

Modelling Valence and Arousal in Facebook Posts

Lyle Ungar

D. Preotiuc-Pietro, H.A. Schwartz

G. Park, J. Eichsteadt, M. Kern, E. Shulman

Positive Psychology Center
University of Pennsylvania

 Penn | World Well-Being Project

16 June 2016

Motivation

Data Sources

amazon yelp 



Product reviews

Opinions towards products,
restaurants, events, etc.

Long, more structured

Affective states

Feelings towards self or
others.

Short, less structured

Models of product sentiment and emotion should be different

Motivation

Models of emotion

Discrete Emotions

Most popular in NLP are Ekman's six emotions: anger, disgust, fear, joy sadness, surprise



Some emotions driven by similar words (*hell, bad* → sadness, fear, anger)

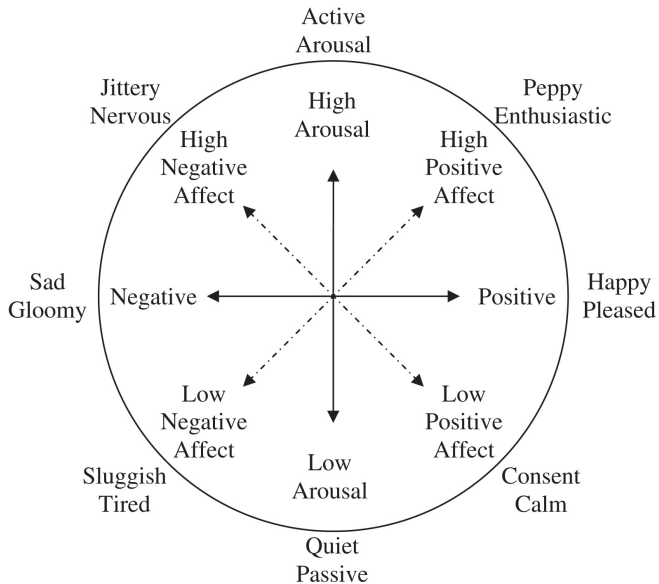
Dimensional Models

Each affective state is a combination of real-valued components

Most popular is the circumplex model (*Russel 1980, Posner 2005*)

Two independent neurophysiological systems: valence (or sentiment) and arousal

Emotion Circumplex



Applications

Goal: Automated large-scale psychological studies

- measuring time-of-day and day-of-week mood swings
 - and what causes them
- mental illness detection
 - bipolar, schizophrenic breaks ...
- analysing movies and books
 - and how they vary in emotion content
- correlating with external effects
 - e.g. weather, sports game outcomes, ...

Measuring Valence and Arousal

- **Valence (or sentiment or polarity)**
 - 1 (very negative) – 5 (neutral/objective) – 9 (very positive)
- **Arousal (or intensity)**
 - 1 (neutral/objective post) – 9 (very high intensity)

Examples

| Message | V | A |
|--|-----|-----|
| Is the one whoz GOing to Light Up your Day!!!!!!!!!!!!!! | 7 | 8 |
| Blessed with a baby boy today ... | 7.5 | 2 |
| the boring life is back :(... | 3 | 2.5 |
| IS SUPER STRESSED AND ITS JUST THE SECOND MONTH OF SCHOOL ..D: | 2.5 | 7 |

Example of posts annotated with average valence (V) and arousal (A) ratings.

Data Source

3120 Facebook posts

Stratified by:

- Age (13-35)
- Gender (M/F)

Each message from a distinct user

All messages from the same time interval

Annotation

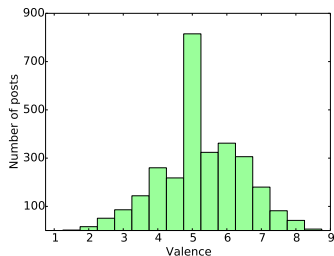
Two annotators:

- psychology students
- received training in annotating these traits, including anchoring
- no distractions that may affect they mood (music, etc.)

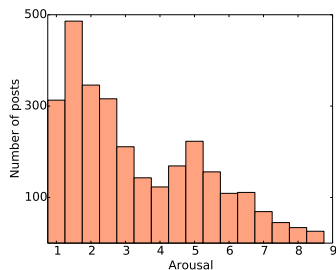
Messages are un-ratable if they are not in English or contain no cues

- 235 messages (~7.5%)
- Cohens Kappa $\kappa = .93$

Annotation Results



Valence



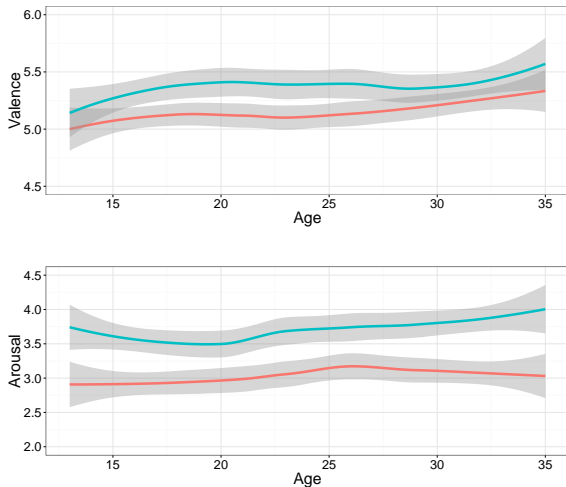
Arousal

Histograms of average rating scores.

