

A Study of the Impact of Persuasive Argumentation in Political Debates

Amparo Elizabeth Cano-Basave and Yulan He

Aston University, UK

a.cano-basave@aston.ac.uk, y.he@cantab.net

Abstract

Persuasive communication is the process of shaping, reinforcing and changing others' responses. In political debates, speakers express their views towards the debated topics by choosing both the content of their discourse and the argumentation process. In this work we study the use of semantic frames for modelling argumentation in speakers' discourse. We investigate the impact of a speaker's argumentation style and their effect in influencing an audience in supporting their candidature. We model the influence index of each candidate based on their relative standings in the polls released prior to the debate and present a system which ranks speakers in terms of their relative influence using a combination of content and persuasive argumentation features. Our results show that although content alone is predictive of a speaker's influence rank, persuasive argumentation also affects such indices.

1 Introduction

In recent years, researchers have studied political texts detecting ideological positions (Sim et al., 2013; Hasan and Ng, 2013), predicting voting patterns (Thomas et al., 2006; Gerrish and Blei, 2011) and characterising power based on linguistic features (Prabhakaran et al., 2013). While there is a vast amount of theoretical research on the rhetoric of politicians, only recently there has been a growing interest in understanding the argumentation processes involved in political communication by means of computational linguistics (Hasan and Ng, 2013; Boltužić and Šnajder, 2014).

During a debate, a speaker tries to convince the audience of a particular point of view. This normally involves an argumentation process, where the structuring of ideas is built upon logical connections between claims and premises, and a persuasive communication style. In this paper, we study the impact of persuasive argumentation in political debates on candidates' power/influence ranking. As opposed to previous approaches, we propose to characterise political debates based on persuasive argumentation modelled through semantic frames.

Previous work (Rosenberg and Hirschberg, 2009) has analysed political speech transcripts identifying prosodic and lexical-syntactic cues which correlate with political personalities. Prabhakaran et al. (2013) proposed interactions within political debates as predictors of a candidate's relative power or influence rank in polls. More recently they also found topic-shifting to be a good indicator of candidate's relative rankings in polls (Prabhakaran et al., 2014). Argumentation in debates has been studied from the perspective of automatic argument extraction (Cabrio and Villata, 2012) and stance classification (Hasan and Ng, 2013). However, to the best of our knowledge, argumentation has not been explored as a influence rank indicator. Moreover the study of persuasion in the NLP community has been so far limited.

The novelty of our work is the proposal of a method to automatically extract persuasive argumentation features from political debates by means of the use of semantic frames as pivoting features. We have trained a rank Support Vector Machine (SVM) model based on content and persuasive ar-

gumentation features in order to rank debate speakers. Our experimental results on the 20 debates for the Republican primary election show that certain types of persuasive argumentation features such as *Premise* and *Support Relation* appear to be better predictors of a speaker's influence rank compared to basic content features such as unigrams. When combining with content-related features, most persuasive argumentation features give superior performance compared to the baselines.

2 Persuasive Argumentation

Argumentation has been defined as a verbal and social activity of reason which aims to increase the acceptability of a controversial standpoint by putting forward a set of connected propositions intending to justify or refute a standpoint before a rational judge (van Eemeren et al., 1996). Different argumentation theories propose various schemes for describing the underlying structure of an argument (Toulmin., 1958; Walton et al., 2008; Freeman, 2011; Peldszus and Stede, 2013). All these theories generally agree in that an argument can be structured by means of two argument components and two argumentative relations. The argument components include *claims* and *premises*. A claim is a central component of an argument and is characterised as being a controversial statement to be judged as true or false. Moreover a claim cannot be accepted by an audience without additional support. Such support is provided in the form of premises underpinning the validity of the claim. The following sentence illustrates an example¹ of an argument highlighting the claim and premises:

“People aren’t investing in America *because this president has made America a less attractive place for investing and hiring than other places in the world.*” (Former Governor Mitt Romney)

While *argumentation* focuses on the rational support structured to justify or refute a standpoint, *persuasion* focuses on language cues aiming at shaping, reinforcing and changing a response. In persuasive communication such response ranges from perceptions, beliefs, attitudes and behaviours.

