Genre Independent Subgroup Detection in Online Discussion Threads: A Pilot Study of Implicit Attitude using Latent Textual Semantics

Pradeep Dasigi
pd2359@columbia.edu

Weiwei Guo
weiwei@cs.columbia.edu

Mona Diab
mdiab@ccls.columbia.edu

Center for Computational Learning Systems, Columbia University

Abstract

We describe an unsupervised approach to the problem of automatically detecting subgroups of people holding similar opinions in a discussion thread. An intuitive way of identifying this is to detect the attitudes of discussants towards each other or named entities or topics mentioned in the discussion. Sentiment tags play an important role in this detection, but we also note another dimension to the detection of people's attitudes in a discussion: if two persons share the same opinion, they tend to use similar language content. We consider the latter to be an implicit attitude. In this paper, we investigate the impact of implicit and explicit attitude in two genres of social media discussion data, more formal wikipedia discussions and a debate discussion forum that is much more informal. Experimental results strongly suggest that implicit attitude is an important complement for explicit attitudes (expressed via sentiment) and it can improve the sub-group detection performance independent of genre.

1 Introduction

There has been a significant increase in discussion forum data in online media recently. Most of such discussion threads have a clear debate component in them with varying levels of formality. Automatically identifying the groups of discussants with similar attitudes, or subgroup detection, is an interesting problem which allows for a better understanding of the data in this genre in a manner that could directly benefit Opinion Mining research as well as Community Mining from Social Networks.

A straight-forward approach to this problem is to apply Opinion Mining techniques, and extract each discussant's attitudes towards other discussants and entities being discussed. But the challenge is that Opinion Mining is not mature enough to extract all the correct opinions of discussants. In addition, without domain knowledge, using unsupervised techniques to do this is quite challenging.

On observing interactions from these threads, we believe that there is another dimension of attitude which is expressed implicitly. We find that people sharing the same opinion tend to speak about the same topics even though they do not explicitly express their sentiment. We refer to this as *Implicit Attitude*. One such example may be seen in the two posts in Table 1. It can be seen that even though discussants *A* and *B* do not express explicit sentiments, they hold similar views. Hence it can be said that there is an agreement in their implicit attitudes.

Attempting to find a surface level word similarity between posts of two discussants is not sufficient as there are typically few overlapping words shared among the posts. This is quite significant a problem especially given the relative short context of posts. Accordingly, in this work, we attempt to model the implicit latent similarity between posts as a means of identifying the implicit attitudes among discussants. We apply variants on Latent Dirichelet Allocation (LDA) based topic models to the problem (Blei et al., 2003).

Our goal is identify subgroups with respect to discussants' attitudes towards each other, the entities and topics in a discussion forum. To our knowledge, this is the first attempt at using text similarity as an indication of user attitudes. We investigate the influence of the explicit and implicit attitudes on two genres of data, one more formal than the other. We find an interesting trend. Explicit attitude alone

65

as a feature is more useful than implicit attitude in identifying sub-groups in informal data. But in the case of formal data, implicit attitude yields better results. This may be due to the fact that in informal data, strong subjective opinions about entities/events or towards other discussants are expressed more explicitly. This is generally not the case in the formal genre where ideas do not have as much sentiment associated with them, and hence the opinions are more "implicit". Finally, we observe that combining both kinds of features improves performance of our systems for both genres.

2 Related Work

Substantial research exists in the fields of Opinion Identification and Community Mining that is related to our current work. (Ganapathibhotla and Liu, 2008) deal with the problem of finding opinions from comparative sentences. Many previous research efforts related to Opinion Target Identification (Hu and Liu, 2004; Kobayashi et al., 2007; Jakob and Gurevych, 2010), focus on the domain of product reviews where they exploit the genre in multiple ways. Somasundaran and Wiebe (2009) used unsupervised methods to identify stances in online debates. They mine the web to find associations indicative of opinions and combine them with discourse information. Their problem essentially deals with the debate genre and finding the stance of an individual given two options. Ours is a more general problem since we deal with discussion data in general and not debates on specific topics. Hence our aim is to identify multiple groups, not just two.

In terms of Sentiment Analysis, the work done by Hassan et al.(2010) in using part-of-speech and dependency structures to identify polarities of attitudes is similar to our work. But they predict binary polarities in attitudes, and our goal of identification of sub-groups is a more general problem in that we aim at identifying multiple subgroups.

