

Identifying Prominent Arguments in Online Debates Using Semantic Textual Similarity

Filip Boltužić and Jan Šnajder

Text Analysis and Knowledge Engineering Lab
FER, University of Zagreb

Second Workshop on Argumentation Mining
NAACL 2015
Denver, Colorado
4 June 2015

Should marijuana be legalized?

User comment 1

No, because marijuana lessen the brain's ability for cognitive thinking.

User comment 2

There have been plenty of highway deaths associated with marajuanna use.

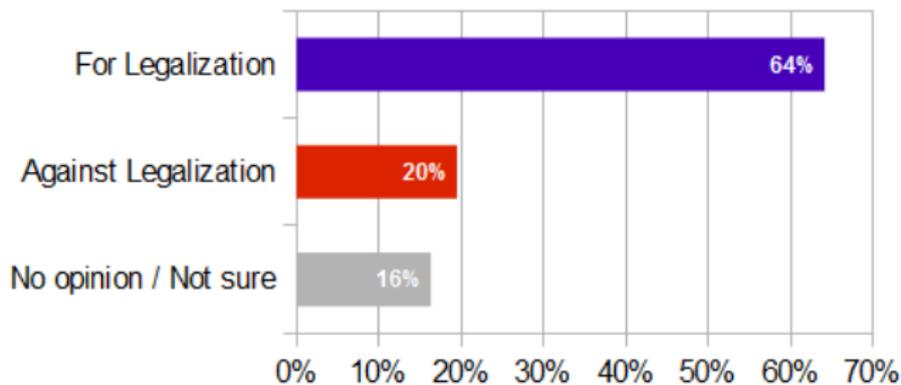
User comment 3

The Legalization of marijuana would lower are crime rates in the United States of America by at least 15 to 20

User comment 4

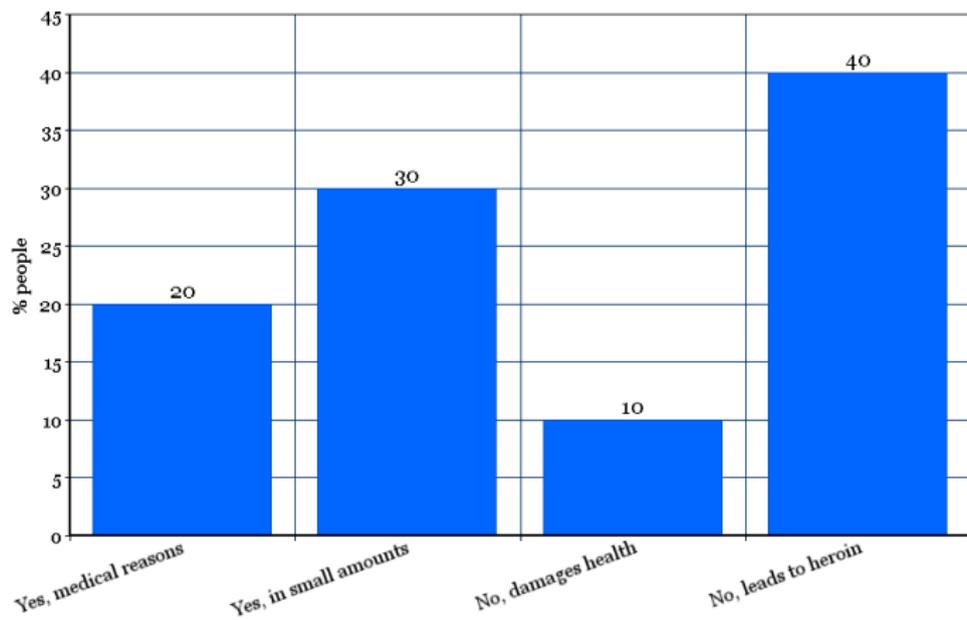
Marijuana is proven to cause depression and change brain patterns in odd ways among other things

Should marijuana be legalized?



Should marijuana be legalized?

Marijuana arguments



Should marijuana be legalized?

User comment 1

No, because marijuana lessen the brain's ability for cognitive thinking.

User comment 4

Marijuana is proven to cause depression and change brain patterns in odd ways among other things

Should marijuana be legalized?

No, damages health

User comment 1

No, because marijuana lessen the brain's ability for cognitive thinking.

