Weakly-Guided User Stance Prediction via Joint Modeling of Content and Social Interaction

Rui Dong
College of Computer and Information Science
Northeastern University
Boston, MA, USA
dongrui@ccs.neu.edu

Yizhou Sun
Computer Science Department
University of California, Los Angeles
Los Angeles, CA, USA
yzsun@cs.ucla.edu

Lu Wang
College of Computer and Information Science
Northeastern University
Boston, MA, USA
luwang@ccs.neu.edu

Yupeng Gu
Computer Science Department
University of California, Los Angeles
Los Angeles, CA, USA
ypgu@cs.ucla.edu

Yuan Zhong
College of Computer and Information Science
Northeastern University
Boston, MA, USA
zhong.yu@husky.neu.edu

ABSTRACT

Social media websites have become a popular outlet for online users to express their opinions on controversial issues, such as gun control and abortion. Understanding users' stances and their arguments is a critical task for policy-making process and public deliberation. Existing methods rely on large amounts of human annotation for predicting stance on issues of interest, which is expensive and hard to scale to new problems. In this work, we present a weakly-guided user stance modeling framework which simultaneously considers two types of information: *what do you say* (via stance-based content generative model) and *how do you behave* (via social interaction-based graph regularization). We experiment with two types of social media data: news comments and discussion forum posts. Our model uniformly outperforms a logistic regression-based supervised method on stance-based link prediction for unseen users on news comments. Our method also achieves better or comparable stance prediction performance for discussion forum users, when compared with state-of-the-art supervised systems [34]. Meanwhile, separate word distributions are learned for users of opposite stances. This potentially helps with better understanding and interpretation of conflicting arguments for controversial issues.

CCS CONCEPTS

• **Information systems → Web applications**;

KEYWORDS

User stance prediction; online behavior mining; social computing; social media.

CIKM'17 , November 6–10, 2017, Singapore, Singapore

ACM ISBN 978-1-4503-4918-5/17/11…$15.00
https://doi.org/10.1145/3132847.3133020

1 INTRODUCTION

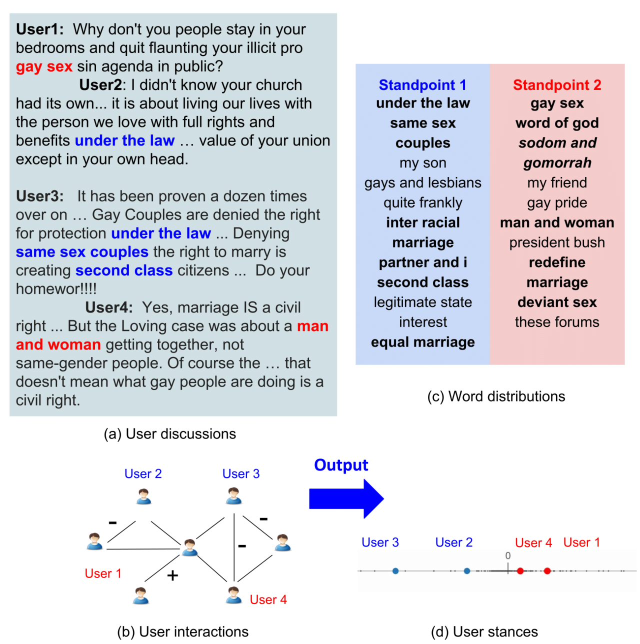

Figure 1: An illustration of our proposed stance prediction model (STML). Users of opposite stances are highlighted in different colors along with representative words in their comments ((a) and (b)). Without knowing any user's stance, STML jointly models the content and the user interactions (e.g., agreement and disagreement marked with "+" and "–" in (b)). It then outputs separate word distributions for opposite stances (as in (c)) as well as predicts users' positions on a specific issue (in real values as in (d)).

Nowadays, it has become popular for people to express and exchange their opinions through social media, such as by posting comments under trendy topics in news commenting systems, or debating with other users on contentious issues in discussion forums. The massive amount of online discussions provide us with valuable resources for studying and understanding public opinions on fundamental societal issues, e.g., abortion or gun rights. Automatically predicting user stance and identifying corresponding arguments are important tasks for improving policy-making process and public deliberation.

Our goal is to develop a weakly-guided stance modeling framework which is able to automatically predict user position on controversial issues and provide insights on the heated arguments. Previous work [3, 14, 25, 32–34, 37] has shown that stance prediction is a challenging task, and would require techniques beyond traditional sentiment analysis methods that only consider textual features or sentiment lexicon [32, 33]. Specifically, user interactions, e.g., rebuttal links between posts, have been demonstrated as important complementary features to content for stance classification [35, 37]. Recently, Sridhar et al. [34, 35] propose to collectively classify both post-level and user-level stances to encourage consistency among the predictions. Stance prediction with regard to a given target has been studied on Twitter [3], which relies on a set of seed tweets that are heuristically labeled based on pre-selected hashtags. Most of the aforementioned work employs supervised methods which require significant amounts of human efforts for annotation, and thus makes it difficult to scale for new issues.

To address the above challenges, we propose a unified model, STML (Stance-based Text Generative Model with Link Regularization), which combines *content generation modeling* and *user interaction modeling*. Intuitively, users with different stances differ in word usages, and they also tend to argue on opposite opinions. For instance, Figure 1 shows two fierce discussions with disagreement among four users on the issue of "same-sex marriage". User 1 and User 4, who are against it, frequently mention "gay sex" and "man and woman". On the contrary, User 2 and User 3, who are on the pro side, focus on arguments containing "under the law". Concretely, STML is built on a novel stance-based content generative model and leverages signed user-user interaction links to further regularize user positions. An interactive learning algorithm is proposed to allow STML to be trained in an unsupervised fashion with weak guidance. The learned STML model is capable of (1) producing word distributions of two opposite stances for a specified issue, and (2) predicting a numerical stance value for either observed user or unseen user, as demonstrated in Figure 1.

We experiment with real-world discussions from two social media — CNN news commenting system, and `4Forums.com` discussion forum — on stance prediction tasks. Experimental results of stance-based link prediction in news commenting system show that our STML model can uniformly outperform a logistic regression-based model which is trained with words and phrases (e.g., an AUC of 0.78 vs. 0.51 on issue "Gaza Israel"). Our model also achieves stance prediction accuracies comparable to a supervised learning-based state-of-the-art model [34] on several issues for discussion forum users (e.g., an accuracy of 0.76 compared to 0.66 on abortion). Qualitative analysis is further carried out on the learned word distributions of opposite stances. We find that the STML model, which considers both content and user interactions, is able to capture the controversial aspects derived from ideology polarization or conflicting arguments. We envision that the output by our model can be used for extracting and summarizing "heated arguments" for disputed issues. Our contributions can be summarized as follows:

(1) Our model is weakly guided by domain independent heuristic rules on user interaction patterns with no supervision from user stances;
(2) We are able to learn user stances and word distributions for opposite stances simultaneously, where the word distributions provide better interpretation for opposite standpoints;
(3) Instead of treating user stance prediction as a binary classification problem, STML learns numerical stances for users to better capture their relative polarity in various issues.

