Emotion Analysis in Natural Language

Pearl Pu
What and why Emotion?
The Distinguished Speakers Program is made possible by

Association for Computing Machinery

Advancing Computing as a Science & Profession

For additional information, please visit http://dsp.acm.org/
ACM, the Association for Computing Machinery (www.acm.org), is the premier global community of computing professionals and students with nearly 100,000 members in more than 170 countries interacting with more than 2 million computing professionals worldwide.

OUR MISSION: We help computing professionals to be their best and most creative. We connect them to their peers, to what the latest developments, and inspire them to advance the profession and make a positive impact on society.

OUR VISION: We see a world where computing helps solve tomorrow’s problems - where we use our knowledge and skills to advance the computing profession and make a positive social impact throughout the world.
What is Emotion?

- emotion is a reaction to events
- emotion leads to changes in multiple organismic subsystems

Emotion is the conduit that connects our mind to our body
https://www.youtube.com/watch?v=9J1lXEXFHog
Why Emotion Matters?
My son in law booked us in here and we were very pleased with his choice. 
Check in was smooth and easy. 
The hotel is smart, trendy and very well situated for exploring the city. We went into town virtually every day and most of the town is walkable even to the castle on the far side of town. 
Our rooms were very large with a spacious bathroom complete with hair dryer. Basic soap and shampoo are provided. 
The wardrobe space is a bit limited but we managed. 
The room was serviced daily and was kept very clean. There is a fridge stocked with the hotels items, all chargeable. 
There is no facility in the room to make tea or coffee but we always pack a mini kettle and put some milk in the fridge. 
We had breakfast only which was OK but at the peak time on the weekends, 8.30 ish it was a bit chaotic and they could do with a second coffee maker. 
All in all I would recommend this hotel and would stay there again.
My son in law booked us in here and we were very **pleased** with his choice.  
Check in was **smooth** and **easy**.  
The hotel is **smart**, **trendy** and **very well situated** for exploring the city. We went into town virtually every day and most of the town is **walkable** even to the castle on the far side of town.  
Our rooms were **very large** with a **spacious** bathroom **complete** with hair dryer. The bed is **comfortable** and so are the **pillows**. The wardrobe space is a bit **limited** but we managed. The room was serviced daily and was kept very **clean**. There is a fridge stocked with the hotels items, all **chargeable**.  
There is **no facility** in the room to make tea or coffee but we always pack a mini kettle and put some milk in the fridge.  
We had breakfast only which was OK but at the peak time on the weekends, 8.30 ish it was a bit **chaotic** and they could do with a second coffee maker.  
All in all I would **recommend** this hotel and would stay there again.
How to model emotions?

Dimensional

Pros
• More universal
• Can describe any experience

Cons
• Difficult to express states linguistically

Categorical

Pros
• Provides linguistic labels
• Allows variety of applications

Cons
• No agreement on the unique set
Categorical - Ekman

Ekman, An argument for basic emotions, 1992
Plutchik's Wheel of Emotions

[Diagram of Plutchik's Wheel of Emotions]

[two-dimensional circumplex model]

[three-dimensional circumplex model]
Geneva Emotion Wheel

GEW, adapted version 2.0
How to model emotions?

**Dimensional**
- Valence
- Arousal
- Dominance

**Categorical**
- Happiness
- Irritation
- Sadness
- Interest
- Fear
- Anxiety

**Pros**
- More universal
- Can describe any experience

**Cons**
- Difficult to express states linguistically

**Pros**
- Provides linguistic labels
- Allows variety of applications

**Cons**
- No agreement on the unique set
How to Detect Emotions?

- facial expression
- voice
- galvanic skin response (GSR)
- skin temperature
- electrocardiogram (ECG)
- electroencephalogram (EEG)
- gesture
- text
Emotion Recognition in Natural Language

Which emotions are expressed?

As a mother I know the pride in one’s child, as an American I know the pride in one’s country. I feel a little both for you.


The emotional life of your brain: how its unique patterns affect the way you think, feel, and live - and how you can change them. By Richard Davidson and Sharon Begley, 2002.

