

A Topic Model for Building Fine-grained Domain-specific Emotion Lexicon

Min Yang[‡] Baolin Peng[§] Zheng Chen[§] Dingju Zhu^{*†,¶} Kam-Pui Chow[‡]

[†]School of Computer Science, South China Normal University, Guangzhou, China
dingjuzhu@gmail.com

[‡]Department of Computer Science, The University of Hong Kong, Hong Kong
{myang, chow}@cs.hku.hk

[§]Department of Computer Science, Beihang University, Beijing, China
b.peng@cse.buaa.edu.cn, tzchen86@gmail.com

[¶]Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, China

Abstract

Emotion lexicons play a crucial role in sentiment analysis and opinion mining. In this paper, we propose a novel Emotion-aware LDA (EaLDA) model to build a domain-specific lexicon for predefined emotions that include anger, disgust, fear, joy, sadness, surprise. The model uses a minimal set of domain-independent seed words as prior knowledge to discover a domain-specific lexicon, learning a fine-grained emotion lexicon much richer and adaptive to a specific domain. By comprehensive experiments, we show that our model can generate a high-quality fine-grained domain-specific emotion lexicon.

1 Introduction

Due to the popularity of opinion-rich resources (e.g., online review sites, forums, blogs and the microblogging websites), automatic extraction of opinions, emotions and sentiments in text is of great significance to obtain useful information for social and security studies. Various opinion mining applications have been proposed by different researchers, such as question answering, opinion mining, sentiment summarization, etc. As the fine-grained annotated data are expensive to get, the unsupervised approaches are preferred and more used in reality. Usually, a high quality emotion lexicon play a significant role when apply the unsupervised approaches for fine-grained emotion classification.

Thus far, most lexicon construction approaches focus on constructing general-purpose emotion lexicons (Stone et al., 1966; Hu and Liu, 2004; Wilson et al., 2005; Dong and Dong, 2006). However, since a specific word can carry various emotions in different domains, a general-purpose emotion lexicon is less accurate and less informative than a domain-specific lexicon (Baccianella et al., 2010). In addition, in previous work, most of the lexicons label the words on coarse-grained dimensions (positive, negative and neutrality). Such lexicons cannot accurately reflect the complexity of human emotions and sentiments. Lastly, previous emotion lexicons are mostly annotated based on many manually constructed resources (e.g., emotion lexicon, parsers, etc.). This limits the applicability of these methods to a broader range of tasks and languages.

To meet the challenges mentioned above, we propose a novel EaLDA model to construct a domain-specific emotion lexicon consisting of six primary emotions (i.e., anger, disgust, fear, joy, sadness and surprise). The proposed EaLDA model extends the standard Latent Dirichlet Allocation (LDA) (Blei et al., 2003) model by employing a small set of seeds to guide the model generating topics. Hence, the topics consequently group semantically related words into a same emotion category. The lexicon is thus able to best meet the user's specific needs. Our approach is a weakly supervised approach since only some seeds emotion sentiment words are needed to launch the process of lexicon construction. In practical applications, asking users to provide some seeds is easy as they usually have a good knowledge what are important in their domains.

*Dingju Zhu is the corresponding author

Extensive experiments are carried out to evaluate our model both qualitatively and quantitatively using benchmark dataset. The results demonstrate that our EaLDA model improves the quality and the coverage of state-of-the-art fine-grained lexicon.

2 Related Work

Emotion lexicon plays an important role in opinion mining and sentiment analysis. In order to build such a lexicon, many researchers have investigated various kinds of approaches. However, these methods could roughly be classified into two categories in terms of the used information. The first kind of approaches is based on thesaurus that utilizes synonyms or glosses to determine the sentiment orientation of a word. The availability of the WordNet (Miller, 1995) database is an important starting point for many thesaurus-based approaches (Kamps et al., 2004; Hu and Liu, 2004; Esuli and Sebastiani, 2006). The second kind of approaches is based on an idea that emotion words co-occurring with each others are likely to convey the same polarity. There are numerous studies in this field (Turney and Littman, 2003; Wiebe and Riloff, 2005; Esuli and Sebastiani, 2006; Barbosa and Feng, 2010).

