 36.16%</b>	

#### 五、 分析與結論

本研究分別以意見極性(正、負向)、意見極性及情感強度(意見有情感、情緒或是不具情感、情緒)，以及情境資訊加入與否以及是否使用詞彙庫來建立主要的實驗項目，共有八種實驗組合，並分別以 precision、recall 及 F<sub>1</sub> Score 作為效能評估的參考依據。 在使用 F<sub>1</sub> Score 為評估指標時得到最佳結果：99.8% 為單獨判斷意見極性時，並同時考慮情境因素，以及使用詞彙庫輔助分類（不使用詞彙庫時最佳結果為 75.42%）；而納入情感意見的綜合分類最佳結果為：89.1% 為同時考慮情境因素及使用詞彙庫輔助（不使用詞彙庫時僅 37.68%）。 以情境輔助的成效上同樣以 F<sub>1</sub> Score 為評估指標，若在各組使用詞彙庫的情況下則績效不顯著，不使用詞彙庫條件下則極性判斷項目有較好的成效（未考慮情境因素：63.65%；有考慮情境因素：75.42%）。 當僅關注在不使用情境資訊、詞彙庫時，意見極性分類往往比較複雜之極性及情感分類於績效上有較好的表現，可見當涉及情感意見時，資料內部的衝突性會比較大，使得分類器無法找出合適的切割點

而最後導致分類效能低落。因此我們可考量其他輔助的方式如本實驗中使用的自建詞庫或是增加有助分類之屬性，在本實驗中可見使用詞彙庫時由於其良好之覆蓋率，使得分類效能之改善比使用情境資訊來得顯著得多。然而雖建立意見領域專屬之詞彙庫有助於分類，但頗受爭議的負面問題如時間效率、標記不客觀、資料更新等議題則會在詞彙庫逐漸增大時趨向顯著，且當意見本身即具有不同意見領域之特性時(例如電影之觀看意見)，效能改善可能也會不如預期(情境資訊應該可於此類狀況下發揮較佳之輔助)。

本研究採用分類器加上情境資訊及自建詞彙庫之輔助分類方法，經由多組實驗證實可得到較佳之分類績效，然而其中使用之線性 SVM (LibLinear) 分類器仍會相當容易發生像是類別不平衡(unbalanced data)、參數調整不易(cost、weight)以及大量衝突性較高資料產生之過度訓練(over-fitting)、訓練不足(under-fitting)的問題。詞彙庫及訓練資料的建立仍仰賴大量的人工標記，此處有可能會有主觀偏差存在(bias)。情境不甚顯著的情況則有可能是因本實驗樣本限定於幾乎相同意見領域(domain)的緣故，意即以旅遊經驗此單一意見領域來看，意見描述用語的語意極性方向大致上是相似的，較無涉及不同領域容易發生之意見極性逆轉現象，但也因為這種單一意見領域的專屬性而造就了詞彙庫相當顯著的績效增長。

## 參考文獻

- [1] Changhua Yang, Kevin Hsin-Yih Lin, Hsin-Hsi Chen, Sentiment Analysis in Weblog Using Contextual Information: A Machine Learning Approach. *Int. J. Comput. Proc. Oriental Lang.* 21(4): 331-345, 2008
- [2] Bing Liu, *Sentiment Analysis and Subjectivity, from Handbook of Natural Language Processing, Second Edition*, 2010
- [3] Bing Liu, *Web Data Mining – Exploring hyperlinks, Contents and Usage Data, Web Crawling & Opinion Mining*, Springer, 2006
- [4] Qi Zhang, Xuanjing Huang, Lide Wu, Sentiment Classification Using Supervised and Semi-Supervised Conditional Maximum Entropy Modeling. *Int. J. Comput. Proc. Oriental Lang. (IJCPOL)* 21(4):295-308, 2008
- [5] Tang, Huifeng and Tan, Songbo and Cheng, Xueqi, A survey on sentiment detection of reviews, *Expert Systems with Applications*, 36-7, 2009
- [6] Bo Pang, Lillian Lee, Opinion mining and sentiment analysis, *Foundations and Trends in Information Retrieval*, Vol. 2, No 1-2, 2008
- [7] Chia-Ying Lee, Ku-Wei Ku and Hsin-Hsi Chen, A Study on Identification of Opinion Holders, *Proceedings of the 21st Conference on Computational Linguistics and Speech Processing*, Taichung, Taiwan, 101-114, September 1-2, 2009