¹This is extracted from our Debate corpus transcripts. Bold letters represent the argument and italics the premises.

Persuasive language is characterised by the use of emotive lexicons (e.g., atrocious, dreadful, sensational, highly effective) where the speaker tries to engage with the audience's emotions (Macagno and Walton, 2014). Often words with emotive meanings can present values and assumptions as uncontroversial, acting therefore as potentially manipulative instruments of argumentation (Macagno, 2010). Other characteristics of persuasive language include the use of alliteration, which is a stylistic device characterised by the repetition of first consonants in series of words. This artistic constraint enables the speaker to sway the audience by feeling an urgency towards a rhetorical situation by intensifying any attitude being signified (Bitzer, 1968; Lanham, 1991). The use of a repeating sounds engages auditory senses leading to the evoking of emotions that engage the audience.

The following is an example of persuasive language²:

“I’m convinced that part of **the divide that we’re experiencing in the United States**, which is **unprecedented**, it’s **unnatural**, and it’s **un-American**, is *because we’re divided economically, too few jobs, too few opportunities*” (Former Governor Huntsman).

To the best of our knowledge however, the study of the relation of persuasion and argumentation in political debates is limited. One of the main challenges is the lack of annotated corpora which include both argument annotations and persuasive messages annotations. While there has recently been released a corpus of persuasive essays (Stab and Gurevych, 2014) containing annotations for both class-level argument components and argument relations, there is yet none annotated corpora for persuasive arguments in political debates. In order to study whether persuasive cues and persuasive argumentation can be used as predictors of speakers' influence ranking on a debate, we propose to bridge between existing persuasive and political corpora through semantic frame features. The following section introduces the proposed strategy to port annotation between two corpora.

²Representing emotive language in italic bold and alliteration in bold underscored.

3 Extracting Persuasive Argumentation Features from Political Debates

In order to study whether persuasive argumentation can be used as predictors of speakers’ influence ranking on a debate, we propose to use the persuasive essays corpus compiled by Stab and Gurevych (2014) to study persuasive argumentation in political debates through the use of semantic frames.

3.1 Persuasive Essays (PE) Corpus

A persuasive essay is an essay written with the aim of convincing a reader on adopting a way of thinking regarding a stance taken on a topic. Unlike speech where an audience can be persuaded by means of social features or speech style, essays only rely on the written word depending therefore solely on the writer’s persuasive style.

The Persuasive Essays (PE) corpus consists of 90 essays comprising 1,673 sentences. It contains annotations for both class-level argument components and argument relations. The class-level annotations include: 1) major claims; 2) claims; 3) premises and 4) the argumentative relations being either “support” or “attack”. Argumentative relations are directed relations between source and target components (e.g., between premises, claims and major claims). The PE argument annotations follows the scheme described in Table 1.

| | |
|---------------|---|
| Claim | Controversial Statement which is either true or false, and which should not be accepted or otherwise without additional support |
| Premise | Justifies the validity of a claim |
| ForStance | Indicates that an argument supports a claim |
| AgainstStance | Indicates that an argument refutes a claim |
| SupportRel. | Indicates which supporting premises belong to a claim |
| AttackRel. | Indicates which refuting premises belong to a claim |

Table 1: Persuasive essays argument annotation scheme.

3.2 Presidential Political Debates (PD) Corpus

Presidential political debates enable candidates to expose and discuss their stances on policy issues contrasting them with other candidates’ stances. During a debate, speakers unveil their discourse style as well as the premises supporting their claims. For our experiments, we collected the manual transcripts of debates for the Republican party presidential primary election from The American Presidency

Project³. This political debates corpus (PD) consists of 20 debates which took place between May 2011 and February 2012. A total of 10 candidates participated in these debates with an average participation of 6.7 candidates per debate. This corpus comprises 30-40 hours of interaction time and an average of 20,466.6 words per debate.