Valence–Arousal $\rightarrow r = 0.222$

Valence–Arousal $\rightarrow r = 0.085$ (ignoring neutral posts)

Gender and Age Differences



Variation in valence and arousal with age in our data set using a LOESS fit. Data is split by gender: **Male** and **Female**.

Predicting Valence & Arousal

Train a classifier for predicting valence and arousal separately

Features: Bag-of-words (only unigrams)

Model: Linear regression with elastic net regularization

Test: 10 fold cross-validation

Baseline Models

1. ANEW

- valence and arousal ratings for ~1400 words (*Bradley and Lang, 1999*)

2. AffNorms

- valence and arousal ratings for ~14000 words (*Warriner et al., 2013*)

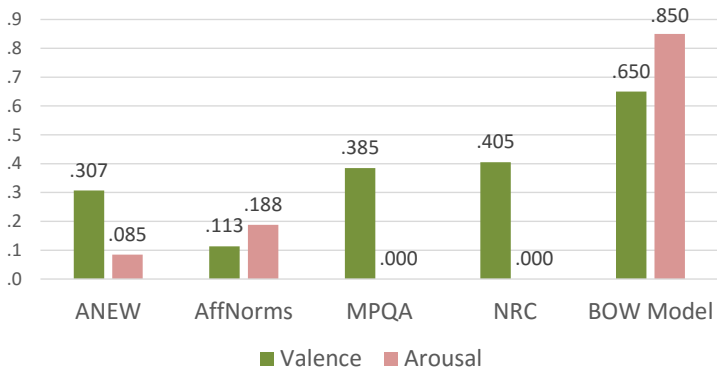
3. MPQA

- 7629 words rated for positive or negative sentiment (*Wilson et al. 2005*)

4. NRC

- Hashtag Sentiment Lexicon adapted to Social Media (*Mohammad et al., 2013*)

Results



Message rating prediction accuracy (in Pearson r).

Results on 10 fold cross-validation.

Quantitative Analysis – Valence

| + Valence | <i>r</i> | - Valence | <i>r</i> |
|------------------|----------|------------------|----------|
| ! | .251 | hate | -.163 |
| :) | .237 | :(| -.159 |
| birthday | .212 | ? | -.117 |
| happy | .197 | sick | -.112 |
| thank | .196 | why | -.102 |
| great | .195 | :'(| -.094 |
| love | .195 | not | -.093 |
| thanks | .179 | bored | -.092 |
| wishes | .170 | stupid | -.089 |
| wonderful | .159 | ... | -.087 |

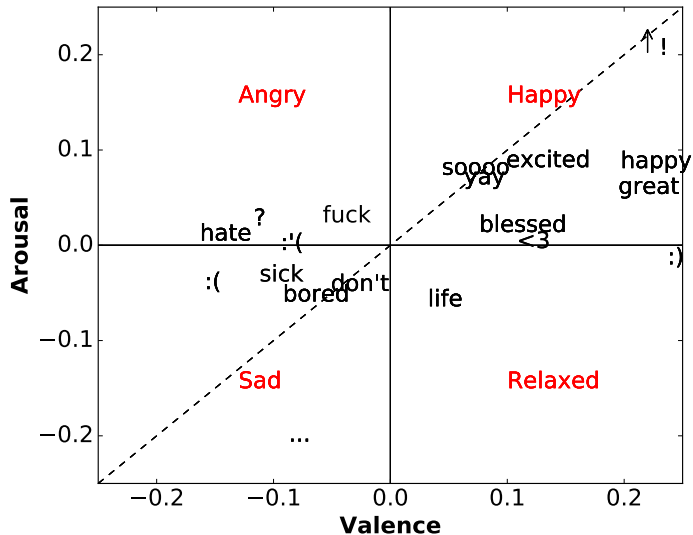
Words most positively and negatively correlated with valence

Quantitative Analysis – Arousal

| + Arousal | <i>r</i> | - Arousal | <i>r</i> |
|------------------|----------|------------------|----------|
| ! | .773 | ... | -.206 |
| birthday | .097 | . | -.164 |
| happy | .081 | status | -.064 |
| its | .079 | life | -.064 |
| wishes | .076 | people | -.060 |
| soooo | .074 | bored | -.059 |
| thanks | .073 | :/ | -.056 |
| christmas | .071 | of | -.056 |
| sunday | .069 | deal | -.056 |
| yay | .064 | every | -.054 |

Words most positively and negatively correlated with arousal

Quantitative Analysis - Circumplex



Reviews \neq Personal Feelings

Valence/Arousal \neq Discrete Emotions

Annotated Facebook data set and bag-of-words model available

<http://wwbp.org/publications.html>

<http://lexhub.org/>

