Ideological Perspective Detection Using Semantic Features

Heba Elfardy
Columbia University
New York, NY
`heba@cs.columbia.edu`

Mona Diab
George Washington University
Washington, DC
`mtdiab@gwu.edu`

Chris Callison-Burch
University of Pennsylvania
Philadelphia, PA
`ccb@cis.upenn.edu`

Abstract

In this paper, we propose the use of word sense disambiguation and latent semantic features to automatically identify a person's perspective from his/her written text. We run an Amazon Mechanical Turk experiment where we ask Turkers to answer a set of constrained and open-ended political questions drawn from the American National Election Studies (ANES). We then extract the proposed features from the answers to the open-ended questions and use them to predict the answer to one of the constrained questions, namely, their preferred Presidential Candidate. In addition to this newly created dataset, we also evaluate our proposed approach on a second standard dataset of "Ideological-Debates". This latter dataset contains topics from four domains: Abortion, Creationism, Gun Rights and Gay-Rights. Experimental results show that using word sense disambiguation and latent-semantics, whether separately or combined, beats the majority and random baselines on the cross-validation and held-out-test sets for both the ANES and the four domains of the "Ideological Debates" datasets. Moreover combining both feature sets outperforms a stronger unigram-only classification system.

1 Introduction

With the pervasiveness of social media and online discussion fora, there has been a significant increase in documented political and ideological discussions. Automatically predicting the perspective or stance of users in such media is a challenging research problem that has a wide variety of applications including recommendation systems, targeted advertising, political polling, product reviews and even predicting possible future events. Ideology refers to the beliefs that influence an individual's goals, expectations and views of the world (Van Dijk, 1998; Ahmed and Xing, 2010). The ideological perspective of a person is often expressed in his/her choice of discussed topics. People with opposing perspectives will choose to make different topics more salient. (Entman, 1993).

From a social-science viewpoint, the notion of "perspective" is related to the concept of "framing". Framing involves making some topics (or some aspects of the discussed topics) more prominent in order to promote the views and interpretations of the writer (communicator). The communicator makes these framing decisions either consciously or unconsciously (Entman, 1993). These decisions are often expressed in the lexical choice. For example, a person who holds anti-abortion views, is more likely to use the terms "life" and "kill" whereas a person who is pro a woman having an option to go for an abortion will often stress on "choice".

From a computational viewpoint, work on perspective-detection is closely related to subjectivity and sentiment analysis. One's perspective normally influences his/her sentiment towards different topics or targets. Conversely identifying the sentiment of a person towards multiple targets can serve as a cue for identifying his/her perspective. The main difference between perspective and sentiment is that unlike sentiment that is more transient, perspective is often more deeply seated and less likely

137

to change. Most of the current perspective-detection work focuses on "Ideological Perspective" by trying to predict a person's stance on controversial topics such as the Palestinian-Israeli conflict, abortion, gay-rights, gun-rights, etc.

In this paper, we are interested in identifying the "Ideological Perspective" of a person using semantic features derived from his/her written text. We use two different sets of semantic features to train several supervised systems that predict different aspects of a person's ideological stance toward specific topics.

We explore the use of Word Sense Disambiguation from the high dimensional space and Latent Semantic models from the low dimensional space on two datasets. We find that explicitly modeling the lexical and contextual semantics to predict a person's perspective outperforms a strong-baseline system trained on standard unigram features.

2   Related Work

Current computational linguistics research on automatic perspective detection uses both supervised and unsupervised techniques. The main task handled by supervised approaches is to perform document (or post) level perspective (or stance) classification, whether binary or multiclass labeling. Unsupervised approaches on the other hand, mainly try to cluster users in a discussion. One of the early works on binary perspective identification is that of Lin et al. (2006) which uses articles from the Bitter-Lemons website –a website that discusses the Palestinian-Israeli conflict from each side's point of view– to train a system for performing automatic perspective detection on the sentence and document levels. On the website, an Israeli editor and a Palestinian editor, together with invited guests, contribute articles to the website on a weekly basis. Lin et al. (2006) use bag-of-words features. They run different experiments where they vary the training and test sets between: (a) editors' articles and (b) guests' articles. The accuracies of the different experimental conditions vary between 86% and 99%. As one might expect the highest accuracy (99%) is that of the system that is trained and tested on the editors' articles. For this system, the classifier is not only capturing the perspective but also the editors' writ-