These debates follow a common structure in which a moderator directly addresses questions to the candidates where disruptions to answers are common due to interruptions from other candidates. In this corpus, each debate transcript lists the speakers including moderator and candidates and questions asked during the debate. Each transcript also clearly delimits turns between speakers and moderators as well as mark-up occurrences of the audience’s reactions such as booing and laughter.

3.3 Semantic Frames

We propose to make use of the persuasion essays corpus annotations to understand persuasive argumentation in political debates by means of the use of semantic frames. A semantic frame is a description of context in which a word sense is used. We make use of FrameNet (Baker et al., 1998), which consist of over 1000 patterns used in English (e.g., Leadership, Causality, Awareness, and Hostile encounter). In this work we extract such patterns using SEMAFOR (Das et al., 2010).

Consider the sentence in Table 2 in which two semantic frames are detected. Each parsed semantic frame consists of {Frame, SemanticRole, label} providing a higher level characterisation of a text, highlighting the semantics of the discourse used in this text. If such semantic frames appear to be some of the most prominent features for a certain persuasive argumentation annotation scheme (e.g., “Claim”), then we can extract persuasive argumentation features from the unlabelled Political Debates corpus using semantic frames as pivoting features.

In this work we propose to port annotations between the Persuasive Essays (PE) and Political Debates (PD) corpora by means of the use of semantic frames as pivoting features.

To represent the PE corpus, let $A = \{a_1, \dots, a_n\}$

³<http://www.presidency.ucsb.edu/debates.php>

| | | |
|-------------------|--|-------------------|
| Sentence: | What we need in this country is to use this issue as a national security tool. | |
| FRAME | SEMANTIC ROLE | LABEL |
| Political_locales | Target | national security |
| Point_of_dispute | Target | this issue |

Table 2: Semantic frames parsed for a sentence extracted from the debates dataset.

be the set of annotation schemes described in Table 1 and let $T_a = \{t_1, \dots, t_n\}$ be the collection of sentences annotated with argument scheme a . To represent the PD corpus, let's $U_D = \{u_1, \dots, u_n\}$ be the set of speakers taking part on a debate D . Let $S_{uD} = \{s_1, \dots, s_n\}$ be the set of sentences generated by speaker u on debate D .

Taking the PE corpus as a reference corpus, we propose to generate a vector representation of each annotation scheme in A for each speaker in each debate of corpus PD by following the steps below: i) Based on tf-idf we extract the most representative semantic frames for each annotation scheme a in PE as the vector SF_a ; ii) We compute the weighted representation of each annotation scheme a in the PD corpus as the vector $f_{d,a}^u$ for each speaker u on each debate d as follows: a) First we compute the bag of semantic frames SF_d^u from speaker u in debate d based on the speaker's content on the debate; then b) For each annotation scheme a we weight vector $f_{d,a}^u$ based on the normalised frequency of each semantic frame element in SF_d^u appearing in SF_a .

3.4 Semantic Frames and Argument Types

The statistics of the extracted semantic frames from PE for each argument type are presented in Table 3.

| Arg. Type | Sentences | Semantic Frames (SF) |
|---------------|-----------|----------------------|
| Claim | 519 | 404 |
| Premise | 1,033 | 518 |
| ForStance | 365 | 369 |
| AgainstStance | 64 | 173 |
| SupportRel. | 1,312 | 535 |
| AttackRel. | 161 | 275 |

Table 3: Number of semantic frames extracted from PE.

Such semantic frames provide a vector representation characterising each persuasive argumentation scheme described in Table 1. Table 4 presents a sample of the top semantic frames representing each ar-

| Arg. Type | Top 5 Semantic Frames |
|---------------|--|
| Claim | Reason, Stage_of_Progress, Evaluative_Comparison, Competition, Cause_to_Change |
| Premise | Removing, Inclusion, Killing, Cognitive_Connection, Causation |
| ForStance | Cause_to_Make_Progress, Collaboration, Purpose, Kinship, Expensiveness |
| AgainstStance | Intentionally_Act, Importance, Capability, Leadership, Usefulness |
| SupportRel. | Dead_or_Alive, Institutions, State_Continue, Taking_Sides, Reliance |
| AttackRel. | Usefulness, Likelihood, Desiring, Importance, Intentionally_Act |

Table 4: Top 5 semantic frames for each argument type of PE.

gumentation type.

