Finding Arguing Expressions of Divergent Viewpoints in Online Debates

Amine Trabelsi
Department of Computing Science
University of Alberta
`atrabels@ualberta.ca`

Osmar R. Zaïane
Department of Computing Science
University of Alberta
`zaiane@ualberta.ca`

Abstract

This work suggests a fine-grained mining of contentious documents, specifically online debates, towards a summarization of contention issues. We propose a Joint Topic Viewpoint model (JTV) for the unsupervised identification and the clustering of arguing expressions according to the latent topics they discuss and the implicit viewpoints they voice. A set of experiments is conducted on online debates documents. Qualitative and quantitative evaluations of the model's output are performed in context of different contention issues. Analysis of experimental results shows the effectiveness of the proposed model to automatically and accurately detect recurrent patterns of arguing expressions in online debate texts.

1 Introduction

This paper addresses the issue of improving the quality of opinion mining from online contentious texts like the posts in debate sites. Mining and summarizing these new resources is crucial, especially when the opinion is related to a subject that stimulates divergent viewpoints within people (e.g. Healthcare Reform, Same-Sex Marriage). We refer to such subjects as issues of contentions. A **contentious issue** is "*likely to cause disagreement between people*" (cf. Oxford Dictionaries). Documents such as debate sites' posts may contain multiple contrastive viewpoints regarding a particular issue of contention. Table 1 presents an example of short-text documents expressing divergent opinions where each is exclusively supporting or opposing a healthcare legislation [1].

---

[1] extracted from a Gallup Inc. survey http://www.gallup.com/poll/126521/favor-oppose-obama-healthcare-plan.aspx

Opinion in contentious issues is often expressed implicitly, not necessarily through the usage of usual negative or positive opinion words, like "bad" or "great". This makes its extraction a challenging task. It is usually conveyed through the arguing expression justifying the endorsement of a particular point of view. The act of arguing is "to give reasons why you think that something is right/wrong, true/not true, etc, especially to persuade people that you are right" (cf. Oxford Dictionaries). For example, the arguing expression "many people do not have healthcare", in Table 1, implicitly explains that the reform is intended to fix the problem of uninsured people, and thus, the opinion is probably on the supporting side. On the other hand, the arguing expression "it will produce too much debt" denotes the negative consequence that may result from passing the bill, making it on the opposing side.

The automatic identification and clustering of these kind of arguing expressions, according to their topics and the viewpoints they convey, is enticing for a variety of application domains. For instance, it can save journalists a substantial amount of work and provide them with drafting elements (viewpoints and associated arguing expressions) about controversial issues. In addition, it would enhance the output quality of the opinion summarization task in general.

The rest of this paper is organized as follows. Section 2 covers the details of the problem statement. Section 3 explains the key issues in the context of recent related work. Section 4 provides the technical details of our model, the Joint Topic Viewpoint model (JTV) . Section 5 describes the clustering task that might be used to obtain a feasible solution. Section 6 provides a description of the experimental set up. Section 7 assesses the adequacy and the performance of our solution. Section 8 concludes the paper.

Support Viewpoint	Oppose Viewpoint
Many people do not have health care	The government should not be involved
Provide health care for 30 million people	It will produce too much debt
The government should help old people	The bill would not help the people

Table 1: Excerpts of support and opposition opinion to a healthcare bill in the USA.

2 Problem Statement

This paper examines the task of mining the topics and the viewpoints of arguing expressions towards the summarization of contentious text. An example of a human-made summary of arguing expressions (Jones, 2010) on, what is commonly known as, the Obama healthcare reform is presented in Table 2. Ultimately, the target is to automatically generate similar summaries given a corpus of contentious documents. However, this paper tackles the sub-problem of identifying recurrent words and phrases expressing arguing and cluster them according to their topics and viewpoints. This would help solve the general problem. We use Table 2's examples to define some key concepts which can help us formulate this latter. Here, the contentious issue yielding the divergent positions is the Obama healthcare. The documents are people's verbatim responses to the question "Why do you favor or oppose a healthcare legislation similar to President Obama's ?".

We define a ***contention question*** as a question that can generate expressions of two or more divergent viewpoints as a response.

While the previous question explicitly asks for the reasons ("why"), we relax this constraint and consider also usual opinion questions like "Is the passing of Obamacare bad for Americans ?" or "Do you favor or oppose Obamacare ?".

