

CU-GWU Perspective at SemEval-2016 Task 6: Ideological Stance Detection in Informal Text

Heba Elfardy

Department of Computer Science
Columbia University
New York, NY
heba@cs.columbia.edu

Mona Diab

Department of Computer Science
The George Washington University
Washington, DC
mtdiab@gwu.edu

Abstract

We present a supervised system that uses lexical, sentiment, semantic dictionaries and latent and frame semantic features to identify the stance of a tweeter towards an ideological target. We evaluate the performance of the proposed system on subtask A in SemEval-2016 Task 6: “Detecting Stance in Tweets”. The system yields an average $F_{\beta=1}$ score of 63.6% on the task’s test set and has been ranked the 6th by the task organizers out of 19 judged systems.

1 Introduction & Background

Automatically identifying a person’s ideological stance (or perspective) from written text is quite a challenging research problem that finds applications spanning from enhancing advertisements targeting to planning political campaigns. From a computational viewpoint, perspective detection is closely related to subjectivity and sentiment analysis. A person’s perspective normally influences his/her sentiment towards different topics or targets. Conversely identifying the sentiment of a person towards multiple targets can serve as a cue for identifying his/her perspective. While from a social-science viewpoint, the notion of “perspective” is related to the concept of “framing”. Framing involves making some topics (or some aspects of the discussed topics) more prominent in order to promote the views and interpretations of the writer (communicator). The communicator can either make these framing decisions consciously or unconsciously (Entman, 1993).

In this paper, we present a system that employs lexical and semantic features to identify the stance of

a person – “Favor”, “Against” or “None” – on topics that are often backed up by one’s ideology or belief system such as abortion, climate change and feminism. We evaluate the performance of the proposed system through the participation in Sub-Task A “Supervised Framework” of SemEval Task 6 “Detecting Stance in Tweets”. (Mohammad et al., 2016)

We build on our previous work on automatic perspective detection (Elfardy et al., 2015) by exploring the use of sentiment, Linguistic Inquiry and Word Count (LIWC) dictionaries, frame and latent semantics in addition to standard lexical features to automatically classify a given tweet according to its stance on the topic of interest.

Most current computational linguistics research on supervised stance detection performs document (or post) level stance classification, whether binary or multiway and use a variety of lexical, syntactic and semantic features to identify the stance of a post towards a specific contentious topic/target such as the Israeli-Palestinian conflict, abortion, creationism, gun-rights, gay-rights, healthcare, legalization of marijuana and death penalty (Lin et al., 2006; Klebanov et al., 2010; Hasan and NG, 2012; Hasan and Ng, 2013; Somasundaran and Wiebe, 2010; Elfardy et al., 2015). For a detailed literature review, we direct the reader to our *Sem paper. (Elfardy et al., 2015)

2 Shared Task Description

SemEval Task 6 “Detecting Stance in Tweets” (Mohammad et al., 2016) aims at evaluating how well an automated system can identify the stance of a

	Training					Held-Out Test				
	Abortion	Atheism	Climate	Clinton	Feminism	Abortion	Atheism	Climate	Clinton	Feminism
Favor	105	92	212	112	210	46	32	123	45	58
Against	334	304	15	361	328	189	160	11	172	183
None	164	117	168	166	126	45	28	35	78	44
All	603	513	395	639	664	280	220	169	295	285

Table 1: Number of tweets per class in the training and test data.

tweeter on several contentious targets. Sub-Task A “Supervised Framework” of the task –in which we participate– focuses on five targets: “Atheism”, “Climate Change is a Real Concern”, “Feminist Movement”, “Hillary Clinton”, and “Legalization of Abortion”. The training data has a total of 2,814 tweets and the test data has 1,249 tweets. While each tweet can have one of three class labels “Favor”, “Against” or “None”, the official metric calculates the performance of each system as the average $F_{\beta=1}$ score of the first two class labels only. Table 1 shows the distribution of the training and test data for each one of the five targets.

3 Approach

Our goal is to see how well lexical and semantic features can help in identifying the stance of a tweeter. We split the data into five subsets according to the target –“Atheism”, “Climate Change is a Real Concern”, “Feminist Movement”, “Hillary Clinton”, and “Legalization of Abortion”– and train a separate classifier for each one of these targets.

3.1 Lexical Features

For lexical features, we use standard n-grams. We apply basic preprocessing to the text by removing all punctuation and numbers and converting all words to lower case. Converting the text to lower case is intended to reduce the sparseness of the data while excluding the punctuation and number is meant to avoid overfitting the training data. We use n-grams having a length between 1 and 3 and exclude the ones that occur in only one training instance.