Using the vector representation of each annotation scheme generated from PE, we computed the persuasive argumentation features for the PD corpus. Table 5 presents a sentence sample for each argument type identified in the PD corpus along with the semantic frames characterising the sentence.

4 Influence Ranking in Political Debates

We study a speaker's influence on an audience based on his/her persuasiveness language and argumentation styles during a political debate. To measure how influential a speaker is on an audience, we make use of the *influence index* (Prabhakaran et al., 2013), which is calculated based on a speakers relative standing on poll released prior to the debate.

Poll scores describe the influence a speaker has to favourably change the electorate position towards his/her campaign. Given a debate D and the set of speakers U_D we retrieve the poll results released prior to the debate and use the percentage of electorate supporting each candidate. If for a given debate there are multiple polls then the index is computed taking the mean of poll scores. Therefore the influence index P of speaker $u \in U_D$ is:

$$P(u) = \frac{1}{|\text{polls}(D)|} \sum_{i=1}^{|\text{polls}(D)|} p_i \quad (1)$$

where p_i is the poll percentage assign to speaker u in poll i in the reference polls.

| Arg. Type | Sentence | Semantic Frames |
|---------------|--|--|
| Claim | If we can turn Syria and Lebanon away from Iran, we finally have the capacity to get Iran to pull back. | Cause_Change, Manipulation, Capability |
| Premise | Because they put that money in, the president gave the companies to the UAW, they were part of the reason the companies were in trouble. | Causation, Predicament, Leadership |
| ForStance | And the reason is because that's how our founding fathers saw this country set up. | Reason, Kinship, Perception.Experience |
| AgainstStance | I was concerned that if we didn't do something, there were some pretty high risks that not just Wall Street banks, but all banks would collapse. | Emotion_Directed, Intentionally_Affect, Daring |
| SupportRel. | I went to Washington, testifying in favor of a federal amendment to define marriage as a relationship between man and a woman. | Taking_Sides, Cognitive_Connection, Evidence |
| AttackRel. | But you can't stand and say you give me everything I want or I'll vote no. | Desiring, Posture, Capability |

Table 5: Example sentence for each argument type and its corresponding semantic frames identified from PD. Note that there is no annotation in PD. The argument types here are assigned manually for easy reference.

4.1 Features

We characterise each speaker in each debate based on the content and emotion cues he/she generated. Specifically we analyse each candidate in three dimensions: i) what they said (content features); ii) the persuasiveness of the language they used including persuasive argumentation features and emotive language; iii) and external emotions evoked during the debates. We described each set of features below.

4.1.1 Content Features

We use a set of features which characterise content of a candidate's participation on a debate (Prabhakaran et al., 2013). These include: 1) Unigrams (UG), which represents lexical patterns by counting frequencies of word occurrences; 2) Question Deviation (QD), difference between observed percentage of questions asked to a candidate and the fair share percentage of questions in the debate; 3) Word Deviation (WD), difference between observed percentage of words spoken by a candidate and the fair share percentage of words in the debate; 4) Mention Percentage (MP), a candidate mention counts normalised based on all candidates' mentions in a debate.

4.1.2 Persuasiveness Features

We represent three types of persuasiveness features as follows:

1) *Persuasive Argumentation Features*. Following the method described in the previous section, we extract the semantic frame feature vector representing each annotation scheme ($f_{d,a}^u$) for each speaker on each debate. These vectors provide information of different argumentation dimensions. We have extracted a total of 710 semantic frames in PD.

2) *Alliteration*. After removing stopwords, we computed alliteration as repetitions of part of a word or a full word within a sentence.

3) *Emotive Language*. To characterise the use of emotive language, we generated a list of emotion-related semantic frames (e.g., emotion_directed, emotions_by_stimulus, emotions_by_possibility)⁴, then for each speaker u in each debate d , we generated an emotion-frame vector weighted by tf-idf.