Thinking, fast and slow. By Daniel Kahneman, 2011.
Emotion Lexicons
Gabby should only feel immense pride in her accomplishments at Rio. Ignore the jealous haters. (Pride/Elation, strong)

Have you seen Chinese swimmer #Fu Yuanhui? Her reactions are infectious, hysterical and really authentic. #Rio2016. (Happiness/Joy, strong)

#realDonaldTrump first president in a long time that has the American people and their interests at heart. Thank you Mr. President. (Love/tenderness, worry/fear, sadness/despair, embarrassment/shame)
Emotion Lexicons

- General Inquirer
- ANEW
- Bing Liu’s Lexicon
- OpinionFinder
- WordNet Affect
- NRC
- GALC (geneva affect label coder)
- LIWC (linguistic inquiry and word count)
爱恨怕悲念恋

412 characters with heart in it, the most complex one has 28 strokes
OlympLex 2013 (EPFL)

• Create an emotion lexicon dedicated to sport events
• Distinguish up to 20 categories of emotions
• Develop a novel method for crowd worker
Crowdsourcing by Darwin

Watch Video
Darwin’s Letters - a Visualisation
Our Human Computation Task

**Task: Annotate this tweet with emotions**

1. Read this tweet and imagine you were the author of it:
   
   Really enjoying watching team GB in the #Gymnastics. They actually doing really good. Impressive moves.
   #london2012

2. What emotion did you feel?
   (Choose a circle of corresponding category. Different circle sizes means different emotion strength)

3. Copy textual indicators of your emotion:
   (Place each expression on a new line, it can be a word or a phrase)
   
   Really enjoying
   really good
   impressive

4. How else would you express this emotion?
   (Please, be creative and place each expression on a new line)
   
   amazing
   fantastic job
   awesome moment
Human Computation Task

Task: Annotate this tweet with emotions

1. Read this tweet and imagine you were the author of it:
   Really enjoying watching team GB in the #Gymnastics. They actually doing really good. Impressive moves. #london2012

2. What emotion did you feel?
   (Choose a circle of corresponding category. Different circle size means different emotion strength)

3. Copy textual indicators of your emotion:
   (Place each expression on a new line, it can be a word or a phrase)
   Really enjoying
   really good
   impressive

4. How else would you express this emotion?
   (Please, be creative and place each expression on a new line)
   amazing
   fantastic job
   awesome moment

Emotion Label
Happiness, Anger, Fear, No emotion...

Emotion Strength
Low, Medium, High

Emotion Polarity
Positive, Negative, Neutral
Human Computation Task

Task: Annotate this tweet with emotions

1. Read this tweet and imagine you were the author of it:

   Really enjoying watching team GB in the #Gymnastics. They actually doing really good. Impressive moves. #london2012

2. What emotion did you feel?
   (Choose a circle of corresponding category. Different circle size means different emotion strength)

3. Copy textual indicators of your emotion:
   (Place each expression on a new line, it can be a word or a phrase)

   Really enjoying
   really good
   Impressive

4. How else would you express this emotion?
   (Please, be creative and place each expression on a new line)

   amazing
   fantastic job
   awesome moment

Tweet Emotion Indicators
n-grams expressing or revealing the chosen emotion in the text
Human Computation Task

Task: Annotate this tweet with emotions

1. Read this tweet and imagine you were the author of it:
   Really enjoying watching team GB in the #Gymnastics. They actually doing really good. Impressive moves. #london2012

2. What emotion did you feel? (Choose a circle of corresponding category. Different circle size means different emotion strength)
   - Contempt
   - Scorn
   - Irritation
   - Anger
   - Involvement
   - Interest
   - Amusement
   - Laughter
   - Pride
   - Elation
   - Happiness
   - Joy
   - Envy
   - Jealousy
   - Disappointment
   - Regret
   - Guilt
   - Remorse
   - Embarrassment
   - Shame
   - Worry
   - Fear
   - Sadness
   - Despair
   - Pity
   - Compassion
   - Longing
   - Nostalgia
   - Astonishment
   - Surprise
   - Feeling disturbed
   - Relief
   - Wonderment
   - Feeling awe
   - Tenderness
   - Feeling love
   - Enjoyment
   - Pleasure
   - Other emotion
   - input name
   - No emotion

3. Copy textual indicators of your emotion: (Place each expression on a new line, it can be a word or a phrase)
   Really enjoying
   really good
   impressive

4. How else would you express this emotion? (Please, be creative! and place each expression on a new line)
   amazing
   fantastic job
   awesome moment

Additional Emotion Indicators
other n-grams expressing or indicating the chosen emotion
SREC (Sport-related Emotion Corpus)

Figure 4.3: Distribution of emotion labels in crowdsourced workers’ answers comprising the SREC data (i.e. after the application of posterior quality control).
Quality of Labels

Observed agreement: \( A = \frac{2}{6} = 0.33 \)

Kappa: \( \kappa = \frac{A - A_c}{1 - A_c} \), where \( A_c \) - chance agreement

<table>
<thead>
<tr>
<th>Agreement</th>
<th>Kappa</th>
</tr>
</thead>
<tbody>
<tr>
<td>Emotion category</td>
<td>29.3%</td>
</tr>
<tr>
<td>Polarity</td>
<td>75.7%</td>
</tr>
<tr>
<td>Strength</td>
<td>43.8%</td>
</tr>
</tbody>
</table>
How did we aggregate results?

