Topic-Based Microblog Polarity Classification Based on Cascaded Model

Quanchao Liu[1,2]([✉]), Yue Hu[1,2], Yangfan Lei[2], Xiangpeng Wei[2], Guangyong Liu[4], and Wei Bi[3]

[1] Institute of Information Engineering, Chinese Academy of Science, Beijing, China
liuquanchao@iie.ac.cn
[2] University of Chinese Academy of Science, Beijing, China
[3] SeeleTech Corporation, San Francisco, USA
[4] Beijing, China

Abstract. Given a microblog post and a topic, it is an important task to judge the sentiment towards that topic: positive or negative, and has important theoretical and application value in the public opinion analysis, personalized recommendation, product comparison analysis, prevention of terrorist attacks, etc. Because of the short and irregular messages as well as containing multifarious features such as emoticons, and sentiment of a microblog post is closely related to its topic, most existing approaches cannot perfectly achieve cooperating analysis of topic and sentiment of messages, and even cannot know what factors actually determined the sentiment towards that topic. To address the issues, MB-LDA model and attention network are applied to Bi-RNN for topic-based microblog polarity classification. Our cascaded model has three distinctive characteristics: (i) a strong relationship between topic and its sentiment is considered; (ii) the factors that affect the topic's sentiment are identified, and the degree of influence of each factor can be calculated; (iii) the synchronized detection of the topic and its sentiment in microblog is achieved. Extensive experiments show that our cascaded model outperforms state-of-the-art unsupervised approach JST and supervised approach SSA-ST significantly in terms of sentiment classification accuracy and F1-Measure.

Keywords: Cascaded model · Attention model · LDA model
Bi-RNN · Sentiment analysis · Microblog topic

1 Introduction

With the fast development of social network, more and more Chinese, especially young people, are enjoying the convenience brought by the social network. Take microblog for example, people have published various topics, such as entertainment news, political events, sports reports, etc. They express their various sentiment and opinions towards the topic with multiple forms of media. However, the unique features appear in microblog, such as the sparsity of topics, contact relation, retweet, the short message, the homophonic words, abbreviations, the network language (the popular words), emoticons, etc. These make it very difficult to analyze microblog's topic and its sentiment.

To address the issues, a new cascade model, which excavates the topic of microblog and takes into account the relationship between the topic and its sentiment, is proposed. Our cascaded model aims to identify microblog topic and its sentiment more automatically and efficiently. It mainly has three distinctive advantages: (i) a novel MB-LDA model, which takes both contact relation and document relation into consideration based on LDA, is introduced to mining microblog topic, and the strong relationship between topic and its sentiment is considered in a model; (ii) attention network is introduced to identifying the factors that affect the topic's sentiment and calculating the degree of influence of each factor; (iii) because both MB-LDA model and attention network are considered when using Bi-RNN to judge the sentiment towards the topic, the synchronized detection of the topic and its sentiment in microblog is achieved.

The rest of our paper will be organized as follows. In Sect. 2, we briefly summarize related works. Section 3 gives an overview of data construction, including the dictionaries of sentiment words, internet slang and emoticons. Section 4 gives an overview for cascaded model, including principles, graph models, related resources needed. The experimental results are reported in Sect. 5. Lastly, we conclude in Sect. 6.

2 Related Works

2.1 Topic Model

The present text topic recognition technologies mainly are: traditional topic mining algorithm, topic mining algorithm based on linear algebra, topic mining algorithm based on probability model. The traditional topic model can be traced back to the algorithm of text clustering, and it maps the unstructured data in the text into the points in the vector space by VSM (vector space model), and then uses traditional clustering algorithm to achieve text clustering. Usually text clustering has division-based algorithm, hierarchical-based algorithm, density-based algorithm and so on. However, these clustering algorithms generally depend on the distance calculation between the text and the distance calculation in the mass text is difficult to define; in addition, the clustering result is to distinguish the categories and doesn't give the semantic information, it is not conducive to people's understanding. LSA (latent semantic analysis) is a new method for mining text topics based on linear algebra, proposed by [1]. LSA uses the dimensionality reduction method of SVD to excavate the latent structure (semantic structure) of documents, and then we query and analyze correlation in low dimensional semantic space. By means of SVD and other mathematical methods, the implicit correlation can be well mined. However, the limitation of LSA is that it does not solve the "polysemous" problem of the text, because a word only has one coordinate in semantic space (that is the average of the word more than one meaning), instead of using multiple coordinate to express more than one meaning, and what's more, SVD involves matrix operations, the computational cost is large, and the calculation results in many dimensions is negative, which makes the understanding of the topic is not intuitive.