Once the features for each speaker have been generated, we followed a supervised learning approach for ranking speakers of a debate based on their influence Index, which can be used to denote how well a speakers participation on a debate has impacted the audience endorsement of his/her campaign.

4.1.3 External Emotion Cues

Previous work (Strapparava et al., 2010) has shown that an audiences' social signal reactions to an idea, such as booing or cheering, are good pre-

⁴FrameNet's frame index, <http://tinyurl.com/q2ytth9>

dictors of hot-spots where persuasion attempts succeeded or at least such attempts were recognised by the audience. In this work, rather than recognising such persuasion hot-spots, we explore these audiences’ reaction cues (e.g applause) as potential predictors of a candidate success on a political debate, we refer to such cues as external emotion cues. For each speaker in a debate, we computed the number of i) applauses (APL); ii) booings (BOO); iii) laughs (LAU); and iv) crosstalks (CRO) he/she received during his/her participation on a debate. These counts were normalised based on the total number of each emotion appeared on the debate.

With these features, we train a supervised learning classifier for ranking speakers of the debates based on their influence indices described in the following section.

4.2 Influence Ranking Approach

In ranking, a training set consists of an ordered data set. Let “A is preferred to B” be denoted as “ $A \succ B$ ”. Let D denote a debate with a set of speakers $U_D = \{u_1, u_2, ..u_n\}$ and influence indexes $P(u_i)$ for $1 < i < n$. We specify a training set for ranking as $R = \{(u_i, \gamma_i), .., (u_n, \gamma_n)\}$ where γ_i is the ranking of u_i based on its $P(u_i)$ so $\gamma_i < \gamma_j$ if $u_i \succ u_j$.

We want to find a ranking function F which outputs a score for each instance from which a global ordering of data is constructed. So the target function $F(u_i)$ outputs a score such that $F(u_i) > F(u_j)$ for any $u_i \succ u_j$. In this work we use the Ranking SVM (Joachims, 2006) to estimate the ranking function F .

5 Experimental Setup

For our experiments we used the Persuasive Essays (PE) and Political Debates (PD) corpora introduced in previous sections. While the PE was used as a reference corpus, all our experiments were performed on the PD corpus.

All features are computed for the aggregation of a candidate’s content in a debate. For content and alliteration features, we first removed stopwords. In particular, for computing unigram features we also stemmed words using a Porter stemmer (Porter, 1997).

To compute persuasive argumentation features we used the collection of semantic frame features for the reference corpus PE.

5.1 Evaluation

In this work, the ranking task evaluation for each debate consists on comparing the generated ranked list of candidates, using the *influence ranking approach* introduced above, against a reference ranked list. Such a reference ranked list corresponds to our gold standard of ranked list of candidates generated based on the polled scores for that debate.

Following a 5-fold cross validation, we report results applying four commonly used evaluation metrics for ranking tasks, *nDCG*, *nDCG-3*, *Kendall’s Tau* and *Spearman* correlations. The discounted cumulative gain metric (*nDCG*) penalises inversions happening at the top n elements⁵ of a ranked list more than those inversions happening at the bottom. While the *nDCG* metric penalises certain elements in the list, Kendall’s tau and Spearman’s rank correlations penalises inversions equally across the ranked list.

6 Results and Discussion

6.1 Correlation Analysis

We performed a correlation analysis for the content and persuasive emotion numeric features⁶. We computed the Pearson’s product correlation between each feature with the candidate’s influence index $P(u)$ derived from the polls. The computed correlations for these features are presented in Figure 1. Darker bars indicate statistical significance correlation at $p < 0.001$; lighter dark bars at $p < 0.05$; and light bars not statistically significant.

These results show that for the content features, both question deviation (QD) and word deviation (WD) correlate moderately with the influence index; while the mention percentage (MP) feature correlates highly with the influence index ($p < 0.05$). For the emotion cues, we obtained statistically significant ($p < 0.05$) moderate correlations between the applause (APL), laugh (LAU), crosstalk (CRO) and

⁵*nDCG-3* therefore assigns a higher penalisation for inversions happening at the top three elements of a ranked list.