Emotion distributions for the tweets with the term

Additional indicators using the same term

Averaging distributions

Term emotion distribution

Example term: hopefully

Tweets

Hopefully she can smash it and bring home medal

They are now watching the #TeamGB boys hopefully get medal

Hopefully #TeamGB will get an #Olympic Gold Soon!

hopefully
OlympLex 2013 (EPFL)

- Number of annotation ~2000 tweets
- Contains 3,193 terms
- Examples (per quadrant)

<table>
<thead>
<tr>
<th>Anger, Disgust, Scorn, ...</th>
<th>Pride, Happiness, Interest, ...</th>
</tr>
</thead>
<tbody>
<tr>
<td>unfair, mad, ugh, annoyed, ticked</td>
<td>bravo, champions, my girl, hero,</td>
</tr>
<tr>
<td>off, idiots, slap, offended, epicfail,</td>
<td>woohoo, sohappy, good job, yessss,</td>
</tr>
<tr>
<td></td>
<td>...</td>
</tr>
<tr>
<td>ouch, noooo, eek, tough to watch,</td>
<td>astounded, luv u, incredible talent,</td>
</tr>
<tr>
<td>heartbroken, feel so bad, fearful,</td>
<td>omg, marry me, desiring, amaze,</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Sadness, Fear, Pity, ...</td>
<td>Love, Surprise, Awe, ...</td>
</tr>
</tbody>
</table>
Lessons Learned

- Crowdsourcing is a viable approach to collect annotated data
- Distinguish speaker’s emotions from the audience’s emotion
- Online inexpert workers require training
- Provide context for emotion labelling
- Consider diversity as an advantage rather than noise
Reference

The Psychological Meaning of Words: LIWC and Computerized Text Analysis Methods. Volume: 29 issue: 1, page(s): 24-54. Article first published online: December 8, 2009; Issue published: March 1, 2010


3
Lexicons -> Classifiers
Why classifiers?

• Scalability
• Domain dependent words
• Fine-grained detection

Classifiers are more suitable for machines, more adapted to big data, capable of integrating the nuances of how humans collectively express emotions.
Supervised learning

Joy

Pride

Anger

Semi-supervised learning

Distant Learning
How does our method work?

GALC Scherer, 2005

<table>
<thead>
<tr>
<th>Happiness</th>
<th>Joy</th>
<th>Interest/Involvement</th>
<th>Surprise</th>
<th>Anger</th>
<th>Sadness</th>
<th>Disgust</th>
<th>Fear</th>
<th>Disappointment</th>
</tr>
</thead>
<tbody>
<tr>
<td>cheer*</td>
<td>ecstatic*</td>
<td>absorb*</td>
<td>amaze*</td>
<td>anger</td>
<td>chagrin*</td>
<td>abhor*</td>
<td>afraid*</td>
<td>comedown</td>
</tr>
<tr>
<td>bliss*</td>
<td>elate*</td>
<td>alert</td>
<td>astonished*</td>
<td>angr*</td>
<td>deject*</td>
<td>avers*</td>
<td>aghast*</td>
<td>disappoint*</td>
</tr>
<tr>
<td>delect*</td>
<td>euphoric*</td>
<td>animat*</td>
<td>dumbfound*</td>
<td>cross*</td>
<td>dole*</td>
<td>detest*</td>
<td>alarm*</td>
<td>discontent*</td>
</tr>
<tr>
<td>delight*</td>
<td>exalt*</td>
<td>ardor*</td>
<td>startl*</td>
<td>enrag*</td>
<td>gloom*</td>
<td>disgust*</td>
<td>discontent*</td>
<td>disenchant*</td>
</tr>
<tr>
<td>enchant*</td>
<td>exhilar*</td>
<td>attenti*</td>
<td>stunn*</td>
<td>furious</td>
<td>glum*</td>
<td>dislik*</td>
<td>dread*</td>
<td>fear*</td>
</tr>
<tr>
<td>enjoy*</td>
<td>exult*</td>
<td>curi*</td>
<td>surpris*</td>
<td>fury</td>
<td>grie*</td>
<td>disrelish</td>
<td>fright*</td>
<td>disgruntl*</td>
</tr>
<tr>
<td>felicit*</td>
<td>flush*</td>
<td>alert*</td>
<td>amaz*</td>
<td>incens*</td>
<td>hopeles*</td>
<td>distast*</td>
<td>fright*</td>
<td>disillusion*</td>
</tr>
<tr>
<td>happ*</td>
<td>glee*</td>
<td>alert*</td>
<td>amaz*</td>
<td>infuriat*</td>
<td>melanco*</td>
<td>loath*</td>
<td>panic*</td>
<td>frustrat*</td>
</tr>
<tr>
<td>merr*</td>
<td>joy*</td>
<td>alert*</td>
<td>amaz*</td>
<td>irate*</td>
<td>mourn*</td>
<td>nause*</td>
<td>panic*</td>
<td>jilt*</td>
</tr>
<tr>
<td>jubil*</td>
<td>overjoyed</td>
<td>alert*</td>
<td>amaz*</td>
<td>mad*</td>
<td>sad*</td>
<td>queas*</td>
<td>scare*</td>
<td>letdown</td>
</tr>
<tr>
<td>ravish*</td>
<td>enthusiast*</td>
<td>ferv*</td>
<td>amaz*</td>
<td>rag*</td>
<td>sorrow*</td>
<td>repugn*</td>
<td>scare*</td>
<td>resign*</td>
</tr>
<tr>
<td>rejoic*</td>
<td>interest*</td>
<td>ferv*</td>
<td>amaz*</td>
<td>resent*</td>
<td>tear*</td>
<td>repuls*</td>
<td>sour*</td>
<td>repugn*</td>
</tr>
</tbody>
</table>

