CMPS 245, Winter 17 Project: Ideology-Backed Stance Classification

Eriq Augustine
University of California, Santa Cruz
eagusti@ucsc.edu

Dhawal Joharapurkar
University of California, Santa Cruz
djoharap@ucsc.edu

Abstract

The determination of stance in two-sided discussions/debates in online debate forums is a new and interesting problem in opinion mining. We target the task of classifying stance in ideological debates on online debate forums, and this is a rather challenging problem due to the nature of debate setting and the language used. (Hasan and Ng,) had introduced the notion of using ideology as an inter-post, extra-linguistic constraint which are implemented using Integer Linear Programming. In this work, we seek to improve (Sridhar et al., 2015) approach on jointly-modeling disagreement and stance in online debate forums by integrating ideology as a latent variable in the joint model. We show improvement of a few percentage points over the baseline set by (Sridhar et al., 2015). Additionally, we have also created a new ideology corpus of around 1.5 million tweets by 645 official members of the U.S. Congress.

1 Introduction

In the early days of opinion-mining from text, there was a lot of work on finding the polarity expressed in the text content. Recent incarnations of the opinion mining tasks include classifying the stance expressed in posts in online debate forums. Topics debated on online debate forums are largely two-sided and posts are either written "for" or "against" the topic under consideration. These debate forums are unlike traditional debate settings such company internal discussions, political debates, etc. Whereby the former contain usage of varied language instruments like sarcasm, emotions, insults, innuendos, questions, retorts, etc. The presence of such linguistic characteristics make prediction of stance in online debate forums more challenging than in traditional debate forums.

A growing body of work on classification for stance has found it to be a challenging problem. As the interactions on these social media debate websites are inherently dialogic in nature, they have also proved useful for the computational modeling of dialogue. (Sridhar et al., 2014) has demonstrated how modeling disagreement and stance jointly leads to improved prediction over just trying to model stance as so many works have tried on the (Walker et al., 2012). This is because the disagreement between authors provides strong evidence that their stance on the topic will be different, given the polar nature of these debates. Prediction of user stance can support the identification of social or political groups, and can be valuable user modeling information for recommender systems.

In this piece of work, we are trying to leverage the use of ideology and the dependency structures present between ideology and stance by jointly modeling ideology, stance and disagreement. Our intuition is that for a person who is "pro" gay-marriage is very likely to be "pro" gun-control or "pro" marijuana-legalization. So, we see that there is a connection between the stances a person has across topics, and is intrinsically determined by the ideology of the person.

2 Related Work

(Hasan and Ng,), in their work try to capture the effect of ideologies on stance. They encode this by defining extra-linguistic inter-post constraints called Ideology Constraints (IC). ICs are cross-domain, author-based constraints and are applicable on a per-author basis across all topics that the author participates in. The IC intends to capture the intuition that, the stance of an author on a topic in some part, is determined by their ideologies and that there might be a correlation between their stances on different topics.

They implement the ICs by first defining conditional probabilities. They find $P(stance(d_q) = s_d|stance(d_p) = s_c)$, where (1) $d_p, d_q \in$ Domains, (2) $s_c, s_d \in$ for, against, and (3) $d_p \neq d_q$. To compute conditional probability $P(stance(d_q) = s_d|stance(d_p) = s_c)$ they take the ratio from the set of authors who posted in both domains d_p and d_q the number of authors who had stance s_d in d_q and s_c in d_p by number of authors who had stance s_c in d_p.

Their work however, uses these probabilities and converts them to hard linear constraints using Integer Linear Programs (ILP) by constraining the probabilities calculated above against a threshold which is tuned using development data.

However, in doing so, they don't capture the correlation between disagreement and ideology, nor are they capable of extrapolating the use of ideology to predict stances on domains not encountered in training data. Thereby, our method of building a HL-MRF (as discussed ahead) allows us greater flexibility in terms of how to model ideology in problem of prediction of stance in online debate forums.