⁶Note that we can perform such correlation analysis only with the numeric features but not with vector features.

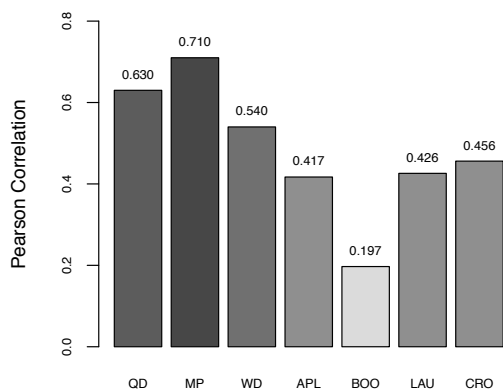


Figure 1: Pearson correlation for content and emotion cues features. Correlation windows: Negligible (0-0.19); Weak (0.2-0.39); Moderate (0.4-0.69); High (>0.69).

the influence index; while the correlation between the booing (BOO) cue and the influence index was not statistically significant. These results indicate that speakers with higher influence index spoke for longer periods of time, in line with existing empirical findings in sociological studies (Ng and Bradac, 1993; Reid and Ng., 2000; Prabhakaran et al., 2013), and were asked a higher number of questions. This analysis also indicates that speakers with higher influence index generated more crosstalk, in line with previous empirical sociological findings (Ng et al., 1995); received more applause and made the audience laugh more often.

6.2 Influence Ranking Results

Following the results of the correlation analysis, we conducted experiments using those content and emotion cue features presenting statistically significant correlations with the influence index. Apart from these features, we also consider the persuasive argumentation features and a combination of features from both content and persuasion categories. Results for the prediction of influence ranking using these features are presented in Table 6.

For the *content features*, using the simple unigrams gives the best results. The mention percentage (MP) feature also attains competitive performance. A combination of word deviation, question deviation and mention percentage (WD+RD+MP) however degrades the performance. This is in contrast to the results reported in (Prabhakaran et al.,

2013) (denoted as [PR13] in Table 6), where the unigram feature gives much worse results and their best results were obtained using WD+RD+MP. One possible reason is that for the unigram feature used in our experiments, we have performed pre-processing by removing stop words and stemming.

For *external emotion cues*, although some emotion cues appeared to be significantly correlated with the influence index in our analysis, they did not outperform the unigram baseline. We suspect that this might be due to the fact that depending on the location of a debate, certain candidates may bring bigger crowds into the debate’s venue, therefore emotion cues can be a deliberate biased way of support. Consequently emotion cues happening within the debate venue may not reflect the emotions induced to the audiences that followed the broadcast of the debate.

When analysing the *persuasion features*, alliteration and emotive language features give better results compared to external emotion cues. But they did not outperform the unigram baseline either.

We find that *persuasive argumentation* features alone provide improvement upon the unigram baseline. In particular, in terms of *nDCG* and *nDCG-3*⁷, the *premise* and *support relation* types provide the best results for this feature category. In terms of *Tau* and *Spearman*⁸ correlations, the *attack relation* type provides the best results. When focusing only on persuasive argumentation features, these results suggest that speakers with higher influence index tend to use well supported arguments (i.e. present more premisses supporting their claims) and/or tend to attack more other candidates by presenting premisses refuting a claim.

When *combining features*, we found that the top 100 persuasive argumentation features ranked by *tfidf* together with word deviations and mention percentages significantly improve upon the baselines for particular argumentation cases including Claim, Premise, ForStance, and SupportRel.

The overall best performing features for predicting influence ranking in terms of *nDCG*, *Tau* and *Spearman* was consistently obtained with the com-

⁷Since *nDCG* and *nDCG-3* penalises inversions at the top of the list, good results in these metrics mean predicting accurately the top of the ranked list.

⁸These metrics provide a general evaluation of the accuracy of the full ranked list.

bined feature for the *Premise* type.

Our results improve upon those recently obtained in (Prabhakaran et al., 2014) in both *nDCG* and *Tau* where topic shift patterns have been added for influence ranking (denoted as [PR14] in Table 6).