3.2 Latent Semantics

Latent Semantics map text from a high-dimensional space such as n-grams to a low-dimensional one

such as topics. Most of these models assign a semantic profile to each given sentence (or document) by considering the observed words and assuming that each given document has a distribution over K topics. We apply Weighted Textual Matrix Factorization (WTMF) (Guo and Diab, 2012) to each tweet. The advantage WTMF offers –over standard topic models– is that in addition to modeling observed words, it models missing ones. WTMF defines missing words as the whole vocabulary of the training data minus the ones observed in the given tweet/post. Modeling missing words is particularly useful when the input is very short in length because in such case, only a limited context of observed words is present.

We use the distributable version of WTMF which is trained on WordNet (Fellbaum, 2010), Wiktionary definitions¹ and the Brown Corpus² and sets the number of topics (K) to 100.

3.3 Sentiment

As mentioned earlier, sentiment analysis is closely related to stance detection. The ideological stance of a person normally influences his/her sentiment towards different ideological topics. Accordingly we decide to identify the overall sentiment of each sentence in the given tweet and use this sentiment as a possible indicator of a tweeter’s stance. Our assumption is that since tweets are inherently short in length, we can assume that the expressed sentiment targets the ideological topic the tweet is discussing. We use Stanford’s sentiment analysis system (Socher et al., 2013) to identify the number of positive, negative and neutral sentences in a given tweet and use these three counts as features for our

¹<http://en.wiktionary.org>

²http://en.wikipedia.org/wiki/Brown_corpus

	Training					Held-Out Test				
	Abortion	Atheism	Climate	Clinton	Feminism	Abortion	Atheism	Climate	Clinton	Feminism
Feeling	8.93	5.83	7.56	8.58	10.21	7.45	6.31	6.43	11.45	7.32
Biological Pr.	2.81	4.08	12.85	4.68	5.41	3.55	4.5	8.77	4.04	5.92
Body	49.59	27.77	18.14	18.1	34.98	54.26	26.58	16.37	9.76	32.4
Health	12.07	8.74	6.05	6.55	12.31	11.7	6.76	6.43	3.7	13.94
Sexual	35.37	10.1	7.56	4.06	6.31	41.13	11.26	2.92	3.37	5.23
Ingestion	30.74	9.13	2.02	7.96	20.87	30.14	10.36	4.09	3.03	16.72
Relativity	1.32	3.5	5.29	2.03	4.05	1.06	1.8	5.85	2.02	4.88
Motion	73.39	76.5	84.89	73.17	69.97	71.99	73.87	83.04	73.74	73.17
Relativity	15.7	22.14	27.71	23.09	17.27	13.48	22.07	36.26	22.22	21.6
Space	48.93	54.37	59.45	46.96	46.1	50.35	50	60.82	44.78	46.34
Time	45.62	41.36	51.64	43.53	42.19	42.55	41.89	49.12	42.42	47.74
Work	16.2	20.39	25.94	30.42	20.42	16.31	21.62	21.64	34.01	14.98
Achievement	20.66	33.98	29.47	27.77	19.97	20.57	27.93	17.54	28.62	19.86
Leisure	6.61	11.65	9.82	10.76	21.92	10.28	13.51	8.77	9.76	17.77
Home	3.14	2.72	6.55	3.28	5.41	2.48	2.7	9.94	4.04	5.57
Money	7.27	6.41	8.82	11.08	7.81	6.38	4.95	14.62	10.44	8.01
Religion	15.7	64.27	7.05	3.28	6.46	14.18	68.02	1.75	6.4	4.18
Death	17.36	4.08	3.27	3.28	3.6	16.31	4.95	4.68	4.38	4.53
Assent	3.97	3.11	3.02	9.2	5.71	4.96	3.15	4.09	3.37	4.88

Table 2: Coverage of the used LIWC categories on the training and test sets of each target/domain

classifier.

3.4 Linguistic Inquiry & Word Count (LIWC)

LIWC (Tausczik and Pennebaker, 2010) uses a set of dictionaries to assign words in a given text to a set of psychologically meaningful categories such as death, science, emotion, space, family, swear words and many others. We apply LIWC to each given tweet in order to estimate the percentage of words belonging to each of the following categories:

Feeling, Biological Processes, Body, Health, Sexual, Ingestion, Relativity, Motion, Space, Time, Work, Achievement, Leisure, Home, Money, Religion, Death and Assent.