3 Problem Description

The objective of this set of experiments is to jointly use ideology, and disagreement to collectively classify stance in online debate forums settings, specially for ideological debate domains. The previous sections have indicated that this is a new way of modeling ideology and is different from the other works which had spun ideologies into a constraint problem, and that this is a more extensible way of modeling ideology. Once we found stance predictions by jointly modeling the stance, disagreement and ideology variables, we used the truth labels in the IAC to compute our model's performance statistics. Most commonly, the performance index used for stance classification tasks is accuracy.

4 Dataset

We are working primarily with two data sets, namely ("Datasets From Transcripts Of US Congressional Floor Debates") (ConVote) and the Internet Argument Corpus (IAC) (Walker et al., 2012) (Abbott et al., 2016). ConVote contains the textual transcripts of the speeches given in the US Congressional Floor debates and information about the authors - which party they associate themselves with.

From the IAC, we choose to work with only ideological domains, namely, Gay Marriage, Abortion, Evolution, and Gun Control. Note that the pro stance on each topic correlates with a liberal ideology while a con stance correlates with a conservative ideology, since the topics are pretty co-linear with respect to the ideology a person could have. We hope to extend this work to incorporate ideologies that lie somewhere in-between fully conservative or fully liberal.

4.1 Twitter

Because the ConVote data is fairly dated, we considered moving away from ConVote to a more modern dataset. For this, we scraped data from political entities off of Twitter. We were able to use lists put out by CSPAN[1] and the Twitter Government project[2] to gather the Twitter handles for current and recent member of the US congress, US governors, as well as congressional committees.

We then labeled all the Twitter handles with the party affiliation (Democrat, Independent, or Republican) associated with each politician. All US politicians at this level have a declared party affiliation.

However, some handles had to be thrown out either because it belonged to a bipartisan committee, such as the Ways and Means Committee (@WaysMeansCmte), or the Twitter account has changed hands between two different parties. For example, the account @GovernorVA is for the governor of Virginia. The account was created in 2010 and was used by Bob McDonnell (a member of the Republican party). However in 2014, Terry McAuliffe (a member of the Democratic party) was elected as the governor of Virginia and started using the @GovernorVA Twitter account. All accounts with this weakness were identified and thrown away.

After all labeling and filtering was finished, we ended up with 647 Twitter handles. We then scraped all tweets for each account resulting in 1,442,468 tweets. Table 1 shows a breakdown of the data by party.

The initial data is publicly available[3]. We have

[1] https://twitter.com/cspan/lists/members-of-congress/members

[2] https://twitter.com/gov/lists/us-house/members

[3] https://linqs-data.soe.ucsc.edu/public/twitter-ideology/

Party	# Users	# Tweets	Mean Tweets Per User
Democrat	296	683342	2308
Republican	346	751846	2172
Independent	3	7280	2426
Overall	647	1442468	2236

Table 1: Twitter data by party

made the data available in two forms. First, the full tweets divided by user containing all the information the Twitter API returns for each tweet. Second, a single tab-separated file that contains just the user handle, tweet id, party of user, and text of the tweet.

5 Methodology

In order to be able to create our joint model, we will need some information about the ideology of authors of posts on the IAC. For this, we train a simple binary text classifier that predicts how liberal or conservative people are on the ConVote dataset. Then, we used this trained classifier to predict an author's ideology from their post on the IAC corpus. This classifier was a fairly simple one - a binary logistic regression. To reduce our vocabulary size and keep the sparsity of our data under check, we limited our vocabulary to the intersection of the vocabulary of ConVote and IAC. The set of features that we used are - i) TF-IDF vectors with n_grams between 1 and 3 ii) count vectors of LIWC categories. We are aware that using context based features will transfer better than using statistical features, but, we only wanted a weak signal for ideology, and hence we didn't build a robust classifier and created a better suite of features. However, as we see the final accuracy statistics, we realized that using synthetic ideology data produced better results than using the ideology labels generated by our weak local classifier by 4.32%. This tells us that we can get by with a weak classifier and rely on the joint modeling to propagate stance information across authors and between topics, however, a stronger, more robust classifier that transfers over better will allow us to get finer improvements in our predictions.

