Learning Domain-specific Sentiment Lexicon with Supervised Sentiment-aware LDA

Min Yang[†]   Dingju Zhu [1][‡,§]   Rashed Mustafa[§]   Kam-Pui Chow[†]

[†] Department of Computer Science, The University of Hong Kong, Hong Kong
{myang,chow}@cs.hku.hk
[‡] School of Computer Science, South China Normal University, Guangzhou, China
dingjuzhu@gmail.com
[§] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China
rashed.m@siat.ac.cn

Abstract. Analyzing and understanding people's sentiments towards different topics has become an interesting task due to the explosion of opinion-rich resources. In most sentiment analysis applications, sentiment lexicons play a crucial role, to be used as metadata of sentiment polarity. However, most previous works focus on discovering general-purpose sentiment lexicons. They cannot capture domain-specific sentiment words, or implicit and connotative sentiment words that are seemingly objective. In this paper, we propose a supervised sentiment-aware LDA model (ssLDA). The model uses a minimal set of domain-independent seed words and document labels to discover a domain-specific lexicon, learning a lexicon much richer and adaptive to the sentiment of specific document. Experiments on two publicly-available datasets (movie reviews and Obama-McCain debate dataset) show that our model is effective in constructing a comprehensive and high-quality domain-specific sentiment lexicon. Furthermore, the resulting lexicon significantly improves the performance of sentiment classification tasks.

1 Introduction

Sentiment analysis has attracted increasing attention recently due to the explosive growth of online opinion data such as online review sites and blogs. Generally speaking, existing sentiment analysis approaches can be divided into two categories by the knowledge they use: lexicon-based approaches and corpus-based approaches [18]. Among these sentiment analysis techniques, a comprehensive and high quality sentiment lexicon plays a crucial role. More specifically, for lexicon-based analysis, a sentiment lexicon that is generated manually or automatically [24, 25] assigns sentiment scores to sentiment indicators (e.g., adjectives, verbs, adverbs and nones) in the document. For corpus-based analysis, using sentiment lexicon helps to achieve a higher overall accuracy than using n-grams or POS features [3]. Actually, feature selection is achieved by dropping the words without sentiment polarity.

To date, many researches are dedicated to construct a general-purpose sentiment lexicon [20, 9, 24, 30, 12]. This general-purpose lexicon can be used in various situations. However, since a specific word can carry opposite sentiment meanings in different domains, a general-purpose sentiment lexicon is less accurate and less informative than a domain-specific lexicon, especially in domain-specific sentiment analysis [2]. For instance, "unpredictable" is negative in the electronics review domain while being positive in the movie review domain. Despite its usefulness, the construction domain-specific lexicon has received little attention in previous study of sentiment analysis. On the other hand, previous literature [10] shows that understanding the connotation of words plays is important in interpreting subtle sentiment of text. This is because that seemingly objective statements often allude nuanced sentiments. For example, the words such as "family" and "friendship" usually carry positive sentiment, but they are treated as neutral words in traditional sentiment lexicon. To our knowledge, there is little work focusing on constructing a rich domain-specific sentiment lexicon, that contains the words seemingly objective but that appears to convey connotative sentiment in specific domains.

To meet the challenge of learning a domain-specific lexicon, in this paper, we focus on the problem of constructing a sentiment lexicon that is not only domain specific but also be able to discover the sentiments beneath the surface meaning. Here, we propose a novel supervised sentiment-aware LDA model (ssLDA). The ssLDA model constructs a sentiment lexicon that captures the implicit and connotative sentiment of the words, even if they seem to be neural without the context. The model thus meets the needs of domain-specific sentiment analysis. In a standard Latent Dirichlet Allocation (LDA) model [5], topics of a document are drawn from a dirichlet distribution. In contrast, the proposed ssLDA model extends the standard LDA model, employing domain-independent seed words and document labels to guide the topic model construction. The topics consequently group semantically related words into a same sentiment category (i.e., positive, negative, or non-sentiment topics). While the standard LDA model is an unsupervised model and only the words in the documents are modeled, our model is a supervised model. We jointly model the documents and the labels, in order to find latent topics that will best predict the sentiment of the words in the documents. ssLDA accepts domain-independent sentiment seed words (e.g., "good" and "wrong") and document labels (positive and negative) as supervising inputs, discovering domain-specific sentiment words like "friendship" (positive) and "snake" (negative) through the model inference.

---

[1] Dingju Zhu is the corresponding author

To evaluate the effectiveness of our proposed model, we conduct experiments on data sets in two different domains: movie reviews and live tweets related to Obama-McCain debate. The experiment results demonstrate that our ssLDA model improves the quality and the coverage of state-of-the-art lexicon constructing techniques. It also significantly improves the performance of sentiment classification algorithms that rely on sentiment lexicons.