2 PRELIMINARIES AND PROBLEM DEFINITION

In this paper, we study the user stance prediction problem for two types of social media websites: news commenting systems (e.g., CNN) and online debate forums (e.g., 4Forums). In a news commenting system, comments under each news article serve as a thread, and users can freely comment to the news article and reply to each other. A post can either be a direct comment to the article or a reply to other existing posts. In an online debate forum, users debate in discussion threads on a variety of issues, such as *gun control*, *abortion*, or *gay marriage*. A post can either be written to initiate a new thread under a certain issue or respond to any previous post in that thread. Our goal is to infer user stance for a given issue according to two types of information: the content of the posts and user interaction behavior via replying. Terminologies that are frequently used are introduced below.

Issues. Each issue denotes a certain topic or event that users are interested in. News articles along with the comments and threads in online forums are extracted based on a set of keywords related to that issue. We typically observe two contrasting standpoints for each issue.

Threads. One thread contains all comments for a news article or represent a discussion thread in an online forum.

Posts. Posts are either comments under each related article or posts in online forums. A post is represented as a bag of words. In order to capture more meaningful statement, we use a phrase mining tool, i.e., SegPhrase [22], to extract high-quality phrases (e.g., "same sex marriage") to further enrich the word vocabulary.

User-User interactions. Users interact via replying function in both news commenting and online forum settings. We aggregate all the post-post level interaction to user-user level interaction for each thread. Note that, in this paper we need to consider the signs of user interactions based on agreement or disagreement. Some social media, such as Epinions and Slashdot, signs of links are explicitly available. Our datasets do not contain this type of information; simple heuristics are utilized to generate the signs (see Section 4).

User stance. User stance denotes a user's position for a particular issue, which is represented as a real number. Different signs imply opposite stances, and larger absolute value denotes more extreme stance. For issues where stances are related to liberal and conservative ideology (e.g., same sex marriage), we use negative scores to

u	a user
d	a post
t	a thread (e.g. all comments under a news article, or all posts under a discussion theme)
$n(d, w)$	frequency of word w in post d
$y(u_i, u_j, t)$	signed label between user u_i and u_j in thread t
β_B	background word distribution
γ^+	positive word distribution
γ^-	negative word distribution
μ_B	prior probability of background topic
x_u	position of user u
$\mathbf{D}$	a set of posts $\{d\}$
$\mathbf{Y}$	a set of links $\{y(u_i, u_j, t)\}$
$\mathbf{x}$	parameter vector $\{x_u\}$

Table 1: Notations used for the proposed STML model.

denote liberal-leaning stance and positive scores for conservative-leaning stance. In general, the signs of the scores can be flipped, and positive and negative signs simply denote two contrasting standpoints.

Problem Definition. Given an online discussion issue, where N_U users are involved in the discussion across N_T threads with N_D posts, our goal is to estimate stance for each user as $X = \{x_u\}_{u=1}^{N_U}$, and learn word distributions γ^+ and γ^- for two opposite standpoints of the issue. Notations of our model are summarized in Table 1.

3 USER STANCE-BASED TEXT GENERATIVE MODEL

In this section, we first design a stance-based text generative model, which describes how posts are generated when users have different user stance for a given issue. We will discuss how user interactions can further enhance the model in Section 4, and propose a joint model in Section 5.

3.1 The Generative Model

As shown in the example of same sex marriage issue, the content information of a post could be very different if the users have different stances (e.g., "same sex marriage" vs. "gay marriage"). We thus design a text generative model based on two intutions:

(1) Users with different standpoints tend to choose different word vocabularies to write their posts; and
(2) Users with more extreme stances have a higher probability to choose more words from vocabulary of her standpoint.

Following intuition 1, we use two word distributions representing two opposite standpoints for each issue, denoted as γ^+ and γ^-. Following intuition 2, given a post d and its author u, the probability of picking a word from each word distribution is determined by the user's stance x_u. In addition, we use a background word distribution β_B to capture the commonality of word usage (i.e., neutral words) from two parties.

Given an issue, formally, the user stance-based text generative model for a post d written by user u is described below.

(1) Decide the standpoint of u in post d, $s(u,d)$, by drawing a binary position from a Bernoulli distribution $Bern(\sigma(v \cdot x_u))$, where x_u is the numerical stance for user u on issue z_c, $\sigma(\cdot)$ is the sigmoid function that turns user stance into a probability, and v is a scaling coefficient.
(2) For each word w in post d:
Decide whether w is sampled from background distribution, by drawing $z_{d,w}^B$ from Bernoulli distribution $Bern(\mu_B)$, where μ_B is a given parameter and is set as 85% in our experiments.
   (a) If $z_{d,w}^B = 1$ (background word), draw w from the background word distribution $Mult(\beta_B)$;
   (b) if $z_{d,w}^B = 0$ (stance-sensitive word):
      (i) if $s(u,d) = +1$ (positive standpoint), draw w from the word distribution for positive standpoint $Mult(\gamma^+)$,
      (ii) otherwise (negative standpoint), draw w from the word distribution for negative standpoint $Mult(\gamma^-)$.

3.2 The Objective Function

According to the above generative process of each post, the probability of observing a post d written by user u is given by

$$\begin{aligned} p(d|x_u, v, \gamma^+, \gamma^-, \beta_B, \mu_B) &= \sum_{s(u,d)} p(d, s(u,d)|x_u, v, \gamma^+, \gamma^-, \beta_B, \mu_B) \\ &= \sum_{s(u,d)} p(s(u,d)|x_u, v) P(d|s(u,d), \gamma^+, \gamma^-, \beta_B, \mu_B) \end{aligned} \tag{1}$$

where $p(s(u,d) = 1|x_u, v) = 1/(1 + \exp(-v \cdot x_u))$ and $p(d|s(u,d), \gamma^+, \gamma^-, \beta_B, \mu_B)$ is the probability of observing all the words in post d. The probability of observing word w in post d is given by

$$\begin{aligned} p(w|s(u,d); \gamma^+, \gamma^-, \beta_B, \mu_B) &= \sum_{z_{d,w}^B} p(w, z_{d,w}^B|s(u,d), \gamma^+, \gamma^-, \beta, \mu_B) \\ &= \mu_B \cdot \beta_B(w) + (1 - \mu_B) \cdot (\gamma^+(w))^{\mathbf{1}(s(u,d)=1)} \cdot (\gamma^-(w))^{\mathbf{1}(s(u,d)=-1)} \end{aligned} \tag{2}$$

where $\mathbf{1}(\cdot)$ is the indicator function which equals 1 if the internal predicate holds. As the post length varies a lot in social media, we normalize the length by considering the geometric mean of probabilities of each word for a given post d. Thus, we have

$$\begin{aligned} & p(d|s(u,d), \gamma^+, \gamma^-, \beta, \mu_B) \\ &= \left(\prod_{w=1}^{N_W} p(w|s(u,d), \gamma^+, \gamma^-, \beta_B, \mu_B, v)^{n(d,w)} \right)^{\frac{1}{n_d}} \end{aligned} \tag{3}$$

where $n(d,w)$ denotes the number of word w in post d, N_W is the size of vocabulary, and n_d is the length of post d.

