Extra-Linguistic Constraints on Stance Recognition in Ideological Debates

Kazi Saidul Hasan and Vincent Ng
Human Language Technology Research Institute
The University of Texas at Dallas

Debate Stance Classification

Determine the stance (i.e., *for* or *against*) of a post written for a *two-sided* topic discussed in an online debate forum

A Sample Debate

Should abortion be allowed?	
Yes (*for*)	**No (*against*)**
Women should have the ability to choose what they do with their bodies.	Technically abortion is murder. They are killing the baby without a justified motive.

Related Work

- Three popular debate settings
  - US congressional floor debates (Thomas et al., 2006; Bansal et al., 2008; Burfoot et al., 2011)
  - Company-internal debates (Murakami and Raymond, 2010)
  - Ideological debates (Somasundaran and Wiebe, 2010; Anand et al., 2011)

Our Setting:
Ideological Debates

- Various social, political, and ideological issues
  - Abortion, gay rights, gun rights, god's existence
- Informal (often include insults)
- Sarcastic comments
- Rhetorical questions

Goal

To improve the state of the art in supervised stance classification of ideological debates

- by proposing two extra-linguistic extensions to state of the art baselines

Plan for the Talk

- Two baseline stance classification systems
- Two extra-linguistic extensions to the baselines
- Evaluation

Plan for the Talk

- Two baseline stance classification systems
- Two extra-linguistic extensions to the baselines
- Evaluation

Baseline 1: Anand et al., 2011

- Supervised approach, one stance classifier per domain
  - SVM in our implementation
  - One training/test instance for each post
  - Two labels – *for* and *against*

Feature Type	Features
Basic	Unigrams, bigrams, syntactic and POS generalized dependencies
Sentiment	LIWC counts, opinion dependencies
Argument	Cue words, repeated punctuation, context

Baseline 2: Anand et al.'s system enhanced with Author Constraints

- Author constraints (ACs)
  - a type of constraints for postprocessing the output of a stance classifier
  - ensure that all test posts written for the same domain by an author have the same stance

- How to postprocess Anand et al.'s output with ACs?
  - For each author, sum up classification values of her test posts
    - Classification value is the signed distance from the hyperplane
  - If sum > 0, assign *for* to all her test posts; else *against*

Plan for the Talk

- Two baseline stance classification systems
- Two extra-linguistic extensions to the baselines
- Evaluation

Our Two Extensions

- Extra-linguistic inter-post constraints
  1. Ideology Constraints (IC)
  2. User-interaction Constraints (UC)

Ideology Constraints (ICs)

- Designed specifically for stance classification of ideological debates

- Cross-domain, author-based constraints
  - Only applicable to debate posts written by the same author in different domains

Ideological Constraints (ICs)

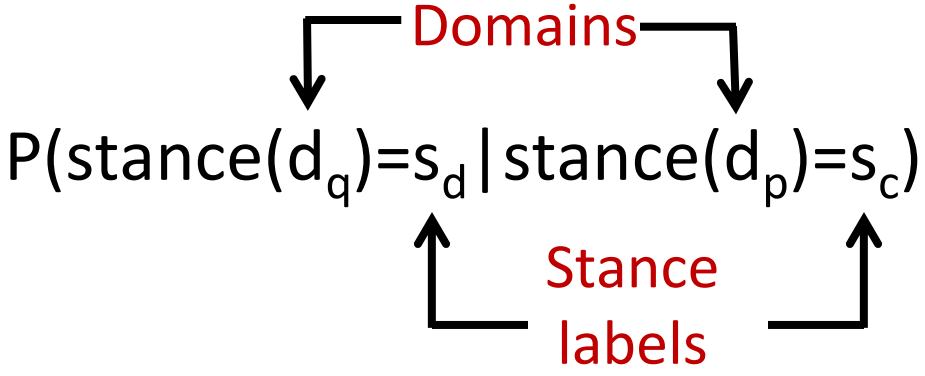
- **Observation**: An author's stance in one domain may be indicative of her stance in another domain.

  Example: An anti-abortion author is likely to be anti-Obama.

- **Goal**: exploit such correlation between the stance labels of the test posts by performing joint inference over them

Implementing ICs: 2 Steps

1. Compute for each pair of domains and each pair of stance labels the conditional probability:



2. If the conditional probability is above a certain threshold (to be determined using development data), we create a hard constraint:

$$stance(d_p)=s_c \rightarrow stance(d_q)=s_d$$

How to employ these Hard ICs?

- As constraints for postprocessing the output of a baseline stance classifier

- But … for Baseline 2, we have two types of postprocessing constraints: ACs and ICs
  - How can they be enforced jointly?
    - Integer Linear Programming (ILP)

Inference Using ILP

maximize: $\sum_n \; p_n x_n + (1-p_n)(1-x_n)$

$p_n = P(for \mid post_n)$

x_n = 1 means *for*; 0 means *against*

subject to ACs and ICs

Implementing ACs as Linear Constraints

For any pair of posts, x_i and x_j, written by the same author in domain d, create linear constraint

$$|x_i - x_j| = 0$$

Implementing ICs as Linear Constraints

- Recall that ICs are cross-domain, author-based constraints of the form:

  stance(d_p)=s_c → stance(d_q)=s_d

  where d_p and d_q are domains; s_c and s_d are stance labels

If	Create
stance(d_p)=*for* → stance(d_q)=*for*	$(1-x_j) \leq (1-x_i)$
stance(d_p)=*against* → stance(d_q)=*against*	$x_j \leq x_i$
stance(d_p)=*for* → stance(d_q)=*against*	$x_j \leq (1-x_i)$
stance(d_p)=*against* → stance(d_q)=*for*	$(1-x_j) \leq x_i$

  x_i and x_j are binary variables for a post in d_p and d_q, respectively

User-interaction Constraints (UC)

[P1: Anti-abortion] There are thousands of people who want to take these children because they cannot have their own. If you do not want a child, have it and put it up for adoption. At least you will be preserving a human life rather than killing one.