The rest of the paper is organized as follows. Section 2 reviews the related work. Section 3 describes the motivation and implementation of ssLDA model in details. In Section 4, we presents the description of dataset, followed by the discussion of the experimental results. Section 5 concludes the paper and indicates future works.

2 Related Work

Sentiment lexicons play a important role in opinion mining and sentiment analysis. The lexicons can be created by using seed words to expand the list of the words automatically [11, 27, 14] or manually [1]. The availability of the WordNet [16] database is an important starting point for many lexicon construction approaches. The synset model is sufficiently simple to provide an intrinsic correlation between a concept and the corresponding words. Moreover, WordNet covers the entire English lexicon and provides an extraordinarily large amount of conceptual distinctions. Turney and Littman [26] present a general strategy for inferring semantic orientation from semantic association. The semantic orientation of a given word is calculated from the strength of its association with the positive paradigm words, minus the strength of its association with the negative paradigm words. Velikovich et al. [29] propose a method that can be seen as a combination of methods for propagating sentiment across lexical graphs and methods for building sentiment lexicons based on distributional characteristics of phrases in raw data. To improve the scalability of lexicon-based sentiment analysis system, Naviarouskaya et al. [17] describe methods to automatically generate and score a new sentiment lexicon, and expand it through direct synonymy and antonymy relations, hyponymy relations, derivation, and compounding with known lexical units.

Most of the early studies focus on constructing general-purpose sentiment lexicons which can be used in any domain. However, the general-purpose sentiment lexicon are not always applicable since the sentiment orientation of some words are domain-dependent and context-aware. Recently, to enhance the domain specific sentiment data, a few researches have been done to generate the semantic orientation of words dependent on domains. Andreevskaia and Bergler [1] propose a sentiment annotation system that integrates a corpus-based classifier trained on a small set of annotated in-domain data and a lexicon-based system trained on a WordNet, to generate a cross-domain sentiment lexicon. Blitzer et al. [6] investigate domain adaption for sentiment classifiers, and extend the structural correspondence learning (SCL) domain adaptation algorithm [7] for use in sentiment classification. Choi and Cardie [8] propose a novel method based on integer linear programming that can adapt an existing lexicon into a new one to reflect the characteristics of the data more directly. In [15], Lu et al. focus on the problem of learning a sentiment lexicon that is not only domain specific but also dependent on aspect in context given an unlabeled opinionated text collection. The proposed framework provides a unified and principled way to combine different sources of information for learning a context-dependent sentiment lexicon.

In recent years, topic models have been used to perform extraction and grouping at the same time [5, 4, 21]. Based on standard LDA model, Xie and Li [31] propose a probabilistic modeling framework to construct a domain-specific sentiment lexicon for sentiment analysis task. The proposed model gives a flexible way to determine the real sentiment role of words over different labels, which gives a flexible way to determine the real sentiment role of words. Although our model is also inspired by LDA model, we don't assume every word in documents has a sentiment polarity, which is impractical in real world corpus. In our model, we assume that each document has both sentiment topics (corresponding to negative and positive emotions) and non-sentiment topics (corresponding to topics that are not associated with any emotion). For each document and for each word in the document, we first generate a sentiment or non-sentiment topic for the word, then sample the word from the associated word distribution. Another similar research to our work is proposed in [10], which describes induction algorithms encoding a diverse set of linguistic insights and prior knowledge drawn from lexical resources. Different from our work, [10] employs graph propagation algorithms that induce positive and negative polarities in isolation via separate graphs. This kind of approaches rely on the lexical relations within the text heavily. Our novel ssLDA model construct a domain-specific sentiment lexicon using a small number of seed words and document labels as prior knowledge, which is able to recognize the implicit and connotative sentiment of the words that are objective on surface.

3 Supervised Sentiment-aware LDA model

The original LDA model is an unsupervised generative model to learn hidden thematic of structures in a large corpus of documents. Because of its unsupervised nature, LDA can be used as a generic tool to analyze arbitrary documents, even without any metadata. But sometimes, we may also have some known relevant metadata about the documents to analyze, such as whether the document's sentiment polarity is positive or negative. And we may also have some prior knowledge about the documents to analyze, like we may know some of the words from the documents show a strong tendency of sentiment polarity, while some other words are not necessarily associated to the document's overall sentiment polarity. These metadata or prior knowledge can and should be accounted for the topic model we want to build to analyze the document.

The author-topic model [22] is an example for incorporating metadata (in this case, the authorship of documents) into the LDA model, where a document with multiple authors is modeled as a distribution over topics that is a mixture of the distributions associated with the authors. Another topic model that is capable to incorporate generic kind of metadata is the Supervised LDA (sLDA) model [4], where an unconstrained real value response variable is associated with each document, and is generated from a Gaussian distribution centered with the mean of all topic assignments.