GALC as the initial classifier
How does our method work?

Balanced Weighted Voting

1. Rebalancing
2. Weighted Voting

So proud 2 be British massive well done 2 all of Team GB! :D
so proud <int> be british massive well done <int> all of team gb ! <emot59>
proud, so proud, proud <int>, massive, well, massive well, done, well done, done <int>, <emot59>

Learn from all tweets with "well done"

Pride

well done
<table>
<thead>
<tr>
<th>Emotion Labeler</th>
<th>Algorithm</th>
<th>macro</th>
<th>micro</th>
<th>rank</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>P</td>
<td>R</td>
<td>F1</td>
</tr>
<tr>
<td>Random</td>
<td></td>
<td>2.5</td>
<td>1.3</td>
<td>1.7</td>
</tr>
<tr>
<td>Initial</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mcl-MNB</td>
<td></td>
<td>21.4</td>
<td>3.6</td>
<td>4.8</td>
</tr>
<tr>
<td>mcl-LogReg</td>
<td></td>
<td>7.5**</td>
<td>23.9**</td>
<td>8.9**</td>
</tr>
<tr>
<td>1vR-MNB</td>
<td></td>
<td>11.8**</td>
<td>17.1**</td>
<td>9.7**</td>
</tr>
<tr>
<td>1vR-LogReg</td>
<td></td>
<td>12.1**</td>
<td>8.8**</td>
<td>8.1**</td>
</tr>
<tr>
<td>PMI-based</td>
<td></td>
<td>12.7**</td>
<td>10.2**</td>
<td>9.3**</td>
</tr>
<tr>
<td>BWV</td>
<td></td>
<td>16.8**</td>
<td>11.5**</td>
<td>9.8**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Olymp-Lex</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial</td>
<td></td>
<td>11.4</td>
<td>9.7</td>
<td>7.1</td>
</tr>
<tr>
<td>mcl-MNB</td>
<td></td>
<td>19.7**</td>
<td>11.2**</td>
<td>6.8</td>
</tr>
<tr>
<td>mcl-LogReg</td>
<td></td>
<td>9.1**</td>
<td>12.4**</td>
<td>7.6**</td>
</tr>
<tr>
<td>1vR-MNB</td>
<td></td>
<td>19.4**</td>
<td>12.3**</td>
<td>7.3</td>
</tr>
<tr>
<td>1vR-LogReg</td>
<td></td>
<td>11.1*</td>
<td>16.5**</td>
<td>9.8**</td>
</tr>
<tr>
<td>PMI-based</td>
<td></td>
<td>15.8**</td>
<td>9.6</td>
<td>7.3</td>
</tr>
<tr>
<td>BWV</td>
<td></td>
<td>17.8**</td>
<td>9.4</td>
<td>6.7</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PMI-Hash</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial</td>
<td></td>
<td>12.1</td>
<td>17.0</td>
<td>11.5</td>
</tr>
<tr>
<td>mcl-MNB</td>
<td></td>
<td>22.8**</td>
<td>15.9**</td>
<td>13.1**</td>
</tr>
<tr>
<td>mcl-LogReg</td>
<td></td>
<td>14.4**</td>
<td>18.7**</td>
<td>14.8**</td>
</tr>
<tr>
<td>1vR-MNB</td>
<td></td>
<td>19.9**</td>
<td>16.7</td>
<td>14.2**</td>
</tr>
<tr>
<td>1vR-LogReg</td>
<td></td>
<td>17.6**</td>
<td>18.9**</td>
<td>16.2**</td>
</tr>
<tr>
<td>PMI-based</td>
<td></td>
<td>22.3**</td>
<td>15.6**</td>
<td>14.4**</td>
</tr>
<tr>
<td>BWV</td>
<td></td>
<td>29.3**</td>
<td>15.5**</td>
<td>13.1**</td>
</tr>
</tbody>
</table>