A ***contentious document*** is a document that contains expressions of one or more divergent viewpoints in response to the contention question. In the context of online debate, a post usually expresses one viewpoint, although it can mention arguing used to justify a different viewpoint.

Table 2 is split into two parts according to the viewpoint: supporting or opposing the healthcare bill. Each row contains one or more phrases, each expressing a reason (or an explanation), e.g. "costs are out of control" and "would help control costs". Though lexically different, these phrases share a common hidden theme (or topic), e.g. insurance's cost, and implicitly convey the same hidden viewpoint's semantics, e.g. support the healthcare bill.

Thus, we define an ***arguing expression*** as the set of reasons (words or phrases) sharing a common topic and justifying the same viewpoint regarding a contentious issue.

We assume that a ***viewpoint*** (e.g. a column of Table 2) in a contentious document is a stance, in response to a contention question, which is implicitly expressed by a set of arguing expressions (e.g. rows of a column in Table 2).

Thus, the arguing expressions voicing the same viewpoint differ in their topics, but agree in the stance. For example, arguing expressions represented by "system is broken" and "costs are out of control" discuss different topics, i.e. healthcare system and insurance's cost, but both support the healthcare bill. On the other hand, arguing expressions of divergent viewpoints may have similar topic or may not. For instance, "government should help elderly" and "government should not be involved" share the same topic, i.e. government's role, while conveying opposed viewpoints.

Our research problem and objectives in terms of the newly introduced concepts are stated as follows. Given a corpus of unlabeled contentious documents $\{doc_1, doc_2, .., doc_D\}$, where each document doc_d expresses one or more viewpoints $\vec{v}^d$ from a set of L possible viewpoints $\{v_1, v_2, .., v_L\}$, and each viewpoint v_l can be conveyed using one or more arguing expressions $\vec{\phi_l}$ from a set of possible arguing expressions discussing K different topics $\{\phi_{1l}, \phi_{2l}, .., \phi_{Kl}\}$, the objective is to perform the following two tasks:

1. automatically extracting coherent words and phrases describing any distinct arguing expression ϕ_{kl};

2. grouping extracted distinct arguing expressions ϕ_{kl} for different topics, $k = 1..K$, into their corresponding viewpoint v_l.

This paper focuses on the first task while laying the ground for solving the second one. In carrying out the first task, we must meet the main challenge of recognizing arguing expressions having

Support Viewpoint	Oppose Viewpoint
People need health insurance/too many uninsured	Will raise cost of insurance/ less affordable
System is broken/needs to be fixed	Does not address real problems
Costs are out of control/would help control costs	Need more information on how it works
Moral responsibility to provide/Fair	Against big government involvement (general)
Would make healthcare more affordable	Government should not be involved in healthcare
Don't trust insurance companies	Cost the government too much

Table 2: Human-made summary of arguing expressions supporting and opposing Obamacare.

the same topic and viewpoint but which are lexically different, e.g. "provide health care for 30 million people" and "many people do not have healthcare". For this purpose we propose a Joint Topic Viewpoint Model (JTV) to account for the dependence structure of topics and viewpoints.

3 Related Work

3.1 Classifying Stances

An early body of work addresses the challenge of classifying viewpoints in contentious or ideological discourses using supervised techniques (Kim and Hovy, 2007; Lin et al., 2006). Although the models give good performances, they remain data-dependent and costly to label, making the unsupervised approach more appropriate for the existing huge quantity of online data. A similar trend of studies scrutinizes the discourse aspect of a document in order to identify opposed stances (Thomas et al., 2006; Park et al., 2011). However, these methods utilize polarity lexicon to detect opinionated text and do not look for arguing expression, which is shown to be useful in recognizing opposed stances (Somasundaran and Wiebe, 2010). Somasundaran and Wiebe (2010) classify ideological stances in online debates using a generated arguing clues from the Multi Perspective Question Answering (MPQA) opinion corpus[2]. Our problem is not to classify documents, but to recognize recurrent pattern of arguing phrases instead of arguing clues. Moreover, our approach is independent of any annotated corpora.