For our application, to analysis the sentiment polarity of documents, we have the hypothesis that only a small subset of words conveys sentiment polarity. If we can find the subset of words in each document with expressed sentiment polarity, these words can form a domain-specific sentiment lexicon. On the other hand, we may have sentiment polarity labels for all or some of the documents, which we want to incorporate into the a sentiment-aware topic model.

Since only a small subset of words conveys sentiment polarity, a word in documents can either be a sentiment word or a non-sentiment word. It's natural to assume that there are only two classes of topics, sentiment topics and non-sentiment topics. In our ssLDA model, we take a simple approach, that only 2 sentiment topics have sentiment polarity and correspond to negative and positive emotions re-

spectively, while other non-sentiment topics are not associated with any emotions.

To incorporate the metadata of sentiment polarity labels, we assume the sentiment polarity label expressed a prior distribution of sentiment topics, but should not affect the prior distribution of non-sentiment topics.

In the following subsections, the supervised sentiment-aware LDA (ssLDA) model and its learning algorithm is rigorously defined. We begin with the detailed model description in Section 3.1. Then, in Section 3.2, we describe a Gibbs sampling algorithm to infer the model parameters. We finally show how to generate a sentiment lexicon based on the output of ssLDA model.

3.1 Model Description

Like the standard LDA model, the ssLDA model is a generative model. Each document has two classes of topics: 2 sentiment topics (corresponding to negative and positive emotions) and K non-sentiment topics (corresponding to topics that are not associated with any emotion). Each topic is represented by a multinomial distribution over words. To prevent conceptual confusion, we use a superscript "(s)" to indicate variables relating to sentiment topics, and use a superscript "(n)" to indicate variables of non-sentiment topics. In addition, we assume that the corpus vocabulary consists of V distinct words indexed by $\{1, \ldots, V\}$.

For each word in the document, we decide whether its topic is a sentiment topic or a non-sentiment topic by flipping a coin with head-tail probability $(p^{(s)}, p^{(n)})$, where $(p^{(s)}, p^{(n)}) \sim \text{Dir}(\alpha)$. Then we generate its sentiment or non-sentiment topic and sample the word from the associated topic-word distribution. The sentiment topic is generated from a biased Dirichlet prior $\text{Dir}(\alpha_y^{(s)})$ according to the document's label y. Intuitively, the prior parameter $\alpha_y^{(s)}$ should be selected to ensure the positive topic is more likely to be selected for positive labeled document, and vice versa.

For sentiment topics, the word distribution is drawn from a biased Dirichlet prior $\text{Dir}(\beta_k^{(s)})$. The vector $\beta_k^{(s)} \in \mathbb{R}^V$ is constructed from a seed dictionary, where $\beta_k^{(s)} := \gamma_0^{(s)}(1 - \omega_k) + \gamma_1^{(s)}\omega_k$ for $k \in \{0, 1\}$. $\omega_{0,w} = 1$ (or $\omega_{0,w} = 1$) if and only if word w is a negative (or positive) seed word. The scalars $\gamma_0^{(s)}$ and $\gamma_1^{(s)}$ are hyperparameters of the model. Intuitively, the biased prior ensures that the seed words are more probably drawn from the associated sentiment topic.

The generative process of word distributions for non-sentiment topics follows the standard LDA definition with a predefined scalar hyperparameter $\beta^{(n)}$.

To summarize the generative process of the ssLDA model:

1. for each sentiment topic $k \in \{0, 1\}$

   (a) generate $\phi_k^{(s)} \sim \text{Dir}(\beta_k^{(s)})$

2. for each non-sentiment topic $k \in \{1, \ldots, K\}$

   (a) generate $\phi_k^{(n)} \sim \text{Dir}(\beta^{(n)})$

3. for each document

   (a) generate $y \sim \text{Bernoulli}(\lambda)$

   (b) generate $\theta^{(s)} \sim \text{Dir}(\alpha_y^{(s)})$

   (c) generate $\theta^{(n)} \sim \text{Dir}(\alpha^{(n)})$

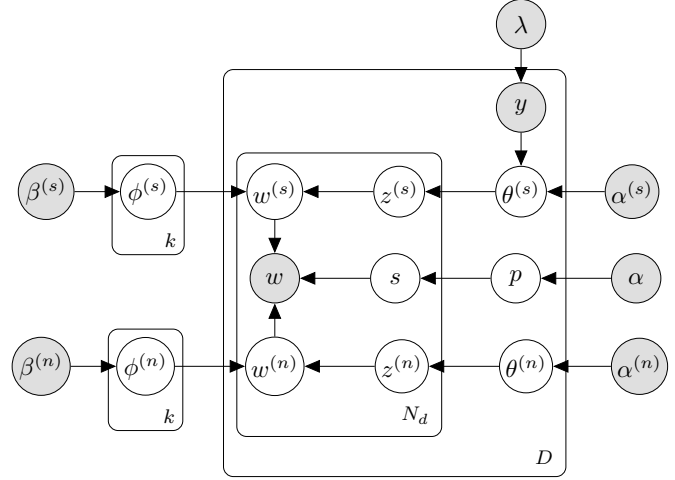   (d) generate $(p^{(s)}, p^{(n)}) \sim \text{Dir}(\alpha)$