User comment 4

Marijuana is proven to cause depression and change brain patterns in odd ways among other things

Online Discussions

- Online discussions growing source of mass opinion
- Expressing opinion varies:
implicit premises, value judgements, irony

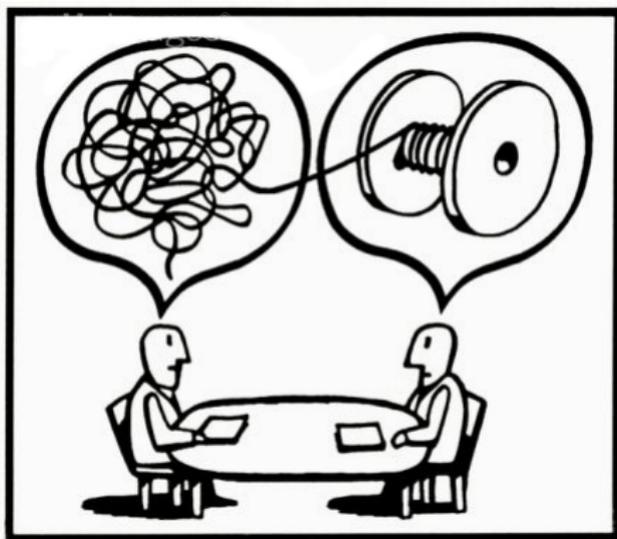
Online Discussions

- Online discussions growing source of mass opinion
- Expressing opinion varies:
implicit premises, value judgements, irony



Arguments from opinions

- Clustering similar opinions gives an argument
- Arguments may be related



Identifying Prominent Arguments

Identifying reasonings and opinions to cluster into arguments.

Identifying Prominent Arguments

Identifying reasonings and opinions to cluster into arguments.

Input:

- 1 **Noisy comments** from online discussions

Output:

- 1 Set of **Argument Clusters**
- 2 **Representative Argument** of each **Cluster**

- Argumentation mining [Palau and Moens, 2009]
- Argument supervised classification
 - Argument recognition [Boltužić and Šnajder, 2014]
 - Reason classification [Hasan and Ng, 2014]
 - Argument tags [Conrad et al., 2012]
- Argument unsupervised topic modeling
 - Identifying arguing expressions [Trabelsi and Zaïane, 2014]
- Stance classification
 - Stance on forum posts [Anand et al., 2011]

Outline

- 1 Corpus
- 2 Model
- 3 Evaluation

Outline

- 1 Corpus
- 2 Model
- 3 Evaluation

- [Hasan and Ng, 2014] annotated threaded debates with arguments
- We extract pairs of gold arguments and comments
- Ignoring non-argumentative content
- Sentence level comments

- [Hasan and Ng, 2014] annotated threaded debates with arguments
- We extract pairs of gold arguments and comments
- Ignoring non-argumentative content
- Sentence level comments

Comment

Medically speaking marijuana is one of the safest and most effective medications for the widest variety diseases known

Gold Argument

Used as a medicine for its positive effects

- Majority pro – 2028 (65%)
- Four topics
 - Should gay marriage be legal?
 - Should marijuana be legalized?
 - Is Obama a good president?
 - Should abortion be legalized?

- Majority pro – 2028 (65%)
- Four topics
 - Should gay marriage be legal?
 - Should marijuana be legalized?
 - Is Obama a good president?
 - Should abortion be legalized?

	GM		MAR		OBA		ABO	
	Pro	Con	Pro	Con	Pro	Con	Pro	Con
#Arguments	5	4	5	5	8	8	7	5
#Comments	639	197	585	239	358	272	446	368

Outline

- 1 Corpus
- 2 Model**
- 3 Evaluation

- Vector-space similarity
 - Bag-of-words (BoW)
 - Inverse sentence frequency weight
 - Neural network skip-gram [Mikolov et al., 2013]
 - Word-vector sum for sentences
 - Cosine distance
- Semantic textual similarity (STS) [Šarić et al., 2012]
 - Text comparison features
 - Output real valued similarity score

Hierarchical agglomerative clustering (HAC) [Xu et al., 2005]

- **Input:** Distance matrix
- **Output:** Hierarchical structures

Linkage criterion

- Complete linkage
- Ward's method

Outline

- 1 Corpus
- 2 Model
- 3 Evaluation**

Evaluation metrics

- Comparison against gold corpus labels
- Hierarchical clustering stopping criteria #gold labels

Supervised measures

- Adjusted Rand Index (ARI)
- V-measure (V)