All performance scores are percentages. The results of learned classifiers are compared with those of corresponding initial classifiers. One asterisk * indicates a p-value ≤ 0.05; two asterisks ** indicate a p-value ≤ 0.01.
Lessons Learned for Emotion Recognition in Tweets

- Distant learning is a viable approach to build emotion classifiers across domains
- Including pseudo-neutral documents avoids over-classifying emotional content
- Can be applied to dialogs, food & mood data, etc.
Supervised:


Semi-supervised using emoticons and hashtags:


Distant learning without labels and structural data:

Modifiers: emotion shifts
Examples

I am not ashamed: the emotion has shifted from shame to pride
I feel so relieved now: intensifier to increase the degree of relief
I feel a little sad: it diminishes the degree of sad
I know i should be happy: shift from happy to sad/regret
I’ll be sad if you leave: fear for event that may happen
Do you love her? interest/involvement, anger
I was happy then: disappointment/regret
Novelty

- treat 6 modifiers simultaneously
- data-driven method
- re-mapping
1. Collect tweets with emotional hashtags

2. Detect lexicon emotional terms and their modifiers

(a) I am **happy** you are here #joy
TERM EMOTION DETECTED MODIFIER HASHTAG EMOTION
Happiness No modifier Happiness
(b) *Not ashamed* to admit it #proud
Shame Negation Pride
(c) I **love** you so much #love
Love Intensifier Love

3. Aggregate distributions of hashtag emotions for each term emotion and modifier class

<table>
<thead>
<tr>
<th>Hashtag Emotions</th>
<th>No modifier</th>
<th>Negation</th>
<th>Intensifier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Happiness</td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>Love</td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
</tr>
<tr>
<td>Sadness</td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
<td><img src="image" alt="Graph" /></td>
</tr>
</tbody>
</table>
Kullback-Leibler divergence $= 1.96$

Smallest KL divergence to $(0.24)$
Shifts of Emotions under Negation

- Happiness
  - Pleasure
    - Regret
      - Involvement
        - Amusement
  - Awe
    - Shame
      - Pride
        - Anger
          - Guilt
            - Envy
  - Surprise
    - Love
      - Sadness
        - Worry
      - Nostalgia
        - 0.53
          - 0.35
            - 0.23
## All 6 Modifiers

<table>
<thead>
<tr>
<th>Modifier</th>
<th>Modified to non-modified average distance</th>
<th>Average Certainty Coefficient</th>
<th>% of Emotion Shift</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensifiers</td>
<td>0.14</td>
<td>1.32</td>
<td>17%</td>
</tr>
<tr>
<td>Past Tense</td>
<td>0.17</td>
<td>0.75</td>
<td>6%</td>
</tr>
<tr>
<td>Modality</td>
<td>0.19</td>
<td>0.74</td>
<td>19%</td>
</tr>
<tr>
<td>Conditionality</td>
<td>0.27</td>
<td>0.82</td>
<td>36%</td>
</tr>
<tr>
<td>Diminishers</td>
<td>0.30</td>
<td>1.17</td>
<td>38%</td>
</tr>
<tr>
<td>Interrogation</td>
<td>0.41</td>
<td>1.51</td>
<td>53%</td>
</tr>
<tr>
<td>Negation</td>
<td>0.80</td>
<td>0.56</td>
<td>75%</td>
</tr>
</tbody>
</table>
References of our work

• **OlympLex**
  Sintsova, Valentina, Claudiu-Cristian Musat, and Pearl Pu. "Fine-Grained Emotion Recognition in Olympic Tweets based on Human Computation." In 4th Workshop on Computational Approaches to Subjectivity, Sentiment and Social Media Analysis (WASSA), 2013.

• **Dystemo**

• **Modifiers Analysis**

• **Incentive Schemes for Inexpert Workers**
Conclusion
EmotionWatch Video

http://ijcai13.org/video/05
Thank You!