3.2 Topic Modeling in Reviews Data

Another emerging body of work applies probabilistic topic models on reviews data to extract appraisal aspects and the corresponding specific sentiment lexicon. These kinds of models are usually referred to as joint sentiment/aspect topic models (Jo and Oh, 2011; Titov and McDonald, 2008;

Zhao et al., 2010). Lin and He (2009) propose the Joint Sentiment Topic Model (JST) to model the dependency between sentiment and topics. They make the assumption that topics discussed on a review are conditioned on sentiment polarity. Reversely, our JTV model assumes that a viewpoint endorsement (e.g., oppose reform) is conditioned on the discussed topic (e.g., government's role) and its application is different from that of JST. Most of the joint aspect sentiment topic models are either semi-supervised or weakly supervised using sentiment polarity words (Paradigm lists) to boost their efficiency. In our case, viewpoints are often expressed implicitly and finding specific arguing lexicon for different stances is a challenging task in itself. Indeed, our model is enclosed in another body of work that based on a probabilistic Topic Model framework to mine divergent viewpoints.

3.3 Topic Modeling in Contentious Text

A recent study by Mukherjee and Liu (2012) examines mining contention from discussion forums data where the interaction between different authors is pivotal. It attempts to jointly discover contention/agreement indicators (CA-Expressions) and topics using three different Joint Topic Expressions Models (JTE). The JTEs' output is used to discover points (topics) of contention. The model supposes that people express agreement or disagreement through CA-expressions. However, this is not often the case when people express their viewpoint via other channels than discussion forums like debate sites or editorials. Moreover, agreement or disagreement may also be conveyed implicitly through arguing expressions rejecting or supporting another opinion. JTEs do not model viewpoints and use the supervised Maximum Entropy model to detect CA-expressions.

Recently, Gottipati et al. (2013) propose a topic model to infer human interpretable text in the do-

---

[2]http://mpqa.cs.pitt.edu/

main of issues using Debatepedia[3] as a corpus of evidence. Debatepedia is an online authored encyclopedia to summarize and organize the main arguments of two possible positions. The model takes advantage of the hierarchical structure of arguments in Debatepedia. Our work aims to model unstructured online data, with unrestricted number of positions, in order to, ultimately, output a Debatepedia-like summary.

The closest work to ours is the one presented by Paul et al. (2010). It introduces the problem of contrastive summarization which is very similar to our stated problem in Section 2. They propose the Topic Aspect Model (TAM) and use the output distributions to compute similarities' scores for sentences. Scored sentences are used in a modified Random Walk algorithm to generate the summary. The assumption of TAM is that any word in the document can exclusively belong to a topic (e.g., government), a viewpoint (e.g., good), both (e.g., involvement) or neither (e.g., think). However, according to TAM's generative model, an author would choose his viewpoint and the topic to talk about independently. Our JTV encodes the dependency between topics and viewpoints.

4   Joint Topic Viewpoint Model

Latent Dirichlet Allocation (LDA) (Blei et al., 2003) is one of the most popular topic models used to mine large text data sets. It models a document as a mixture of topics where each topic is a distribution over words. However, it fails to model more complex structures of texts like contention where viewpoints are hidden.

We augment LDA to model a contentious document as a pair of dependent mixtures: a mixture of arguing topics and a mixture of viewpoints for each topic. The assumption is that a document discusses the topics in proportions, (e.g. 80% government's role, 20% insurance's cost). Moreover, as explained in Section 2, each one of these topics can be shared by divergent arguing expressions conveying different viewpoints. We suppose that for each discussed topic in the document, the viewpoints are expressed in proportions. For instance, 70% of the document's text discussing the government's role expresses an opposing viewpoint to the reform while 30% of it conveys a supporting viewpoint. Thus, each term in a document is assigned a pair topic-viewpoint label (e.g.
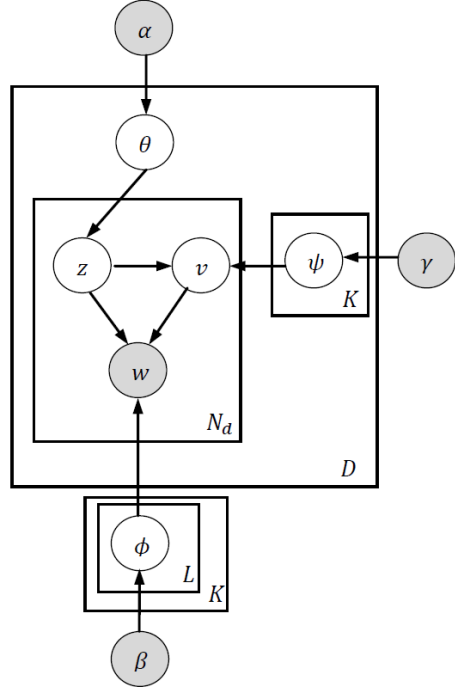
Figure 1: The JTV's graphical model (plate notation)