Most of the previous studies for emotion lexicon construction are limited to positive and negative emotions. Recently, to enhance the increasingly emotional data, a few researches have been done to identify the fine-grained emotion of words (Strapparava and Mihalcea, 2007; Gill et al., 2008; Rao et al., 2012). For example, Gill et al. (2008) utilize computational linguistic tools to identify the emotions of the words (such as, joy, sadness, acceptance, disgust, fear, anger, surprise and anticipation). While, this approach is mainly for public use in general domains. Rao et al. (2012) propose an method of automatically building the word-emotion mapping dictionary for social emotion detection. However, the emotion lexicon is not outputted explicitly in this paper, and the approach is fully unsupervised which may be difficult to be adjusted to fit the personalized data set.

Our approach relates most closely to the method proposed by Xie and Li (2012) for the construction of lexicon annotated for polarity based on LDA model. Our approach differs from (Xie and Li, 2012) in two important ways: first, we do not address the task of polarity lexicon construction, but

instead we focus on building fine-grained emotion lexicon. Second, we don't assume that every word in documents is subjective, which is impractical in real world corpus.

3 Algorithm

In this section, we rigorously define the emotion-aware LDA model and its learning algorithm. We describe with the model description, a Gibbs sampling algorithm to infer the model parameters, and finally how to generate a emotion lexicon based on the model output.

3.1 Model Description

Like the standard LDA model, EaLDA is a generative model. To prevent conceptual confusion, we use a superscript “(e)” to indicate variables related to emotion topics, and use a superscript “(n)” to indicate variables of non-emotion topics. We assume that each document has two classes of topics: M emotion topics (corresponding to M different emotions) and K non-emotion topics (corresponding to topics that are not associated with any emotion). Each topic is represented by a multinomial distribution over words. In addition, we assume that the corpus vocabulary consists of V distinct words indexed by $\{1, \dots, V\}$.

For emotion topics, the EaLDA model draws the word distribution from a biased Dirichlet prior $\text{Dir}(\beta_k^{(e)})$. The vector $\beta_k^{(e)} \in \mathbb{R}^V$ is constructed with $\beta_k^{(e)} := \gamma_0^{(e)}(1^V - \Omega_k) + \gamma_1^{(e)}\Omega_k$, for $k \in \{1, \dots, M\}$. $\Omega_{k,w} = 1$ if and only if word w is a seed word for emotion k , otherwise $\Omega_{k,w} = 0$. The scalars $\gamma_0^{(e)}$ and $\gamma_1^{(e)}$ are hyperparameters of the model. Intuitively, when $\gamma_1^{(e)} > \gamma_0^{(e)}$, the biased prior ensures that the seed words are more probably drawn from the associated emotion topic.

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

For each word in the document, we decide whether its topic is an emotion topic or a non-emotion topic by flipping a coin with head-tail probability $(p^{(e)}, p^{(n)})$, where $(p^{(e)}, p^{(n)}) \sim \text{Dir}(\alpha)$. The emotion (or non-emotion) topic is sampled according to a multinomial distribution $\text{Mult}(\theta^{(e)})$ (or $\text{Mult}(\theta^{(n)})$). Here, both $\theta^{(e)}$ and $\theta^{(n)}$ are document-level latent variables. They are generated from Dirichlet priors $\text{Dir}(\alpha^{(e)})$ and $\text{Dir}(\alpha^{(n)})$ with $\alpha^{(s)}$ and $\alpha^{(n)}$ being hyperparameters.