These categories convey the discussed topics in the given tweet. Table 2 shows the coverage of each of these categories. The coverage for each category corresponds to the percentage of tweets –out of all tweets– that have at least one word in this LIWC category. We calculate the coverage of the used LIWC categories on both the training and test sets for all five targets. By analyzing the numbers in Table 2, we find that “Atheism” dataset has the highest coverage for “Religion” category, “Abortion” has the highest coverage for “Death” category while “Cli-

mate Change” has the highest coverage for “Motion”. Moreover for “Abortion”, 25.8% of the tweets opposing it use words in the “death” category as opposed to only 8.6% for the tweets favoring the topic. For “Atheism”, the use of “religion” category is high among both opposing and favoring tweets where 78.26% of the tweets that favor atheism use words in the “religion” category and 80.26% of the ones that oppose it use this category. While LIWC will potentially help in identifying the stance from which a given tweet or text in general is written, a main drawback to it is that it only performs a dictionary look-up on each word hence judging words in isolation of their context.

3.5 Frame Semantics

Frame semantics assemble the meanings of different elements in a given piece of text to model the meaning of the whole text (Baker et al., 1998). The basic semantic unit in frame semantics theory is the “frame”. A frame is a conceptual structure that refers to a group of related concepts (or elements) where understanding one of these concepts requires an understanding of its whole structure. When any of these structures is present, it automatically triggers all of the other ones in the reader’s mind (Fillmore, 2006). The word (or phrase) that triggers the

	Abortion	Atheism	Climate	Clinton	Feminism	All-Targets
Majority-BL	35.87	37.04	2.5	36	33	32.2
n-grams	58.04	66.67	40.48	58.9	47.82	62.39
WTMF	54.39	53.59	35	42.03	51.25	57.22
Sentiment	35.87	37.04	35	36	33	54.21
Frames	48.67	37.18	35.42	33.68	45.56	54.34
LIWC	35.29	36.11	34.48	36.36	33	53.36
Semantic	60.61	52.6	36.84	34.48	51.38	57.84
All	66.06	64.62	41.46	55.45	51.58	63.98

Table 3: Development Set Results (measured in average $F_{\beta=1}$ score of “Favor” and “Against” classes)

	Abortion	Atheism	Climate	Clinton	Feminism	All-Targets
Majority-BL	40.3	42.11	6.11	36.83	39.1	36.41
n-grams	57.43	54.98	36.71	51.59	48.37	60.41
WTMF	48.08	57.73	39.83	45.81	52.25	62.27
Semantic	57.25	48.77	39.48	39.89	54.29	63.1
All	60.52	56.39	38.89	55.82	51.93	63.6

Table 4: Held-Out Test Set Results (measured in average $F_{\beta=1}$ score of “Favor” and “Against” classes)

frame is called the “frame-target”. For example, the target for the frame “Killing” can be any kill verb and the frame elements will include the killer and the victim.

We use SEMAFOR (Das et al., 2010; Chen et al., 2010), a publicly available frame-semantic parser, to identify all the semantic frames in each given tweet. For example, in the tweet “*Because I want young American women to be able to be proud of the 1st woman president.*”, SEMAFOR identifies the following frames, “*Leadership Target:president*”, “*Capability Target:able*”, “*Origin Target:American*”, “*Desiring Target:want*”, “*People Target:women*” and “*Age Target:young*”. We create a list of all the frames that occur in the training data and use binary features to indicate the presence/absence of each of them in each given tweet. This set of features provides yet another abstraction in order to infer the topics discussed in the given text.

3.6 Classifier Training

Using WEKA toolkit (Hall et al., 2009) and the derived lexical and semantic features, we train an SVM classifier for each one of the five targets. We use a

radial basis kernel and set the cost parameter C to 1.

4 Experiments & Results

For tuning, we split the training data for each target into two subsets: 90% training and 10% development. We evaluate different configurations of the proposed features on the development set and apply the ones that yield best tuning results to the held-out test set. Specifically we experiment with the following configurations: (1) Lexical Features (n-grams), (2) WTMF, (3) Sentiment, (4) Frames, (5) LIWC, (6) All Semantic features, (7) All features. We compare the proposed approach against a majority baseline which assigns all tweets to the most frequent class in the data (“Against”). Table 3 shows the results on the development set. N-grams outperform using any of the semantic features separately which is quite expected given how well n-grams generally perform on text classification tasks. Moreover WTMF outperforms all other semantic features except on “Climate Change” dataset. Overall we find “Climate Change” to show different trends and a much lower performance than those of the other four targets. We believe that this is at-

tributed to the small size of this set and to its very skewed distribution. Only 3.8% of its tweets belong to “against” class which is the most frequent class in the whole dataset. Combining all semantic features yields better results than using any of them separately for three out of the five targets “Abortion”, “Climate Change” and “Feminist Movement” while for the other two targets using WTMF only outperforms using the full set of semantic features. For the same two targets – “Atheism” and “Hillary Clinton” – using only n-grams outperforms combining n-grams with all-semantic features. Overall, combining lexical and semantic features yields the best results on the majority of tweets. Accordingly we used this setup in our official task submission.