Figure 1. The graphical model representation of the Supervised Sentiment-aware LDA model

(e) for each word in document

   i. generate topic class indicator $s \sim \text{Bernoulli}(p_s)$

   ii. if s is a sentiment topic

   A. generate $z^{(s)} \sim \text{Mult}(\theta^{(s)})$

   B. generate $w \sim \text{Mult}(\phi_{z^{(s)}}^{(s)})$, emit word w

   iii. otherwise

   A. generate $z^{(n)} \sim \text{Mult}(\theta^{(n)})$

   B. generate $w \sim \text{Mult}(\phi_{z^{(n)}}^{(n)})$, emit word w

As an alternative representation, the graphical model of the the generative process is shown by Figure 1.

3.2 Inference Algorithm

Given the set of hyperparameters $\{\alpha, \alpha^{(s)}, \alpha^{(n)}, \beta^{(s)}, \beta^{(s)}\}$ and the whole corpus of documents, our goal is to estimate the set of latent variables $\{p^{(s)}, p^{(n)}, \theta^{(s)}, \theta^{(n)}, \phi^{(s)}, \phi^{(n)}\}$ in the ssLDA model. We use a collapsed Gibbs sampling algorithm, which iteratively takes a word w from a document with label y and sample the topic z that this word belongs to.

Let $q_i^{(s)}$ and $q_j^{(n)}$ represent the probability that the current word belongs to sentiment topic $i^{(s)}$ or non-sentiment topic $j^{(n)}$. Then, with respect to the above notation, we have

$$q_i^{(s)} \propto p^{(s)}\theta_i^{(s)}\phi_{i,w}^{(s)} \qquad (1)$$

$$q_j^{(n)} \propto p^{(n)}\theta_j^{(n)}\phi_{j,w}^{(n)} \qquad (2)$$

In order to compute posterior topic distribution of the current word, it is sufficient to compute the posterior expectation of $q_i^{(s)}$ and $q_j^{(n)}$, thus reducing to the problem of computing the the posterior distribution of $p^{(s)}$, $\theta_i^{(s)}$ and $\phi_{i,w}^{(s)}$ (or $p^{(n)}$, $\theta_j^{(n)}$ and $\phi_{j,w}^{(n)}$).

To introduce some auxiliary notations, let the whole corpus excluding the current word be denoted by $\mathcal{D}$. Let $n_{i,w}^{(s)}$ (or $n_{j,w}^{(n)}$) indicate the number of occurrences of topic $i^{(s)}$ (or topic $j^{(n)}$ for non-sentiment topics) with word w in $\mathcal{D}$. Let $m_i^{(s)}$ (or $m_j^{(n)}$) indicate the number of occurrence of topic $i^{(s)}$ (or topic $j^{(n)}$ for non-sentiment topics) in the

current document. All these counts are defined excluding the current word. Using the definition of the ssLDA model and the Bayes Rule, we find that the joint density of these random variables are equal to

$$
\begin{aligned}
&\Pr\left(p^{(s)}, p^{(n)}, \theta^{(s)}, \phi^{(s)}, \theta^{(n)}, \phi^{(n)} | \mathcal{D}\right) \\
\propto\ & \Pr\left(p^{(s)}, p^{(n)}, \theta^{(s)}, \phi^{(s)}, \theta^{(n)}, \phi^{(n)}\right) \\
& \quad\cdot \Pr\left(\mathcal{D} | p^{(s)}, p^{(n)}, \theta^{(s)}, \phi^{(s)}, \theta^{(n)}, \phi^{(n)}\right) \\
\propto\ & \left(p^{(s)}\right)^{\alpha + \left(\sum_{i=0}^{1} m_i^{(s)}\right)} \cdot \left(p^{(n)}\right)^{\alpha + \left(\sum_{j=1}^{K} m_j^{(n)}\right)} \\
& \cdot \prod_{i=0}^{1}\left(\theta_i^{(s)}\right)^{\alpha_{y,i}^{(s)} + m_i^{(s)} - 1} \cdot \prod_{j=1}^{K}\left(\theta_j^{(n)}\right)^{\alpha^{(n)} + m_j^{(n)} - 1} \\
& \cdot \prod_{i=0}^{1}\prod_{w=1}^{V}\left(\phi_{i,w}^{(s)}\right)^{\beta_{i,w}^{(s)} + n_{i,w}^{(s)} - 1} \cdot \prod_{j=1}^{K}\prod_{w=1}^{V}\left(\phi_{j,w}^{(n)}\right)^{\beta^{(n)} + n_{j,w}^{(n)} - 1}
\end{aligned}
$$
(3)