3 Approach

We tackle the problem using Vector Space Modeling techniques to represent the discussion threads. Each vector represents a discussant in the thread creating an *Attitude Profile* (AP). We use a clustering algorithm to partition the vector space of APs into multiple sub-groups. The idea is that resulting clusters would comprise sub-groups of discussants with similar attitudes.

3.1 Basic Features

We use two basic features, namely Negative and Positive sentiment towards specific discussants and entities like in the work done by (Abu-Jbara et al., 2012). We start off by determining sentences that express attitude in the thread, attitude sentences (AS). We use OpinionFinder (Wilson et al., 2005) which employs negative and positive polarity cues. For determining discussant sentiment, we need to first identify who the target of their sentiment is: another discussant, or an entity, where an entity could be a topic or a person not participating in the discussion. **Sentiment toward another discussant:** This is quite challenging since explicit sentiment expressed in a post is not necessarily directed towards another discussant to whom it is a reply. It is possible that a discussant may be replying to another poster but expressing an attitude towards a third entity or discussant. However as a simplifying assumption, similar to the work of (Hassan et al., 2010), we adopt the view that replies in the sentences that are determined to be attitudinal and contain second-person pronouns (you, your, yourself) are assumed to be directed towards the recipients of the replies. **Sentiment toward an entity:** We again adopt a simplifying view by modeling all the named entities in a sentence without heeding the roles these entities play, i.e. whether they are targets or not. Accordingly, we extract all the named entities in a sentence using Stanford's Name Entity Recognizer (Finkel et al., 2005). We only focus on Person and Organization named entities.

3.2 Extracting Implicit Attitudes

We define implicit attitudes as the semantic similarity between texts comprising discussant utterances or posts in a thread. We cannot find enough overlapping words between posts, since some posts are very short. Hence we apply LDA (Blei et al., 2003) on texts to extract latent semantics of texts. We split text into sentences, i.e., each sentence is treated as a single document. Accordingly, each sentence is represented as a K-dimension vector. By computing the similarity on these vectors, we obtain a more accurate semantic similarity.

A: *There are a few other directors in the history of cinema who have achieved such a singular and consistent worldview as Kubrick.*
His films are very philosophically deep, they say something about everything, war, crime, relationships, humanity, etc.

B: *All of his films show the true human nature of man and their inner fights and all of them are very*
philosophical. Alfred was good in suspense and all, but his work is not as deep as Kubrick's

Table 1: Example of Agreement based on Implicit Attitude

	WIKI	CD
Median No. of Discussants (n)	6	29
Predicted No. of Clusters ($\lceil \sqrt{\frac{n}{2}} \rceil$)	2	4
Median No. of Actual Classes	3	3

Table 2: Number of Clusters

Property	WIKI	CD
Threads	117	34
Posts per Thread	15.5	112
Sentences per Post	4.5	7.7
Tokens per Post	78.9	118.3
Word Types per Post	11.1	10.6
Discussants per Thread	6.5	34.15
Entities Discovered per Thread	6.15	32.7

Table 3: Data Statistics

3.3 Clustering Attitude Space

A tree-based (hierarchical) clustering algorithm, SLINK (Sibson, 1973) is used to cluster the vector space. Cosine Similarity between the vectors is used as the inter-data point similarity measure for clustering.[1] We choose the number of clusters to be $\lceil \sqrt{\frac{n}{2}} \rceil$, described as the rule of thumb by (Mardia et al., 1979), where n is the number of discussants in the group. This rule seems to be validated by the fact that in the data sets with which we experiment, we note that the predicted number of clusters according to this rule and the classes identified in the gold data are very close as illustrated in Table 2. On average we note that the gold data has the number of classes per thread to be roughly 2-5.

4 Data

We use data from two online forums - Create Debate [CD][2] and discussions from Wikipedia [WIKI][3]. There is a significant difference in the kind of discussions in these two sources. Our WIKI data comprises 117 threads crawled from Wikipedia. It is relatively formal with short threads. It does not have much negative polarity and discussants essentially discuss the Wikipedia page in question. Hence it is closer to an academic discussion forum. The threads are manually annotated with sub-group information. Given a thread, the annotator is asked to identify if there are any sub-groups among the discussants with similar opinions, and if yes, the membership of those

subgroups.

