SENTIMENT ANALYSIS OF CHINESE MICROBLOG MESSAGE USING NEURAL NETWORK-BASED VECTOR REPRESENTATION FOR MEASURING REGIONAL PREJUDICE

Yung-Chun Chang, Department of Information Management, National Taiwan University, and Institute of Information Science, Academia Sinica, Taiwan.
changyc@iis.sinica.edu.tw

Chin-Shun Chou, Institute of Information Science, Academia Sinica, Taiwan.
chouchris@iis.sinica.edu.tw

Yang Zhang, Department of Political Science, University of Iowa, USA.
yang-zhang@uiowa.edu

Xi Wang, Interdisciplinary Graduate Program in Informatics, University of Iowa, USA.
xi-wang-1@uiowa.edu

Wen-Lian Hsu, Institute of Information Science, Academia Sinica, Taiwan.
hsu@iis.sinica.edu.tw

Abstract

Regional prejudice is prevalent in Chinese cities in which native residents and migrants lack a basic level of trust in the other group. Like Twitter, Sina Weibo is a social media platform where people actively engage in discussions on various social issues. Thus, it provides a good data source for measuring individuals' regional prejudice on a large scale. We find that a resentful tone dominates in Weibo messages related to migrants. In this paper, we propose a novel approach, named DKV, for recognizing polarity and direction of sentiment for Weibo messages using distributed real-valued vector representation of keywords learned from neural networks. Such a representation can project rich context information (or embedding) into the vector space, and subsequently be used to infer similarity measures among words, sentences, and even documents. We provide a comprehensive performance evaluation to demonstrate that by exploiting the keyword embeddings, DKV paired with support vector machines can effectively recognize a Weibo message into the predefined sentiment and its direction. Results demonstrate that our method can achieve the best performances compared to other approaches.

Keywords: Sentiment Analysis, Neural Network, Regional Prejudice, Distributed Word Representation, Text Classification.

1 INTRODUCTION

On the new year's eve of 2015, 36 visitors died and 49 injured in a stampede when more than a million visitors rushed to the viewing deck at the Bund in Shanghai. This accident caught great attention around China and the criticisms were mostly cast toward local officials for their inadequate preventive actions. But on Weibo (a Chinese Twitter-equivalent), many Shanghai natives blamed the non-local visitors for the tragedy and complained that a large non-native population brings a number of social problems to Shanghai. Shanghai is one of the most developed cities in China but is also notorious for its regional prejudice. Regional prejudice is based on distrust and conflicts between insiders and outsiders, which is also prevalent in Beijing and many other Chinese cities. In late 1980s, China started to seek market-driven economic development and allow people to leave their birth towns for work opportunities in another city. Chinese people subsequently began to move on an unprecedented scale. The number of migrants grows fast over time and they now make up a considerable part in every large city. The National Health and Family Planning Commission provided a conservative estimate that the migrant population reached 253 million in 2014[1]. Although visitors and migrant workers are an important force driving local economic growth, native urban residents tend to focus more on the negative side of population inflow. They blame migrants for increasing crime rate, crowding living space, stealing local jobs, and taking scarce public resources. Some of these concerns are reasonable, but many are rooted in stereotypes and prejudice toward outsiders. Local resentment, along with discriminative policies, create great barriers for migrants and their children to settling down and getting assimilated in a new city. The Chinese national government calls for a fair treatment of native residents and migrants. However, without breaking regional prejudice, it is difficult to allocate a just portion of public resources to migrants and prevent the urban society from falling into segregation.

Microblogs are a good subject for studying various social issues including regional prejudice. Microblogs are a popular communication tool for Internet users. Millions of messages are posted daily on popular microblog websites such as Weibo[2], Twitter[3] and Facebook[4]. The authors post their life stories and exchange opinions on a variety of topics and issues. Because of microblogs' user-friendly features and easy access, many Internet users shift from traditional communication tools (such as traditional blogs and emails) to microblog services. Since more and more users become actively engaged in microblogging, microblog websites are a valuable data source of public opinion and sentiment. Microblog data are useful for marketing and social studies. Therefore, we use Sina Weibo as our data source for measuring regional prejudice. Like Twitter, Weibo is an important social media platform in China where people publicize short and instant messages on various topics on the Internet. We focus on the Weibo messages relevant to migrants in cities. We propose a new method for sentiment analysis which provides reliable estimates of regional prejudice. Our big-data-based measures of prejudice complement more common measures from surveys and experiments. In addition, based on the content, we attempt to identify whether a Weibo message centers on migrants or native residents.
There are three contributions of our paper:

- Our research systematically examines regional prejudice in China by analyzing numerous discussions on a social media platform. Our big-data-based measurement of regional prejudice does not exist in previous study.
- Our proposed method is based on distributed real-valued vector representation of keywords learned from neural networks. Such a representation can project rich context information onto the vector space, and subsequently be used to infer similarity measures among words, sentences, and even documents.
- We measure Weibo messages' sentiment and the subject (migrants/native residents). Thus, we can extract more details in Weibo texts and improve the accuracy of our regional prejudice measures.