Given the collection of all the posts in the issue, the goal is to maximize the average log likelihood of observing all the posts.

$$l(\mathbf{x}, \gamma^+, \gamma^-, \beta_B, v|\mathbf{D}, \mu_B) = \frac{1}{N_D} \sum_{d=1}^{N_D} \log p(d|\mathbf{x}, v, \gamma^+, \gamma^-, \beta_B, \mu_B) \tag{4}$$

where $\mathbf{D}$ denotes a collection of posts, and N_D is the number of posts in the corpus.

Note that, the current model is unidentifiable for x_u and v. In other words, we have the same objective if we multiple user stance by a constant and divide v by the same constant. Therefore, we fix the scale of user stance x_u by introducing a length constraint, i.e., $\mathbf{x}^T\mathbf{x} = \sum_u x_u^2 = N_U$, where N_U is the total number of users in the dataset. By doing so, we can expect the average absolute value of user stances is around 1. The scaling factor v thus controls the sensitivity of an issue. For a larger v, more users are with an extreme stance; and for a smaller v, more users are with neutral stance.

The generative model can well explain how posts are generated given users' stances. But unfortunately, according to the objective function described above, we can hardly learn two opposite standpoints and thus the user stance correctly, as shown in Table 9 in the Experiment section. The main reason is that the objective function can lead to many local optimum points, and they do not have to correspond to the desired contrasting standpoints. To overcome this issue, we introduce the link regularization in the next section.

4 USER INTERACTION-BASED REGULARIZATION

As we discussed in the last section, user stance-based text generative model itself does not suffice to identify user stances correctly. Fortunately, user interaction links, especially when associated with signs, will be helpful in addressing the issue, due to following observations:

(1) Users with opposite stances tend to disagree with each other in their interaction; and
(2) Users with the same stance tend to agree with each other in their interaction.

Thus, we design a link-based regularizer accordingly to find the desired user stance.

Note that, signs of links play a very important role here, as interactions do not necessarily imply homophily, which makes our regularizer different from most of the existing graph regulariziers, e.g., graph Laplacian-based regularization [2]. In the real world, signs of interations can be explicitly found in some social media, such as in Epinions and Slashdot; but most likely they are implicit. Therefore, we propose to use simple heuristics to generate signs for these user interactions.

4.1 Rule-based Sign Generation for User-User Interaction Links

We now describe the strategies in determining the signs of a user-user interaction. First, we consider user-user interaction at the thread level, and the goal is to determine the sign of the interaction of two users u and v in thread t, denoted as $y(u,v,t)$. If u agrees with v in thread t, $y(u,v,t) = 1$; otherwise, $y(u,v,t) = -1$. We choose to use thread-level interaction, as it (1) allows user stance variation between threads and (2) avoids noisy signals at the post level. Second, we design a set of heuristics that assign a sign to each reply link from u to v in thread t. By aggregating all the post-post level interaction in the thread, the sign of thread level ineraction can be easily determined. For example, if we have observed 10 replies from u to v in thread t and all of them are "disagree", we will assign -1 to $y(u,v,t)$.

The set of rules are described below, and the goal is to label an interaction only when we are very confident.

Number of turns of discussion. According to [39], the number of turns of discussion is very useful to detect the disagreement between the users. Therefore, if the number of replying links between two users in thread t exceeds a predefined threshold, e.g., 10, we label this interaction as -1 (i.e., disagree).

# Total user-user interaction links	682,903
# Total labeled links	91,083
# Positive labels	8,017
# Negative labels	83,066

Table 2: Statistics of thread-level user-user interaction links and their labels on CNN data.

Question mark. Since people tend to ask rhetorical questions when they disagree with each other, a reply link is labeled as -1 if more than one sentence ends with a question mark in the replying post.
Text-based agree signals. We check the first sentence of the reply post for the text-based agree signals, as people usually state their stance at the beginning of their posts. If the post starts with strong agree signals like "agree[d]", "totally agree[d]" or "I [do] agree", we label the replying link as +1. Next, we check the whole post with sentences like "I agree", "I'm with you" or phrases like "fully agree" or "100% agree", if they are not with any negation word such as "but" or "however", then we label this reply link as +1 (i.e., argee).
Text-based disagree signals. If the post starts with "disagree", "I disagree", or contains "disagree" without negation words such as "not", we label the replying link as -1. If the first sentence starts with words like "no", "nope" or contains phrases like "not true", we label the replying link as -1. Finally, if disagree signals like "I disagree", "try", "read", "loser", "rubbish", "garbage", or "pathetic" appear in the post, we label this reply link as -1.

All the above-mentioned rules are topic-independent and carry only relative stance information. These rules are inspected in order, namely if one rule is satisfied, then the remaining rules will not be examined. Then the labels for reply links are aggregated to thread-level user-user interactions (u,v,t). We will not assign any label to user interaction link (u,v,t), if their reply links in thread t receive both -1 and +1 labels or receive no labels. Otherwise, if all labels between two users are -1 or +1 in the same thread (i.e., consensus reached), we assign the corresponding label to them.

The statistics of the user-user interaction sign labels generated by these rules for CNN dataset are listed in Table 2. We find that only 13.33% of the user interaction links are assigned with a label, and positive labels are much fewer than negative labels. These labeled links are with high quality due to the rigid rules, and will be used for link regularization. The side effect of such rules is that only a small portion of high-quality links will be utilized. Fortunately, this will be covered up by the text generative model.

4.2 User Stance-based Network Regularizer

Once the labels for user interactions are obtained, either through heuristics designed above or explicit labels provided in the system, a regularizer is then designed according to the intuitions mentioned earlier this section. For each interaction link (u,v,t), we hope its sign $y(u,v,t)$ is consistent with the sign of the dot product of x_u and x_v, i.e., the stances of the two involved users. In other words, if u and v are with different signs, they tend to disagree with each other; and if u and v are with the same sign, they tend to agree with each other. Also, if the two users are with more extreme stances (i.e., bigger absolute values for x_u and x_v), $x_u \cdot x_v$ tends to have a higher absolute value, indicating a higher confidence of the sign prediction for the interaction. Thus our objective is to maximize the following regularization term on all the labeled user-user interactions:

$$R(\mathbf{x}, \mathrm{Y}) = \frac{1}{N_L} \sum_{u=1}^{N_U} \sum_{v=1}^{N_U} \sum_{t=1}^{N_T} y(u, v, t) x_u x_v \tag{5}$$

where $\mathbf{x}$ is the user stance vector for all users, Y is the collection of all the labeled links, and N_L is the total number of user interaction labels. Similarly, the length constraint over $\mathbf{x}$ is also applied here:

$$\mathbf{x}^T \mathbf{x} = N_U \tag{6}$$

where N_U is the number of users in total.

5 THE JOINT MODEL AND LEARNING ALGORITHM

According our discussions in Sections 3 and 4, it is clear that (1) merely using text information may lead to local optimums of user stances that do not reflect opposite standpoints of issues and (2) merely using links may not have a good coverage of users. Thus we propose a joint model, STML (Stance-based Text Generative Model with Link Regularization) that combines the two components together in this section.

