

## 經濟史導讀

學號：R05323048 經研所碩一蘇哲生

題目：Emotion Analysis for Big Data

### 1 What is the question (of the paper)?

ANS：

如何在有限的成本並追求最高的準確率下，利用各種網路上的資料進行大數據的情緒分析？

### 2 Why should we care about it?

ANS：

若能成功分析人們的情緒，則可以藉此預測某人是否具有邊緣性人格或是躁鬱症等症狀，近一步提早治療，減少社會成本；亦可以分析消費者在購買產品之後對於該產品的評價及回饋，使廠商了解如何改進該產品，提高廠商利潤。

### 3 What is your (or the author's) answer?

ANS：

相較於傳統的分析需要了解每個字的字義，且換種語言就要全部重分析相比，使用群眾的淺意識可以節省大量的金錢與時間成本。

大部分的語句都蘊藏著說話者的情緒，藉由這些大量的資料加以整理，去除掉不重要的訊息後再來進行分析，可以在有限的資金下得到一樣甚至更好的成果。

### 4 How did you (or the author's) get there?

ANS：

利用公開合法且可免費獲得的網站來搜集資料，如 FB、twitter...等等；並利用發文者的標籤（hashtag）和表情符號來當作特殊的情緒標語來記錄發文者當下的情緒，然後分別建立實驗組與對照組進行實驗。

之後再拿掉不重要（如客觀的新聞）或難以處理的資料，比較實驗組與對照組之間的相異處。最後將其畫成圖表來觀察，去掉兩組相同特徵的資料並分析哪些文字常被使用，剩餘的字詞便是能夠代表該情緒的重要字詞。

完成後再比較每種情緒之間的頻率、獨特性與多元性進行排序，最後再來試驗正確率，判斷是否足以支持此方法。