According to equation (3), we see that $\{p^{(s)}, p^{(n)}\}$, $\{\theta_i^{(s)}, \theta_j^{(n)}\}$, $\{\phi_{i,w}^{(s)}\}$ and $\{\phi_{j,w}^{(n)}\}$ are mutually independent sets of random variables. Each of these random variables satisfies Dirichlet distribution with a specific set of parameters. By the mutual independence, we decompose the probability of the topic z for the current word as

$$
\Pr\left(z = i^{(s)} | D\right) = \mathbb{E}[q_i^{(s)}] \propto \mathbb{E}[p^{(s)}] \cdot \mathbb{E}[\theta_i^{(s)}] \cdot \mathbb{E}[\phi_{i,w}^{(s)}]
$$
(4)

$$
\Pr\left(z = j^{(n)} | D\right) = \mathbb{E}[q_j^{(n)}] \propto \mathbb{E}[p^{(n)}] \cdot \mathbb{E}[\theta_i^{(n)}] \cdot \mathbb{E}[\phi_{j,w}^{(n)}]
$$
(5)

Then, by examining the property of Dirichlet distribution, we can compute expectations on the right hand side of equation (4) and equation (5). More specifically, we have

$$
\mathbb{E}[p^{(s)}] = \frac{\alpha + \sum_{i=0}^{1} m_i^{(s)}}{2\alpha + \sum_{i=0}^{1} m_i^{(s)} + \sum_{j=1}^{K} m_j^{(n)}}
$$
(6)

$$
\mathbb{E}[p^{(n)}] = \frac{\alpha + \sum_{j=1}^{K} m_j^{(n)}}{2\alpha + \sum_{i=0}^{1} m_i^{(s)} + \sum_{j=1}^{K} m_j^{(n)}}
$$
(7)

$$
\mathbb{E}[\theta_i^{(s)}] = \frac{\alpha_y^{(s)} + m_i^{(s)}}{\sum_{y'=0}^{1} \alpha_{y'}^{(s)} + \sum_{i'=0}^{1} m_{i'}^{(s)}}
$$
(8)

$$
\mathbb{E}[\theta_j^{(n)}] = \frac{\alpha^{(n)} + m_j^{(n)}}{K\alpha^{(n)} + \sum_{j'=1}^{K} m_{j'}^{(n)}}
$$
(9)

$$
\mathbb{E}[\phi_{i,w}^{(s)}] = \frac{\beta_{i,w}^{(s)} + n_{i,w}^{(s)}}{\sum_{w'=1}^{V}\left(\beta_{i,w'}^{(s)} + n_{i,w'}^{(s)}\right)}
$$
(10)

$$
\mathbb{E}[\phi_{j,w}^{(n)}] = \frac{\beta_{j,w}^{(n)} + n_{j,w}^{(n)}}{V\beta^{(n)} + \sum_{w'=1}^{V} n_{j,w'}^{(n)}}
$$
(11)

Combining the above results with equations (4) and (5), we can sample the random topic z for each word iteratively. In the Gibbs sampling procedure, we only need to maintain the counters $n^{(s)}$, $n^{(n)}$, $m^{(s)}$ and $m^{(n)}$, which takes $O(1)$ time to update for each iteration.

3.3 Constructing Sentiment Lexicon

After we estimated the parameters of the ssLDA model, our final step is to generate the domain-specific sentiment lexicon from the esti-

mates. Specifically, we use the estimates $\phi_0^{(s)}$ and $\phi_1^{(s)}$ to construct the negative and positive sentiment lexicons.

The intuition is, if a word is a major component of the topic-word distribution for negative sentiment topic, the word should be part of the negative sentiment lexicon. In order to differentiate the words with no sentiments attached at all, we use the weight of words in non-sentiment topics as a baseline.

Following the above intuition, the lexicon construction is performed in the following way: for each word w in the vocabulary, we compare the three values $\phi_{0,w}^{(s)}$, $\phi_{1,w}^{(s)}$ and $\frac{1}{K}\sum_{i=1}^{K} \phi_{i,w}^{(n)}$. If $\phi_{0,w}^{(s)}$ is the largest, then the word w is added to the negative sentiment lexicon. If $\phi_{1,w}^{(s)}$ is the largest, then the word w is added to the positive sentiment lexicon. Otherwise, $\frac{1}{K}\sum_{i=1}^{K} \phi_{i,w}^{(n)}$ is the largest among the three, which suggests that the word w is more probably drawn from a non-sentiment topic. Thus, the word is considered neutral and not included in the sentiment lexicon.

4 Experiments

In this section, we report empirical evaluations of our proposed model. Since there is no metric explicitly measuring the quality of a sentiment lexicon, we demonstrate the performance of our algorithm in two ways: (1) we perform a case-study for the lexicon generated by our algorithm, and (2) we solve sentiment classification tasks using different sentiment lexicons, and demonstrate the advantage of our lexicon over other popular and standard lexicons.