5.1 The Joint Model: STML

We now combine Eq. 4 and 5 together, and seek to maximize the following objective function:

$$J = \lambda \cdot l(X, v, \gamma^+, \gamma^-, \beta_B | D, \mu_B) + (1 - \lambda) \cdot R(\mathbf{x}, \mathrm{Y}) - \frac{1}{2\epsilon^2} v^2 \tag{7}$$

$s.t.$

$$0 \le \gamma^+(w), \gamma^-(w) \le 1, \sum_{w=1}^{N_W} \gamma^+(w) = 1, \sum_{w=1}^{N_W} \gamma^-(w) = 1$$

$$v > 0, \mathbf{x}^T \mathbf{x} = N_U$$

where λ is a trade-off parameter which controls the effect of the post content and the user interaction part, and ϵ^2 is a weight on the regularization term on v, which is set as 10^6 in experiments.

5.2 The Learning Algorithm

In this section, we introduce the learning algorithm to estimate the parameters in STML. Maximizing the objective function (Eq. 7) w.r.t. γ^+, γ^-, β_B, $\mathbf{x}$, and v involves the latent parameters $\{z^B_{d,w}\}$ and $\{s(u,d)\}$, therefore, we introduce a two-level expectation-maximization algorithm to estimate the parameters. In particular, the parameters are learned by running the following two steps iteratively until the parameters converge:

(1) E-step: We update the probability of the latent variables $s(u,d)$ when other parameters are fixed.
(2) M-step: We update the remaining parameters as follows:
   (a) E-step: update $p(z^B_{d,w})$, the probability of the word w in post d belonging to background.
   (b) M-step: (1) update word distribution related parameters β_B, γ^+ and γ^- with other parameters fixed; and (2) update user stance parameter $\mathbf{x}$ and v using gradient ascent with other parameters fixed.

Due to space limit, the detailed formulas are not included here.

5.3 Inference On New Users

Given a newly coming user x^{new}_u who has never been observed in the dataset, our model could also predict her stance according to her posts with our learned parameters. In this case, we fix all the parameters γ^+, γ^-, β_B and v and infer the stance for x^{new}_u by maximizing Eq. 4 with gradient ascent algorithm.

6 EXPERIMENTS

In this section, we first describe the two datasets used for experiments. We then introduce two major tasks for evaluating our models: *Relative Stance-based Link Sign Prediction* and *User Stance Prediction*, as well as methods for comparison in Section 6.2. Experimental results are reported and explained in Sections 6.3 and 6.4. Further discussions on the joint effect of content and social interaction, and other aspects of the model are described in Section 6.5.

6.1 Datasets

We evaluate our proposed approach on five popular issues from CNN news commenting system and four issues from 4Forums[1].

We crawled news articles on CNN from 05/27/2014 to 12/04/2014[2] and collected corresponding user comments to these articles via Disqus[3]. News related to 5 hot issues are selected using keywords search, and all the posts under these news articles are retrieved. For each issue, user-user interaction labels are generated according to our proposed heuristic rules in Section 4. For 4Forums, we select discussions related to four issues from the Internet Argument Corpus (IAC) [36], where binary users' stances are annotated for 269 discussions. Table 3 shows the detailed statistics for the two datasets.

6.2 Evaluation Tasks and Comparisons

Here we introduce two evaluation tasks and the experimental setup for our model, followed by the description of compared methods.

6.2.1 Task One: User Stance-based Link Sign Prediction. The first task we consider is predicting whether two users tend to agree with each other for a certain issue. Given two users u_i and u_j, a prediction score is computed as $x_i \cdot x_j$, where x_i and x_j are position variables inferred as described in Section 5.3. A pair of nodes (u_i, u_j) with score above a certain threshold indicates a positive link; otherwise, it is a negative link.

Experiment are carried out on CNN comments data, where 80% of the users are randomly selected to learn the model parameters. The rest 20% are used as held-out users. Predictions are made between observed and held-out users, and among held-out users.

Since gold-standard interaction labels for CNN users are not available, we will first show evaluation results based on the rule-induced labels for held-out data. We further evaluate on human annotated labels for a subset of user pairs, which is described in Section 6.3. Area Under the ROC Curve (AUC) is utilized as the evaluation metric.

6.2.2 Task Two: User Stance Prediction. We then evaluate our model on user stance prediction task, i.e., whether a user is of

[1]http://4forums.com
[2]CNN gradually turned off the comments section at the end of 2014.
[3]https://disqus.com/

issue	source	# threads	# posts	# users	# unique words	# reply links	# user-user interaction	# user-user interaction labels
Bowe Bergdahl	CNN	54	254,235	18,593	30,956	181,935	96,290	12, 094
Gaza Israel Conflict	CNN	83	573,705	28,775	113,437	377,474	202,718	25,840
Immigration	CNN	97	440,677	26,655	54, 292	317,643	160,332	21,637
Hobby Lobby	CNN	41	225,021	16,549	25,252	175,238	84,336	12,984
MH17 Crash	CNN	95	388,467	23,232	53,786	288,169	140,759	18,796
Gun Control	4Forum	464	21,850	447	15,051	20,361	6,546	2,566
Abortion	4Forum	392	31,864	700	20,638	30,530	8,540	3,503
Gay Marriage	4Forum	193	14,343	408	10,145	13,785	4102	1,579
Evolution	4Forum	569	33,060	686	20,662	30,647	10,731	4,201

Table 3: Statistics for the issues from CNN news commenting board and discussion forum 4Forums.com.

Model	**Rule-Based Labels**										**Human Labels**		
	Between Observed and Held-out Users					**Between Held-out Users**							
	Bowe Bergdal	Gaza Israel	Immigrant	Hobby Lobby	MH17	Bowe Bergdal	Gaza Israel	Immigrant	Hobby Lobby	MH17	Bowe Bergdal	Gaza Israel	MH17
LogReg	*78.3*	*79.9*	*76.8*	*72.8*	*77.0*	*75.7*	*77.7*	*78.9*	*68.0*	*77.6*	52.4	50.5	53.2
STML_text	50.9	50.4	51.9	50.1	50.3	56.4	50.1	50.5	**57.6**	52.1	51.7	56.2	57.7
STML_link	52.5	51.5	51.8	53.4	51.0	52.7	52.3	51.6	52.2	55.1	51.0	62.4	59.9
STML	**68.8**	**72.0**	**64.4**	**65.0**	**66.3**	**65.2**	**76.7**	**65.1**	54.3	**70.3**	**69.7**	**79.5**	**71.2**

Table 4: Link sign prediction results by AUC on CNN users. Our STML model outperforms the other two variations where only user interaction (STML_link) is considered or only text is considered (STML_text). STML also achieves comparable performance of LogReg model, which is trained on content of users' posts and biased towards to rule-based labels.

"pro" side or "con" side on a certain issue. Experiments are carried out on IAC discussion forum dataset, and accuracy is reported in accordance with previous work [34, 37].

We first evaluate our model by treating positive stance as "pro" side and negative score as "con". Then another accuracy score is computed by using the alternative assignment. Better accuracy score of the two is reported for our model.

6.2.3 Comparison. For link sign prediction, we consider comparing with a logistic regression model (LogReg)[4]. Unigrams and phrases of the comments from pairwise users are extracted (as in our model) and concatenated as feature vector. Upsampling is utilized to resolve the imbalance issue of the training data.