We summarize the generative process of the EaLDA model as below:

1. for each emotion topic $k \in \{1, \dots, M\}$, draw $\phi_k^{(e)} \sim \text{Dir}(\beta_k^{(e)})$
2. for each non-emotion topic $k \in \{1, \dots, K\}$, draw $\phi_k^{(n)} \sim \text{Dir}(\beta_k^{(n)})$
3. for each document
 - (a) draw $\theta^{(e)} \sim \text{Dir}(\alpha^{(e)})$
 - (b) draw $\theta^{(n)} \sim \text{Dir}(\alpha^{(n)})$
 - (c) draw $(p^{(e)}, p^{(n)}) \sim \text{Dir}(\alpha)$
 - (d) for each word in document
 - i. draw topic class indicator $s \sim \text{Bernoulli}(p_s)$
 - ii. if $s = \text{“emotion topic”}$
 - A. draw $z^{(e)} \sim \text{Mult}(\theta^{(e)})$
 - B. draw $w \sim \text{Mult}(\phi_{z^{(e)}}^{(e)})$, emit word w
 - iii. otherwise
 - A. draw $z^{(n)} \sim \text{Mult}(\theta^{(n)})$
 - B. draw $w \sim \text{Mult}(\phi_{z^{(n)}}^{(n)})$, emit word w

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

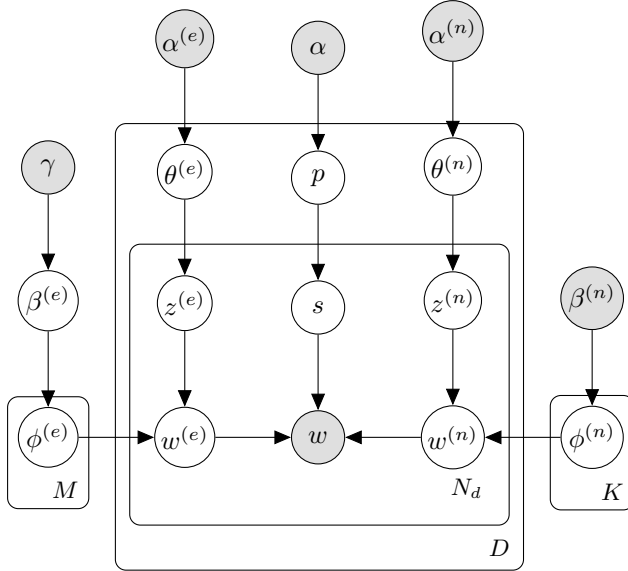


Figure 1: The Emotion-aware LDA model.

3.2 Inference Algorithm

Assuming hyperparameters $\alpha, \alpha^{(e)}, \alpha^{(n)}$, and $\beta^{(e)}, \beta^{(n)}$, we develop a collapsed Gibbs sampling algorithm to estimate the latent variables in the EaLDA model. The algorithm iteratively takes a word w

from a document and sample the topic that this word belongs to.

Let the whole corpus excluding the current word be denoted by D . Let $n_{i,w}^{(e)}$ (or $n_{j,w}^{(n)}$) indicate the number of occurrences of topic $i^{(e)}$ (or topic $j^{(n)}$) with word w in the whole corpus. Let $m_i^{(e)}$ (or $m_j^{(n)}$) indicate the number of occurrence of topic $i^{(e)}$ (or topic $j^{(n)}$) in the current document. All these counts are defined excluding the current word. Using the definition of the EaLDA model and the Bayes Rule, we find that the joint density of these random variables are equal to

$$\begin{aligned}
& \Pr(p^{(e)}, p^{(n)}, \theta^{(e)}, \phi^{(e)}, \theta^{(n)}, \phi^{(n)} | D) \\
& \propto \Pr(p^{(e)}, p^{(n)}, \theta^{(e)}, \phi^{(e)}, \theta^{(n)}, \phi^{(n)}) \\
& \quad \times \Pr(D | p^{(e)}, p^{(n)}, \theta^{(e)}, \phi^{(e)}, \theta^{(n)}, \phi^{(n)}) \\
& \propto (p^{(e)})^{\alpha + (\sum_{i=1}^M m_i^{(e)})} \cdot (p^{(n)})^{\alpha + (\sum_{j=1}^K m_j^{(n)})} \\
& \quad \cdot \prod_{i=1}^M (\theta_i^{(e)})^{\alpha^{(e)} + m_i^{(e)} - 1} \cdot \prod_{j=1}^K (\theta_j^{(n)})^{\alpha^{(n)} + m_j^{(n)} - 1} \\
& \quad \cdot \prod_{i=0}^1 \prod_{w=1}^V (\phi_{i,w}^{(e)})^{\beta_{i,w}^{(e)} + n_{i,w}^{(e)} - 1} \\
& \quad \cdot \prod_{j=1}^K \prod_{w=1}^V (\phi_{j,w}^{(n)})^{\beta_{j,w}^{(n)} + n_{j,w}^{(n)} - 1}
\end{aligned} \tag{1}$$