"government's role-oppose reform"). A term is a word or a phrase i.e. n-grams ($n>1$). For each topic-viewpoint pair, the model generates a topic-viewpoint probability distribution over terms. This topic-viewpoint distribution would corresponds to what we define as an arguing expression in Section 2, i.e. a set of terms sharing a common topic and justifying the same viewpoint regarding a contentious issue. The Joint Topic Viewpoint (JTV), is similar to the Joint Sentiment Topic model (JST) (Lin and He, 2009), as it models documents as two dependent mixtures. However, here we condition viewpoints on topics instead of conditioning topics on sentiment. Moreover, the application is different from that of JST which intend to model reviews data.

Formally, assume that a corpus contains D documents $d_{1..D}$, where each document is a term's vector $\vec{w}_d$ of size N_d; each term w_{dn} in a document belongs to the corpus vocabulary of distinct terms of size V. Let K be the total number of topics and L be the total number of viewpoints. Let θ_d denote the probabilities (proportions) of K topics under a document d; ψ_{dk} be the probability distributions (proportions) of L viewpoints for a topic k in the document d (the number of viewpoints L is the same for all topics); and ϕ_{kl} be the multinomial probability distribution over terms associated with a topic k and a viewpoint l. The generative

process (see. the JTV graphical model in Figure 1) is the following:

- for each topic k and viewpoint l, draw a multinomial distribution over the vocabulary V: $\phi_{kl} \sim Dir(\beta)$;

- for each document d,

  draw a topic mixture $\theta_d \sim Dir(\alpha)$

  for each topic k, draw a viewpoint mixture $\psi_{dk} \sim Dir(\gamma)$

  for each term w_{dn}, sample a topic assignment $z_{dn} \sim Mult(\theta_d)$; sample a viewpoint assignment $v_{dn} \sim Mult(\psi_{dz_{dn}})$; and sample a term $w_{dn} \sim Mult(\phi_{z_{dn}v_{dn}})$.

We use fixed symmetric Dirichlet's parameters γ, β and α. They can be interpreted as the prior counts of: terms assigned to viewpoint l and topic k in a document; a particular term w assigned to topic k and viewpoint l within the corpus; terms assigned to a topic k in a document, respectively.

In order to learn the hidden JTV's parameters ϕ_{kl}, ψ_{dk} and θ_d, we draw on approximate inference as exact inference is intractable (Blei et al., 2003). We use the collapsed Gibbs Sampling (Griffiths and Steyvers, 2004), a Markov Chain Monte Carlo algorithm. The collapsed Gibbs sampler integrate out all parameters ϕ, ψ and θ in the joint distribution of the model and converge to a stationary posterior distribution over viewpoints' assignments $\vec{v}$ and all topics' assignments $\vec{z}$ in the corpus. It iterates on each current observed token w_i and samples each corresponding v_i and z_i given all the previous sampled assignments in the model $\vec{v}_{\neg i}$, $\vec{z}_{\neg i}$ and observed $\vec{w}_{\neg i}$, where $\vec{v} = \{v_i, \vec{v}_{\neg i}\}$, $\vec{z} = \{z_i, \vec{z}_{\neg i}\}$, and $\vec{w} = \{w_i, \vec{w}_{\neg i}\}$. The derived sampling equation is:

$$p(z_i = k, v_i = l | \vec{z}_{\neg i}, \vec{v}_{\neg i}, w_i = t, \vec{w}_{\neg i}) \propto$$

$$\frac{n_{kl,\neg i}^{(t)} + \beta}{\sum\limits_{t=1}^{V} n_{kl,\neg i}^{(t)} + V\beta} \cdot \frac{n_{dk,\neg i}^{(l)} + \gamma}{\sum\limits_{l=1}^{L} n_{dk,\neg i}^{(l)} + L\gamma} \cdot n_{d,\neg i}^{(k)} + \alpha \quad (1)$$

where $n_{kl,\neg i}^{(t)}$ is the number of times term t was assigned to topic k and the viewpoint l in the corpus; $n_{dk,\neg i}^{(l)}$ is the number of times viewpoint l of topic k was observed in document d; and $n_{d,\neg i}^{(k)}$ is the number of times topic k was observed in document d. All these counts are computed excluding the current token i, which is indicated by the symbol $\neg i$.