[P2: Pro-abortion] I agree that if people don't want their babies, they should have the choice of putting it up for adoption. But it should not be made compulsory, which is essentially what happens if you ban abortion.

[P3: Anti-abortion] Why should it not be made compulsory? Those children have as much right to live as you and I. Besides, no one loses with adoption, so why wouldn't you utilize it?

User-interaction Constraints (UC)

- Regularities in user interaction
  - Training data shows that stances alternate 80% of the time in a post sequence
  - Agrawal et al.'s (2003) simple assumption – a reply post indicates a disagreement

- **Aim**: model regularities in how users interact as soft constraints. How?
  - Recast stance classification as a sequence labeling task

Stance Classification as Sequence Labeling

- Input: debate post sequence
- Output: stance label sequence

- Supervised sequence learning using CRFs

Stance Classification as Sequence Labeling

- How to generate post sequences from debate post threads?
  - A post sequence is a path from the root of the thread to one of its leaves
- How to assign a final stance label to a post that appears in multiple post sequences?
  - We simply take the average of the classifier output values for all the instances of the post

Plan for the Talk

- Two baseline stance classification systems
- Two extra-linguistic extensions to the baselines
- Evaluation

Experimental Setup

- 4 Datasets
  - Collected from http://www.createdebate.com

Domain	Posts	"for" %	Thread Length
ABO (support abortion?)	1741	54.9	4.1
GAY (support gay rights?)	1376	63.4	4.0
OBA (support Obama?)	985	53.9	2.6
MAR (legalize marijuana?)	626	69.5	2.5

Experimental Setup

- Performance metric – accuracy
- 5-fold cross validation
- Development set for parameter tuning

Results

System	ABO	GAY	OBA	MAR
Anand	61.4	62.6	58.1	66.9

Results

System	ABO	GAY	OBA	MAR
Anand	61.4	62.6	58.1	66.9
Anand+AC	72.0	64.9	62.7	67.8

- Anand+AC significantly outperforms Anand by 4.6 points

Results

System	ABO	GAY	OBA	MAR
Anand	61.4	62.6	58.1	66.9
Anand+AC	72.0	64.9	62.7	67.8
Anand+AC+UC	73.7	69.9	64.1	75.4

- Incorporating UCs yields a significant improvement of 3.9 points over Anand+AC

Results

System	ABO	GAY	OBA	MAR
Anand	61.4	62.6	58.1	66.9
Anand+AC	72.0	64.9	62.7	67.8
Anand+AC+UC	73.7	69.9	64.1	75.4
Anand+AC+UC+IC	74.9	70.9	72.7	75.4

- Incorporating ICs yields another significant improvement of 2.7 points over UCs
- Overall improvement over Anand+AC is 6.6 points
- No ICs for MAR; ICs between ABO and GAY in 2 folds

Results

System	ABO	GAY	OBA	MAR
Anand	61.4	62.6	58.1	66.9
Anand+AC	72.0	64.9	62.7	67.8
Anand+AC+UC	73.7	69.9	64.1	75.4
Anand+AC+UC+IC	74.9	70.9	72.7	75.4

- Incorporating ICs yields another significant improvement of 2.7 points over UCs
- Overall improvement over Anand+AC is 6.6 points
- No ICs for MAR; ICs between ABO and GAY in 2 folds

Results

System	ABO	GAY	OBA	MAR
Anand	61.4	62.6	58.1	66.9
Anand+AC	72.0	64.9	62.7	67.8
Anand+AC+UC	73.7	69.9	64.1	75.4
Anand+AC+UC+IC	74.9	70.9	72.7	75.4

- Incorporating ICs yields another significant improvement of 2.7 points over UCs
- Overall improvement over Anand+AC is 6.6 points
- No ICs for MAR; ICs between ABO and GAY in 2 folds

Results

System	ABO	GAY	OBA	MAR
Anand	61.4	62.6	58.1	66.9
Anand+AC	72.0	64.9	62.7	67.8
Anand+AC+UC	73.7	69.9	64.1	75.4
Anand+AC+UC+IC	74.9	70.9	72.7	75.4

- Incorporating ICs yields another significant improvement of 2.7 points over UCs
- Overall improvement over Anand+AC is 6.6 points
- No ICs for MAR; ICs between ABO and GAY in 2 folds

Conclusions

- Proposed two types of extra-linguistic constraints for stance classification
  1. Ideology Constraints (IC)
  2. User-interaction Constraints (UC)

- Outperformed an improved version of Anand et al.'s approach by 2.9–10 accuracy points