According to equation (1), we see that $\{p^{(e)}, p^{(n)}\}$, $\{\theta_i^{(e)}, \theta_j^{(n)}\}$, $\{\phi_{i,w}^{(e)}\}$ 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

$$\Pr(z = i^{(e)} | D) \propto \mathbb{E}[p^{(e)}] \cdot \mathbb{E}[\theta_i^{(e)}] \cdot \mathbb{E}[\phi_{i,w}^{(e)}] \tag{2}$$

$$\Pr(z = j^{(n)} | D) \propto \mathbb{E}[p^{(n)}] \cdot \mathbb{E}[\theta_j^{(n)}] \cdot \mathbb{E}[\phi_{j,w}^{(n)}] \tag{3}$$

Then, by examining the property of Dirichlet distribution, we can compute expectations on the right hand side of equation (2) and equation (3) by

$$\mathbb{E}[p^{(e)}] = \frac{\alpha + \sum_{i=0}^1 m_i^{(e)}}{2\alpha + \sum_{i=1}^M m_i^{(e)} + \sum_{j=1}^K m_j^{(n)}} \tag{4}$$

$$\mathbb{E}[p^{(n)}] = \frac{\alpha + \sum_{j=1}^K m_j^{(n)}}{2\alpha + \sum_{i=1}^M m_i^{(e)} + \sum_{j=1}^K m_j^{(n)}} \tag{5}$$

$$\mathbb{E}[\theta_i^{(e)}] = \frac{\alpha^{(e)} + m_i^{(e)}}{M\alpha^{(e)} + \sum_{i'=1}^M m_{i'}^{(e)}} \quad (6)$$

$$\mathbb{E}[\theta_j^{(n)}] = \frac{\alpha^{(n)} + m_j^{(n)}}{K\alpha^{(n)} + \sum_{j'=1}^K m_{j'}^{(n)}} \quad (7)$$

$$\mathbb{E}[\phi_{i,w}^{(e)}] = \frac{\beta_{i,w}^{(e)} + n_{i,w}^{(e)}}{\sum_{w'=1}^V (\beta_{i,w'}^{(e)} + n_{i,w'}^{(e)})} \quad (8)$$

$$\mathbb{E}[\phi_{j,w}^{(n)}] = \frac{\beta_{j,w}^{(n)} + n_{j,w}^{(n)}}{V\beta^{(n)} + \sum_{w'=1}^V n_{j,w'}^{(n)}} \quad (9)$$

Using the above equations, we can sample the topic z for each word iteratively and estimate all latent random variables.

3.3 Constructing Emotion Lexicon

Our final step is to construct the domain-specific emotion lexicon from the estimates $\phi^{(e)}$ and $\phi^{(n)}$ that we obtained from the EaLDA model.

For each word w in the vocabulary, we compare the $M + 1$ values $\{\phi_{1,w}^{(e)}, \dots, \phi_{M,w}^{(e)}\}$ and $\frac{1}{K} \sum_{i=1}^K \phi_{i,w}^{(n)}$. If $\phi_{i,w}^{(e)}$ is the largest, then the word w is added to the emotion dictionary for the i th emotion. Otherwise, $\frac{1}{K} \sum_{i=1}^K \phi_{i,w}^{(n)}$ is the largest among the $M + 1$ values, which suggests that the word w is more probably drawn from a non-emotion topic. Thus, the word is considered neutral and not included in the emotion dictionary.

4 Experiments

In this section, we report empirical evaluations of our proposed model. Since there is no metric explicitly measuring the quality of an emotion 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 compare the results of solving emotion classification task using our lexicon against different methods, and demonstrate the advantage of our lexicon over other lexicons and other emotion classification systems.