The third topic model is generative probability model. It assumes that the topic can generate words according to certain rules. When text words are known, the topic distribution of text set can be calculated by probability. The most representative topic model are PLSA (probabilistic latent semantic analysis) and LDA (latent dirichlet allocation). Based on the study of LSA, PLSA is proposed by [2], which combines the maximum likelihood method and the generation model. It follows the dimension reduction of LSA: the text is a kind of high dimensional data when it is represented with TF·IDF, the number of topics is limited and the topic corresponds to the low dimensional semantic space, the topic mining is to project the document from the high dimensional space to the semantic space by reducing the dimension. LDA is a breakthrough extension of the PLSA by adding a priori distribution of Dirichlet on the basis of PLSA. The founder of LDA [3] point out that PLSA does not use a unified probability model in the probability calculation of the document corresponding to the topic, too many parameters will lead to overfitting, and it is difficult to assign a probability to a document outside the training set. Based on these defects, LDA introduces the super parameters and form a Bayesian model with 3 layers "document-topic-word", and then the model is derived by using the probability method to find the semantic structure of the text and to mine the topic of the text.

In recent years, the research on topic model has been deepened, and a variety of models have been derived, such as Dynamic topic model [4], Syntactic topic model [5] and so on. There are also models that consider the relationships between texts, such as Link-PLSA-LDA and HTM (Hypertext Topic Model). Link-PLSA-LDA is a topic model proposed by [6] for citation analysis. In this model, the quoted text is generated by PLSA, and the citation text is generated by LDA, and the model assumes that the two has the same topic. HTM is a topic model proposed by [7] for hypertext analysis. In the process of generating text, HTM adds the influence factors of hyperlinks to mine the topic and classify the text for the hypertext.

2.2 Microblog Sentiment Analysis

Sentiment analysis is one of the fastest growing research areas in computer science, making it challenging to keep track of all the activities in the area. In the research domain of sentiment analysis, polarity classification for twitter has been concerned for some time, such as Tweetfeel, Twendz, Twitter Sentiment. In previous related work, [8] use distant learning to acquire sentiment data. They use tweets ending in positive emoticons like ":)" as positive and negative emoticons like ":(" as negative. They build models using Naives Bayes (NB), MaxEnt (ME) and Support Vector Machines (SVM), and they report SVM outperforms other classifiers. In terms of feature space, they try a Unigram, Bigram model in conjunction with parts-of-speech (POS) features. They note that the unigram model outperforms all other models. However, the unigram model isn't suitable for Chinese microblog, and we make full use of new emoticons which appear frequently in Chinese microblog.

Another significant effort for sentiment classification on Twitter data is by [9]. They use polarity predictions from three websites as noisy labels to train a model. They propose the use of syntax features of tweets like retweet, hashtags, link, punctuation and exclamation marks in conjunction with features like prior polarity of words and

POS of words. In order to improve target-dependent twitter sentiment classification, [10] incorporate target-dependent features and take the relations between twitters into consideration, such as retweet, reply and the twitters published by the same person. We extend their approach by adding a variety of Chinese dictionaries of sentiment, internet slang, emoticons, contact relation and document relation (forwarding), and then by using attention network and Bi-RNN to achieve the sentiment towards the topic.

The problem we address in this paper is to identify microblog topic and its sentiment automatically and synchronously. So the input of our task is a collection of microblogs and the output is topic labels and sentiment polarity assigned to each of the microblogs.

3   Data Description

Microblog allows users to post real time messages and are commonly displayed on the Web as shown in Fig. 1. "# #" identifies the microblog topic, "//" labels user's forwarding relation (document relation), "@" specified the user who we speak to (contact relation).