ing styles. In Klebanov et al. (2010), the authors tackle the same problem of binary-perspective detection and experiment with four corpora; Bitter-Lemons, Bitter-Lemons-International, Partial-Birth-Abortion and Death-Penalty. They show that using term-frequencies does not improve over using binary bag-of-words features and that using only the best 1-4.9% features is sufficient to achieve high accuracy. They achieve the highest accuracy (97%) on the Partial-Birth-Abortion dataset and the lowest accuracy (73%) on the Death-Penalty dataset.

Hasan and Ng (2012) also tackle the problem of binary perspective detection but using Integer Linear Programming (ILP) to perform joint inference over the predictions made by a post-stance classifier and several topic-stance classifiers. The authors use n-grams, sentence-type and opinion-dependencies as features to train their classifiers. They collect debate posts discussing Abortion and Gun-Rights and achieve an $F_{\beta=1}$ score of 61.1% on the Abortion dataset and 57.8% on the Gun-Rights dataset. In Hasan and Ng (2013), they extend their previous work by incorporating two soft-constraints that treat the task of post-stance classification as a sequence-labeling problem and ensure that the topic-stance of each author is consistent across all posts.

Somasundaran and Wiebe (2010) employ the notion of *"arguing"* to identify a person's stance (supporting or opposing) towards a topic. Arguing can be indicated by using either positive lexical cues such as *"actually"* or negative ones such as *"certainly not"*. They construct an arguing lexicon and use it to derive features for their classifier. They experiment with both arguing and sentiment features on four datasets; Abortion, Creationism, Gun-Rights and Gay-Rights. They show that combining arguing and sentiment features outperforms a unigram baseline on Abortion, Gay-Rights and Gun-Rights datasets while the unigram system performs best on the Creationism dataset.

A closely related work is that of Al Khatib et al. (2012). In this work, the authors use a set of Arabic and English Wikipedia articles about Arab and Israeli public figures to explore the differences in point of view between the Arabic and English articles about each figure. They assign a point-of-view score to each article in each language, and use these scores to train a classifier to predict the difference in

PCC		Obama	Romney	Neither
	Train	62.9	25.3	11.8
	Test	67.6	18.5	13.9

Table 1: Class Distribution of Presidential Candidate Choice (PCC) in the ANES dataset

		Pro	Against
Train	Abortion	55.3	44.7
	Creationism	35.8	64.2
	Gay-Rights	64	36
	Gun-Rights	74	26
Test	Abortion	50.4	49.6
	Creationism	27.9	72.1
	Gay-Rights	63.6	36.4
	Gun-Rights	57.5	42.5

Table 2: Class Distribution across the four domains of the Online Debates Dataset

point of view between each article-pair.

For unsupervised approaches, two of the most recent works are those of Abu-Jbara et al. (2012) and Dasigi et al. (2012). In Abu-Jbara et al. (2012), the authors perform subgroup detection by clustering authors according to their sentiment towards topics, Named-Entities as well as other discussants. Dasigi et al. (2012) extend the previous work by introducing the notion of implicit attitude which models the similarity between the topics discussed by a pair of people. They note that people that share the same opinion tend to discuss similar topics, thereby their texts tend to have a high semantic similarity. By adding implicit attitude, namely by explicitly modeling latent sentential semantics, they achieve an $F_{\beta=1}$ score improvement of 3.83% and 2.12% on "Wikipedia-Discussions" and "Online-Debates" datasets, respectively.