	AW		GM		ObCare	
View pt	allow	not	illegal	not	bad	not
#doc	213	136	44	54	129	54
tot.#toks	44482		10666		22733	
avg.#toks. doc.	127.45		108.83		124.22	

Table 3: Statistics on the three used data sets

After the convergence of the Gibbs algorithm, the parameters ϕ, ψ and θ are estimated using the last obtained sample.

5 Clustering Arguing Expressions

Although we are not tackling the task of clustering arguing expressions according to their viewpoints in this paper (Task 2 in Section 2), we explain how the structure of JTV lays the ground for performing it. We mentioned in the previous Section that an inferred topic-viewpoint distribution ϕ_{kl} can be assimilated to an arguing expression. For convenience, we will use "arguing expression" and "topic-viewpoint" interchangeably to refer to the topic-viewpoint distribution.

Indeed, two topic-viewpoint ϕ_{kl} and $\phi_{k'l}$, having different topics k and k', do not necessarily express the same viewpoint, despite the fact that they both have the same index l. The reason stems from the nested structure of the model, where the generation of the viewpoint assignments for a particular topic k is completely independent from that of topic k'. In other words, the model does not trace and match the viewpoint labeling along different topics. Nevertheless, the JTV can still help overcome this problem. According to the JTV's structure, a topic-viewpoint ϕ_{kl}, is more similar in distribution to a divergent topic-viewpoint $\phi_{kl'}$, related to the same topic k, than to any other topic-viewpoint $\phi_{k'*}$, corresponding to a different topic k'. Therefore, we can formulate the problem of clustering arguments as a constrained clustering problem (Basu et al., 2008). The goal is to group the similar topics-viewpoints ϕ_{kl}s into L clusters (number of viewpoints), given the constraint that the ϕ_{kl}s of the same topic k should not belong to the same cluster. The similarity between the topic-viewpoint distributions can be measured using the Jensen-Shannon Divergence (Bishop, 2006).

6 Experimental Set up

In order to evaluate the performances of the JTV model, we experiment with three different corpora of contentious documents. Recall, we assume that any input document to the JTV is answering a contentious question which makes it contentious according to the definitions stated in Section 2. Posts in online debate websites, like "createdebate.com" or "debate.org", match this requirement. They correspond to online users' takes on a clearly stated contention question making them more adequate for our matter than debate forums' posts. These latter contain online interactions between users where the objective is not necessarily answering a contention question but rather discussing a contentious topic. Classifying a document as contentious or not is not an issue considered in this paper but can be explored in our future work. Table 3 describes the used data sets.

Assault Weapons (AW) [4]: includes posts extracted from "debate.com". The contention question is "Should assault weapons be allowed in the United States as means of allowing individuals to defend themselves?". The viewpoints are either "should be allowed" or "should not be allowed".

Gay Marriage (GM) [5]: contains posts from "debate.com" related to the contention question "Should gay marriage be illegal?". The posts' stance are either "should be illegal" or "should be legal".

Obama Healthcare (ObCare) [6]: includes posts from "debate.org" responding to the contention question "Is the passing of ObamaCare bad for the American public?". Stances are either "bad" or "not bad".

Paul et al. (2010) stress out the importance of negation features in detecting contrastive viewpoints. Thus, we performed a simple treatment of merging any negation indicators, like "nothing", "no one", "never", etc., found in text with the following occurring word to form a single token. Moreover, we merge the negation "not" with any Auxiliary verb (e.g., is, was, could, will) preceding it. Then, we removed the stop-words.

[4]http://www.debate.org/opinions/should-assault-weapons-be-allowed-in-the-united-states-as-means-of-allowing-individuals-to-defend-themselves

[5]http://www.debate.org/opinions/should-gay-marriage-be-illegal

[6]http://www.debate.org/opinions/is-the-passing-of-obamacare-bad-for-the-american-public

Throughout the experiments below, the JTV's hyperparameters are set to fixed values. The γ is set, according to Steyvers and Griffiths's (Steyvers and Griffiths, 2007) hyperparameters settings, to $50/L$, where L is the number of viewpoints. β and α are adjusted manually, to give reasonable results, and are both set to 0.01. Along the experiments, we try different number of topics K. The number of viewpoints L is equal to 2. The TAM model (Paul et al., 2010) (Section 3.3) is run as a means of comparison during the evaluation procedure. Its default parameters are used.