4.1 Datasets

We conduct experiments to evaluate the effectiveness of our model on SemEval-2007 dataset. This is an gold-standard English dataset used in the 14th task of the SemEval-2007 workshop which focuses on classification of emotions in the text. The attributes include the news headlines, the score of

emotions of anger, disgust, fear, joy, sad and surprise normalizing from 0 to 100. Two data sets are available: a training data set consisting of 250 records, and a test data set with 1000 records. Following the strategy used in (Strapparava and Mihalcea, 2007), the task was carried out in an unsupervised setting for experiments.

In experiments, data preprocessing is performed on the data set. First, the texts are tokenized with a natural language toolkit NLTK¹. Then, we remove non-alphabet characters, numbers, pronoun, punctuation and stop words from the texts. Finally, Snowball stemmer² is applied so as to reduce the vocabulary size and settle the issue of data sparseness.

4.2 Emotion Lexicon Construction

We first settle down the implementation details for the EaLDA model, specifying the hyperparameters that we choose for the experiment. We set topic number $M = 6$, $K = 4$, and hyperparameters $\alpha = 0.75$, $\alpha^{(e)} = \alpha^{(n)} = 0.45$, $\beta^{(n)} = 0.5$. The vector $\beta^{(e)}$ is constructed from the seed dictionary using $\gamma = (0.25, 0.95)$.

As mentioned, we use a few domain-independent seed words as prior information for our model. To be specific, the seed words list contains 8 to 12 emotional words for each of the six emotion categories.³ However, it is important to note that the proposed models are flexible and do not need to have seeds for every topic.

Example words for each emotion generated from the SemEval-2007 dataset are reported in Table 1. The judgment is to some extent subjective. What we reported here are based on our judgments what are appropriate and what are not for each emotion topic. From Table 1, we observe that the generated words are informative and coherent. For example, the words “flu” and “cancer” are seemingly neutral by its surface meaning, actually expressing fear emotion for SemEval dataset. These domain-specific words are mostly not included in any other existing general-purpose emotion lexicons. The experimental results show that our algorithm can successfully construct a fine-grained domain-specific emotion lexicon for this corpus that is able to understand the connotation of the words that may not be obvious without the context.

¹<http://www.nltk.org>

²<http://snowball.tartarus.org/>

³<http://minyang.me/acl2014/seed-words.html>

Anger	Disgust	Fear	Joy	Sadness	Surprise
attack	mar	terror	good	kill	surprise
warn	sex	troop	win	die	first
gunman	lebanon	flu	prize	kidnap	jump
baghdad	game	dead	victory	lose	marijuana
immigration	gaze	die	adopt	confuse	arrest
hit	cancer	cancer	madonna	crach	sweat
kidnap	amish	kidnap	celebrity	leave	find
kill	imigration	force	boost	cancer	attack
alzheim	sink	iraq	ship	flu	hiv
iraqi	force	fear	star	kidnap	discover

Table 1: Part of Emotion example words

Algorithm	Anger	Disgust	Fear	Joy	Sadness	Surprise
WordNet-Affect	6.06%	-	-	22.81%	17.31%	9.92%
SWAT	7.06%	-	18.27%	14.91%	17.44%	11.78%
UA	16.03%	-	20.06%	4.21%	1.76%	15.00%
UPAR7	3.02%	-	4.72%	11.87%	17.44%	15.00%
EaLDA	16.65%	10.52%	26.21%	25.57%	36.85%	20.17%

Table 2: Experiment results for emotion classification in term of F1 score

4.3 Document-level Emotion Classification

We compare the performance between a popular emotion lexicon WordNet-Affect (Strapparava and Valitutti, 2004) and our approach for emotion classification task. We also compare our results with those obtained by three systems participating in the SemEval-2007 emotion annotation task: SWAT, UPAR7 and UA. The emotion classification results is evaluated for each emotion category separately. For each emotion category, we evaluates it as a binary classification problem. In the evaluation of emotion lexicons, the binary classification is performed in a very simple way. For each emotion category and each text, we compare the number of words within this emotion category, and the average number of words within other emotion categories, to output a binary prediction of 1 or 0. This simple approach is chosen to evaluate the robustness of our emotion lexicon.