Fig. 1. Chinese microblog example

People usually use sentiment words, internet slang and emoticons to express their opinions and sentiment in microblog. According to [11], the sentiment word is one of the best sentiment features representations of text, and the rich sentiment words can be conductive to improving sentiment analysis. Internet slang that more and more people use in social network is also important factor for polarity classification. The constructions of them are not only a significant foundation, but also a time-consuming, labor-intensive work. In order to obtain sentiment polarity on microblog topic, we use the same method to construct some dictionaries based on [12].

3.1   The Dictionary of Sentiment Words

In order to obtain more abundant sentiment words, we regard these sentiment words provided by HowNet[1] and National Taiwan University Sentiment Dictionary (NTUSD)[2] as the foundation, and then use lexical fusion strategy to enrich the dictionary of sentiment words.

[13] uses lexical fusion strategy to compute the degree of correlation between test word and seed words that have more obvious sentiment polarity, and then obtain sentiment polarity of test word. We respectively take 20 words as seed words in this paper, as shown in Tables 1 and 2.

Table 1. Seed words with positive polarity

辉煌	美妙	漂亮	俱佳	动听
体面	淳美	良好	出色	完美
美丽	精英	优秀	高手	先进
快乐	最佳	优质	幸福	积极

Table 2. Seed words with negative polarity

罪恶	诅咒	责备	丑陋	丑恶
藐小	累赘	错误	失败	麻烦
不良	恶意	色情	暴力	讨厌
魔鬼	野蛮	腐败	流氓	残酷

So emotional orientation of the test word is computed as follows:

$$SO(word) = \sum_{pword \in Pset} PMI(word, pword) - \sum_{nword \in Nset} PMI(word, nword) \quad (1)$$

where *pword* and *nword* are positive seed word and negative seed word, *Pset* and *Nset* are positive seed words collection and negative seed words collection respectively. *PMI (word₁, word₂)* is described in formula (2), *P(word₁&word₂)*, *P(word₁)* and *P(word₂)* are probabilities of *word₁* and *word₂* co-occurring, *word₁* appearing, and *word₂* appearing in a microblog respectively. When *SO(word)* is greater than zero, sentiment polarity of word is positive. Otherwise it is negative.

$$PMI(word_1, word_2) = \log\left(\frac{P(word_1 \& word_2)}{P(word_1)\, P(word_2)}\right) \quad (2)$$

3.2   The Dictionary of Internet Slang

People usually use homophonic words, abbreviated words and network slang to express their opinions in social network, and [14] has analysed the sentiment of twitter data. Sometimes new words, produced by important events or news reports, are used to express their opinions. So we use the dictionary of internet slang appeared in [12] to support microblog topic polarity classification, containing homophonic words, abbreviated words, network slang and many new words. Table 3 shows part of the dictionary.

Table 3. Part of the dictionary of internet slang

Internet slang	Meaning	Polarity
达人	高人	Positive
狂顶	强烈支持	Positive
萝莉	16岁以下的可爱小女孩	Positive
灰常桑心	非常伤心	Negative
菜鸟	网上低手	Negative

3.3   The Dictionary of Emoticons

We construct the dictionary of emoticons by combining emotional symbol library in microblog with other statistical methods. The former is used to select obvious emotion symbols in microblog, such as Sina, Tencent microblog et al. The latter chooses emoticons used in other social network, containing user-generated emoticons.

Firstly, two laboratory personnel obtain emotional symbol library, and keep the emoticons with the same sentiment polarity after their analysis, and then get rid of emotional symbols with ambiguous polarity, the result is described in Table 4.

Table 4. Part of the dictionary of emoticons

Emoticons	Meaning	Polarity
	Good	Positive
	给力	Positive
	鄙视	Negative
	怒骂	Negative

Secondly, in order to enrich the dictionary of emoticons, especially user-generated emoticons in social network, two laboratory personnel collect and analyse sentiment polarity, and finally obtain the result shown in Table 5.