Yano et al. (2010) study the linguistic cues for bias in political blogs. The authors draw sentences from American political blogs and annotate them for bias on Amazon Mechanical Turk. They explore whether the Turkers' decisions are influenced by their perspectives, for example whether a self proclaimed liberal Turker is more likely to view sentences written by a conservative as biased and vice versa.

3 Datasets

We use two datasets to evaluate our approach.

3.1 ANES Dataset

We create this dataset by drawing a set of questions from the American National Election Studies (ANES) survey questions.[1] ANES conducts various surveys in order to provide better explanations and analysis of the outcomes of USA Presidential elections. While the officially administered ANES survey contains both constrained multiple choice

questions and open-ended (free form essay style) questions, the answers to the open-ended questions, which are more interesting from an NLP perspective, are not made publicly available in order to protect the privacy of respondents. In this work, we run an Amazon Mechanical Turk annotation experiment where we ask Amazon Mechanical Turk annotators (aka Turkers) to answer a large set of constrained and open-ended questions drawn from ANES.[2] The constrained questions may be considered a form of self labeling indicating the respondent/Turker's background or perspective on specific issues. All Turkers participating in the experiment were required to be from the US. Moreover, we added seven quality-control questions with a correct (and obvious) answer in order to identify spam Turkers. All submissions that rendered more than one of these questions wrong were automatically rejected.

The first set of questions that required constrained answers, such as multiple choice or binary responses as true or false can be binned into the following categories:

- Background Questions: A person's age, gender, educational level, income, marital-status, social-status, how often he/she follows the news, what news sources he/she follows, etc.;
- Opinion of Political Parties: Democratic and Republican parties and their respective public figure representatives;
- Opinion on major economic and political problems facing the USA;

---

[1] electionstudies.org/studypages/2010_2012EGSS/2010_2012EGSS

[2] Please contact the authors to obtain the dataset

Q1	I approve of Obama's and the Democrats' position on abortion and gay marriage and their tendency to favor programs that help the poor and working class. They seem more compassionate and more socially progressive.
Q2	Neither Obama nor the Democrats seems able to get a hold on spending, the deficit or help the economy and unemployment. They seem to spend too much time criticizing their opponents rather than work toward viable solutions and seem to distort facts against the other party more.
Q3	I think Mitt Romney and the republicans in general would do a better job at lowering the deficit and stimulating the economy and reducing unemployment. I also agree with their position of less government involvement in some areas.
Q4	I dislike Mitt Romney's plans to eliminate funding for Planned Parenthood and the republicans stand on social issues such as abortion and gay rights, especially gay marriage. I feel Republicans have been taken over by the religious right and are socially regressive.

Table 3: Sample answers provided by one Turker to the first four essay questions in the ANES dataset.

- Ideology Questions: Importance of religion, political-party-affiliation, presidential candidate choice, etc.;
- Opinion on contentious issues: Such as Race (White, Black, Asian and Hispanic Americans), same-sex marriage, gun-control, universal health-care, etc.

The second set of questions ask about a person's opinion of certain ideological topics. The responses are not constrained in any manner.

Since our main objective is to study whether a person's perspective can be automatically identified using NLP techniques applied to his/her written text, we choose to predict the answer to one of the constrained ideological questions, "Presidential Candidate Choice" (PCC), based on the answers to the following open ended questions: (Table 1 shows the distribution of PCC in the training and test sets)

- **Q1**: Is there something that would make you vote for a Democratic presidential candidate?
- **Q2**: Is there something that would make you vote against a Democratic presidential candidate?
- **Q3**: Is there something that would make you vote for a Republican presidential candidate?
- **Q4**: Is there something that would make you vote against a Republican presidential candidate?
- **Q5**: If you said there is something you like about the Democratic Party: What is that?
- **Q6**: If you said there is something you dislike about the Democratic Party: What is that?
- **Q7**: If you said there is something you like about the Republican Party: What is that?