7 Model Evaluation

7.1 Qualitative Evaluation

Tables 4 and 5 present the inferred topic-viewpoints words, i.e. arguing expressions, by JTV for the Obama Healthcare and Gay Marriage data sets, respectively. We set a number of topics of $K = 3$ for the former and $K = 2$ for the latter. The number of viewpoints is $L = 2$ for both data sets. For the Obamacare data set, we run the model with balanced number of posts from "bad" and "not bad" stances. Each topic-viewpoint pair (e.g. Topic 1-view 1) is represented by the set of top terms. The terms are sorted in descending order according to their probabilities. Inferred probabilities over topics, and over viewpoints for each topic, are also reported. We try to qualitatively observe the distinctiveness of each arguing (topic-viewpoint) and assess the coherence in terms of the topic discussed and the viewpoint conveyed and its divergence with the corresponding pair-element.

In both Tables 4 and 5, most of the topic-viewpoint pairs, corresponding to a same topic, are conveying opposite stances. For instance, taking a closer look to the original data suggests that Topic3-view5 (Table 4) criticizes the healthcare system and compares it to the other countries (e.g. a sample from the original documents:"*revise our **healthcare system** with the most efficient systems in the **world**"). On the other side, Topic 3-view 6 explains the negative consequence of obamacare on middle class, e.g. "**ObamaCare** was supposed to help the poor and the **middle class**. In the end, the businesses fire all the people because of the ObamaCare **taxes** and then IT IS THE **MIDDLE CLASS** PEOPLE WHO SUFFER!*". Similarly, Topic1-view1 advances the question of the costs that the bill will cause at the level of people

Topic 1	0.328	Topic 2	0.334	Topic 3	0.337
view 1 0.64	view 2 0.36	view 3 0.59	view 4 0.41	view 5 0.63	view 6 0.37
pay	universal	people	insurance	healthcare	obamacare
people	care	insurance	health	obamacare	healthcare
make	life	good	companies	system	government
money	law	free	medicare	americans	class
costs	act	health	doctors	affordable	taxes
government	poor	work	plan	country/world	middle

Table 4: JTV's generated topics-viewpoints (arguing expressions) from Obamacare data set

Topic 1	0.50	Topic 2	0.50
view 1 0.47	view 2 0.53	view 3 0.60	view 4 0.40
marriage	marriage	people	gay
love	man	gay	children
life	woman	religion	people
couples	god	shouldnt	sex
person	bible	wrong	parents
legal	illegal	rights	natural
married	wrong	government	human
happy	love	marry	population
samesex	homosexual	freedom	opposite
illegal	word	argument	race

Table 5: JTV's generated topics-viewpoints (arguing expressions) from Gay Marriage data set

and government, e.g. *"The **government** doesn't even have enough **money** to **pay** of a fraction of the towering debt that we've accrued"*, "forcing **people** to buy insurance or **pay** an even higher tax will **make** more families poverty stricken". However, Topic1-view2 stresses out the importance of having a universal healthcare, e.g. *"ObamaCare certainly has problems, but just like any **law**, we can work on these problems and make the **law** better (..). The fundamental goal is **Universal Health-care** (...)"*, *"If you were **poor** and had a hernia that needed surgery, you need money to pay for it. Denying Obama's Plan for a health **care** system means you cannot pay for it which means you will DIE."*. Similar pattern is observed in Topic 2.

The results on Gay Marriage 1 dataset (Table 5) encompass the notion of shared topic between divergent arguing expressions (Section 2) more clearly than the results obtained from Obamacare. This may be related to the nature of the contention. For instance, Topic 1 in Table 5 is "the concept of marriage" and it is shared by both view 1 and view 2. However, the concept is perceived differently according to the stance. The terms in view 1 (not illegal) suggest that **marriage** is about **love**, **hap-**

piness and it wouldn't disturb anyone's **life** (as it may be read from original data). The view 2 (illegal) may emphasize the notion of a **marriage** as a union between **man** and **woman** and the sacredness aspect of it (**god**, **bible**). Similarly, Topic 2 is about "**people** who are **gay**". The terms in view 3 (not illegal) may advocate that **religious arguments** from opposing stance do not make sense and that **gay people** are **free** and have the same **rights** as other people. Moreover, the **government** should not interfere in this matter. View 4 (illegal) suggests that **gay people** can not have **children** which raises the problem of **population** decrease. It also casts doubt on their ability to be **parents**.