Table 5. Part of the dictionary of user-generated emoticons

Emoticons	Meaning	Polarity
:o)	大笑	Positive
:)	微笑	Positive
T_T	哭泣	Negative
:(	伤心	Negative

In order to deal with the content conveniently, we pre-process all the microblogs and replace all the emoticons with their "Meaning" by looking up the dictionary of emoticons.

4   The Cascaded Model

4.1   MB-LDA Model for Microblog Topic Mining

MB-LDA is based on the research of LDA, and makes unified modeling for microblog's contact relation and text relation. It is suitable for microblog topic mining. The parameters of the model are shown in Table 6.

Table 6. Parameter definition description

Id	Parameter	Definition
1	α, α_c	Hyperparameters for θ_d and θ_c
2	β	Hyperparameter for φ
3	c	Contactor in conversation message (@)
4	θ_c	Topic distribution associated with contactor c
5	θ_d	Topic distribution over microblog d
6	$\theta_{d_{RT}}$	Topic distribution over retweet microblog d
7	λ	Weight parameters for retweet microblog
8	φ	Word distribution over topics
9	r	Retweet relation in conversation message (//)
10	φ	Word distribution over topics
11	w	Word in microblogs
12	z_i	Topic of word i
13	π_c	Bool parameters to decide specific microblogs

Bayesian network diagram of MB-LDA is shown in Fig. 2, c and r are used to represent the relation of the contact and the retweet respectively. At first, MB-LDA extracts the relation φ between the words and the topic which follows the Dirichlet distribution of the parameter β. Usually conversation message in microblog begins with

"@", it is difficult to judge whether it is conversation message when "@" appears in other positions. In this paper we only consider contact relation in microblog beginning with "@". When MB-LDA generates a microblog, we regard the microblog beginning with "@" as conversation message and set $\pi_c = 1$, and then extract the relation θ_c between each topic and the contact c which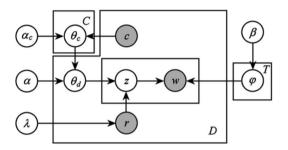 follows the Dirichlet distribution of the parameter α_c, and assign α_c to the relation θ_d between the microblog d and each topic; Otherwise set $\pi_c = 0$, directly extract the relation θ_d between each topic and the microblog d which follows the Dirichlet distribution of the parameter α.

Fig. 2. Bayesian network of MB-LDA

Throughout the microblog sets, the topic probability distribution θ is defined as follows:

$$P(\theta|\alpha, \alpha_c, c) = P(\theta_c|\alpha_c)^{\pi_c} P(\theta_d|\alpha)^{1-\pi_c} \tag{3}$$

Secondly, how to identify retweet relation? If microblog contains "//", we regard the relation between retweet microblog d_{RT} and each topic as $\theta_{d_{RT}}$, and extract r from the Bernoulli distribution with parameter λ, as well as extract the topic probability z_{dn} of the current word from the polynomial distribution with parameters $\theta_{d_{RT}}$ or θ_d. However, we set $r = 0$ when "//" doesn't exist in microblog, and extract the topic probability z_{dn} of the current word from the polynomial distribution with parameter θ_d. Finally, the specific words are extracted from the polynomial distribution with the parameter $\varphi_{z_{dn}}$. More the details about MB-LDA model, see [15].

In microblog, the joint probability distribution of all the words and their topics is shown as follows:

$$P(w, z|\lambda, \theta, \beta) = P(r|\lambda)P(z|\theta)P(w|z, \beta) = P(r|\lambda)P(z|\theta_d)^{1-r}P(z|\theta_{d_{RT}})^{r}P(w|z, \beta) \tag{4}$$

4.2   Hierarchical Attention Network

Traditional approaches of text polarity classification represent documents with sparse lexical features, such as n-grams, and then use a linear model or kernel methods on this

representation. More recent approaches used deep learning, such as convolutional neural networks and recurrent neural networks based on long short-term memory (LSTM) to learn text representations. However, a better sentiment representation can be obtained in this paper by incorporating knowledge of microblog structure in the attention network. We know that not all parts of a microblog are equally relevant for judging the microblog polarity and that determining the relevant sections involves modeling the interactions of the words, not just their presence in isolation.