4.1 Datasets

We conduct experiments on two publicly-available datasets: MR and OMD data sets.

Movie Review (MR) The movie reviews were selected when the author rating was stars or a numerical score. The ratings are automatically extracted and converted into three categories: positive, negative and neutral. In this paper, we use the Movie Review dataset containing 1000 positive examples and 1000 negative examples. All reviews were written before 2002, as used in the past work [19].

Obama-McCain Debate (OMD) The original Obama-McCain Debate (OMD) dataset was constructed from 3,269 tweets crawled during the first U.S. presidential TV debate in September 2008 [23]. Sentiment labels were acquired for these tweets using Amazon Mechanical Turk, where each tweet was rated by at least three annotators as either positive, negative, mixed, or other. Following the strategy used of past work [13], we use majority votes to label the sentiment of tweets. This results in a dataset of 1,906 tweets, in which 710 tweets have positive label and 1,196 tweets have negative label.

In experiments, data preprocessing was performed on both data sets. First, the texts are tokenized with a natural language toolkit NLTK[2]. Then, we remove non-alphabet characters, numbers, pronoun, punctuation and stop words from the texts. Finally, WordNet stemmer[3] is applied so as to reduce the vocabulary size and settle the issue of data spareness.

4.2 Evaluation of Lexicon Quality

We first settle down the implementation details for the ssLDA model, specifying the hyperparameters that we choose for the experiment.

---

[2] http://www.nltk.org
[3] http://wordnet.princeton.edu

For the MR dataset, we choose topic number $K = 20$, and hyperparameters $\alpha = 0.5$, $\alpha^{(s)} = \alpha^{(n)} = 0.45$, $\beta^{(n)} = 0.5$. The vector $\beta^{(s)}$ is constructed from the seed dictionary using $\gamma = (0.25, 0.75)$. For OMD dataset, we use cross-validation to set the topic number $K = 7$, while other hyperparameters remain the same. We also observe that the ssLDA model is not sensitive to the hyperparameters.

As mentioned, we use a few domain-independent seed words as prior information for our model. To be specific, the seed words list containing seven positive words (i.e., excellent, good, nice, positive, fortunate, correct, and superior) and seven negative words (i.e., nasty, bad, poor, negative, unfortunate, wrong, and inferior). The fourteen words are chosen for their lack of sensitivity to text, and are the same as the seed words list used in [28].

The top 20 positive and negative words generated from the MR and the OMD dataset are reported in Table 1. From Table 1, we observe that the generated words are informative and coherent. We use bold font to indicate the sentiment words that are not in existing popular sentiment lexicons but included in our generated lexicon. For example, the words "Obama", "job", "cut" are seemingly neutral by its surface meaning, actually expressing positive sentiment for the Obama-McCain Debate. At the same time, the words "McCain", "Iraq", "Republican" express negative sentiment. These domain-specific words are not included in any other sentiment lexicons. It is worth noting that users label the sentiment largely based on the final presidential election result, and to some extent subjective. On the other hand, for Movie Review dataset, the sentiments are labeled in align with the popularity of movies and actors. We observe similar phenomenons for the MR dataset as we observe in the OMD dataset. The experimental results show that our algorithm can successfully construct a domain-specific sentiment lexicon for this corpus that is able to understand the connotation of the words that may not be obvious without the context.

MR dataset		OMD dataset	
Positive	Negative	Positive	Negative
good	bad	**obama**	**mccain**
nice	wrong	good	**john**
excellent	**gang**	win	**doe**
american	**spice**	**job**	**iraq**
tale	nasty	**tax**	**bush**
superior	**fight**	**back**	hate
beauty	unfortunate	great	attack
bridge	awful	**cut**	**policy**
princess	**damme**	won	wrong
effective	**wrestling**	**candidate**	**republican**
family	worst	nice	**iran**
artist	mess	**cnn**	long
disney	**showgirl**	**lehrer**	**house**
british	**wcw**	agree	**pow**
dude	**crystal**	better	**personal**
positive	**seagal**	**vote**	exactly
prof	terrible	wow	**drilling**
brook	unfunny	**energy**	**senator**
friendship	snake	**moderator**	bad
freedom	waste	**presidential**	**stop**
color	negative	**already**	defense

Table 1. Top 20 sentiment words generated from MR and OMD dataset

4.3 Document-level sentiment classification

To further quantitatively evaluate the lexicon quality, we perform binary sentiment classification task with four wildly used sentiment lex-

icons: General Inquirer (GI) [24], HowNet [9], MPQA [30] and Liu's Lexicon [12]. We also compare our results with three state-of-the-art sentiment lexicon construction models, i.e., OPT [15], TSTM [31] and C-LP [10]. The classification problem is solved on both the MR dataset and the OMD dataset. Both lexicon-based and corpus-based sentiment classification are tested. For lexicon-based classification, we count the number of positive words and negative words in the review, and predict the sentiment of the review by comparing two word counts. For corpus-based sentiment classification, an ℓ_1-regularized logistic regression is employed to train a classifier using the sense-based features extracted from the sentiment lexicon. We use the LIBLinear package[4] to implement the training. In all experiments, the training parameters are set to the default parameter values of LIBLinear, and the results are evaluated using classification accuracy.