- **Q8**: If you said there is something you dislike about the Republican Party: What is that?
- **Q9**: What has been the most important issue to you personally in this election?
- **Q10**: What has been the second most important issue to you personally in this election?
- **Q11**: What do you think is the most important political problem facing the United States today?
- **Q12**: What do you think is the second most important political problem facing the United States today?
- **Q13**: What do you think the terrorists were trying to accomplish by September 11th attacks?

Table 3 shows the answers provided by a Turker to the first four of these questions.

In order to simulate user generated content where people are not providing answers to a predefined set of questions but are rather discussing current events or topics, we decide to combine the answers to all of these questions in one document per Turker and use this combined resulting document to derive features (as opposed to deriving features from the answer to each question separately). In order to reduce ambiguity, we perform a quasi co-reference resolution step on pronouns. Prior to combining the answers to all 13 questions, we perform a "pronoun-rewriting" step where we replace the sentence initial pronouns with the topic the question is about. For example, for Q3, *"Is there something that would make you vote for a Republican presidential candidate?"*, and the answer provided is *"They are against voting rights for illegal immigrants. They want to balance the budget and find a way to slowly reduce the national*

Domain	Stance	Post
Abortion	Pro	So abortion is okay in areas where more people like it than don't?
Abortion	Against	your exact words "But successful abortion carries a 100% rate of risk of death to the child" no duh, that's the whole point of abortion, is to KILL THE BABY. well actually that's MURDER
Creationism	Pro	You cant make nothing out of nothing!!!
Creationism	Against	It is only belief. No one has any real evidence.
Gay-Rights	Pro	This post is almost insulting in its complete lack of evidence or even a reasoned argument. Merely dismissing the other side is not an argument
Gay-Rights	Against	Compared to children with a father and a mother married to each other and getting along with each other, the answer is yes. Compared to children living in an orphanage, it's hard to say.
Gun-Rights	Pro	An assault weapon ban violates the second amendment
Gun-Rights	Against	Dude. Are you home all the time? Is this secured? Do you have a lot of fire extinguishers?

Table 4: Sample posts from "Ideological Debates" dataset.

	Train					Test				
	Posts	Tokens	Token/Post	Types	Types/Post	Posts	Tokens	Tokens/Post	Types	Types/Post
ANES	965	437,080	453	14,410	15	108	61,414	569	5,487	51
Abortion	1,036	154,929	150	10,192	10	115	18,299	159	2,734	24
Creationism	1,108	217,262	196	13,466	12	122	10,756	88	2,160	18
Gay-Rights	1,858	312,900	168	16,742	9	206	17,400	84	2,842	14
Gun-Rights	963	146,886	153	10,969	11	106	7,142	67	1,588	15

Table 5: Statistics of the training and test sets for both the ANES and the four domains of the Ideological-Debates datasets.

debt.", we replace "They" with "Republicans".

3.2 Ideological Debates Dataset

This dataset was collected by Somasundaran and Wiebe (2010) . It contains debate posts from six domains; (a) Abortion, (b) Creationism, (c) Gay-Rights, (d) Gun-Rights, (e) Healthcare and (f) Existence of God. Each domain represents an ideological topic with two possible perspectives, pro and against. Similar to the work of (Somasundaran and Wiebe, 2010), we use the first four domains to evaluate our approach. Table 2 shows the class distribution in each of these four domains while table 4 lists some sample posts. It should be noted that our results are not comparable to those obtained by (Somasundaran and Wiebe, 2010), since they used a subset of the posts in each domain and the split was not publicized.

Table 5 shows the size of the training and test data in the ANES and Ideological-Debates datasets.

4 Approach

Our goal is to determine whether semantic features help in identifying a person's ideological perspective as determined by his/her answer to the PCC constrained question in the "ANES" dataset and his/her stance towards the ideological-topics discussed in the "Ideological-Debates" dataset independently.

4.1 Preprocessing

We apply basic preprocessing to the text by separating punctuation and numbers from words. All punctuation and numbers are then ignored when training the classifier for all of our systems including the unigram baseline. The intuition behind this is that punctuation and numbers do not capture the perspective of a person but rather the writing style. Moreover, by ignoring them, we avoid overfitting the training data.