On the other hand, CD is a forum where people debate a specific topic. The CD data we use comprises 34 threads. It is more informal (with pervasive negative language and personal insults) than WIKI and has longer threads. It is closer to the debate genre. It has a poll associated with every debate. The votes cast by the discussants in the poll are used as the class labels for our experiments. Detailed statistics related to both the data sets and a comparison can be found in Table 3.

5 Experimental Conditions

The following three features represent discussant attitudes:

• Sentiment towards other discussants (SD) - This corresponds to $2 * n$ dimensions in the *Attitude Profile* (AP) vector, n being the number of discussants in the thread. This is because there are two polarities and n possible targets. The value representing this feature is the number of sentences with the respective polarity – negative or positive – towards the particular discussant.

• Sentiment towards entities in discussion (SE) - Number of dimensions corresponding to this feature is $2 * e$, where e is the number of entities discovered. Similar to SD, the value taken by this feature is the number of sentences in which that specific polarity is shown by the discussant towards the entity.

• Implicit Attitude (IA) - $n * t$ dimensions are expressed using this feature, where t is the number of topics that the topic model contains. This means that the AP of every discussant contains the topic model distribution of his/her interactions with every other member in the thread. Hence, the topics in the interation between the given discussant and other members in the thread are being modeled here. Accord-

---

[1] We also experimented with K-means (MacQueen, 1967) and found that it yields worse results compared to SLINK. There is a fundamental difference between the two algorithms. Where as K-Means does a random initialization of clusters, SLINK is a deterministic algorithm. The difference in the performance may be attributed to the fact that the number of initial data points is too small for random initialization. Hence, tree based clustering algorithms are more well suited for the current task.

[2] http://www.createdebate.com

[3] en.wikipedia.org

ingly, high vector similarity due to IA between two members in a thread means that they discussed similar topics with the same people in the thread. In our experiments, we set $t = 50$. We use the Gibbs sampling based LDA (Griffiths and Steyvers, 2004). The LDA model is built on definitions of two online dictionaries WordNet, and Wiktionary, in addition to the Brown corpus (BC). To create more context, each sentence from BC is treated as a document. The whole corpus contains 393,667 documents and 5,080,369 words.

The degree of agreement among discussants in terms of these three features is used to identify subgroups among them. Our experiments are aimed at investigating the effect of explicit attitude features (SD and SE) in comparison with implicit feature (IA) and how they perform when combined. So the experimental conditions are: the three features in isolation, each of the explicit features SD and SE together with IA, and then all three features together.

SWD-BASE: As a baseline, we employ a simple word frequency based model to capture topic distribution, Surface Word Distribution (SWD). SWD is still topic modeling in the vector space, but the dimensions of the vectors are the frequencies of all the unique words used by the discussant in question.

RAND-BASE: We also apply a very simple baseline using random assignment of discussants to groups, however the number of clusters is determined by the rule of thumb described in Section 3.3.

6 Results and Analysis

Three metrics are used for evaluation, as described in (Manning et al., 2008): Purity, Entropy and F-measure. Table 4 shows the results of the 9 experimental conditions. The following observations can be made: All the individual conditions SD, SE and IA clearly outperform SWD-BASE. All the experimental conditions outperform RAND-BASE which indicates that using clustering is contributing positively to the problem. SE performs worse than SD across both datasets CD and WIKI. This may be due to two reasons: Firstly, since the problem is of clustering the discussant space, SD should be a better indicator than SE. Secondly, as seen from the comparison in Table 5, there are more polarized sentences indicating SD than SE. IA clearly outperforms SD, SE and SD+SE in the case of WIKI. In

Property	WIKI	CD
Positive Sentences towards Discussants	5.15	17.94
Negative Sentences towards Discussants	6.75	40.38
Positive Sentences towards Entities	1.65	8.85
Negative Sentences towards Entities	1.59	8.53

Table 5: Statistics of the Attitudinal Sentences per each Thread in the two data sets

the case of CD, it is exactly the opposite. This is an interesting result and we believe it is mainly due to the genre of the data. Explicit expression of sentiment usually increases with the increase in the informal nature of discussions. Hence IA is more useful in WIKI which is more formal compared to CD, where there is less overt sentiment expression. We note the same trend with the SWD-BASE where performance on WIKI is much better than its performance on CD. This also suggests that WIKI might be an easier data set. A qualitative comparison of the inter-discussant relations can be gleaned from Table 5. There is significantly more negative language than positive language in CD when compared with the ratios of negative to positive language in WIKI, which are almost the same. The best results overall are yielded from the combination of IA with SD and SE, the implicit and explicit features together for both data sets, which suggests that Implicit and explicit attitude features complement each other capturing more information than each of them individually.

