情緒分析的應用與實作

系所/資訊工程學系 指導老師/顏秀珍 組員/張鐘元、李旺哲、曹文杰

隨著近年來人們大量在網路上發 表自己的想法及言論,因考慮到這些大 量網路言論的價值,部份學者開始對這 些資料做情感分析;在這個資訊爆炸的 時代,現代人往往因為繁忙而忽略了 章所要表達的含意,當我們閱讀一篇 章或評論時,如何能夠以最快的速度了 解此文章所表達的含意是正面還是負 面則成為了閱讀文章的關鍵。

本研究以蒐集正面與負面各一千 則酒店評論為例,將兩千則評論加入情 感辭典,並將其結果以結巴斷詞進行斷 詞。接著對其進行情感傾向分析,以此 得到每則評論的特徵對應的情感分數, 最終之結果能讓我們快速判別此評論 所表達之情緒。

目前專題成果如圖 1、圖 2 所示。 圖 1 為輸入一個正面情緒語句,可由中 間的預測按鈕中按下預測,得到的結果 就如圖 1 所示。在圖中,綠色表格為預 測結果為正面的評論,並顯示該語句的 情感權重。而紅色表格為預測結果為負 面的評論,也顯示出該語句的情感權 重。圖 2 為讀取任意五篇評論的文字檔 預測顯示。顯示結果能看到灰色表格為 文字檔內的文字內容。綠色表格為預測 結果為正面的文章,並顯示該評論正面 的機率值。紅色表格為預測結果為負面 的文章,也顯示出該評論負面的機率 值。

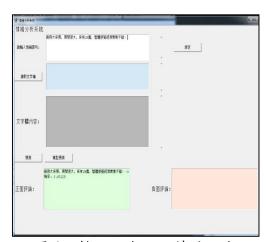


圖 1:輸入一句正面情緒語句

商報入保護部 句:		^:	#0
接取文字權	C:/Python/pos/new pos/pos.4.tat C:/Python/pos/new pos/pos.4.tat	•	
	C:/Python/meg/mew meg/meg.3.txt C:/Python/meg/mew meg/meg.5.txt	1.	
文字權內容;		- :	
	位置解我們幫信保近,從價格來說、性價於依高、稅費的大保期,148元,前 報員限度等好,開閱榜件一般,但是想想價格也就這樣了。這里影響,就是 士婦	分類 R整子 =	
	大校。 eng:沒有腦接電話,這個便好		
	美元監評 2000年3月4日 : 對了,這種發病免費上網,期好上網加班一 不確	700	
预用	模型開発		1
正面評論:	I-GREF, KEFETAN.		接近楼里間: 任角兩點 (為什麼媒捏訂的都是任角型) ,兩點 小、隔身起事。展開非常冷、交通幾千不好用,產中間更治、例
	BTW:沒有程接電話,這個例好 補充點評 2000年3月4日 : 對了,這裡能與免費上級,剛好上級	負面評論:	千法法准治 (三國字"不堪意" (-> 號率: 0.637309 頭紅標準間: 抗角開閉(為什麼護程釘的都是抗角間),開閉6
	開工的計 200年3月1日 - 到1,24日本時代代工時,2007年3 随程一般,但不確 -> 原本: 0.487309 距離川沙公路軌近,但是公交指示不動,如果是"客陸線"的話。音 車里無理,建議局別問發後,展開教為禁煙。-> 健生: 0.63730	共田計劃	1、「陳音越差」 開閉非常冷,空頭幾乎不好用,衛生間更冷。 利 千浪法凍沼 (三語字"不滿思" (-> 幾至: 0.437309

圖 2: 讀取任意五篇評論文字檔