Cluster evaluation

Model (linkage)	OBA		MAR		GM		ABO	
	V	ARI	V	ARI	V	ARI	V	ARI
STS (Complete)	.11	.02	.05	.03	.05	.01	.06	.02
BoW (Complete)	.15	.03	.04	.00	.04	.01	.04	.01
BoW (Ward's)	.27	.04	.17	.02	.15	.04	.24	.07
Skip-gram (Complete)	.21	.04	.13	.02	.10	.04	.20	.03
Skip-gram (Ward's)	.30	.10	.25	.19	.15	.07	.23	.08
Skip-gram (Ward's) pro/con	.24	.08	.25	.20	.16	.07	.20	.07

- Ward's linkage best performance
- Word embeddings best performance
- Stance separated improves performance on two topics

Cluster matching

- Manual cluster matching to **gold arguments** on MAR topic
- Mediod cluster representative
- Compare medoid to **gold label**

Cluster matching

- Manual cluster matching to **gold arguments** on MAR topic
- Mediod cluster representative
- Compare medoid to **gold label**



Cluster matching example

Example 1

Cluster medoid

the economy would get billions of dollars in a new industry if it were legalized (...) no longer would this revenue go directly into the black market.

Gold argument

Legalized marijuana can be controlled and regulated by the government

Cluster matching example

Example 1

Cluster medoid

the economy would get billions of dollars in a new industry if it were legalized
(...) no longer would this revenue go directly into the black market.

Gold argument

Legalized marijuana can be controlled and regulated by the government

Example 2

Cluster medoid

There are thousands of deaths every year from tobacco and alcohol, yet there has never been a recorded death due to marijuana.

Gold argument

Does not cause any damage to our bodies

Main problems identified

- Background knowledge
- Idiomatic language
- Grammatical errors
- Fine/coarse arguments



Comment

Pot is also one of the most high priced exports of Central American Countries and the Carribean

Comment

Pot is also one of the most high priced exports of Central American Countries and the Carribean

- Not addictive

Comment

Pot is also one of the most high priced exports of Central American Countries and the Caribbean

- Not addictive
- Legalized marijuana can be controlled and regulated by the government

Error analysis: Argument granularity

Specific

Damages our bodies | Responsible for brain damage

Damaging our bodies

General

the economy would get billions of dollars
(...) no longer would this revenue go di-
rectly into the black market.

Economy profits

If the tax on cigarettes can be
\$5.00/pack imagine what we could
tax pot for!

Tax benefits

Legalized marijuana can be controlled and regulated by the government

- Baseline unsupervised identification of prominent arguments
- Hierarchical clustering
 - Textual similarity measure
 - 0.15 to 0.30 V-measure

- Baseline unsupervised identification of prominent arguments
- Hierarchical clustering
 - Textual similarity measure
 - 0.15 to 0.30 V-measure
- Future work
 - Semi-supervised approach
 - Argument hierarchy analysis

References I

-  Anand, P., Walker, M., Abbott, R., Tree, J. E. F., Bowmani, R., and Minor, M. (2011).
Cats rule and dogs drool!: Classifying stance in online debate.
In Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, pages 1–9.
-  Boltužić, F. and Šnajder, J. (2014).
Back up your stance: Recognizing arguments in online discussions.
In Proceedings of the First Workshop on Argumentation Mining, pages 49–58.
-  Conrad, A., Wiebe, J., et al. (2012).
Recognizing arguing subjectivity and argument tags.
In Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics, pages 80–88.

References II



Hasan, K. S. and Ng, V. (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.



Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013).

Efficient estimation of word representations in vector space.

In Proceedings of ICLR, Scottsdale, AZ, USA.



Palau, R. M. and Moens, M.-F. (2009).

Argumentation mining: The detection, classification and structure of arguments in text.

In Proceedings of the 12th International Conference on Artificial Intelligence and Law, pages 98–107. ACM.

-  Šarić, F., Glavaš, G., Karan, M., Šnajder, J., and Dalbelo Bašić, B. (2012).
Takelab: Systems for measuring semantic text similarity.
In Proceedings of the Sixth International Workshop on Semantic Evaluation (SemEval 2012), pages 441–448, Montréal, Canada.
-  Trabelsi, A. and Zaïane, O. R. (2014).
Finding arguing expressions of divergent viewpoints in online debates.
In Proceedings of the 5th Workshop on Language Analysis for Social Media (LASM)@ EACL, pages 35–43.
-  Xu, R., Wunsch, D., et al. (2005).
Survey of clustering algorithms.
Neural Networks, IEEE Transactions on, 16(3):645–678.