7 Conclusions

We proposed the use of LDA based topic modeling as an implicit agreement feature for the task of identifying similar attitudes in online discussions. We specifically applied latent modeling to the problem of sub-group detection. We compared this with explicit sentiment features in different genres both in isolation and in combination. We highlighted the difference in genre in the datasets and the necessity for capturing different forms of information from them for the task at hand. The best yielding condition in both the dat sets combines implicit and explicit features suggesting that there is a complementarity between the two tpes of feures.

Acknowledgement

Condition	WIKI			CD		
	Purity	**Entropy**	**F-measure**	**Purity**	**Entropy**	**F-measure**
RAND-BASE	0.6745	0.5629	0.6523	0.3986	0.9664	0.407
SWD-BASE	0.7716	0.4746	0.6455	0.4514	0.9319	0.4322
SD	0.8342	0.3602	0.667	0.8243	0.3942	0.5964
SE	0.8265	0.3829	0.6554	0.7933	0.4216	0.5818
SD+SE	0.8346	0.3614	0.6649	0.82	0.3851	0.6039
IA	0.8527	0.3209	0.6993	0.787	0.3993	0.5891
SD+IA	0.8532	0.3199	0.6977	0.8487	0.3328	0.6152
SE+IA	0.8525	0.3216	0.7015	0.7884	0.3986	0.591
SD+SE+IA	**0.8572**	**0.3104**	**0.7032**	**0.8608**	**0.3149**	**0.6251**

Table 4: Experimental Results

References

Amjad Abu-Jbara, Pradeep Dasigi, Mona Diab, and Dragomir Radev. 2012. Subgroup detection in ideological discussions. In *Proceedings of the 5oth Annual Meeting of ACL.*

David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3.

Jenny Rose Finkel, Trond Grenager, and Christopher Manning. 2005. Incorporating non-local information into information extraction systems by gibbs sampling. In *Proceedings of the 43nd Annual Meeting of the Association for Computational Linguistics.*

Murthy Ganapathibhotla and Bing Liu. 2008. Mining opinions in comparative sentences. In *Proceedings of the 22nd International Conference on Computational Linguistics (Coling 2008).*

Thomas L. Griffiths and Mark Steyvers. 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101.

Ahmed Hassan, Vahed Qazvinian, and Dragomir Radev. 2010. What's with the attitude? identifying sentences with attitude in online discussions. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing,.*

Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining.*

Niklas Jakob and Iryna Gurevych. 2010. Using anaphora resolution to improve opinion target identification in movie reviews. In *Proceedings of the ACL 2010 Conference Short Papers.*

Nozomi Kobayashi, Kentaro Inui, and Yuji Matsumoto. 2007. Extracting aspect-evaluation and aspect-of relations in opinion mining. In *Proceedings of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning.*

J. MacQueen. 1967. Some methods for classification and analysis of multivariate observations. In *Proceedings of Fifth Berkeley Symposium on Mathematical Statistics and Probability.*

Christopher D. Manning, Prabhakar Raghavan, , and Hinrich Schtze. 2008. . *2008. Introduction to Information Retrieval.* Cambridge University Press, New York, NY,USA.

K. V. Mardia, J. T. Kent, and J. M. Bibby. 1979. *Multivariate Analysis.* Publisher.

R. Sibson. 1973. Slink: An optimally efficient algorithm for the single-link cluster method. In *The Computer Journal (1973) 16 (1): 30-34.*

Swapna Somasundaran and Janyce Wiebe. 2009. Recognizing stances in online debates. In *Proceedings of the Joint Conference of the 47th Annual Meeting of the ACL and the 4th International Joint Conference on Natural Language Processing of the AFNLP.*

Theresa Wilson, Paul Hoffmann, Swapna Somasundaran, Jason Kessler, JanyceWiebe, Yejin Choi, Claire Cardie, Ellen Riloff, and Siddharth Patwardhan. 2005. Opinionfinder: A system for subjectivity analysis. In *Proceedings of HLT/EMNLP 2005 Demonstration.*