4.2 Word Sense Disambiguation (WSD)

We use WN-Sense-Relate (Patwardhan et al., 2005) to perform word sense disambiguation. Sense-Relate uses WordNet (Miller, 1995) to tag each word with the part-of-speech and sense-id. The only parts of speech that are handled by WN-Sense-Relate are adjectives (a), adverbs (r), verbs (v) and nouns (n). In addition to the part-of-speech and sense-id, WN-Sense-Relate also identifies and tags compounds. The word sense tagging process can be either contextual or can rely on the most frequent sense. We experiment with both variants.

4.2.1 Contextual WSD (WSD-CXT)

In this variant of WSD, in addition to tagging compounds, we contextually disambiguate each word and tag it with its sense-id and part-of-speech. We use the default setting of SenseRelate which employs a modified version of the Lesk algorithm (Banerjee and Pedersen, 2002) to perform the disambiguation. This version of the Lesk algorithm measures the similarity between the WordNet gloss of each sense of the target word and those of its surrounding context words in the text. It then chooses the sense whose gloss is most similar to the surrounding words. We use a window of size three which uses one word before and one word after the target word.

ex.
"The Democratic Party supports women 's equality , including equal pay , access to health care and other issues ."
becomes:
"the#ND democratic_party#n#1 supports#n#10 women#n#1 's#ND equality#n#1 including#v#3 equal#a#1 pay#v#1 access#n#2 to#ND health_care#n#1 and#ND other#a#1 issues#n#7"[3]

We then use the tagged-words to retrieve the Synonym-Set (Synset) of this sense of the word using WN-QueryData (Pedersen et al., 2004). We assign each Synset an ID and whenever any of the words in the retrieved Synset is seen in the input text, we replace it with this Synset-ID.

For example, the Synset of *"issues#n#7"* is identified as *"issue#n#7, consequence#n#1, effect#n#1,*

---

outcome#n#2, result#n#1, event#n#4, issue#n#7, upshot#n#1". We assign this Synset a unique ID, for example *"Synset-100"*, and any occurrence of any of *"issue#n#7, consequence#n#1, effect#n#1, outcome#n#2, result#n#1, event#n#4, issue#n#7, and upshot#n#1"* is replaced by *"Synset-100"*.

4.2.2 Most Frequent Sense WSD (WSD-MFS)

In this variant of WSD, instead of performing the disambiguation contextually, we rely on the most frequent sense. Using this scheme, *"issues"* is tagged as *"issues#n#1"* whose Synset is *"issue#n#1"*, while *"support"* is tagged as *"support#v#1"* whose Synset is *support#v#1', back_up#v#1*.

4.3 Latent Semantics

The next set of features relies on "Latent Semantics" which maps text from a high-dimensional space such as unigrams to a low-dimensional one such as topics. Most of these models assign a semantic profile to each given sentence (or document) by considering the observed words and assuming that each given document has a distribution over "K" topics. We apply (1) Latent Dirichlet Allocation (LDA) (Blei et al., 2003) as implemented in MALLET toolkit (McCallum, 2002), and (2) Weighted Textual Matrix Factorization (WTMF) (Guo and Diab, 2012) to each post. In addition to observed words, WTMF also models missing ones namely explicitly modeling what the post is not about. WTMF defines missing words as the whole vocabulary of the training data minus the ones observed in the given document.

4.3.1 Number of Topics

We vary the number of topics (K) between 100 and 500 (with a step-size of 100) and use the best "K" for each dataset. We define the best K, for each of LDA and WTMF, as the one that yields the best cross-validation results when combined with unigram features. The best K value for LDA is 400 for PCC and Abortion, 500 for Creationism, 300 for Gay-Rights and 100 for Gun-Rights. For WTMF, the best K is 500 for PCC, and Gun-Rights and 100 for Abortion, Creationism and Gay-Rights.

