HW4

Rahul Huilgol (rrh2226)

1 Reviewing Faulkner 2014

1.1 Introduction

In this report, I discuss the work on 'Automated classification of stance in student essays: An approach using stance target information and the Wikipedia link-based measure' by Adam Faulkner [1]. Stance classification is the problem of automatically identifying from text whether the stance of the author towards some target or proposition is for, against or neither. Faulkner looks at the subtask of document-level stance detection, where he studies argumentative student essays and tries to determine the stance of student towards the target prompt.

Stance taking is an essential part of online debates, social media interaction, opinion news, etc. Building systems that automatically detect and understand the stance of the author has numerous applications like information retrieval, text summarization, textual entailment, identify social groups, recommendations, etc.

Faulkner used the International Corpus of Learner English (2003), a collection of argumentative essays written by non-native speakers of English. This comprised of around 1300 essays each responding to one of 7 prompts on topics like television, feminism, science, etc. These essays were then annotated using the crowdsourcing platform Crowdflower.

1.2 Methodology

Faulkner builds two sets of features which are the strong points of his appraoch. These features are then input to a classifier.

Part of speech generalized dependency subtrees

Faulkner builds a stance lexicon to identify phrases that are indicative of stance. It comprises of two lexicons. The first is a lexicon of arguing words each assigned a probability of polarity, *for* or *against*. This lexicon comprises 2166 *for* and 513 *against* ngrams and was built from MPQA corpus annotate for stance [3]. The second is a selection of metadiscourse markers [2], taken from the following categories: boosters (clearly, decidedly), hedges(claim, estimate), and engagement markers(demonstrate, evaluate). This gave a lexicon of 373 for and 80 against unigrams.

Targets are the concepts on which the author takes a stance. To capture the targets of the stancetaking language, the following approach was used. The dependency parse of each sentence is created using the Stanford parser. Stance taking and opinion bearing language in the parse is located using the stance lexicon above and the MPQA subjectivity lexicon. If the immediate neighboring node of a stance word node contains a negator, the polarity of stance word is reversed by appending *not* to the word. Now, starting at the stance word node, undirected version of dependency tree is traversed in breadth-first manner until a node containing an opinion-bearing word is found. The phrase structure parse is examined to see if the opinion bearing word is located in the immediate or embedded clause of the stance word. If so, this opinion bearing word is considered a good proxy for the proposition targeted by the stance word. The subtree containing the stance and opinion bearing nodes is returned. The midnodes are then POS generalized. For example, consider the sentence *'So we can infer that the statement is very true'*. *Can* is a stance word in the lexicon, *true* is a positive opinion word. The subtree between *can* and *true* turns out to be *can infer true*. After POS generalizing this is converted to a feature*can-V-true*, where V stands for verb. An unordered set of POS-generalized stance proposition subtrees is used as first set of features.

Relationship between words in prompt and essay

The second set of features captures the relationship between languages of prompt and essay. This tries to account for the differences in the words used to refer to similar concepts. For example consider a prompt which talks about money. If the student uses words like rich and poor, the model should consider that rich is associated to money. To capture these associations, Faulkner uses a Wikipedia link-based measure (WLM). Wikipedia pages contain large network of cross-references to external and internal hyperlinks. WLM is defined in [4] as where a and b are Wikipedia article titles, A and B

$$wlm(a,b) = \frac{log(max(|A|,|B|) - log(|A \cap B|)}{log(|W|) - log(min(|A|,|B|))}$$

are sets of articles that backlink to a and b, W is the count of all articles in Wikipedia. $wlm(a,b) = 0$ imples a and b are semantically similar. $wlm(a,b) \geq 1$ means they are semantically dissimilar. The phrase structure representations created above were used to identify propositions in the immediate or embedded clause of the stance word. The WLM scores of content words in prompt with content words in proposition were calculated. Each topic word is mapped to the most similar prompt word. An unordered collection of stemmed WLM-scored topic words is the second set of features.

2 Relation to our project

The aim of our project is to detect the stance of author in tweets towards some target. Given a tweet and a target, the goal is to identify whether the tweet is in favor of the target, against the target or neither. This work discussed above was one of the few papers closest to this goal, but the primary difference from our project is that it tries to detect stance of a long essay rather than short text. Till date, there has been no work on stance detection of short text. Tweets are challenging because they are short, informal, have mispellings and slang.

Despite this major difference, Faulkner's work has several interesting ideas that can be employed for tweets as well. Most important is how he used WLM score to determine the association of target word and the words in essay. This technique would be really useful for tweets as well. Since a tweet is short, such topic features can capture subtleties in variation of the words used to refer to the target or words topically related to the target. Even in his results, this set of features has significantly higher contribution to both precision and recall than the other set of features.

Secondly, his method of using the dependency parse to identify stancetaking, opinion bearing and proposition words is interesting. This however may not be very useful because tweets lack proper grammatical structure. But a dependency parse does help identify if there is a negator that affects the stance, in which case the stance can be reversed. It may also help identify the target.

Our model would have additional aspects. Since we also need to determine the case where the tweet does not talk about the target, we plan to have a classifier that tries to predict whether the tweet refers to the target. Apart from ngrams constructed from the tweet this classifier would greatly benefit from the WLM scores based features. We also plan to assign a sentiment to the tweet using a sentiment lexicon. The intuition behind this is that, if the tweet mentions the target, does not have a negator, and has positive sentiment, it is likely to be in favor of the target. We hope to build an ensemble of the above or stack different classifiers on top of one another.

3 Conclusion

Faulkner's work has valuable lessons even when we change the domain of stance classification from essays to tweets. I tried to explain the main ideas behind his work, how closely it relates to our project on tweets and how various ideas can be used in our project.

References

[1] Adam Faulkner. Automated classification of stance in student essays: An approach using stance target information and the wikipedia link-based measure. In *The Twenty-Seventh International Flairs Conference*, 2014.

[2] Gavin Melles and KEN HYLAND. Metadiscourse: Exploring interaction in writing, 2008.

[3] Swapna Somasundaran and Janyce Wiebe. Recognizing stances in ideological on-line debates. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 116–124. Association for Computational Linguistics, 2010.

[4] Ian Witten. An effective, low-cost measure of semantic relatedness obtained from wikipedia links. 2008.