Words form sentences, sentences form a document. In the application of micro-blog's polarity classification, we introduce hierarchical attention network created by Zichao Yang into our cascaded model. Our intention is to let the network to pay more or less attention to individual emotional factor when constructing microblog's polarity classifier. The overall architecture is shown in Fig. 3. It consists of five parts: a word sequence encoder, a word-level attention layer, a sentence encoder, a sentence-level attention layer and softmax layer. The details of different parts have been described in [16], we don't introduce them anymore.
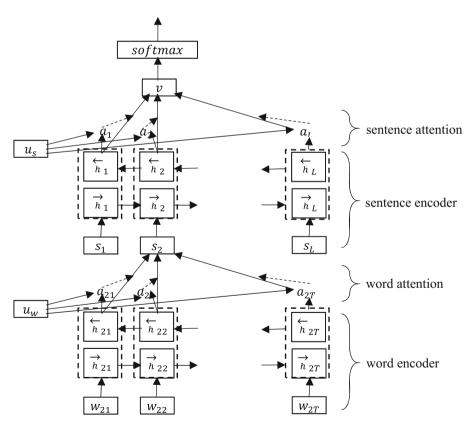

Fig. 3. Hierarchical attention network

4.3 The Cascaded Model Architecture for Topic Polarity Classification

Although attention-network-based approaches to polarity classification have been quite effective, it is difficult to identify the topic and give the polarity towards that topic synchronously. We combine the MB-LDA model and attention network to generate the cascaded model. The overall architecture of the cascaded model is shown in Fig. 4. T_{w_i} expresses the probability of the word w_i belongs to the topic T, where i $\in$ [1, T]. The advantages of this architecture are as follows: (i) polarity classification is carried out on the basis of the results of topic recognition; (ii) the information input into the neural network takes into account the probability T_{w_i}.

The processing steps are as follows:

(i)   The MB-LDA model is used to obtain the topics of microblog data sets and the top 50 sentiment words in each topic. These sentiment words are selected from the topic according to the dictionary of sentiment words.
(ii) Both the microblogs and the topic probabilities of each sentiment words from the same topic are used as the input of hierarchical attention network.
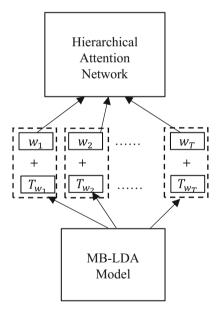(iii) The polarity classification of each microblog of each topic is achieved in the softmax layer.

Fig. 4. The cascaded model architecture

5 Experiments and Results

In order to quantitatively analyze the performance of the cascade model, we use 4 different real microblog topic datasets to do experiments, and analyze the accuracy of polarity classification, the influence of topic number on accuracy, and the influence of emoticons on accuracy.

5.1 Data Sets

The labeled data sets in NLP&CC 2012[3] & 2013[4], a total of 405 microblogs, are provided by *Tencent Weibo*, including four topics: *hui_rong_an*, *ipad*, *kang_ri_shen_ju_sample* and *ke_bi_sample*. We reserve the microblog labeled with "*opinionated = Y*" and "*forward*" on behalf of "*//*" (retweet) in a microblog. When the number of "polarity = 'POS'" in microblog is more than or equal to the number of "polarity = 'NEG'", we think that microblog is positive. Otherwise, it is negative, and according to the polarity tagging, we randomly add the corresponding emoticons to microblog to enrich the emotional characteristics of the data sets.

In order to avoid over-fitting or under-fitting, we adopt 10-fold cross-validation in the experiments. Namely data sets would be randomly divided into 10 parts, 9 parts of them are used as training sets and the others are used to test. We repeat the process for 10 times and finally take the average value.

In addition, in order to encode emoticons, such as "👍", "T_T", and so on, we carry out the corresponding string processing "Good" and "哭泣".