We summarize the results on both data sets in Table 2 and Table 3, and highlight in bold font the best performance under each measure. As an easy observation, the lexicon generated by the ssLDA model consistently and significantly outperforms other lexicons and models, on both datasets. Particularly, the performance gain of ssLDA is even more significant on OMD dataset. It has improved the accuracy from 0.701 to 0.716 in corpus-based setting, and from 0.692 to 0.722 in lexicon-based setting of sentiment analysis classification tasks, respectively. The advantages of our model comes from its capability of exploring domain-specific sentiments and the sentiments beneath the surface meaning.

Lexicon	Data	
	MR dataset	OMD dataset
GI	0.537	0.558
Liu's Lexicon	0.702	0.688
MPQA	0.651	0.617
HowNet	0.527	0.585
OPT	0.685	0.693
TSTM	0.701	0.655
C-LP	0.708	0.695
ssLDA	**0.721**	**0.728**

Table 2. Sentiment Classification Accuracy (Lexicon-based)

Lexicon	Data	
	MR dataset	OMD dataset
GI	0.744	0.673
Liu's Lexicon	0.816	0.682
MPQA	0.808	0.686
HowNet	0.780	0.663
OPT	0.819	0.752
TSTM	0.769	0.742
C-LP	0.813	0.771
ssLDA	**0.835**	**0.784**

Table 3. Sentiment Classification Accuracy (Corpus-based)

5 Conclusions and Future Work

In this paper, we have presented a supervised sentiment-aware LDA model for domain-specific sentiment lexicon discovery. The proposed ssLDA model essentially extended a standard LDA model, employing domain-independent seed words and document labels to guide the model generating topics. The topics consequently group

---

[4] http://www.csie.ntu.edu.tw/~cjlin/liblinear/

semantically related words into a same sentiment category (i.e., positive, negative, or neutral). Thus, it makes the sentiment lexicon containing much richer and adaptive domain-specific words. We have demonstrated that our method is capable of learning new sentiment words from domain knowledge and discovering subtle sentimental meaning of the words. Further quantitative evaluation against widely used lexicons and state-of-the-art methods shows that sentiment lexicon generated by our algorithm is of high quality, and can assist both lexicon-based and corpus-based sentiment classification tasks.

For feature works, we hope to extend the proposed sentiment-aware LDA model by exploiting discourse structure knowledge, which has been shown significant in identifying the polarity of content-aware words. We also hope to leverage document sentiment label information to make it a semi-supervised or supervised model, to enhance the domain adaption.

In addition, we have plans to combine the lexicon learned in ssLDA model with corpus-based classification algorithms in co-training setting, to achieve higher sentiment analysis accuracy. We also have plans to evaluate the performance of our ssLDA model in other languages (e.g., Chinese and Japanese) as well as in other sentiment related applications, such as opinion summarization and question answering.

References

[1] Alina Andreevskaia and Sabine Bergler, 'When specialists and generalists work together: Overcoming domain dependence in sentiment tagging.', in *ACL*, pp. 290–298, (2008).

[2] Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani, 'Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining.', in *LREC*, volume 10, pp. 2200–2204, (2010).

[3] AR Balamurali, Aditya Joshi, and Pushpak Bhattacharyya, 'Cross-lingual sentiment analysis for indian languages using linked wordnets.', in *COLING (Posters)*, pp. 73–82, (2012).

[4] David M Blei and Jon D McAuliffe, 'Supervised topic models.', in *NIPS*, volume 7, pp. 121–128, (2007).

[5] David M Blei, Andrew Y Ng, and Michael I Jordan, 'Latent dirichlet allocation', *the Journal of machine Learning research*, **3**, 993–1022, (2003).

[6] John Blitzer, Mark Dredze, and Fernando Pereira, 'Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification', in *ACL*, volume 7, pp. 440–447, (2007).

[7] John Blitzer, Ryan McDonald, and Fernando Pereira, 'Domain adaptation with structural correspondence learning', in *Proceedings of the 2006 conference on empirical methods in natural language processing*, pp. 120–128. Association for Computational Linguistics, (2006).

[8] Yejin Choi and Claire Cardie, 'Adapting a polarity lexicon using integer linear programming for domain-specific sentiment classification', in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 2-Volume 2*, pp. 590–598. Association for Computational Linguistics, (2009).

[9] Zhendong Dong and Qiang Dong, *HowNet and the Computation of Meaning*, World Scientific, 2006.