These results suggest the relevance of “what they said”, the “persuasiveness style of their arguments” and the relative importance given by others by means of mentions are good predictors of influence ranking in political debates. In particular when combining the lexical content of candidates’ discourse with their persuasive argumentation style, our results indicate that candidates with higher influence ranking tend to present more premises while clearly stating their stance (i.e. supporting a claim) on a particular topic.

7 Conclusions and Future Work

In this paper, we have studied the impact of argumentation in speaker’s discourse and their effect in influencing an audience on supporting their candidature. In particular, we have conducted the study in the domain of political debates. In order to extract persuasive argumentation features from political debates, we have proposed a novel method to port annotations from a persuasive essay corpus using semantic frames as pivot features.

Our experimental results on the 20 debates for the Republican primary election show that when combined with word deviations and mention percentages, most persuasive argumentation features give superior performance compared to the baselines. Particularly with the *Premise* and *SupportRel* types appear to be better predictors of a speaker’s influence rank. In future work, we will aim to improve the accuracy of the extracted persuasive argumentation features by exploring other methods for identifying persuasive argumentations from text.

Acknowledgments

This work was supported by the EU-FP7 project SENSE4US (grant no. 611242)

References

C. F. Baker, C. J. Fillmore, and J. B. Lowe. 1998. The Berkeley FrameNet project. In *Proceedings of the*

| Feature | nDCG | nDCG3 | Tau | Spearman |
|---|--------------|--------------|--------------|--------------|
| Content Features | | | | |
| UN | 0.964 | 0.924 | 0.597 | 0.723 |
| MP | 0.962 | 0.915 | 0.597 | 0.712 |
| WD+QD+MP | 0.957 | 0.905 | 0.571 | 0.689 |
| UN[PR13] | 0.860 | 0.733 | 0.250 | - |
| WD+QD+MP[PR13] | 0.961 | 0.921 | 0.470 | - |
| WD+QD+MP+TopicShift [PR14] | 0.970 | 0.937 | 0.600 | - |
| External Emotion Cues | | | | |
| Applause | 0.886 | 0.746 | 0.373 | 0.481 |
| Laughter | 0.870 | 0.721 | 0.335 | 0.423 |
| Crosstalk | 0.836 | 0.641 | 0.314 | 0.416 |
| Persuasion Features | | | | |
| Alliteration | 0.929 | 0.856 | 0.379 | 0.494 |
| Emo Lan- guage | 0.928 | 0.846 | 0.478 | 0.588 |
| Persuasive Argumentation | | | | |
| Claim | 0.957 | 0.902 | 0.559 | 0.691 |
| Premise | 0.968 | 0.927 | 0.600 | 0.731 |
| ForStance | 0.959 | 0.906 | 0.557 | 0.690 |
| AgainstStance | 0.951 | 0.897 | 0.512 | 0.643 |
| SupportRel. | 0.967 | 0.921 | 0.600 | 0.728 |
| AttackRel. | 0.964 | 0.910 | 0.632 | 0.739 |
| Combined Features | | | | |
| Persuasive Argumentation + WD + MP | | | | |
| Claim | 0.965 | 0.922 | 0.613 | 0.734 |
| Premise | 0.973 | 0.931 | 0.639 | 0.766 |
| ForStance | 0.965 | 0.924 | 0.590 | 0.724 |
| AgainstStance | 0.953 | 0.894 | 0.524 | 0.669 |
| SupportRel. | 0.971 | 0.929 | 0.619 | 0.753 |
| AttackRel. | 0.963 | 0.920 | 0.582 | 0.726 |

Table 6: Influence ranking results for baselines, persuasion and combined features. Statistically Significant at $p < 0.05$.

17th International Conference on Computational Linguistics (COLING), pages 86–90.

Lloyd Bitzer. 1968. *The Rhetorical Situation*. Philosophy and Rhetoric.

Filip Boltužić and Jan Šnajder. 2014. Back up your stance: Recognizing arguments in online discussions. In *Proceedings of the 1st Workshop on Argumentation Mining*, pages 49–58.