In the experiments, performance is evaluated in terms of F1-score. We summarize the results in Table 2. As an easy observation, the emotion lexicon generated by the EaLDA model consistently and significantly outperforms the WordNet-Affect emotion lexicon and other three emotion classification systems. In particular, we are able to obtain an overall F1-score of 10.52% for disgust classification task which is difficult to work out using pre-

viously proposed methods. The advantage of our model may come from its capability of exploring domain-specific emotions which include not only explicit emotion words, but also implicit ones.

5 Conclusions and Future Work

In this paper, we have presented a novel emotion-aware LDA model that is able to quickly build a fine-grained domain-specific emotion lexicon for languages without many manually constructed resources. The proposed EaLDA model extends the standard LDA model by accepting a set of domain-independent emotion words as prior knowledge, and guiding to group semantically related words into the same emotion category. Thus, it makes the emotion lexicon containing much richer and adaptive domain-specific emotion words. Experimental results showed that the emotional lexicons generated by our algorithm is of high quality, and can assist emotion classification task.

For future works, we hope to extend the proposed EaLDA model by exploiting discourse structure knowledge, which has been shown significant in identifying the polarity of content-aware words.

References

- Stefano Baccianella, Andrea Esuli, and Fabrizio Sebastiani. 2010. Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In *LREC*, volume 10, pages 2200–2204.
- Luciano Barbosa and Junlan Feng. 2010. Robust sentiment detection on twitter from biased and noisy data. In *Proceedings of the 23rd International Conference on Computational Linguistics: Posters*, pages 36–44. Association for Computational Linguistics.
- David M Blei, Andrew Y Ng, and Michael I Jordan. 2003. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022.
- Zhendong Dong and Qiang Dong. 2006. *HowNet and the Computation of Meaning*. World Scientific.
- Andrea Esuli and Fabrizio Sebastiani. 2006. Sentiwordnet: A publicly available lexical resource for opinion mining. In *Proceedings of LREC*, volume 6, pages 417–422.
- Alastair J Gill, Robert M French, Darren Gergle, and Jon Oberlander. 2008. The language of emotion in short blog texts. In *CSCW*, volume 8, pages 299–302.
- Minqing Hu and Bing Liu. 2004. Mining and summarizing customer reviews. In *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 168–177. ACM.
- Jaap Kamps, MJ Marx, Robert J Mokken, and Maarten De Rijke. 2004. Using wordnet to measure semantic orientations of adjectives.
- George A Miller. 1995. Wordnet: a lexical database for english. *Communications of the ACM*, 38(11):39–41.
- Yanghui Rao, Xiaojun Quan, Liu Wenyin, Qing Li, and Mingliang Chen. 2012. Building word-emotion mapping dictionary for online news. In *SDAD 2012 The 1st International Workshop on Sentiment Discovery from Affective Data*, page 28.
- Philip J Stone, Dexter C Dunphy, and Marshall S Smith. 1966. The general inquirer: A computer approach to content analysis.
- Carlo Strapparava and Rada Mihalcea. 2007. Semeval-2007 task 14: Affective text. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 70–74. Association for Computational Linguistics.
- Carlo Strapparava and Alessandro Valitutti. 2004. Wordnet affect: an affective extension of wordnet. In *LREC*, volume 4, pages 1083–1086.
- Peter D Turney and Michael L Littman. 2003. Measuring praise and criticism: Inference of semantic orientation from association. *ACM Transactions on Information Systems (TOIS)*, 21(4):315–346.
- Janyce Wiebe and Ellen Riloff. 2005. Creating subjective and objective sentence classifiers from unannotated texts. In *Computational Linguistics and Intelligent Text Processing*, pages 486–497. Springer.
- Theresa Wilson, Janyce Wiebe, and Paul Hoffmann. 2005. Recognizing contextual polarity in phrase-level sentiment analysis. In *Proceedings of the conference on human language technology and empirical methods in natural language processing*, pages 347–354. Association for Computational Linguistics.
- Rui Xie and Chunping Li. 2012. Lexicon construction: A topic model approach. In *Systems and Informatics (ICSAI), 2012 International Conference on*, pages 2299–2303. IEEE.