

Obtaining Reliable Human Ratings of Valence, Arousal, and Dominance for 20,000 English Words

論文作者：Saif M. Mohammad

報告人：曹容



大綱

情感分析的三大要素

01

人工標記

03

結論

05

02

VAD的應用

04

人為因素的影響



情感分析的 三大要素(VAD)



Valence

正向—負向
宴會 vs. 葬禮



Arousal

積極—消極
緊張 vs. 懶惰



Dominance

主導—順從
爭論 vs. 柔和

VAD的應用



自然語言處理

文字的自動情緒
和情感分析

心理學

了解人們如何看待
周圍的世界

認知科學

了解人類如何表達
和使用語言

社會科學

了解人與人之間的關係

人工標記

作者預先標註了約2%的數據。這些問題稱為黃金問題。

黃金問題上的準確性低於80%，則將丟棄該測試者所有的標記。

同時填寫一份簡短的人口統計問卷，使我們能夠確定年齡、性別和個性等屬性是否會影響我們的標記結果。



人為因素的影響

性別

男 – Valence 和 Dominance 較高
女 – Arousal 較高

年齡

35歲以下 – Valence和Arousal較一致
35歲以上 – Dominance較一致

個性

外向比起內向一致性較高



我們分析了人口統計信息，顯示出儘管註釋總體上導致重複註釋中的評分保持一致，但不同人口群體（例如，男性和女性，35歲以上的年齡段以及35歲或35歲以上的人群）的協議在統計上存在顯著差異在人格維度下和跨人格維度（性格外向，性格內向，神經質和安全感等）。這些結果表明，某些人口統計屬性會根據其中的概念的相對價，喚醒和主導地位影響我們如何看待周圍的世界。

結論





THANKS