[10] Song Feng, Jun Sak Kang, Polina Kuznetsova, and Yejin Choi, 'Connotation lexicon: A dash of sentiment beneath the surface meaning', in *Proceedings of the 51th Annual Meeting of the Association for Computational Linguistics*, volume 2, (2013).

[11] Vasileios Hatzivassiloglou and Kathleen R McKeown, 'Predicting the semantic orientation of adjectives', in *Proceedings of the 35th Annual Meeting of the Association for Computational Linguistics and Eighth Conference of the European Chapter of the Association for Computational Linguistics*, pp. 174–181. Association for Computational Linguistics, (1997).

[12] Minqing Hu and Bing Liu, 'Mining and summarizing customer reviews', in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 168–177. ACM, (2004).

[13] Xia Hu, Lei Tang, Jiliang Tang, and Huan Liu, 'Exploiting social relations for sentiment analysis in microblogging', in *Proceedings of the sixth ACM international conference on Web search and data mining*, pp. 537–546. ACM, (2013).

[14] Alistair Kennedy and Diana Inkpen, 'Sentiment classification of movie reviews using contextual valence shifters', *Computational Intelligence*, **22**(2), 110–125, (2006).

[15] Yue Lu, Malu Castellanos, Umeshwar Dayal, and ChengXiang Zhai, 'Automatic construction of a context-aware sentiment lexicon: an optimization approach', in *Proceedings of the 20th international conference on World wide web*, pp. 347–356. ACM, (2011).

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

[17] Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka, 'Sentiful: Generating a reliable lexicon for sentiment analysis', in *Affective Computing and Intelligent Interaction and Workshops, 2009. ACII 2009. 3rd International Conference on*, pp. 1–6. IEEE, (2009).

[18] Bo Pang and Lillian Lee, 'Opinion mining and sentiment analysis', *Foundations and trends in information retrieval*, **2**(1-2), 1–135, (2008).

[19] Bo Pang, Lillian Lee, and Shivakumar Vaithyanathan, 'Thumbs up?: sentiment classification using machine learning techniques', in *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pp. 79–86. Association for Computational Linguistics, (2002).

[20] James W Pennebaker, Martha E Francis, and Roger J Booth, 'Linguistic inquiry and word count: Liwc 2001', *Mahway: Lawrence Erlbaum Associates*, **71**, 2001, (2001).

[21] Daniel Ramage, David Hall, Ramesh Nallapati, and Christopher D Manning, 'Labeled lda: A supervised topic model for credit attribution in multi-labeled corpora', in *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing: Volume 1-Volume 1*, pp. 248–256. Association for Computational Linguistics, (2009).

[22] Michal Rosen-Zvi, Thomas Griffiths, Mark Steyvers, and Padhraic Smyth, 'The author-topic model for authors and documents', in *Proceedings of the 20th conference on Uncertainty in artificial intelligence*, pp. 487–494. AUAI Press, (2004).

[23] David A Shamma, Lyndon Kennedy, and Elizabeth F Churchill, 'Tweet the debates: understanding community annotation of uncollected sources', in *Proceedings of the first SIGMM workshop on Social media*, pp. 3–10. ACM, (2009).

[24] Philip J Stone, Dexter C Dunphy, and Marshall S Smith, 'The general inquirer: A computer approach to content analysis.', (1966).

[25] Richard M Tong, 'An operational system for detecting and tracking opinions in on-line discussion', in *Working Notes of the ACM SIGIR 2001 Workshop on Operational Text Classification*, volume 1, p. 6, (2001).

[26] Peter Turney and Michael L Littman, 'Unsupervised learning of semantic orientation from a hundred-billion-word corpus', (2002).

[27] Peter D Turney, 'Thumbs up or thumbs down?: semantic orientation applied to unsupervised classification of reviews', in *Proceedings of the 40th annual meeting on association for computational linguistics*, pp. 417–424. Association for Computational Linguistics, (2002).

[28] Peter D Turney and Michael L Littman, 'Measuring praise and criticism: Inference of semantic orientation from association', *ACM Transactions on Information Systems (TOIS)*, **21**(4), 315–346, (2003).

[29] Leonid Velikovich, Sasha Blair-Goldensohn, Kerry Hannan, and Ryan McDonald, 'The viability of web-derived polarity lexicons', in *Human Language Technologies: The 2010 Annual Conference of the North American Chapter of the Association for Computational Linguistics*, pp. 777–785. Association for Computational Linguistics, (2010).

[30] Theresa Wilson, Janyce Wiebe, and Paul Hoffmann, 'Recognizing contextual polarity in phrase-level sentiment analysis', in *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pp. 347–354. Association for Computational Linguistics, (2005).

[31] Rui Xie and Chunping Li, 'Lexicon construction: A topic model approach', in *Systems and Informatics (ICSAI), 2012 International Conference on*, pp. 2299–2303. IEEE, (2012).