5.2 The Evaluation of Microblog Topic Polarity Classification

Polarity classification on microblog topic is evaluated by *Precision*, *Recall* and *F-measure*.

$$\Pr ecision = \frac{\#system\_correct}{\#system\_proposed} \tag{5}$$

$$\mathrm{Re}call = \frac{\#system\_correct}{\#person\_correct} \tag{6}$$

$$F - measure = \frac{2 \times \Pr ecision \times \mathrm{Re}call}{\Pr ecision + \mathrm{Re}call} \tag{7}$$

Where *#system_correct* is the correct result from system, *#system_proposed* is the whole number of microblogs from system, *#person_correct* is the number of microblogs that has been annotated correctly by people, *#weibo_topic* is the number of microblogs containing topic words, *#weibo_total* is the whole number of microblogs in the collection.

---

[3] http://tcci.ccf.org.cn/conference/2012/pages/page04_eva.html.

[4] http://tcci.ccf.org.cn/conference/2013/pages/page04_eva.html.

5.3   Results

In order to evaluate microblog topic polarity recognition ability, considering the semi-supervised learning of the cascaded model, we compare it with the most representative unsupervised learning model JST [17], semi-supervised learning model SSA-ST [18] and supervised learning model SVM in four data sets for microblog topic polarity classification. The results of the experiment are shown in Table 7. The value in the table shows the average value of the correct rate of each group of data.

Table 7. The comparison of polarity classification in 4 data sets

Model name	Precision	Recall	F-measure
JST	71.09	62.3	66.41
SSA-ST	78.9	74.32	76.54
SVM	89.1	85.19	87.1
Cascaded model	**86.74**	**81.35**	**83.96**

From the above table, we can see that the precision of polarity classification in cascaded model is higher than that of unsupervised model JST and semi-supervised model SSA-ST, while our result is similar to that of supervised model SVM. The reason is that our cascaded model has strong ability to identify emotional characteristics, and we find that the attention network has higher weight in features' calculation. This helps us quickly identify key elements that affect microblog's topic polarity. Although the experimental results of the cascaded model are lower than SVM, the cascaded model can discover topics and achieve higher polarity classification with fewer training sets.

Because the cascaded model can synchronously detect the topic and its polarity in microblog data sets, it is necessary to explore the interaction between polarity classification and topic detection. We carry out an experimental analysis that how does the number of topics affect the precision of polarity classification. The results of the experiment are shown in Fig. 5.
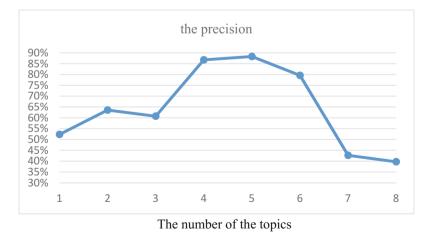

Fig. 5. The influence of the number of topics on the precision of polarity classification

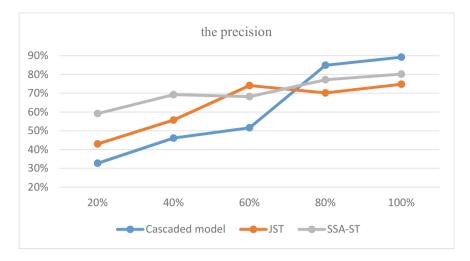The precision chart title: "the precision"

Fig. 6. The influence of the proportion of emoticons on the precision of polarity classification

As shown in Fig. 5, the influence of different numbers of topics generated by the cascaded model is different on the same data sets. The inappropriate number of topics will reduce the precision of microblog's polarity classification. Too little number of topics can reduce the correlation between the topic and its polarity. Too much number of topics can make the complete topic fragmented, which improves the noises of polarity classification and reduces the precision.

At the same time, we know that usually emoticons can effectively improve the effect of polarity classification, so what is the quantitative correlation between the two? We are gradually raising the number of microblogs containing emoticons in four data sets, that is to increase the proportion of microblogs with emoticons. The results of the experiment are shown in Fig. 6.

Figure 6 shows that with the increase of the number of emoticons in microblog, the precision of polarity classification is also increasing. From the trend of the precision, different polarity classification models have different promotion when we increase the proportion of emoticons in data sets, the precision of all classification models and the proportion of emoticons is linear positive correlation, and based on the topic identified, the polarity classification performance of our cascade model is better obviously.