---

[3]#ND indicates a non-defined word

	PCC $F_{\beta=1}$ score
Train	74.65
Test	74.81
All	74.69

Table 6: Performance of human annotators in predicting "PCC" of a person from his/her responses to ANES essay questions.

4.3.2 Training Data

We collected our training data for topic modeling from Facebook comments of renowned American politicians such as Joe-Biden, Chris-Christie, George W. Bush, Michelle-Obama, etc. We trained LDA and WTMF using a subset of 100,000 comments (corresponding to ~5,000,000 tokens and ~265,000 types.

4.4 Classifier Training

Using WEKA toolkit (Hall et al., 2009) and the derived features, we train Sequential Minimal Optimization (SMO) SVM classifiers (Platt, 1998) for each of "ANES" and the four domains of "Ideological-Debates" datasets. We use a normalized quadratic kernel, set the parameter C to 100 and apply a 10-fold cross validation on the training sets.

5 Experiments

5.1 Baselines

We compare our approach to three baselines;

- **Majority Baseline *(MAJ-BL)***: which assigns all posts to the most frequent class-label;
- **Random Baseline *(RAND-BL)***: which randomly chooses the class-label;
- **Unigram Baseline *(UNI-BL)***: a strong baseline that uses standard unigram features.

In addition to these three baselines, we do a human-evaluation for the ANES dataset in order to assess the difficulty of the task and in order to get an upper-bound on how well we can do in predicting PCC. We run an Amazon Mechanical Turk experiment where we ask Turkers to read each post (constructed by combining the answers to the open-ended questions of each record) and ask them to guess the PCC of the person who wrote that text along with the reason for their answer. We found
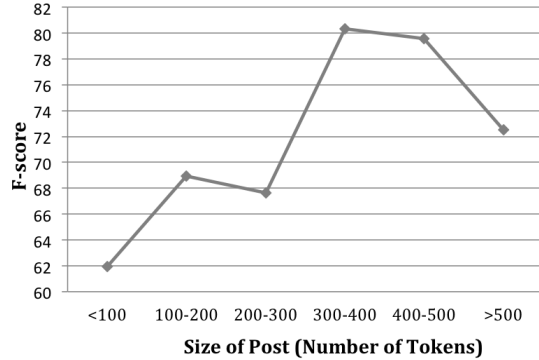

Figure 1: $F_{\beta=1}$ score of human judgments in predicting "PCC" from the answers to the essay questions in the ANES dataset across different post-sizes

that Turkers were able to predict the PCC with an average $F_{\beta=1}$ score of ~75% on both the cross-validation and test-sets. We also found that the task is particularly difficult for very short (< 100 words) documents. Table 6 and Figure 1 show the results of this qualitative assessment.

5.2 Experimental Setup

We first evaluate each variant of the proposed features separately and then we combine the latent-semantics features with unigram-features and the two variants of WSD.

Tables 7 and 8 show the cross-validation results on the training data and the results on the held-out test sets respectively.

6 Discussion

6.1 Cross Validation Results

For the cross-validation results, all configurations of the proposed features outperform the majority and random baselines. Moreover using WSD-MFS, either separately or combined with LDA or WTMF, outperforms the unigram-baseline. Overall WSD-MFS performs better than WSD-CXT, except on the Abortion dataset.

For Latent-Semantics, even though using either of LDA or WTMF separately, without unigram-features, does not outperform the unigram baseline, combining each of them with unigrams outperforms the unigram-only setup. When combined with unigram or WSD features, WTMF outperforms LDA on PCC, Creationism and Gun-Rights while LDA out-