Next, we use these predictions of ideologies for online debate forum authors as seed ideologies in our joint inference model that collectively classifies stance, disagreement, and ideology. Our intuition behind this, as stated earlier, is that a person's stance on various topics are correlated with their ideology and people of similar ideology take similar stances on congruent issues.

6 Hinge-loss Markov Random Fields and Probabilistic Soft Logic

We used a spacial class of Markov random field (MRF) known as a hinge-loss Markov random field (HL-MRF). HL-MRFs model all MRF potential functions as a linear hinge loss, thereby making MAP inference a convex optimization problem (Bach et al., 2015). The variables in our HL-MRF are all defined over a continuous range between 0 and 1. The functional definition for a HL-MRF joint probability density is

$$P(Y|X) = \frac{1}{Z}exp(-\sum_{r=1}^{M}\lambda_r\phi_r(Y,X))$$

where X is the set of observed variables, and Y is the set of target variables. Both X and Y are defined over $[0, 1]$. λ is a vector of weights for each potential function. Z is a normalization constant. ϕ is a hinge-loss potential specified by some linear function l_r and an exponent $p_r \in 1, 2$.

$$\phi_r(Y,X) = (max(l_r(Y,X),0))^{p_r}$$

Probabilistic Soft Logic (PSL) provides a first order logic-like language to provide templates for a HL-MRF. PSL allows us to define *rules* that combine soft-logic predicates with soft-logic operations. PSL uses a Lukasiewicz interpretation (Klir and Yuan, 1995) of some of the first order logic operations to forulate a linear potential function for use in a HL-MRF. Lukasiewicz logic provides a continuous representation of many first order logic operators:

$$x_1 \wedge x_2 = Max(0, x_1 + x_2 - 1)$$
$$x_1 \vee x_2 = Min(1, x_1 + x_2)$$
$$\neg x = 1 - x$$

7 PSL Model

In this section, we will describe the components of our PSL program and the function that they serve.

7.1 Predicates

7.1.1 Observed Predicates

Observed predicates are predicates that get their value directly from data and do not change over the course of the program.

LocalDisagree(Author1, Author2, Topic)
The output from a local disagreement classifier.

This predicate is a measure of the initial disagreement between authors *Author1* and *Author2* over *Topic*.

LocalIdeology(Author, Ideology)
The output from a local ideology classifier. This predicate is a measure of the initial affinity *Author* has for *Ideology*. The local classifier treats ideology as a binary choice representing the liberal or conservative. Therefore, each author will have two instantiations of this predicate. Note however, that the latent third ideology will not be used here since it is not considered by the local classifier.

LocalStance(Author, Topic, Stance)
The output from a local stance classifier. This predicate is a measure of the initial confidence we hold for a specific author having a specific stance for a specific topic. *Stance* will be either "PRO" or "CON".

Participates(Author, Topic)
A discrete observation of whether or not *Author* has participated in any debates focused on *Topic*.

Responds(Author1, Author2, Topic)
A discrete observation of whether or not *Author2* has directly responded to *Author1* on a debate about *Topic*.

7.1.2 Inferred Predicates

Inferred predicates are the target of joint inference. Every instantiation of these predicates will get a continuous value in the range $[0, 1]$.

Stance(Author, Topic, Stance)
The inferred value for a specific author's stance on a specific topic.

Ideology(Author, Ideology)
The inferred value for an author's ideology. Unlike *LocalIdeology*, this predicate will have three values representing liberal, conservative, and other.

Disagree(Author1, Author2, Topic)
The inferred disagreement between two authors on a specific topic.

7.2 Constraints

Constraints in a PSL program allow us to encode domain constraints or desired behavior into a PSL program.

7.2.1 Symmetry Constraints

Symmetry constraints are a convenience utility that allows us to specify less rules by allows certain arguments in a predicate to swap places.

Disagree(A1, A2, T) = Disagree(A2, A1, T)
This constraint allows disagreement to be symmetric. If author *A1* and *A2* disagree, then *A2* and *A1* also disagree.

