Signals from Text

Stephen Pulman

Dept. of Computer Science, Oxford University stephen.pulman@cs.ox.ac.uk
and TheySay Ltd, www.theysay.io

@sgpulman

March 4, 2016
Signals from text: sentiment, emotion, etc.

Overview
Sentiment analysis proper: classify text into positive, negative or neutral. But there are many other signals we can get from text:

- Future intent, risk...
- Demographic profile: age, gender, politics, religion...
- Deception: how well can we tell if someone is lying?
- Emotion: joy, sadness, fear, anger...
Linguistically based pattern matching (information extraction) e.g.

<PERSON> resign/be (sacked/fired/moved) from <COMPANY-ROLE>

Machine learning methods: train a classifier using annotated data: Naive Bayes, Support Vector Machine, Averaged Perceptron, Neural Networks etc.

Currently fashionable: Deep Learning methods - Convolutional Neural Net, Long Short Term Memory models etc.

Choice usually depends on the availability of annotated data: expensive and time-consuming to acquire.
What is sentiment analysis?

The term has been around since 2000ish, and has been used to cover a variety of different phenomena:

**Sentiment proper**

Positive, **negative**, or neutral attitudes expressed in text:

*Suffice to say, Skyfall is one of the best Bonds in the 50-year history of moviedom’s most successful franchise.*

*Skyfall abounds with bum notes and unfortunate compromises.*

*There is a breach of MI6. 007 has to catch the rogue agent.*
Building a sentiment analysis system

- Version 1: (cheap and cheerful) collect lists of positive and negative words, classify a text based on proportion of pos/neg.
- Version 2: (what most commercial systems do) get a training corpus of texts human annotated for sentiment (e.g. pos/neg/neut); represent each text as a vector of counts of words or successive pairs of words; train your favourite classifier on these vectors

Problems:
- if number of positive = number of negatives?
- bag-of-words means structure is ignored:
  “Airbus: orders slump but profits rise” wrongly = “Airbus: orders rise but profits slump”
- **Compositional** effects will be missed:
  - clever, too clever, not too clever
  - bacteria, kill, kill bacteria, fail to kill bacteria, never fail to kill bacteria
Version 3: use linguistic analysis

- do as full a parse as possible on input texts
- use the syntax to do ‘compositional’ sentiment analysis:

```
S
   /\              
  NP    VP
     /\     /\     
    Adv   VP  VP
       /\  /\  /\  
      never V  V  NP
       /\  /\  /\  
      fails to kill bacteria
```
Sentiment logic rules

- \( \text{kill} + \text{negative} \rightarrow \text{positive} \) (kill bacteria)
- \( \text{kill} + \text{positive} \rightarrow \text{negative} \) (kill kittens)
- \( \text{kill} + \text{neutral} \rightarrow \text{neutral} \) (kill time)
- \( \text{too} + \text{anything} \rightarrow \text{negative} \) (too clever, too red, too cheap)
- etc. In our system (www.theysay.io) we have 75,000+ rules...

Problems:

- still need extra work for context-dependence (‘cold’, ‘wicked’, ‘sick’...)
- can’t deal with reader perspective: “Oil prices are down” is good for me, not for investors.
- can’t deal with sarcasm or irony: “Oh, great, they want it to run on Windows”
Intent and Risk detection

- Few corpora: mismatch between small number of risk/intent sentences and very large number of non-risk/intent sentence.
- But we can treat it as an information extraction problem, by looking out for patterns like:
  ...hopes/intends/plans/believes that...
  ...risk/likelihood/threat/promise/danger/intention/hope that...

- Building such a system involves a certain amount of trial and error, tuning the patterns to get good recall and minimise false positives.
- Why would you want to recognise and capture expressions of intent, and risk?
Intent and Risk

- detection of future predictions or commitments in financial reports (Wikipedia):

  ‘In United States business law, a forward-looking statement is (one) that cannot sustain itself as merely a historical fact... predicts, projects, or uses future events as expectations or possibilities.’

  Company X plans to build a new facility ...
  We anticipate the value to continue to rise...

- Company reports often show high positive sentiment, but this should probably be discounted if they also show high forward looking language.

- Get early warnings of customer actions, particularly ‘intent to churn’ signals in blogs or CRM messages:
  “Terrible service... Paypal should take responsibility for accounts which have been hacked into ... Very disappointed and will never use Paypal again.”
Recognising deception in text

- What is deception?
  A believes X to be false.
  A causes B to (continue to) believe X via Y where B takes Y as evidence for X.

- Difficult to detect in text, because vocal, eye gaze and posture cues are absent, and these have proved quite reliable in e.g. border security checks.