For user stance prediction, we compare with a logistic regression model (LOGREG) and the state-of-the-art method [34] based on probabilistic soft logic (henceforth PSL). Both methods are supervised which require annotated user stance labels. Features such as *n*-grams, lexical category counts and text lengths are utilized to train the LOGREG model and the local classifier of the PSL model.

For both tasks, we consider two variations of our models: (1) STML_text, which comes from the generative module of STML when $\lambda = 1$, and (2) STML_link, which is the user interaction-based regularization module of STML when $\lambda = 0$.

In our pilot study, we also compared with text similarity-based clustering models, including constrained k-means [5] and constrained spectral clustering [41] to cluster users into 2 groups with opposite stances. However, both systems performed poorly — with AUC scores slightly above 0.5 on CNN comments. The inferior performance is due to link sparsity and unreliable text representation (i.e. TFIDF) and their similarity over noisy text. We thus omit their results here.

For all the STML model used in the experiment, we set the trade-off parameter λ as 0.8, and background work probability μ_B as 0.85, unless stated otherwise.

6.3 Link Sign Prediction

We first experiment with CNN comments data for the task of link sign prediction for pairwise users, with at least one user from the held-out data, against our rule-induced labels. As shown in Table 4, our STML model which considers both text and user interaction information uniformly outperforms the other two model variations that consider only one factor over almost all issues.

It is noteworthy that LogReg performs better when evaluated against user links labeled by our rules. This is because LogReg captures the signal phrases used for heuristic labeling process very well, and thus yields superior performance. For instance, we find that features with highest positive weights from LogReg are "agree", "true", "totally agree" while the ones with highest negative weights are "actually", "nope", "disagree". However, rule-based labels only cover 13.33% user-user interactions in the CNN dataset as shown in Table 2, and LogReg might not be applicable to cases when signal phrases are not observed.

Therefore, we further carry out experiments on human annotated link labels. 400 pairs of users are randomly sampled from the ones with interactions but are not labeled by our rules for each issue of "Bowe Bergdahl", "Gaza Israel" and "MH17" respectively. We recruit two fluent English speakers to read interactions between each pair of users, and ask them to annotate their relative stance as on the same side (+1), on different sides (-1) or not sure (0). Pairs with consistent nonzero labels by our annotators are retained as our test

[4] We use scikit-learn: http://scikit-learn.org/.

set. In total, we have 224 samples for issue "Bowe Bergdahl", 239 samples for issue "Gaza Israel" and 173 samples for issue "MH17". 20% of them are labeled as on the same side, and the rest are on different sides.

From Table 4, we can see that our STML model significantly outperforms LogReg on all of the three topics when evaluated against human annotations. It also leads to better AUC scores than the variations that only consider text or link information. This implies that combining content and user interaction information can better identify user stance on a given topic.

6.4 User Stance Prediction

Model	Accuracy			
	Gun Control	Abortion	Gay Marriage	Evolution
LogReg [34]	67.1	64.9	74.5	77.3
PSL[34]	67.1	65.8	**77.1**	**78.7**
STML_text	54.1	51.7	52.3	54.5
STML_link	**67.8**	73.0	68.4	64.3
STML	66.3	**75.6**	68.6	64.7

Table 5: Accuracy with standard deviation on user stance prediction for discussion forum data (IAC) on four popular issues. Our STML obtains the best accuracy on "abortion" and comparable performance on "gun control" issues with supervised PSL-based Method and Logistic Regression [34].

Here we evaluate our model on user stance prediction task on IAC discussion forum dataset. Following previous work [34], 10 fold cross-validation is used. We report accuracy on four popular issues in Table 5.

As can be seen, our STML model achieves better accuracy on the "abortion" issue compared to the supervised methods (LogReg and PSL-based method). It also achieves comparable performance on "gun control". There are two main reasons why LOGREG and PSL-based method [34] performs better on the other topics. Firstly, it is easier to distinguish between contrasting opinions according to text information on the issue "gay marriage" and "evolution". Thus the human annotated stance labels are of higher quality on these two issues, leading to better performance of supervised methods on these two issues than the others. Moreover, both LOGREG and PSL utilize richer features and are trained on high-quality user stance labels, while our model only utilizes the weak guidance from rule-induced user interaction labels for learning.

Furthermore, both STML and STML_link achieve significantly better performance than STML_text for discussion forum users. This is due to the reason that users on discussion forums interact more than on news commenting systems, and this also leads to richer interaction information that can be leveraged to enhance our models.

6.5 Discussions

In this section, we provide further analysis on different aspects of our model. We first analyze whether user activity level affects the prediction performance. Furthermore, we present a study on whether leveraging more user interaction information would further facilitate with user stance prediction. Lastly, we comment on the trade-off between content and user interaction.

Does Activity Level Tell About User Stance? We start with investigating whether our model will achieve better performance for users with higher activity level, e.g., constructing more posts. Users in CNN data are divided into 3 groups according to their post number: users with less than 10 posts, at least 10 but less than 50 posts, and with at least 50 posts. Then our model is evaluated on link sign prediction for unobserved labels between users in each group and all the other users. Table 6 demonstrates the performance of our STML method for different groups when evaluated against rule-induced labels. Our model achieves the best AUC on group of users with at least 50 posts on almost every issue.

	Btw Observed and Held-out Users			Btw Held-out Users		
# Posts	[0, 10)	[10, 50)	[50, ∞)	[0, 10)	[10, 50)	[50, ∞)
Bowe Bergdahl	60.0	67.3	*70.3*	60.6	*65.4*	65.3
Gaza Israel Conflict	63.0	67.8	*72.8*	56.9	*78.1*	78.0
Immigrant	58.0	63.3	*64.8*	54.4	59.1	*66.0*
Hobby Lobby	58.7	64.2	*65.5*	51.9	*54.9*	51.5
MH17 Crash	57.4	61.4	*68.5*	69.6	55.4	*73.1*

Table 6: Link sign prediction results by AUC for STML model on CNN users with different activity level on rule-based labels. Users are divided into three groups according to their post numbers which fall in $[0, 10)$, $[10, 50)$, **or** $[50, \infty)$.

We further experiment with different variations of STML model with regard to human annotated link sign labels on the issue of "Bowe Bergdahl", "Gaza Israel" and "MH17". Table 7 shows that our models achieves better AUC results for the user group with higher activity level (i.e., more posts). Nevertheless, this is not observed for other comparisons.

Can More Interactions Further Benefit Learning? We further examine whether changing the amount of user interaction labels used for learning will affect performance on relative stance prediction. We thus vary the percentage of observed user interaction labels in learning, and the performance by AUC is displayed in Figure 2. Our STML model achieves stable performance when 60% or more of the interaction labels are applied for learning for all five issues. This suggests that with a modest amount of user interaction information, our model is able to learn word distributions for users of different stances and thus predict with reasonable performance.