6   Conclusions and Future Work

With the popularity of microblog services, people can see and share reality events on microblog platform. Mining the topic sentiment hidden in massive microblog messages can effectively assist users in making decisions. [19, 20] have introduced a number of different sentiment analysis methods for twitter, but our approach is also suitable for twitter. In this paper, MB-LDA model and attention network are applied to Bi-RNN for topic-based microblog polarity classification, and the synchronized detection of the topic and its sentiment in microblog is achieved.

References

1. Deerwester, S., Dumais, S.T., Furnas, G.W., et al.: Indexing by latent semantic analysis. J. Assoc. Inf. Sci. Technol. **41**(6), 391–407 (1990)
2. Hofmann, T.: Probabilistic latent semantic indexing. In: International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 50–57. ACM (1999)
3. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent dirichlet allocation. J. Mach. Learn. Res. Arch. **3**, 993–1022 (2003)
4. Blei, D.M., Lafferty, J.D.: Dynamic topic models. In: International Conference, DBLP, pp. 113–120 (2006)
5. Boydgraber, J., Blei, D.M.: Syntactic topic models. In: Advances in Neural Information Processing Systems, pp. 185–192 (2008)
6. Nallapati, R., Cohen, W.: Link-PLSA-LDA: a new unsupervised model for topics and influence of blogs. In: ICWSM (2008)
7. Sun, C., Gao, B., Cao, Z., et al.: HTM: a topic model for hypertexts. In: Conference on Empirical Methods in Natural Language Processing, pp. 514–522. Association for Computational Linguistics (2008)
8. Go, A., Bhayani, R., Huang, L.: Twitter sentiment classification using distant supervision. CS224N Project report, Stanford (2009)
9. Barbosa, L., Feng, J.: Robust sentiment detection on Twitter from biased and noisy data. In: Proceedings of COLING 2010 Beijing, China, pp. 36–44 (2010)
10. Long, J., Yu, M., Zhou, M., et al.: Target-dependent Twitter sentiment classification. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, Portland, Oregon, pp. 151–160 (2011)
11. Du, W., Tan, S., Yun, X., et al.: A new method to compute semantic orientation. J. Comput. Res. Dev. **46**(10), 1713–1720 (2009)
12. Liu, Q., Feng, C., Huang, H.: Emotional tendency identification for micro-blog topics based on multiple characteristics. In: 26th Pacific Asia Conference on Language, Information and Computation (PACLIC 26), pp. 280–288 (2012)
13. Wang, S., Li, D., Wei, Y.: A method of text sentiment classification based on weighted rough membership. J. Comput. Res. Dev. **48**(5), 855–861 (2011)
14. Agarwal, A., Xie, B., Vovsha, I., et al.: Sentiment analysis of Twitter data. In: Proceedings of the Workshop on Language in Social Media (LSM 2011), Portland, Oregon, pp. 30–38 (2011)
15. Zhang, C., Sun, J., Ding, Y.: Topic mining for microblog based on MB-LDA model. J. Comput. Res. Dev. **48**(10), 1795–1802 (2011)

16. Yang, Z., Yang, D., Dyer, C., et al.: Hierarchical attention networks for document classification. In: Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 1480–1489 (2017)
17. Lin, C., He, Y., Everson, R., et al.: Weakly supervised joint sentiment-topic detection from text. IEEE Trans. Knowl. Data Eng. **24**(6), 1134–1145 (2012)
18. Hu, X., Tang, L., Tang, J., et al.: Exploiting social relation for sentiment analysis in microblogging. In: Proceedings of the 6th International Conference on Web Search and Data Mining. Rome, Italy, pp. 537–546 (2013)
19. Nakov, P.: Semantic sentiment analysis of Twitter data. arXiv preprint arXiv:1710.01492 (2017)
20. Wang, B., Liakata, M., Tsakalidis, A., et al.: TOTEMSS: topic-based, temporal sentiment summarisation for Twitter. In: Proceedings of the IJCNLP 2017, System Demonstrations, pp. 21–24 (2017)