

# 爬文PTT八卦版留言之情緒分析

—以2019年12月為例

第十二組 徐榆婷 林蕎安 黃欒雅 劉育致 林家好

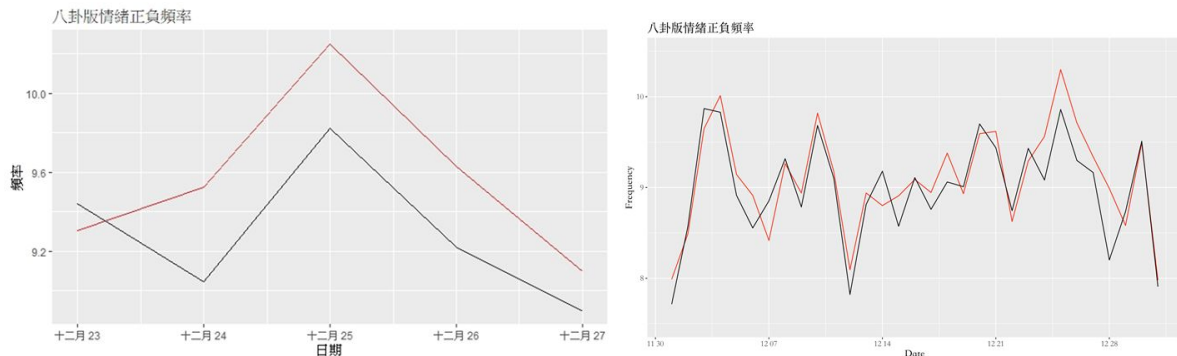
研究目的：

希望藉由分析留言中正負向情緒用詞的頻率比較，能了解人們的情緒變動與時間的關聯性。

研究方法：

透過python爬PTT八卦版在2019年12月份的所有文章留言以取得數據，並使用jiebaR斷詞，再將詞語與情緒辭典(Liang-Chih Yu et al. 2016)做比對，以日期為x軸，正/負向頻率（次數/當日貼文數）為y軸，畫出折線圖。

研究結果與討論：



結果顯示出正負詞情緒頻率趨於重疊，於12/25聖誕節正向詞才明顯高於負向詞，但由於我們的研究方法為比對詞語與情緒辭典，一是情緒辭典可能無法包含所有的情緒詞，且PTT用戶有發展出獨特的PTT用詞，但本次研究未能涵括PTT鄉民辭典中的詞語，二是單指比對，無法藉由上下文意分析情緒，故可能有缺漏。

研究的未來發展除了可以將時間軸拉長以外，另可分析情緒頻率高低與日期及發生事件的關聯性，此外也可以細部分析特定情緒頻率高的詞語內容，並將PTT常用詞語編列情緒正負向詞辭典，以利後續研究分析使用。

參考文獻：

Liang-Chih Yu, Lung-Hao Lee, Shuai Hao, Jin Wang, Yunchao He, Jun Hu, K. Robert Lai, and Xuejie Zhang (2016). Building Chinese Affective Resources in Valence-Arousal Dimensions. In Proceedings of the 15th Annual Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies (NAACL-HLT'16), San Diego, California, USA, 12-17 June, 2016.