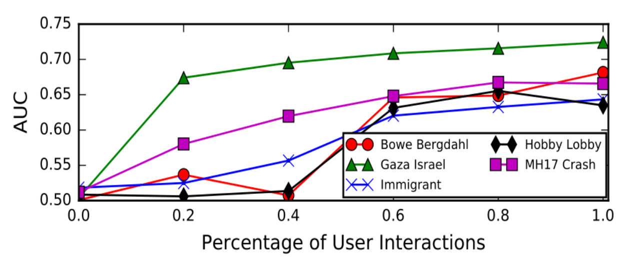

Figure 2: Our STML model learned with varying amount of user interactions on link sign prediction on CNN comments.

More about the Joint Effect of Content and User Interaction. Here we want to study the effect of model parameter λ, which

	Bowe Bergdahl			Gaza Israel			MH17		
# Posts	[0, 10)	[10, 50)	[50, ∞)	[0, 10)	[10, 50)	[50, ∞)	[0, 10)	[10, 50)	[50, ∞)
LogReg	52.2	52.2	50.7	50.0	51.7	50.1	51.4	50.0	54.0
STML_text	55.3	52.3	54.9	55.5	57.7	55.5	60.0	62.2	56.3
STML_link	66.2	52.0	58.5	61.8	63.2	61.7	63.6	51.6	62.5
STML	66.4	67.3	**71.8**	70.4	76.2	**81.0**	62.6	51.4	**72.4**

Table 7: Link sign prediction results by AUC on CNN users with different activity level on the human annotated labels. Our models have even better stance prediction performance when more posts are available for users.

controls the trade-off between content and user interaction. We vary the value of λ, and compute the average AUC value for the task of relative stance prediction over five issues in CNN data. From Figure 3, we can see that the best performance is obtained when λ falls in the range of $[0.7, 0.9]$. In general, our model benefits from learning based on both content information and user interaction (i.e., λ is between 0.1 and 0.9).

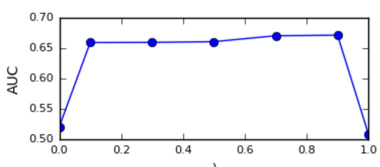

Figure 3: Trade-off study on content and user interaction by using different λ value. AUC is reported for our STML model on issues of CNN dataset.

7 CASE STUDY ON WORD DISTRIBUTIONS

In this section, we investigate whether our models can learn meaningful word distributions for users of different stances given a specific issue.

We list the top 10 words or phrases from the word distributions learned for opposite stances by STML in Table 8 along with sample posts. From the representative terms, our model is able to capture the *word usage difference derived from ideology*. For instance, on the issue of "Bowe Bergdahl", the usage of "liberals" and "democrats" is signified by the side who opposes prisoner exchange and blames "Obama". Meanwhile, the other side fights back via attacking "republicans" and "conservatives". Similar observation is found on "Immigration" as well due to ideological polarization.

In addition to discovering ideology of users, our model also captures *salient aspects of arguments delivered when users confront each other.* For issue "Gaza Israel Conflict", where religion has been an important factor to provoke disagreement among users, we see arguments concerning "Muslims" or "Jews" from conflicting standpoints. Similarly, posts for "MH17 Crash" show that some users attribute the crash of the Malaysia Airlines Flight 17 to "Putin" and "Russian", while users on the other side blame "USA".

Compared with STML, we also display the output word distributions by STML_text model which only considers content information. From Table 9, we can see that STML_text separates terms mainly based on topic, rather than the conflicting arguments. This means that by introducing user interactions, our model better captures the controversial aspects, which facilitates stance prediction for unseen users only according to their posts.

8 RELATED WORK

Stance Classification on Debate Forums. Recently, stance classification attracts a significant amount of research attention from both text mining and natural language processing communities [1, 7, 9, 15, 25, 32–35, 37, 38]. Previous work mostly employs supervised methods with rich feature set, which requires large amount of human annotation. Conditional random fields (CRFs) is also investigated to jointly determine the stances of both the post and its sentences [14] . Recent work by [30] presents a statistical model for stance classification based on the extracted arguments. Qiu et al. [28] propose a unified model combining user profiling, user post and interaction modeling and user stance modeling to infer user stances on a set of issues, where high quality annotations for user stance, user interactions and fine-grained user attributes are required. Fang et al. [10] aim to find the contrastive opinions on political texts with no social interaction information. Most of the aforementioned models are not applicable to our problem since they are supervised methods that need either human annotations or highly domain specific knowledge guided annotations of user stances. This is the gap we aim to fill in this work.

Stance Classification on Twitter In addition to previous work on stance prediction for debate forums, stance prediction on Twitter [3, 4, 9, 17, 18, 25, 26, 31] has also gained increasing popularity in recent years. There has been work that detects user's political stances purely from links on Twitter [12]. For methods that also utilize text, the SemEval 2016 Stance Detection for Twitter shared task [25] aims to detect the stance of individual tweets on given opinion targets. Ebrahimi et al. [9] use relational boostrapping to classify user stances on given targets with weakly supervision from a set of stance-indicative patterns of tweets. The task of stance prediction on previously unseen targets is studied in [3], where bidirectional conditional LSTM [11] encodings are generated for both tweets and opinion targets for stance classification. Lukasik et al. [23] classify the stances of tweets with respect to rumours by exploiting both temporal and textual information. Different from our problem setting, most of these methods focused on predicting whether a tweet is "for" or "against" a given opinion target. Also, domain knowledge for the opinion targets are often required by those weakly supervised methods for tweet stance prediction, which limits their applicability to other online discussion dataset.

Topic Sentiment Mining. Our work is also in line with topic sentiment analysis. Joint topic and sentiment modeling is first studied in [24], who propose to automatically learn the latent topical facets of the Weblog collection and the associated sentiments. In order to detect the topics and associated sentiments from the text, LDA-based

Bowe Bergdahl		Gaza Israel		Immigrant		MH17	
Stance 1	Stance 2	Stance 1	Stance 2	Stance 1	Stance 2	Stance 1	Stance 2
republicans	**obama**	hamas	**jew**	**republicans**	**obama**	**putin**	**usa**
gop	**liberal**	**muslim**	**jews**	**the gop**	**liberals**	**russian**	**kiev**
the gop	**deserter**	**free palestine from arad terror**	netanyahu	**gop**	**democrats**	russians	ukrainian
conservatives	**liberals**	egypt	**isreal**	boehner	**illegals**	the russians	iraq
allegedly	a deserter	yawn	israeli	**republican**	liberal	**kremlin**	**americans**
Exu **republican**	arabs	hitler	**conservatives**	**conservatives**	the illegals	russia	american
reagan	traitor	hamass	aipac	congress	illegal aliens	vodka	**cia**
conservative	he deserted	syria	part of this genocide	perry	obama is	comrade	poroshenko
fox news	obama is	hamas_is	**cut all aid to israel**	the republicans	**dems**	**russian troll**	com watch
right wing	susan rice	**muslims**	zionists	conservative	citizens	huh	youtube

Issue	Discussion
Bowe Bergdahl	**dragonemp:** So when you **liberals** can no longer deny the truth that **obama**'s newest hero was **deserter**, you guys simply resort to mindless personal attack? typical **liberal**. **disqus_jstf5729hx:** Now his family is receiving death threats this is unacceptable. What type of country is this? Where's the humanity? **Republican** lawmakers, the **republican** party, **Fox news**, **conservative** pundits and even **republican** voters are scums of the earth. And all other american citizens who are judging this soldier and his family without even knowing all the facts are pathetic hypocrites.
Gaza Israel	**princeduomarr:** **Israel**'s actions only make the world hate **jews** more. **disqus_v1TJlyTz8e**: The same applies to **muslims**. Islam's actions will only cause the world to hate **muslims** more and more.
Immigrant	**disqus_DyF69AkLq3**: Why are **liberals** too stupid to understand the difference between legal and **illegal**? **pkmyt1**: Why does the **GOP**/TP **conservatives** think the constitution and laws are only for them?
MH17	**disqus_aEmd4h22ye**: **kiev** shot it down **cia** gaves the order—soon world will hate **usa** again a little bit more then now **disqus_mBjaqI1BJ9** : **Russia** invaded a Europea country and recently shot down a plane with Europeans in it. Well done **Russia** for creating enemies.