7.2.2 Functional Constraints

Functional constraints allow us to specify a limit on the possible values something can take. In the discrete case this would be saying that some relation is a function, ie a value in the domain maps to exactly one value in the co-domain. In the continuous case, we just enforce that all values for some entity sum to one. Without functional constraints, it may be possible for the optimizer to trivially satisfy the problem by setting all values to one or zero.

Stance(A, T, +S) = 1
This constraint states that for a specific topic, and author can only have one stance. In the discrete case, this would push either "PRO" or "CON" to one and the other to zero. In the continuous case, it just ensure that the confidence in "PRO" and "CON" sum to one.

Ideology(A, +I) = 1
This constraint states that an author can only have one ideology.

7.3 Rules

The rules have been grouped into logical units that can be turned on and off for evaluation.

7.3.1 Initialization

These rules load data from local classification into the inferred predicates.

LocalStance(A, T, S) $\rightarrow$ Stance(A, T, S)
LocalDisagree(A, T, S) $\rightarrow$ Disagree(A, T, S)
!LocalDisagree(A, T, S) $\rightarrow$!Disagree(A, T, S)

7.3.2 Disagreement Affects Stance

Disagrees(A1, A2, T) & Stance(A1, T, S1) → Stance(A2, T, S2)
States that authors who disagree on a topic should have differing stances on that topic. Note that because of symmetry in disagreement, all variations of this rule are covered.

!Disagrees(A1, A2, T) & Participates(A2, T) & Stance(A1, T, S) → Stance(A2, T, S)
States that if two authors do not disagree on a topic, then they have the same stance. Note that *Participates* here is acting as a scoping predicate and ensures that author *A2* has also taken part in a debate focused on the topic at hand.

7.3.3 Stance Affects Disagreement

These rules are the inverse of the "Disagreement Affects Stance" rules.

Stance(A1, T, S) & Stance(A2, T, S) → !Disagrees(A1, A2, T)
States that if two authors have the same stance on a topic, then they are likely to not disagree on that topic.

Stance(A1, T, S1) & Stance(A2, T, S2) → Disagrees(A1, A2, T)
States that if two authors have differing stances on a topic, then they are likely to disagree on that topic.

7.3.4 Correlated Stances

These rules capture the assumption that am author will tend to be consistent with their stances across different topics. For example, an author that is pro gay marriage will often also be pro choice.

Stance(A, T1, S) & Participates(A, T2) → Stance(A, T2, S)

7.3.5 Ideology Disagreement

These rules capture the relationship between author disagreement and the stance of those authors.

Ideology(A1, I1) & Ideology(A2, I2) & Participates(A1, T) & Participates(A2, T) & Responds(A1, A2, T) → Disagrees(A1, A2, T)
States that authors with differing ideologies who participate on the same topic tend to disagree on that topic.

Ideology(A1, I) & Ideology(A2, I) & Participates(A1, T) & Participates(A2, T) → !Disagrees(A1, A2, T)
States that authors with the same ideology who participate on the same topic tend to agree on that topic.

Disagrees(A1, A2, T) & Ideology(A1, I) → !Ideology(A2, I)
States that authors that disagree on any topic tend to have differing ideologies.

7.3.6 Ideology Implies Stance

These rules capture the prior biases for democrats and republicans. Recall that for all the topics, a pro stance is consistent with the platform of the Democratic party and a con stance is consistent with the platform of the Republican party.

Ideology(A, 'DEMOCRAT') & Participates(A, T) → Stance(A, T, 'PRO')
Ideology(A, 'REPUBLICAN') & Participates(A, T) → Stance(A, T, 'CON')
On any topic the author participates in, they will tend to choose the stance aligned with their political party.

7.3.7 Stance Affects Ideology

These rules capture the when two authors agree they tend to have the same ideology and when they disagree they tend to have differing stances. Note that these rules bypass the *Disagree* predicate and looks directly at the inferred stance.

Stance(A1, T, S) & Stance(A2, T, S) & Ideology(A1, I) → Ideology(A2, I)
States that two authors that share the same stance on a topic tend to share the same ideology.