---

[1]http://politics.people.com.cn/n/2015/1112/c1001-27805401.html
[2]http://weibo.com
[3]http://twitter.com
[4]http://facebook.com

This paper is organized as follows. The following section reviews related works on sentiment analysis. Next, in section 3, we introduce a neural-network-based approach for sentiment classification regarding regional prejudice. In section 4, we evaluate the performance of this approach. Finally, we conclude our paper and discuss future research directions.

2 SENTIMENT ANALYSIS FOR MEASURING REGIONAL PREJUDICE

People often convey their feelings through words. Identifying essential factors that affect sentiment transition is important for human language understanding. With rapid growth of computer-mediated communication applications, such as social websites and microblogs, research on sentiment analysis attracts increasing attention from enterprises for business intelligence (Chen et al., 2010; Purver and Battersby, 2012). For instance, Pang et al. (2002) designed an algorithm to classify movie reviews into positive and negative emotions. Mishne (2005), and Yang et al. (2006) used emoticon as tag to train SVM (Cortes and Vapnik, 1995) classifiers at the document and sentence level, respectively. In their studies, emoticon was treated as mood or emotion tag, and textual keywords were considered as features. Moreover, Wu et al. (2006) proposed a sentence-level emotion recognition method using dialogs as corpus in which each sentence was classified into “Happy”, “Unhappy” or “Neutral” category. Some microblog websites, such as Sina Weibo, incorporate the Web 2.0 technologies that allow users to express their opinions toward anything. Companies and media organizations are increasingly interested in seeking ways to mine microblogs for information about what people think and feel about their products and services. While there has been a fair amount of research on expressed sentiments in online reviews and news articles, microblogs with informal language and message-length constraint have been studied by just a few researchers. Given the length limit, classifying sentiment of Weibo messages is similar to sentence-level sentiment analysis (e.g., (Yu and Hatzivassiloglou, 2003; Kim and Hovy, 2004)); however, the jargons used in Weibo and the special nature of microblogging make Weibo sentiment analysis a different task.

To understand regional prejudice in Chinese cities, the necessary initial step is to have a good measure of regional prejudice. We collected 4.6 million messages from Sina Weibo over four months (December 14, 2014 to April 15, 2015) that discussed migrant-related issues, and measured Weibo users’ sentiment toward migrants. Weibo, namely microblog, is a Chinese equivalent of Twitter. Sina Weibo is the most popular microblog in China and has attracted a huge number of users. A Weibo message can have 140 Chinese characters. Since a Chinese character is like an English word, a Weibo message can contain rich content. Like Twitter, Weibo does not only reflect public opinion, but also shapes public opinion and society. People are more likely to convey their true opinions on Weibo than in a face-to-face social setting. Prejudice, discrimination, and intolerance, these concepts are hard to directly measure with precision, because they are subject to social desirability effects. For example, traditional social surveys always underestimate racism in the United States. The reason is that racism is socially and politically incorrect and respondents tend to hide their racist beliefs (Kuklinski et al., 1997a,b). In contrast, Weibo users can post controversial or provocative messages under an anonymous identity, so they are less restrained to say what they truly believe. In our data, derogatory and offending words are common, but these words are not often heard during survey interviews. For example, below are two messages from our collected Weibo sample that express deep resentment toward migrants and compassion, respectively:

“早高峰高速路匝道堵了，都是外地车害的，大楼着火了，都是外地临时工害的，外滩踩踏了，估计也有不少外地人，外地人、外地车你们什么时候不是害上海之马！？” (*The traffic jams on highway ramps during morning rush hours is because of outside cars, the building caught on fire because of outsider temporary workers, a lot of outsiders were involved in the trampling on the Bund too, outsiders and outside cars, when can you stop bringing Shanghai down?*)

“外地人，本地人，来到温江一家人。” (*Whatever you are outsiders or local residents, we are families when you come to Wenjiang.*)