7.2 Quantitative Evaluation

We assess the ability of the model to fit the online debate data and generate distinct topic-viewpoint pairs by comparing it with TAM which models also the topic-viewpoint dimension.

7.2.1 Held-Out Perplexity

We use the perplexity criterion to measure the ability of the learned topic model to fit a new held-out data. Perplexity assesses the generalization performance and, subsequently, provides a com-
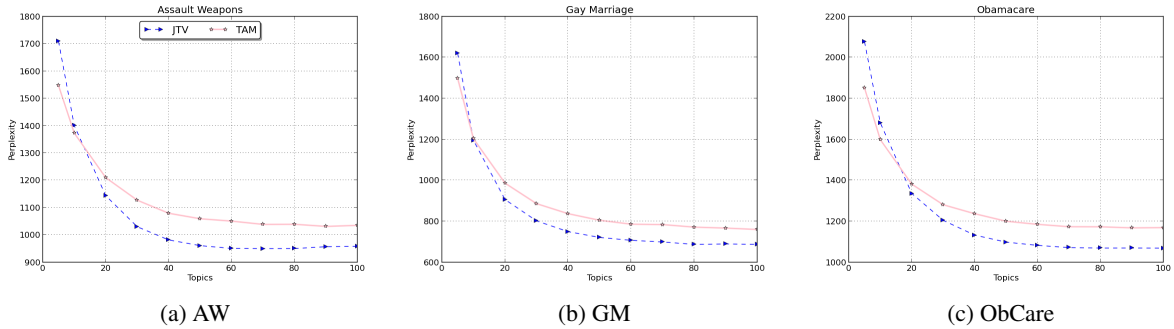
41

<div align="center">

(a) AW   (b) GM   (c) ObCare

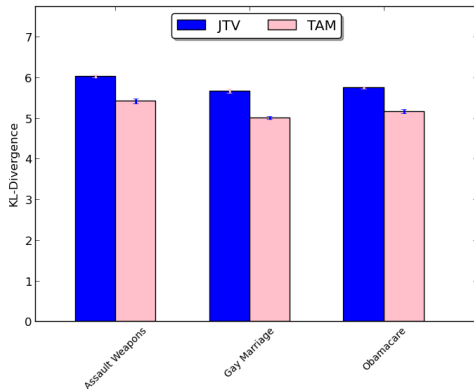Figure 2: JVT and TAM's perplexity plots for three different data sets

</div>

Figure 3: Average of overall topic-viewpoint divergences of JTV and TAM

paring framework of learned topic models. The lower the perplexity, the less "perplexed" is the model by unseen data and the better the generalization. It algebraically corresponds to the inverse geometrical mean of the test corpus' terms likelihoods given the learned model parameters (Heinrich, 2009). We compute the perplexity under estimated parameters of JTV and compare it to that of TAM for our three unigrams data sets (Section 6).

Figure 2 exhibits, for each corpus, the perplexity plot as function of the number of topics K for JTV and TAM. Note that for each K, we run the model 50 times. The drawn perplexity corresponds to the average perplexity on the 50 runs where each run compute one-fold perplexity from a 10-fold cross-validation. The figures show evidence that the JTV outperforms TAM for all data sets, used in the experimentation.

7.2.2 Kullback-Leibler Divergence

Kullback-Leibler (KL) Divergence is used to measure the degree of separation between two probability distributions. We utilize it to assess the dis-

tinctiveness of generated topic-viewpoint by JTV and TAM. This is an indicator of a good aggregation of arguing expressions. We compute an overall-divergence quantity, which is an average KL-Divergence between all pairs of topic-viewpoint distributions, for JTV and TAM and compare them. Figure 3 illustrates the results for all datasets. Quantities are averages on 20 runs of the models. Both models are run with a number of topics $K = 5$. Comparing JTV and TAM, we notice that the overall-divergence of JTV's topic-viewpoint is significantly ($p - value < 0.01$) higher for all data sets. This result reveals a better quality of our JTV extracting process of arguing expressions (the first task stated in Section 2)

8 Conclusion

We suggested a fine grained probabilistic framework for improving the quality of opinion mining from online contention texts. We proposed a Joint Topic Viewpoint model (JTV) for the unsupervised detection of arguing expressions. Unlike common approaches the proposed model focuses on arguing expressions that are implicitly described in unstructured text according to the latent topics they discuss and the implicit viewpoints they voice. The qualitative and quantitative analysis of the experimental results show the effectiveness of our (JTV) model in generating informative summaries of recurrent topics and viewpoints patterns in online debates' texts. Future study needs to give more insights into the clustering of arguing expressions according to their viewpoints, as well as their automatic extractive summary.

