Words with Attitude

Jaap Kamps

Maarten Marx

Paper's Goal

- Judge the emotive or affective meaning of a text
- Use WordNet to determine values of words with Osgood's semantic differential technique

Osgood's Semantic Differential Technique

- Judge words, phrases, texts by asking subjects to rate them on scales of bipolar adjectives
- A subject might be asked to rate "proper" on scales like optimistic-pessimistic, serious-humorous, and active-passive.
- It turns out that good-bad, strong-weak, and active-passive values account for most variance in judgment

Using WordNet with Osgood's theory

- Authors want to get values for words from WordNet
- They define MPL(w_1,w_2) as the minimal path length between w_1 and w_2, using only same-synset relations
- Allowing more than just same-synset damages metric

MPL Examples

- MPL(good, proper) = 2
  - (good,right,proper)
- MPL(good, neat) = 3
- MPL(good, noble) = 4
- Can we use this to rate “proper”, “neat”, and “noble” on a good-bad scale?

MPL

- MPL(good, bad) = 4
- If we just look at MPLs, “noble” is as good as “bad”
- We need to do something a bit more complicated

TRI

$$TRI(w_i; w_j, w_k) = \frac{MPL(w_i, w_k) - MPL(w_i, w_j)}{MPL(w_k, w_j)}$$

- To determine the good-bad ("evaluative") value of w_i, examine TRI(w_i;good,bad)
- Define EVA(w) = TRI(w;good,bad)

EVA results

$$EVA(proper) = TRI(proper; good, bad) = \frac{MPL(proper, bad) - MPL(proper, good)}{MPL(good, bad)} = \frac{6-2}{4} = 1$$

$$EVA(neat) = \frac{3-3}{4} = 0$$

$$EVA(noble) = \frac{5-4}{4} = 0.25$$

$$EVA(good) = \frac{4-0}{4} = 1$$

$$EVA(bad) = \frac{0-4}{4} = -1$$

- There are 5410 adjectives linked to “good” or “bad”.
- Average value of EVA for these 5410 words is –0.0089

Other scales

- Define POT as TRI(w;strong,weak)
- Define ACT as TRI(w;active,passive)
- EVA, POT, ACT are well-defined for exactly the same set of 5410 adjectives.

EVA*, POT*, ACT*

- Define EVA*(w) to be EVA(w) if a path exists between w and “good”, and 0 if it doesn’t
- This gives us a well-defined function for all w
- Do the same thing to get POT* and ACT*

Application

- We can now take the sum of EVA*, POT*, ACT* for all words in a text to get an idea of the good-bad, strong-weak, active-passive values for the text as a whole

Accuracy

- No corpus existed that had already been rated for these values, so accuracy could not be tested on a large scale
- Tests on small numbers of Internet discussions show correspondence between results of this method and actual value of texts, but questionable accuracy for short texts
- Works better for long texts

Accuracy problems

- With longer texts, false positives and false negatives cancel each other out; doesn’t help for shorter texts
- Longer texts yield scores of higher magnitude, in general – need to normalize scores
- Apparent bias to positive words (positive opinions more extensively elaborated, affecting a text’s score more than negative opinions)

Author's closing notes

- Authors of texts on Internet discussion sites must be less subtle about good/bad
- Little NLP research addresses subjective aspects; this paperhelps fill the gap