

Stance Classification Using Semantic Features in an Online Debate Corpus

C.J Hsu and Ryan Bielby

Outline

- Motivation
- Reference Papers
- Corpus and Lexicon
- Methods
- Findings
 - Argument Features like appendix
 - Semantics as features vs Semantics as filters
- Future work (Linguistic Analysis)

Motivation

- In this class, we studied almost 20 papers which address different domains, but all of them have some aspects in common.
- We are tired of this SOP; is there any other way to facilitate the information of semantics?

Reference Papers

Somasundaran, Swapna and Janyce Wiebe.
2010. "Recognizing Stances in Ideological On-Line Debates".

- Constructs an "Arguing Lexicon" from MPQA to predict the stances of the online debate posts

Wilson, et al. 2005. "Recognizing Contextual Polarity in Phrase-Level Sentiment Analysis".

- Provides Subjectivity Lexicon built from news articles

Corpus

- Somasundaran and Wiebe paper provides six categories of online debate posts: abortion, creation, guns, gay rights, god, and healthcare
- Total number of posts: 3167
- Although these posts belong to different categories, the authors have unified the class label (positive and negative argument) for different categories. We could merge all these six categories as single one.

Building an Arguing Lexicon

- In the MPQA corpus, some text spans are marked with `attitude-type="arguing-pos"` or `attitude-type="arguing-neg"`
- Generate the unigrams, bigrams, and trigrams starting at the text spans which are marked `"arguing-pos"` or `"arguing-neg"`

Building an Arguing Lexicon

- Remove these n-grams which are already presented in Subjective Lexicon
- Calculate the two conditional probabilities for each entry
 - i.e., $P(\text{type} = \text{"arguing-pos"} \mid \text{n-gram})$ and $P(\text{type} = \text{"arguing-neg"} \mid \text{n-gram})$.

Subjectivity Lexicon

- 15,991 subjective expressions from 425 docs
 - devset: 66 docs, 2808 subjective expressions
 - 10-fold cross-validation: 359 docs, 7611 expressions
- Prior-polarity subjectivity lexicon (8,000 words)
 - Riloff and Wiebe, 2003
 - Hatzivassiloglou and McKeown, 1997
 - General Inquirer, 2000
 - Reliability tags: *strongsubj* and *weaksubj*
 - 33.1% *positive*, 59.7% *negative*, 0.3% *both*, 6.9% *neutral*

Method: Unigram vs Arguing

- **Classifier:** Chose SVM over MaxEnt
- **Unigram Features:** unigram, non-stemmed; negate the unigram which appears after negator
- **Arguing Features:** Break each sentence from each post into trigrams, bigrams, unigrams; check if n-gram (starting with trigrams) is in arguing lexicon; find 'overall' sentiment (sentiment with greatest # arguing features); mark each word (sans stop words) in the sentence as such; e.g., nobody_neg thinks_neg

Methods Results

- Results:
 - Unigrams: 10-fold, accuracy: 62.2198%
 - Arguing: 10-fold, accuracy: 58.236%

Findings: Is an Arguing Lexicon Useful?

- Arguing Lexicon is like human appendix

Domain (#posts)	Distribution	Unigram	Sentiment	Arguing	Arg+Sent
Overall (2232)	50	62.50	55.02	62.59	63.93
Guns Rights (306)	50	66.67	58.82	69.28	70.59
Gay Rights (846)	50	61.70	52.84	62.05	63.71
Abortion (550)	50	59.1	54.73	59.46	60.55
Creationism (530)	50	64.91	56.60	62.83	63.96

Table 4: Accuracy of the different systems

- Almost 60% of the entries in "arguing-negative" have the token "not"
- Once we negate the word appearing after negator, the unigram feature could almost capture the essence just as the arguing feature does

Findings

Sentiment as Feature vs Sentiment as Filter

- Given a semantic lexicon, building the semantic features by counting and voting seems become a SOP in this field
- We think some online posts are suitable for this shallow processing based framework, however, some posts are not
- Could we identify those posts which are not suitable for this framework and perform additional analysis on them?

Findings

Sentiment as Feature vs Sentiment as Filter

- 817 posts have no any clue word of semantic lexicon and 2874 posts have at least one clue word of semantic lexicon
- The result of 10 fold C.V on the 2874 posts by unigram features is 59.53%
- The result of 10 fold C.V on those 817 posts by unigram features is 52.02%
- These posts have no pattern at all in unigram features! What causes this?

Findings

Three categories for the 817 posts

- **Response:** this type of post does not propose any significant supporting points; just tries to deny others' points.
 - e.g., "You should spend more time thinking about what you say before you type ."
- **A/V response:** people are lazy and just post YouTube or other URL to argument their point.
 - e.g., "<http://americansfortruth.com/issues/the-agenda-glbtc-activist-groups/national-glbtc-activist-groups/sisters-of-perpetual-indulgence/page/2>"

Findings

Three categories for the 817 posts

- **Negated Negatives:** author negates negative terms, but then alludes that they are true.
 - e.g., "Mark, you're not an asshole. You're just trying so hard to be!"

Future Work

- For any semantics application, a two-stage framework deserves a try!
- Identify those sentences which are not suitable for shallow processing.
- Incorporate audio and video sentiment analysis to complement the text analysis.
- The semantic lexicons are mostly built on newspapers! They do not have slang words and other casual speech.