Stance(A1, T, S1) & Stance(A2, T, S2) & Ideology(A1, I) → !Ideology(A2, I)
States that two authors with differing stances on a topic tend to have differing ideologies.

7.3.8 Other Ideology

This rule allows for some slack in ideology inference. By introducing a third, latent ideology we can allow for people whose ideology does not neatly fall into Democratic or Republican. In all other rules whenever a negation is used on an *Ideology* predicate, that does not mean the other major party, ie *!Ideology(A, 'Democrat')* does not mean *!Ideology(A, 'Republican')*. This means that ideology can flow into either the other major party, or into this additional ideology.

Stance(A, T1, S1) & Stance(A, T2, S2) → Ideology(A, 'OTHER')

8 Evaluation

8.1 Results

The evaluation metric that is most frequently used in stance classification tasks is accuracy of classification which is percentage of test instances correctly classified. For our baseline, we look at the average of the accuracies that (Sridhar et al., 2015) report for the topics under consideration - abortion, evolution, gay marriage, gun control, which turns out to be 72.18 ± 2.2. The results we obtained are represented in 2. Column headings have the following meanings:

- DS - The inclusion of the "Disagreement Affects Stance" rules.
- SD - The inclusion of the "Stance Affects Disagreement" rules.
- CS - The inclusion of the "Correlated Stances" rules.
- ID - The inclusion of the "Ideology Disagreement" rules.
- IS - The inclusion of the "Ideology Implies Stance" rules.
- SI - The inclusion of the "Stance Ideology" rules.
- OI - The inclusion of the "Other Ideology" rules.

8.2 Synthetic Data

To explore the potential of the downstream joint classification without relying on the noisy upstream ideology classifier, we have constructed synthetic ideology classification results using the observed stance information. The purpose of this data is to see if a classifier that produces the best possible results could help in this task.

CS	DS	SD	ID	IS	SI	OI	S - A	D - A
T	T	T	F	F	F	F	**0.7894**	0.6319
T	T	F	F	F	F	F	0.7882	0.6319
T	T	T	T	F	T	F	0.7869	0.6319
T	T	F	F	F	F	T	0.7857	0.6324
T	T	T	F	F	F	T	0.7857	0.6324
T	T	F	T	F	F	T	0.7857	0.6324
T	T	F	T	F	T	F	0.7857	0.6324
T	T	T	T	F	F	F	0.7845	0.6324
T	T	T	T	F	T	T	0.7833	0.6324
T	T	F	T	F	T	T	0.7709	0.6406

Table 2: Results

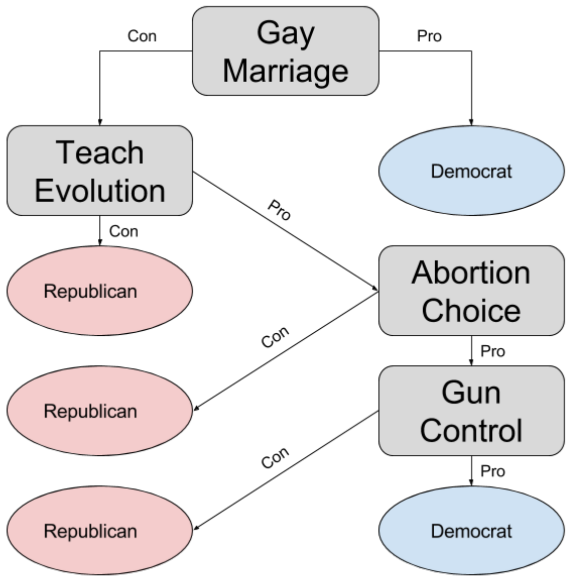

Figure 1: Decision tree for synthetic ideology classification.

8.2.1 Creation

The first step in creating the synthetic data was to get the count of pro and con stances an author took for each topic (since an author can make multiple posts per topic). We then subtracted the author's con count from their pro count to get an overall pro or con stance for the author on the topic. Intuitively, we are checking the stance the the author takes on average for a topic. A positive count means the author is typically pro, a zero mean the author is neutral, and a negative count means the author is typically con.