- Can we train a classifier to recognise deceptive language?
  Obviously impossible in the general case because truth depends on non-linguistic facts, but perhaps there are characteristics of deceptive or truthful language use?
Linguistic characteristics of deceptive vs. truthful text

Zhou, Burgoon et al (2004) found that deceivers sent longer messages, were more informal, showed less lexical diversity, used fewer ‘I, me’, more ‘we, us’, more positive sentiment than truthtellers. Other studies (Hauch et al (2012)) confirm that:

- Liars tend to use more emotion words, more negation words, and more motion verbs (lead, go).
- Truthtellers use more self and other references, exclusive words (except, but, without), tentative words and time related words.
- Possible that liars exaggerate certainty and positive/negative aspects, and do not associate themselves closely with the content...
- ...whereas truthtellers are more cautious, accept alternatives, and do associate themselves with the content.
Financial applications: lying CEOs

- Larcker and Zakolyukina (2012)\(^1\) looked at the language used by CEOs and CFOs talking to analysts in conference calls about earnings announcements.

- Looking at subsequent events:
  - discovery of ‘accounting irregularities’
  - restatements of earnings
  - changes of accountants
  - exit of CEO and/or CFO

- you can identify retrospectively who was telling the truth or not.

- This gives us a corpus of transcripts which can be labelled as ‘true’ or ‘deceptive’: training data.

---

\(^1\)D. F. Larcker, A. A. Zakolyukina, *Detecting Deceptive Discussions in Conference Calls*, Journal of Accounting Research, Volume 50:2, 495-540
Detecting deception

- Training a classifier on features like those just described, Larcker and Zakolyukina were able to get up to 66% prediction accuracy.
- Building a portfolio of deceptive companies will lose you 4 to 11% per annum...
- We (TheySay) have build a general ‘verisimilitude’ classifier which looks out for linguistic indicators of deception...
- ... and also measures clarity and readability (vs. obfuscation, hedging, etc.)
- We tried this on speeches by the 4 UK party leaders prior to the 2015 election.
- Guess what...
Emotion detection

A confusing variety of different theories of emotional state: Johnson-Laird and Oatley (1989)\(^2\) distinguish five **basic** emotions: **anger, disgust, fear, happiness, and sadness**. All others are analysed as:

- generic: e.g. feeling, affect
- relational: e.g. abhor, love
- caused: e.g. afraid, amused
- causative: e.g. affront, offend
- goal: e.g. covet, curious
- complex: e.g. ashamed, assured

\(^2\)http://goo.gl/chf5yW
The Ekman classification

- **Ekman**: anger, disgust, fear, happiness, sadness PLUS surprise.\(^3\)
- Seems to be a correspondence with facial expressions, and emoticons:

![Facial expressions and emoticons](http://www.paulekman.com/wp-content/uploads/2013/07/An-Argument-For-Basic-Emotions.pdf)
Universality of emotion classes

- But whereas Japanese and US subjects agree on classification of expressions of **happiness, surprise, and sadness**...
- ... agreement is significantly lower for **anger, disgust and fear**
- Is this emotion taxonomy universal? Possibly not - in Ifaluk, spoken on a small group of islands in the Pacific:
  - “...*fago* is felt when someone dies, is needy, is ill, or goes on a voyage...”
- So *fago* looks like **sadness** so far. But “…*fago* is also felt when in the presence of someone admirable or when given a gift.”
Emotion classification is difficult

**Emotion:** human annotation usually gives an upper limit on what is possible. For Ekman emotion labels the results are not consistent across emotions: this table\(^4\) shows level of agreement between 6 annotators on 1250 news headlines drawn from major newspapers such as New York Times, CNN, and BBC News, as well as from the Google News search engine.

<table>
<thead>
<tr>
<th>Emotions</th>
<th>human agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>anger</td>
<td>49.55</td>
</tr>
<tr>
<td>disgust</td>
<td>44.51</td>
</tr>
<tr>
<td>fear</td>
<td>63.81</td>
</tr>
<tr>
<td>joy</td>
<td>59.91</td>
</tr>
<tr>
<td>sadness</td>
<td>68.19</td>
</tr>
<tr>
<td>surprise</td>
<td>36.07</td>
</tr>
</tbody>
</table>

(Pearson r correlation, here -100 to +100)

\(^4\)web.eecs.umich.edu/~mihalcea/papers/strapparava.acm08.pdf
Emotion in social media

Supervised vs unsupervised learning: annotated vs unannotated data. Getting large amounts of high quality annotated data is expensive. Purver and Battersby\textsuperscript{5} experimented with ‘distant supervision’, i.e. using indirect proxies for annotations. In tweets: use hashtags and emoticons as emotion labels:

Best day in ages! #Happy :)
Gets so #angry when tutors don’t email back...
Do you job idiots!