	PCC	Abortion	Creationism	Gay-Rights	Gun-Rights
MAJ-BL	48.6	39.4	50.2	50	63
RAND-BL	36.7	47.6	50	52.9	52.6
UNI-BL	66.3	63.1	58.7	67.1	72.7
WTMF	62.9	56.9	57.1	57.1	72.6
LDA	62.4	57.4	58.7	63.6	71
Unigram+WTMF	68.8	62.7	62	67.2	**75.7**
Unigram+LDA	66.6	64.5	60	67.4	72.9
WSD-CXT	65.8	64.8	59.4	69.6	73.8
WSD-MFS	67.5	64	61.1	**69.7**	75
WSD-CXT + WTMF	68	64.1	61.6	68.1	74.8
WSD-CXT + LDA	65	64.3	59.3	69.1	73.8
WSD-MFS + WTMF	**69.2**	64.7	**62.7**	67.6	**75.7**
WSD-MFS + LDA	67.3	**65.1**	62.1	69.5	75

Table 7: 10-fold cross-validation results (measured in $F_{\beta=1}$ score) of using WSD and Latent-Semantics against the baselines.

	PCC	Abortion	Creationism	Gay-Rights	Gun-Rights
MAJ-BL	54.5	33.8	60.5	49.4	42
RAND-BL	46.6	47.3	54.9	41.9	45.5
UNI-BL	68	54.3	67.1	52.4	48.1
WTMF	69.1	58.1	60.1	49.2	43.7
LDA	60.4	55.3	58.7	**58.1**	**62.4**
Unigram+WTMF	68.9	54.2	**71.2**	56.8	58.7
Unigram+LDA	68	58.9	70.7	56.4	48.1
WSD-CXT	66.8	52.8	66.2	53.4	44.1
WSD-MFS	64.2	54.6	69.8	56.3	46.2
WSD-CXT + WTMF	66.1	52.9	71	55.1	53.7
WSD-CXT + LDA	67.6	57	68.5	55.6	46.2
WSD-MFS + WTMF	**71.6**	55.7	69.1	56.7	52.7
WSD-MFS + LDA	65.3	**62.6**	69.6	56.4	44.1

Table 8: Held-out test-set results (measured in $F_{\beta=1}$ score) of using WSD and Latent-Semantics against the baselines.

performs WTMF on the other two datasets. Combining WSD-MFS with LDA for Abortion and Gay-Rights and with WTMF for PCC, Creationism and Gun-Rights yields the best (or close to the best) results.

6.2 Held-out Test-Sets Results

Unlike the cross-validation results, using latent-semantics features separately improves over the unigram baseline for four out of the five datasets and in some cases adding unigrams to latent-semantics

features actually hurts the performance. This suggests that latent-semantics are less likely to overfit the training data.

Table 9 shows examples of the posts that were misclassified by the majority and unigram baselines and correctly classified by the best semantic model for each dataset.

6.3 General Observations

We investigated the data to identify the different challenges faced when trying to identify a person's

Domain	Stance	Post
Abortion	Against	Yes, all innocent life. But that depends on how you define innocent life.
Creationism	Pro	There's a definite difference between micro-evolution and macro-evolution in the sense that with the moths and the finches, there are minor changes that happen. It's kind of like a pendulum. It swings far to the right and to the left, but in the end, it's right in the center again, if you understand me correctly.
Gay-Rights	Against	Not necessarily the question of whether or not same-sex couples' marriages specifically are recognized by the government. As for my personal views on the issue I honestly think the best solution is for the government to simply call all civil unions precisely what they are: civil unions. Leave it to individuals and churches to determine the definition of "marriage."
Gun-Rights	Against	I agree that gun ownership should be strictly controlled. Put a gun in the hands of a crackpot and there's going to be a problem.

Table 9: Examples of the posts that were misclassified by the unigram baseline and were correctly classified by the right semantic model.

perspective and found the following:

1. In ANES dataset, due to the structure of the questions, some Turkers were trying to be objective which makes it difficult even for a human evaluator to identify the political leaning of the person who wrote the text. The example in Table 3 illustrates such a case where it is not easy to detect the PCC of the Turker from the provided answers.