Elena Cabrio and Serena Villata. 2012. Natural language arguments: A combined approach. In *Proceedings*

- of the *European Conference on Artificial Intelligence (ECAI)*, pages 205–210.
- D. Das, N. Schneider, D. Chen, and N. A. Smith. 2010. Semafor 1.0: A probabilistic frame-semantic parser. Technical report, Carnegie Mellon University Technical Report CMU-LTI-10-001.
- James B. Freeman. 2011. *Argument Structure: Representation and Theory*, volume 18. Springer.
- Sean Gerrish and David Blei. 2011. Predicting legislative roll calls from text. In *In Lise Getoor and Tobias Scheffer, editors, Proceedings of the 28th International Conference on Machine Learning (ICML)*, pages 489–496.
- Kazi Saidul Hasan and Vincent Ng. 2013. Stance classification of ideological debates: Data, models, features, and constraints. In *Proceedings of the Sixth International Joint Conference on Natural Language Processing (IJCNLP)*, pages 1348–1356.
- T. Joachims. 2006. Training linear SMMs in linear time. In *Proceedings of the ACM Conference on Knowledge Discovery and Data Mining (KDD)*, pages = 217–226.
- Richard Lanham. 1991. *A Handlist of Rhetorical Terms*. Los Angeles: University of California Press.
- F. Macagno and D. Walton. 2014. *Emotive Language in Argumentation*. Cambridge Press.
- Fabrizio Macagno. 2010. The argumentative uses of emotive language. *Revista Iberoamericana de Argumentacion*, pages 1–33.
- S. H. Ng and J. J. Bradac. 1993. *Power in language: Verbal communication and social influence*. Sage Publications, Inc.
- S. H. Ng, M Brooke, and M. Dunne. 1995. Interruption and influence in discussion groups. *Journal of Language and Social Psychology*, 14(4):369–381.
- Andreas Peldszus and Manfred Stede. 2013. From argument diagrams to argumentation mining in texts: A survey. *International Journal of Cognitive Informatics and Natural Intelligence (IJCINI)*, 7(1):1–31.
- M. F. Porter. 1997. Readings in information retrieval. chapter An Algorithm for Suffix Stripping, pages 313–316.
- Vinodkumar Prabhakaran, Ajita John, and D. Dorée Seligmann. 2013. Who had the upper hand? ranking participants of interactions based on their relative power. In *Proceedings of the 6th International Joint Conference on Natural Language Processing (IJCNLP)*, pages 365–373.
- Vinodkumar Prabhakaran, Ashima Arora, and Owen Rambow. 2014. Staying on topic: An indicator of power in political debates. In *Proceedings of the conference on Empirical Methods for Natural Language Processing (EMNLP)*.
- S. A. Reid and S. H. Ng. 2000. Conversation as a resource for influence: evidence for prototypical arguments and social identification processes. *European Journal of Social Psych.*, (30):83–100.
- Andrew Rosenberg and Julia Hirschberg. 2009. Charisma perception from text and speech. *Speech Communication*, 51(7):640–655.
- Yanchuan Sim, Brice D. L. Acree, Justin H. Gross, and Noah A. Smith. 2013. Measuring ideological proportions in political speeches. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 91–101.
- Carlo Strapparava, Marco Guerini, and Oliviero Stock. 2010. Predicting persuasiveness in political discourses. In *Proceedings of the 7th International Conference on Language Resources and Evaluation (LREC)*.
- Matt Thomas, Bo Pang, and Lillian Lee. 2006. Get out the vote: Determining support or opposition from congressional floor-debate transcripts. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 327–335.
- Stephen E. Toulmin. 1958. *The uses of Argument*. Cambridge University Press.
- van Eemeren et al. 1996. *Fundamentals Of Argumentation Theory: A Handbook Of Historical Backgrounds And Contemporary Developments*. Hillsdale, NJ, England: Lawrence Erlbaum Associates, Inc, PsycINFO, EBSCOhost.
- Douglas Walton, Chris Reed, and Fabrizio Macagno. 2008. *Argumentation Schemes*. Cambridge University Press.