Lexical clues often coincide, as above, but sometimes only the hashtag disambiguates:

Still trying to recover from seeing the #bluewaffle on my TL #disgusted #sick

\textsuperscript{5}http://www.aclweb.org/anthology/E/E12/E12-1049.pdf
Using hashtags and emoticons

happy

sad

anger

fear

surprise

disgust

happy  #happy  #happiness
sad  #sad  #sadness
anger  #angry  #anger
fear  #scared  #fear
surprise  #surprised  #surprise
disgust  #disgusted  #disgust

Note that for many emoticons there is poor agreement on their interpretation.
Hashtag vs. emoticon

(Unigram SVM, 10-fold cross validation) Best results if some human annotation on 1k tweets supplements:

<table>
<thead>
<tr>
<th>Train</th>
<th>Class</th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>emoticon</td>
<td>happy</td>
<td>79.4%</td>
<td>75.6%</td>
<td>77.5%</td>
</tr>
<tr>
<td>emoticon</td>
<td>sad</td>
<td>43.5%</td>
<td>73.2%</td>
<td>54.5%</td>
</tr>
<tr>
<td>emoticon</td>
<td>anger</td>
<td>62.2%</td>
<td>37.3%</td>
<td>46.7%</td>
</tr>
<tr>
<td>emoticon</td>
<td>fear</td>
<td>6.8%</td>
<td>63.6%</td>
<td>12.3%</td>
</tr>
<tr>
<td>emoticon</td>
<td>surprise</td>
<td>15.0%</td>
<td>90.0%</td>
<td>25.7%</td>
</tr>
<tr>
<td>emoticon</td>
<td>disgust</td>
<td>8.3%</td>
<td>25.0%</td>
<td>12.5%</td>
</tr>
<tr>
<td>hashtag</td>
<td>happy</td>
<td>78.9%</td>
<td>51.9%</td>
<td>62.6%</td>
</tr>
<tr>
<td>hashtag</td>
<td>sad</td>
<td>47.9%</td>
<td>81.7%</td>
<td>60.4%</td>
</tr>
<tr>
<td>hashtag</td>
<td>anger</td>
<td>58.2%</td>
<td>76.0%</td>
<td>65.9%</td>
</tr>
<tr>
<td>hashtag</td>
<td>fear</td>
<td>10.1%</td>
<td>81.8%</td>
<td>18.0%</td>
</tr>
<tr>
<td>hashtag</td>
<td>surprise</td>
<td>5.9%</td>
<td>60.0%</td>
<td>10.7%</td>
</tr>
<tr>
<td>hashtag</td>
<td>disgust</td>
<td>6.7%</td>
<td>66.7%</td>
<td>11.8%</td>
</tr>
</tbody>
</table>
Distant supervision helps, but...

- Hashtags are useful sources of ‘distant supervision’, emoticons less so. Better if some human annotation too.
- Text classification accuracy highest on happy, sad, anger (both human and machine)
- Note different results for fear category than earlier work.
- But this kind of ‘distant supervision’ seems promising: in Oxford (TheySay) we have used distant supervision and human annotated data to build a multi-dimensional emotion classifier with acceptable accuracy...
- ... and apparently useful applications in finance, health care monitoring and in politics.
Applications in investment

Markets are driven partly by emotion: Bollen\textsuperscript{6} claimed that ‘calmness’ predicted Dow Jones Industrial Index (much disputed!)

Example from Richard Peterson (MarketPsych): fear and surprise:

- 25/04/2009 WHO declares Swine Flu “public health emergency”
- Anxiety and fear about air travel on blogs etc
- 27/04/2009 WHO pandemic alert
- More fear, selloff of airline stocks
- 30/04/2009 Buy American Airline at $4.81
- 06/05/2009 Sell at $5.95

Peterson has claimed that different asset classes respond to different emotions.

\textsuperscript{6}Johan Bollen, Huina Mao, and Xiao-Jun Zeng. Twitter mood predicts the stock market. Journal of Computational Science, 2(1), March 2011, Pages 1-8
What do the voters think?

We can detect which topics arouse strong emotions among the voters:

**Candidate Jones:**

<table>
<thead>
<tr>
<th>Topic</th>
<th>Economy</th>
<th>Immigration</th>
<th>Education</th>
<th>Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>fear</td>
<td>8</td>
<td>6</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>anger</td>
<td>2</td>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>happiness</td>
<td>1</td>
<td>1</td>
<td>5</td>
<td>6</td>
</tr>
</tbody>
</table>

**Candidate Smith:**

<table>
<thead>
<tr>
<th>Topic</th>
<th>Economy</th>
<th>Immigration</th>
<th>Education</th>
<th>Health</th>
</tr>
</thead>
<tbody>
<tr>
<td>fear</td>
<td>2</td>
<td>7</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>anger</td>
<td>1</td>
<td>9</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>happiness</td>
<td>7</td>
<td>1</td>
<td>5</td>
<td>3</td>
</tr>
</tbody>
</table>
Conclusions

- We can detect several different types of signals in text: sentiment, intent, risk, humour, as well as demographic properties like gender, age, political or religious orientation.

- But we need larger, well annotated corpora to make progress here, particularly with other dimensions like deception and emotion.

- Emotion classification gives a finer grained analysis of opinion, and more insight and explanation than traditional sentiment analysis.

- Deception and, more generally, “information quality” is a potentially very powerful tool.

Try our web demo: apidemo.theysay.io