References

Sugato Basu, Ian Davidson, and Kiri Wagstaff. 2008. *Constrained Clustering: Advances in Algorithms,*

Theory, and Applications. Chapman & Hall/CRC, 1 edition.

Christopher M. Bishop. 2006. *Pattern Recognition and Machine Learning (Information Science and Statistics).* Springer-Verlag New York, Inc., Secaucus, NJ, USA.

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

Yi Fang, Luo Si, Naveen Somasundaram, and Zhengtao Yu. 2012. Mining contrastive opinions on political texts using cross-perspective topic model. In *Proceedings of the fifth ACM international conference on Web search and data mining*, WSDM '12, pages 63–72, New York, NY, USA. ACM.

Swapna Gottipati, Minghui Qiu, Yanchuan Sim, Jing Jiang, and Noah A. Smith. 2013. Learning topics and positions from debatepedia. In *Proceedings of Conference on Empirical Methods in Natural Language Processing*, EMNLP '13.

Thomas L. Griffiths and Mark Steyvers. 2004. Finding scientific topics. *Proceedings of the National Academy of Sciences of the United States of America*, 101(Suppl 1):5228–5235.

Gregor Heinrich. 2009. Parameter estimation for text analysis. Technical report, Fraunhofer IGD, September.

Yohan Jo and Alice H. Oh. 2011. Aspect and sentiment unification model for online review analysis. In *Proceedings of the fourth ACM international conference on Web search and data mining*, WSDM '11, pages 815–824, New York, NY, USA. ACM.

Jeffrey M. Jones. 2010. In u.s., 45% favor, 48% oppose obama healthcare plan. *Gallup*, March.

Soo-Min Kim and Eduard H Hovy. 2007. Crystal: Analyzing predictive opinions on the web. In *EMNLP-CoNLL*, pages 1056–1064.

Chenghua Lin and Yulan He. 2009. Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management*, CIKM '09, pages 375–384, New York, NY, USA. ACM.

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*, CoNLL-X '06, pages 109–116, Stroudsburg, PA, USA. Association for Computational Linguistics.

Arjun Mukherjee and Bing Liu. 2012. Mining contentions from discussions and debates. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*,

KDD '12, pages 841–849, New York, NY, USA. ACM.

Souneil Park, KyungSoon Lee, and Junehwa Song. 2011. Contrasting opposing views of news articles on contentious issues. In *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1*, HLT '11, pages 340–349, Stroudsburg, PA, USA. Association for Computational Linguistics.

Michael J. Paul, ChengXiang Zhai, and Roxana Girju. 2010. Summarizing contrastive viewpoints in opinionated text. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, EMNLP '10, pages 66–76, Stroudsburg, PA, USA. Association for Computational Linguistics.

Swapna Somasundaran and Janyce Wiebe. 2010. Recognizing stances in ideological on-line debates. In *Proceedings of the NAACL HLT 2010 Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, CAAGET '10, pages 116–124, Stroudsburg, PA, USA. Association for Computational Linguistics.

Mark Steyvers and Tom Griffiths. 2007. Probabilistic topic models. *Handbook of latent semantic analysis*, 427(7):424–440.

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 2006 Conference on Empirical Methods in Natural Language Processing*, EMNLP '06, pages 327–335, Stroudsburg, PA, USA. Association for Computational Linguistics.

Ivan Titov and Ryan McDonald. 2008. Modeling online reviews with multi-grain topic models. In *Proceedings of the 17th international conference on World Wide Web*, WWW '08, pages 111–120, New York, NY, USA. ACM.

Wayne Xin Zhao, Jing Jiang, Hongfei Yan, and Xiaoming Li. 2010. Jointly modeling aspects and opinions with a maxent-lda hybrid. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, EMNLP '10, pages 56–65, Stroudsburg, PA, USA. Association for Computational Linguistics.