Table 8: Upper table: Top words and phrases discovered by our STML model which utilizes both content and user interaction. Stance 1 and Stance 2 denote the learned opposite viewpoints. Representative words and phrases frequently used by each stance are highlighted in bold with the same color of corresponding stance. Lower table: Sample posts from users of different stances. Stance-specific words or phrases are in bold of different color.

Bowe Bergdahl		Gaza Israel Conflict		Immigration		MH17 Crash	
Stance 1	Stance 2	Stance 1	Stance 2	Stance 1	Stance 2	Stance 1	Stance 2
mr president you freed a desert	liberal	womenshealthmag	youtube	mexico	obama	eu	plane
the gop	allegedly	schools	com watch	usa	president	europe	evidence
mr president	republicans	hamass	unrwa	mexicans	congress	china	rebels
genius	the rmy	tunnels	free palestine from arab terror	immigrants	the gop	country	missile
american	near	occupation	yawn	countries	vote	world	separatists
barry	unit	islam is garbage	https	pay	boehner	money	shot down
gop	black	asshat	part of this genocide	food	bill	germany	youtube
anyone miss w	taliban	egypt	disqus	parents	republicans	sanctions	flight
iraq	im	translation	cut all aid to israel	canada	senate	gas	video
rwnj	bowe	body armor	spam bot	mexican	dems	america	investigation

Table 9: Top words and phrases discovered by our STML_TEXT model which only considers content of users' posts. Stance 1 and Stance 2 denote the learned opposite viewpoints.

Joint Sentiment/Topic (JST) model is studied to allow each document has topic distributions for different sentiments [20, 21]. Both models use some prior information generated either from online sentiment retrieval services or sentiment lexicons. One important category of topic sentiment analysis is aspect-sentiment mining, which aims to identify the sentiment of online reviews with respect to aspects of reviewed objects [29, 45]. This line of work mainly focuses on inferring user's sentiment on various topics, which is different from the setting of our stance prediction problem.

Joint Content and Social Interaction Modeling. Existing joint content and interaction models mainly focus on bibliographic networks and social networks, where a link reflects the proximity of two entities. Proximity is often explained by nodes with similar latent attributes, such as people belonging to similar communities [43], pieces of text that share similar topic distributions [6, 27], and actors that are alike in terms of tastes or behaviors [8, 16, 19, 40, 42–44]. Those nodes are assumed to be connected with high probability, and the nodes are then inferred from the data by optimizing a global objective. Gu et al. [13] incorporate topic models with matrix factorization on legislators' voting records to estimate their stances on various issues. While those methods target network structure detection and latent user attribute understanding via node membership distribution over various topics/communities, our goal is to associate every user with a signed numeric value indicating their relative polarity on each topic. Moreover, in our model, user polarities are regularized by their signed interactions carrying either agreement or disagreement attitudes, while the interactions in social networks often imply homophily.

9 CONCLUSIONS

In this paper, we have presented a Stance-based Text Generative Model with Link Regularization (STML) for user stance prediction in online discussions. On the converse to most of the existing approaches for stance prediction favoring supervised learning, STML aims at predicting user stances and producing word distributions for two opposite stances on a specific issue simultaneously in a weakly guided fashion. Moreover, different from the methods based on domain specific knowledge for stance-indicative pattern detection, STML depends on simpler guidance from user interaction patterns, which makes it generalizable to online discussions of other domains of interest. Experimental results on users in news commenting system and online discussion forum show that our model can outperform non-trivial comparisons, and produce comparable results when compared with state-of-the-art supervised learning-based methods for stance prediction tasks.

ACKNOWLEDGEMENTS

This work was supported by National Science Foundation grants IIS-1566382, IIS-1705169, CAREER Award 1741634, and NVIDIA GPU grants. We thank Bingyu Wang for his help on data collection. We would also like to thank the anonymous reviewers for their valuable comments.

REFERENCES

[1] Pranav Anand, Marilyn Walker, Rob Abbott, Jean E Fox Tree, Robeson Bowmani, and Michael Minor. 2011. Cats rule and dogs drool!: Classifying stance in online debate. In *Proc. of the 2nd workshop on computational approaches to subjectivity and sentiment analysis*. Association for Computational Linguistics, 1–9.

[2] Rie Kubota Ando and Tong Zhang. 2007. Learning on graph with Laplacian regularization. (2007).

[3] Isabelle Augenstein, Tim Rocktäschel, Andreas Vlachos, and Kalina Bontcheva. 2016. Stance detection with bidirectional conditional encoding. *arXiv preprint arXiv:1606.05464* (2016).

[4] Roy Bar-Haim, Indrajit Bhattacharya, Francesco Dinuzzo, Amrita Saha, and Noam Slonim. 2016. Stance Classification of Context-Dependent Claims. (2016).

[5] PS Bradley, KP Bennett, and Ayhan Demiriz. 2000. Constrained k-means clustering. *Microsoft Research, Redmond* (2000), 1–8.

[6] Jonathan Chang and David M Blei. Relational Topic Models for Document Networks.

[7] Wei-Fan Chen and Lun-Wei Ku. 2016. UTCNN: a Deep Learning Model of Stance Classification on Social Media Text. *arXiv preprint arXiv:1611.03599* (2016).

[8] Yoon-Sik Cho, Greg Ver Steeg, Emilio Ferrara, and Aram Galstyan. 2016. Latent space model for multi-modal social data. In *Proceedings of the 25th International Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 447–458.

[9] Javid Ebrahimi, Dejing Dou, and Daniel Lowd. 2016. Weakly Supervised Tweet Stance Classification by Relational Bootstrapping.. In *EMNLP*. 1012–1017.

[10] Yi Fang, Luo Si, Naveen Somasundaram, and Zhengtao Yu. 2012. Mining contrastive opinions on political texts using cross-perspective topic model. In *Proc. of the fifth ACM Int. Conf. on Web search and data mining*. 63–72.

[11] Alex Graves and Jürgen Schmidhuber. 2005. Framewise phoneme classification with bidirectional LSTM and other neural network architectures. *Neural Networks* 18, 5 (2005), 602–610.

[12] Yupeng Gu, Ting Chen, Yizhou Sun, and Bingyu Wang. 2017. Ideology Detection for Twitter Users via Link Analysis. In *Int. Conf. on Social Computing, Behavioral-Cultural Modeling and Prediction and Behavior Representation in Modeling and Simulation*. 262–268.