We apply the top four development-set configurations – “n-grams”, “WTMF”, “all semantic” and “all features” – to the test set. Table 4 shows the held-out test results. While using the combination of lexical and all-semantic features yields best overall performance, using only the semantics features achieves close to best results. Additionally, using only WTMF yields best performance on two targets – “Atheism” and “Climate Change”. We also find that while the best configuration for each target varies across the development and test sets, the best overall configuration for the majority of tweets remains the same.

5 Conclusion

In this paper, we explore the use of lexical and several semantic feature sets within a supervised framework to identify the stance of a tweet towards an ideological topic. We evaluate the performance of the proposed approach by participating in SemEval-2016 Task 6: “Detecting Stance in Tweets” which aims at identifying the stance of a given tweet towards five targets; “Abortion”, “Atheism”, “Climate Change”, “Hillary Clinton” and “Feminist Movement”. We find that WTMF, a latent semantics model that incorporates both observed and missing words, yields the best results among semantic features. Moreover, combining lexical and all semantic features results in the best performance.

References

Collin F Baker, Charles J Fillmore, and John B Lowe. 1998. The berkeley framenet project. In *Proceedings*

of the 17th international conference on Computational linguistics-Volume 1. Association for Computational Linguistics.

Desai Chen, Nathan Schneider, Dipanjan Das, and Noah A Smith. 2010. Semafor: Frame argument resolution with log-linear models. In *Proceedings of the 5th international workshop on semantic evaluation*. Association for Computational Linguistics.

Dipanjan Das, Nathan Schneider, Desai Chen, and Noah A Smith. 2010. Semafor 1.0: A probabilistic frame-semantic parser. *Language Technologies Institute, School of Computer Science, Carnegie Mellon University*.

Heba Elfardy, Mona Diab, and Chris Callison-Burch. 2015. Ideological perspective detection using semantic features. *Lexical and Computational Semantics (*SEM 2015)*.

Robert M Entman. 1993. Framing: Toward clarification of a fractured paradigm. volume 43. Wiley Online Library.

Christiane Fellbaum. 2010. Wordnet: An electronic lexical database. 1998. *WordNet is available from <http://www.cogsci.princeton.edu/wn>*.

Charles J Fillmore. 2006. Frame semantics. *Cognitive linguistics: Basic readings*, 34.

Weiwei Guo and Mona Diab. 2012. Modeling sentences in the latent space. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*, pages 864–872. Association for Computational Linguistics.

Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, and Ian H Witten. 2009. The weka data mining software: an update. *ACM SIGKDD Explorations Newsletter*, 11(1):10–18.

Kazi Saidul Hasan and Vincent NG. 2012. Predicting stance in ideological debate with rich linguistic knowledge. In *Proceedings of the 24th International Conference on Computational Linguistics*.

Kazi Saidul Hasan and Vincent Ng. 2013. Extralinguistic constraints on stance recognition in ideological debates. In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics.

Beata Beigman Klebanov, Eyal Beigman, and Daniel Diermeier. 2010. Vocabulary choice as an indicator of perspective. In *ACL*, pages 253–257. Association for Computational Linguistics.

Wei-Hao Lin, Theresa Wilson, Janyce Wiebe, and Alexander Hauptmann. 2006. Which side are you on?: identifying perspectives at the document and sentence levels. In *Proceedings of the Tenth Conference on Computational Natural Language Learning*, pages 109–116. Association for Computational Linguistics.

- Saif M. Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. SemEval-2016 Task 6: Detecting stance in tweets. In *Proceedings of the International Workshop on Semantic Evaluation*, SemEval '16, June.
- Richard Socher, Alex Perelygin, Jean Y Wu, Jason Chuang, Christopher D Manning, Andrew Y Ng, and Christopher Potts. 2013. Recursive deep models for semantic compositionality over a sentiment treebank. In *Proceedings of the conference on empirical methods in natural language processing (EMNLP)*, volume 1631. Citeseer.
- Swapna Somasundaran and Janyce Wiebe. 2010. Recognizing stances in ideological online 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.
- Yla R Tausczik and James W Pennebaker. 2010. The psychological meaning of words: Liwc and computerized text analysis methods. *Journal of language and social psychology*, 29(1).