After we have the average stance of the author for each topic, we use a manually designed decision tree to predict the ideology of an author. The decision tree can be seen in Figure 1. In the case that the author was neutral on the final decision in the tree, we just assigned the majority class from our dataset.

CS	DS	SD	ID	IS	SI	OI	S - A	D - A
F	T	T	T	T	T	F	**0.8325**	0.6406
F	T	F	T	T	T	F	**0.8325**	0.6406
F	T	T	T	T	T	T	0.8300	0.6406
T	T	T	T	T	T	F	0.8300	0.6406
F	T	F	T	T	T	T	0.8300	0.6406
T	T	F	T	T	T	F	0.8288	0.6406
T	T	T	T	T	T	T	0.8276	0.6406
F	T	T	T	T	F	F	0.8276	0.6406
F	T	T	T	T	F	T	0.8276	0.6406
T	T	T	T	T	F	F	0.8264	0.6406

Table 3: Synthetic Data Results

8.2.2 Results

Using the synthetic ideology data, we were able to gain a boost of 4.32%. This indicates that ideology can help improve stance classification in the ideal case. The results for the top ten configurations can be see in Table 3.

9 Future Work

There are lot of directions for future work in this space. We can break them down into two subparts: improvements to the local classifier and improvements to the joint modeling. Improvements to the local classifier can mean better feature engineering and improvements to make it more cross-domain robust.

Having syntactic features help across domains rather than statistical features like TF-IDF (Hasan and Ng, 2014). We can also encode LIWC category information is a more sophisticated way than just simply taking their counts. We could also possibly employ word embeddings such as word2vec and combine LIWC and dependency parses of text. There are several avenues that we intend to explore. We also intend to train the classifier, on the much larger corpus we collected from Twitter, as described above. This is in the hopes that the language characteristics, and temporality of information has more of an overlap with the data present in the online debate forums.

On the modeling side of things, we could come up to model more ideologies than just two since most people are not completely either or, but, somewhere in-between when it comes to being opinionated. Since we have a constraint on the ideology mixture that a person could have, it might be interesting to see what kinds of patterns emerge. We could also use other indicators, such as people's personality types, etc, which could be indicators of their ideology composition. We also intend to use this work as a ground basis for modeling latent variables by propagating light signal about them through a Markov random field to help enhance the task at hand - in this case, stance classification.

References

[Abbott et al.2016] Robert Abbott, Brian Ecker, Pranav Anand, and Marilyn Walker. 2016. Internet argument corpus 2.0: An sql schema for dialogic social media and the corpora to go with it. In *Language Resources and Evaluation Conference, LREC2016*.

[Bach et al.2015] Stephen H Bach, Matthias Broecheler, Bert Huang, and Lise Getoor. 2015. Hinge-loss markov random fields and probabilistic soft logic. *arXiv preprint arXiv:1505.04406*.

[Hasan and Ng] Kazi Saidul Hasan and Vincent Ng. Extra-linguistic constraints on stance recognition in ideological debates.

[Hasan and Ng2014] Kazi Saidul Hasan and Vincent Ng. 2014. Why are you taking this stance? identifying and classifying reasons in ideological debates. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 751–762, Doha, Qatar, October. Association for Computational Linguistics.

[Klir and Yuan1995] George J. Klir and Bo Yuan. 1995. *Fuzzy Sets and Fuzzy Logic: Theory and Applications*. Prentice-Hall, Inc., Upper Saddle River, NJ, USA.

[Sridhar et al.2014] Dhanya Sridhar, James Foulds, Bert Huang, Marilyn Walker, and Lise Getoor. 2014. Collective classification of stance and disagreement in online debate forums. In *Bay Area Machine Learning Symposium (BayLearn)*.

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

[Walker et al.2012] Marilyn A Walker, Jean E Fox Tree, Pranav Anand, Rob Abbott, and Joseph King. 2012. A corpus for research on deliberation and debate. In *LREC*, pages 812–817.