[13] Yupeng Gu, Yizhou Sun, Ning Jiang, Bingyu Wang, and Ting Chen. 2014. Topic-factorized ideal point estimation model for legislative voting network. In *Proc. of the 20th ACM SIGKDD Int. Conf. on Knowledge discovery and data mining*. 183–192.

[14] Kazi Saidul Hasan and Vincent Ng. 2013. Stance Classification of Ideological Debates: Data, Models, Features, and Constraints.. In *IJCNLP*. 1348–1356.

[15] Kazi Saidul Hasan and Vincent Ng. 2014. Why are You Taking this Stance? Identifying and Classifying Reasons in Ideological Debates.

[16] Xin Huang, Hong Cheng, and Jeffrey Xu Yu. 2015. Dense community detection in multi-valued attributed networks. *Information Sciences* 314 (2015), 77–99.

[17] Kristen Johnson and Dan Goldwasser. "All I know about politics is what I read in Twitter": Weakly Supervised Models for Extracting Politicians' Stances From Twitter.

[18] Kristen Johnson and Dan Goldwasser. 2016. Identifying Stance by Analyzing Political Discourse on Twitter. *NLP+ CSS 2016* (2016), 66.

[19] Jiwei Li, Alan Ritter, and Dan Jurafsky. 2015. Learning multi-faceted representations of individuals from heterogeneous evidence using neural networks. *arXiv preprint arXiv:1510.05198* (2015).

[20] Chenghua Lin and Yulan He. 2009. Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management*. ACM, 375–384.

[21] Chenghua Lin, Yulan He, Richard Everson, and Stefan Rüger. 2012. Weakly supervised joint sentiment-topic detection from text. *Knowledge and Data Engineering, IEEE Transactions on* 24, 6 (2012), 1134–1145.

[22] Jialu Liu, Jingbo Shang, Chi Wang, Xiang Ren, and Jiawei Han. 2015. Mining quality phrases from massive text corpora. In *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*. ACM, 1729–1744.

[23] Michal Lukasik, PK Srijith, Duy Vu, Kalina Bontcheva, Arkaitz Zubiaga, and Trevor Cohn. 2016. Hawkes processes for continuous time sequence classification: an application to rumour stance classification in twitter. In *Proceedings of 54th Annual Meeting of the Association for Computational Linguistics*. Association for Computational Linguistics, 393–398.

[24] Qiaozhu Mei, Xu Ling, Matthew Wondra, Hang Su, and ChengXiang Zhai. 2007. Topic sentiment mixture: modeling facets and opinions in weblogs. In *Proceedings of the 16th international conference on World Wide Web*. ACM, 171–180.

[25] Saif M Mohammad, Svetlana Kiritchenko, Parinaz Sobhani, Xiaodan Zhu, and Colin Cherry. 2016. Semeval-2016 task 6: Detecting stance in tweets. *Proceedings of SemEval* 16 (2016).

[26] Saif M Mohammad, Parinaz Sobhani, and Svetlana Kiritchenko. 2016. Stance and sentiment in tweets. *arXiv preprint arXiv:1605.01655* (2016).

[27] Ramesh M Nallapati, Amr Ahmed, Eric P Xing, and William W Cohen. 2008. Joint latent topic models for text and citations. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 542–550.

[28] Minghui Qiu, Yanchuan Sim, Noah A Smith, and Jing Jiang. Modeling User Arguments, Interactions, and Attributes for Stance Prediction in Online Debate Forums. In *Proceedings of 2015 SIAM International Conference on Data Mining*.

[29] Arti Ramesh, Shachi H Kumar, James Foulds, and Lise Getoor. 2015. Weakly supervised models of aspect-sentiment for online course discussion forums. In *Annual Meeting of the Association for Computational Linguistics (ACL)*.

[30] Parinaz Sobhani, Diana Inkpen, and Stan Matwin. 2015. From argumentation mining to stance classification. *NAACL HLT 2015* (2015), 67.

[31] Parinaz Sobhani, Saif M Mohammad, and Svetlana Kiritchenko. 2016. Detecting Stance in Tweets And Analyzing its Interaction with Sentiment. *cÑĄ2016 The* SEM 2016 Organizing Committee. All papers cÑĄ2016 their respective authors. This proceedings volume and all papers therein are licensed under a Creative Commons Attribution 4.0 International License. License details: http://creativecommons.org/licenses/by/4.0* (2016), 159.

[32] Swapna Somasundaran and Janyce Wiebe. 2009. Recognizing stances in online debates. In *Proc. of the Joint Conf. of the 47th Annual Meeting of the ACL and the 4th Int. Joint Conf. on Natural Language Processing of the AFNLP: Volume 1-Volume 1*. Association for Computational Linguistics, 226–234.

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

[34] Dhanya Sridhar, James Foulds, Bert Huang, Lise Getoor, and Marilyn Walker. 2015. Joint models of disagreement and stance in online debate. In *Annual Meeting of the Association for Computational Linguistics (ACL)*.

[35] Dhanya Sridhar, Lise Getoor, and Marilyn Walker. 2014. Collective stance classification of posts in online debate forums. *ACL 2014* (2014), 109.

[36] Marilyn A Walker, Rob Abbott, and Joseph King. A Corpus for Research on Deliberation and Debate.

[37] Marilyn A Walker, Pranav Anand, Robert Abbott, and Ricky Grant. 2012. Stance classification using dialogic properties of persuasion. In *Proceedings of the 2012 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*. Association for Computational Linguistics, 592–596.

[38] Marilyn A Walker, Pranav Anand, Rob Abbott, Jean E Fox Tree, Craig Martell, and Joseph King. 2012. That is your evidence?: Classifying stance in online political debate. *Decision Support Systems* 53, 4 (2012), 719–729.

[39] Lu Wang and Claire Cardie. 2014. A Piece of My Mind: A Sentiment Analysis Approach for Online Dispute Detection. In *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics*, Vol. 2. 693–699.

[40] Xiao Wang, Di Jin, Xiaochun Cao, Liang Yang, and Weixiong Zhang. 2016. Semantic Community Identification in Large Attribute Networks.

[41] Xiang Wang, Buyue Qian, and Ian Davidson. 2014. On constrained spectral clustering and its applications. *Data Mining and Knowledge Discovery* (2014), 1–30.

[42] Jaewon Yang, Julian McAuley, and Jure Leskovec. 2013. Community detection in networks with node attributes. In *Data Mining (ICDM), 2013 IEEE 13th international conference on*. IEEE, 1151–1156.

[43] Tianbao Yang, Rong Jin, Yun Chi, and Shenghuo Zhu. 2009. Combining link and content for community detection: a discriminative approach. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*. ACM, 927–936.

[44] Yuan Zhang, Elizaveta Levina, and Ji Zhu. 2015. Community detection in networks with node features. *arXiv preprint arXiv:1509.01173* (2015).

[45] Wayne Xin Zhao, Jing Jiang, Hongfei Yan, and Xiaoming Li. 2010. Jointly modeling aspects and opinions with a MaxEnt-LDA hybrid. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*. Association for Computational Linguistics, 56–65.