2. The use of sarcasm, which can be easily detected by human evaluators but not by an automated system. For example, in Abortion dataset, a participant who does not oppose abortion wrote *"Why should people use reason and logic to discover right and wrong when a priest can decide for them?"*

3. Misspelled words such as writing *"Romeny"* instead of *"Romney"*

4. In each domain of the Ideological Debates dataset, the posts were collected from different discussion fora pertaining to the domain of interest. For example, in the Abortion dataset, posts were collected from "Can Catholics Vote For Pro-Choice Politicians", "Should South Dakota pass the Abortion Ban", "Should abortion be legal" and other fora. For some of the posts, the participants provided very short answers such as *"Once they take the booth who they vote for is supposed to be secret."* which makes it almost impossible to identify their stance without knowing the exact question the forum posed.

7 Conclusion

In this paper, we explore the use of semantic features to perform automatic detection of ideological-perspective from written text. Using Word Sense Disambiguation and Latent Semantics features, we trained several SVM classifiers that predict different aspects of the ideological-perspective of a person. We evaluated the presented approach on two datasets. The first of which comprises answers to questions about American politics collected from an Amazon Mechanical Turk experiment while the second one consists of four subsets of a standard dataset, discussing Abortion, Creationism, Gay-Rights and Gun-Rights. Results show that using the proposed features outperforms a system that relies on standard unigram features on all datasets. On the cross-validation sets, combining word sense disambiguation with latent semantics performs best while on the held-out test sets, the best configuration various across the different domains.

We plan to explore other methods for performing word sense disambiguation in addition to using semantic-role-labeling and modeling sarcasm.

8 Acknowledgment

References

Amjad Abu-Jbara, Mona Diab, Pradeep Dasigi, and Dragomir Radev. 2012. Subgroup detection in ideological discussions. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Long Papers-Volume 1*, pages 399–409. Association for Computational Linguistics.

Amr Ahmed and Eric P Xing. 2010. Staying informed: supervised and semi-supervised multi-view topical analysis of ideological perspective. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1140–1150. Association for Computational Linguistics.

Khalid Al Khatib, Hinrich Schütze, and Cathleen Kantner. 2012. Automatic detection of point of view differences in wikipedia. In *COLING*, pages 33–50.

Satanjeev Banerjee and Ted Pedersen. 2002. An adapted lesk algorithm for word sense disambiguation using wordnet. In *Computational linguistics and intelligent text processing*, pages 136–145. Springer.

David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022.

Pradeep Dasigi, Weiwei Guo, and Mona Diab. 2012. Genre independent subgroup detection in online discussion threads: a pilot study of implicit attitude using latent textual semantics. In *Proceedings of the 50th Annual Meeting of the Association for Computational Linguistics: Short Papers-Volume 2*, pages 65–69. Association for Computational Linguistics.

Robert M Entman. 1993. Framing: Toward clarification of a fractured paradigm. *Journal of communication*, 43(4):51–58.

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

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

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

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

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

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

Andrew Kachites McCallum. 2002. Mallet: A machine learning for language toolkit. http://mallet.cs.umass.edu.

George A Miller. 1995. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.

Siddharth Patwardhan, Satanjeev Banerjee, and Ted Pedersen. 2005. Senserelate:: Targetword: a generalized framework for word sense disambiguation. In *Proceedings of the ACL 2005 on Interactive poster and demonstration sessions*, pages 73–76. Association for Computational Linguistics.

Ted Pedersen, Siddharth Patwardhan, and Jason Michelizzi. 2004. Wordnet:: Similarity: measuring the relatedness of concepts. In *Demonstration Papers at HLT-NAACL 2004*. Association for Computational Linguistics.

John C. Platt. 1998. Sequential minimal optimization: A fast algorithm for training support vector machines. *Technical Report MSR-TR-98-14, Microsoft Research*.

Swapna Somasundaran and Janyce Wiebe. 2010. Recognizing stances in ideological online debates. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pages 116–124. Association for Computational Linguistics.

Teun A Van Dijk. 1998. *Ideology: A multidisciplinary approach*. Sage.

Tae Yano, Philip Resnik, and Noah A Smith. 2010. Shedding (a thousand points of) light on biased language. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 152–158. Association for Computational Linguistics.