3 SENTIMENT ANALYSIS OF WEIBO USING NEURAL NETWORK-BASED VECTOR REPRESENTATION

Distribute word representation, or word embedding, has been used on a wide spectrum of applications in NLP. One of the earlier well-known methods for developing word embedding (Bengio, 2003) estimated a statistical language model and formalized a feed-forward neural network for predicting possible words in future context while inducing word embeddings as a by-product. The work inspired other efforts to develop methods for learning latent semantic and syntactic regularities, including representative methods such as continuous bag-of-word (CBOW) and the skip-gram model (Mikolov, 2013). We model sentiment analysis as a classification problem, and define the task as the following. Let $W=\{w_1, w_2, \ldots, w_k\}$ be a set of words, $M=\{m_1, m_2, \ldots, m_j\}$ be a set of messages, S={*positive*, *neutral*, and *negative*} be a set of sentiments, and D={*local*, *non-local*, *multi-direction*, and *untargeted*} be a set of directions. Each message m is a set of words such that $m \subseteq W$. The goal is to decide the most appropriate sentiment S for a message m_i, although one or more sentiments can be associated with a message. We further predict the direction of sentiment according to the recognized sentiment polarity. We will describe these models briefly in the following sections.

Recently, neural network-based approaches have been proposed for learning vector representations for longer text segments such as paragraphs or documents (Le and Mikolov, 2014), namely, DBOW and DM. In essence, the DBOW model is an extension of SG, and the DM of CBOW. Originally, the word vectors are learned using context words. This idea can be extended for document vectors in a similar manner. If every document is mapped to a unique vector, which can be thought of as a special word, we can use it to predict other words in the same document. During training, the word vectors will be learned first. Then, a sliding window over the whole document is used to sample every word. Eventually, we can obtain a vector for each document that contains information of embeddings in the whole document. More specifically, document vectors (and word vectors) are learned with a stochastic gradient descent obtained via back-propagation. For each iteration, the vector for a document is fed through the neural network, then we can compute the error gradient from the network and use the gradient to update the document vector. Moreover, document vectors have some advantages over traditional bag-of-words models. First, since they are based on word vectors, the semantics of the words can also be incorporated. Second, they can include information from a much broader context, i.e., the whole document. Such feature usually requires a very large n in n-gram models, hence a heavy toll on the memory.

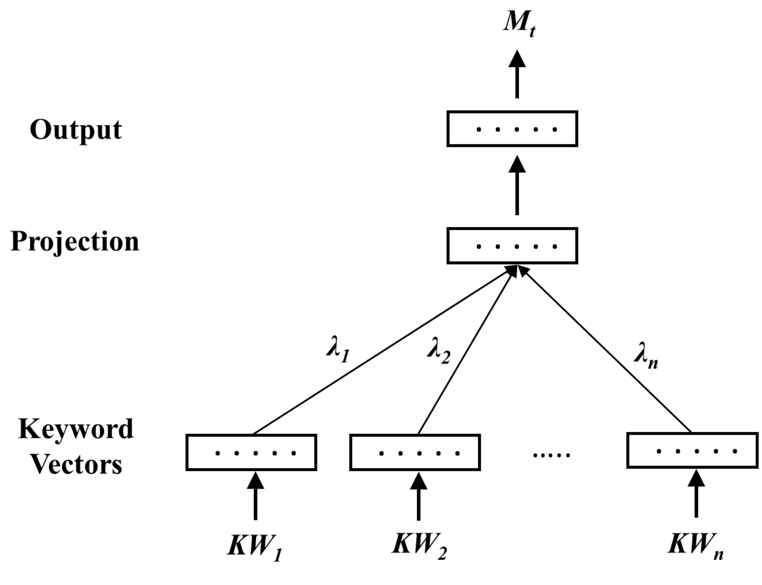

Figure 1. The DKV model estimates the Weibo message representation M_t using keyword vectors KW_1, ..., KW_n in the message. λ_1, ..., λ_n Denotes the weighted LLR scores of respective keywords.

In this work, we propose the Distributed Keyword Vectors (DKV) for representation and classification of a Weibo text. Previous research has found that keyword information is very effective in text classification (Chang et al., 2015). This motivates us to use only keyword vectors to represent a Weibo message in the sentiment classification task. First, sets of sentiment-specific keywords are identified using log likelihood ratio (LLR) (Manning and Sch¨utze, 1999), an effective feature selection method. It calculates the likelihood of the assumption that the occurrence of a word *w* in sentiment *S* is not random, and a higher LLR value indicates that *w* is closely associated with the sentiment. LLR value of each word *w* is calculated as follows. Given a training dataset, we first obtain four frequencies {*k*, *l*, *m*, *n*} defined as:

$$k = N(w \wedge S) \qquad l = N(w \wedge \neg S)$$

$$m = N(\neg w \wedge S) \qquad n = N(\neg w \wedge \neg S)$$

where $N(w \wedge S)$ denotes the number of Weibo messages that contain *w* and belong to sentiment *S*, $N(w \wedge \neg S)$ denotes the number of Weibo messages that contain *w* but does not have sentiment *S*, and so on. Then, we employ Eq. (6) to calculate the likelihood of the assumption that the occurrence of a word *w* with sentiment *S* is not random. In Eq. (1), the probabilities $p(w)$, $p(w|S)$, and $p(w|\neg S)$ are estimated using maximum likelihood estimation. We rank the words in the training dataset based on their LLR values and select those with highest LLR values to compile a sentiment keyword list.

$$LLR(w,S) = 2\log\left(\frac{p(w|S)^k (1-p(w|S))^m \, p(w|\neg S)^l (1-p(w|\neg S)^n)}{p(w)^{k+1}(1-p(w))^{m+n}}\right) \tag{1}$$

Next, word embeddings are learned on the training corpus using *CBOW*. Finally, a Weibo message is represented also as a vector using a weighted average of the vectors which correspond to words in the keyword list. Fig. 1 illustrates the DKV model, in which, the Weibo message M_t is represented as a weighted average of the keyword vectors, and the weight λ_i for a keyword KW_i is determined by its LLR value. If there is no keyword in a Weibo message, we calculate the mean of all word vectors in this message and compute cosine similarity over all keyword vectors to find the closest ones to represent this Weibo message. Conceptually, we are projecting each Weibo message onto a high dimensional vector space constructed from keywords to which subsequent clustering or classifier can be applied.

4 EXPERIMENT

4.1 Dataset and Experiment Setup

To the best of our knowledge, there is no publicly available corpus for measuring regional prejudice. Therefore, we compiled a corpus from Sina Weibo for performance evaluation. To gather the Weibo messages and user information, we built a crawler based on open Weibo API. We rely on 13 keywords including native (本地), permanent population(常住人口), census register (户籍), registered permanent residence (户口), residence permit (居住证), floating population(流动人口), settle in a new place (落户), peasant workers (农民工), non-native (外地), from a different town(外来), people from other provinces(外省人), transient population (暂住人口), temporal residential permit (暂住证) to filter the messages over four months between December 14, 2014 and April 15, 2015. The key words can be divided into two groups. One group of words such as census register, residence permit, and temporal residential permit, are related to internal migration policies. The other group of words are labels indicating residence status. Some of the labels, such as permanent population, floating population, transient population, were created by the Chinese government. Other labels, such as native, non-native, and peasant workers, are common words people use when they talk about migrants. We kept the messages which match at least one of the key words. In total, we collected 4,641,398 Sina Weibo messages from 1,143,698 users.

As it is almost impossible to label all 4.6 million messages manually, we need classification algorithms to decide the types of messages automatically. To train the classifier, we leverage human annotated data. We randomly selected about 5,000 messages out of all original Weibo sample. In this paper, we engage in two tasks: (1) What is the sentiment polarity of each message, positive, negative or neutral? (2) What is the emotion direction of each message, toward migrants or native residents? Table 1 shows the summary statistics of the sample from the collected Weibo messages for performance evaluation. We asked five experts to annotate the Weibo messages via Crowdsdom[5] which is an annotation platform similar to Amazon Mechanical Turk. The Kappa statistic of the labeling process is 0.63, which means that our dataset is reliable.

Dataset	Category	# Dev.	# Test	*Total*
Sentiment (4,490)	*Positive*	124	124	248
	Negative	1137	1137	2274
	Neutral	984	984	1968
Direction (4,742)	Local	368	369	737
	Non-local	427	428	855
	Multi-direction	340	340	680
	Untargeted	1235	1235	2470

Table 1. Descriptive statistics of the datasets.

4.2 Results and Discussion

We use micro-average of accuracy to measure the performance, which is derived from a contingency table of predictions for a target sentiment S_k. A comprehensive performance evaluation of the DKV with other methods is conducted. The first model is an emotion keyword-based model which is trained by SVM to demonstrate the effect of our keyword extraction approach (denoted as *KW-SVM*). Next, we compare a neural network-based document representations, (denoted as *DM*), proposed in (Le and Mikolov, 2013). They are trained using the dimensionality setting of 100, i.e., every document is represented as a vector containing 100 elements. We also included Naive Bayes (McCallum and Nigam, 1998) as the baseline model (denoted as *NB*) Weibo texts' sentiment. We employed Jieba[6] for Chinese word segmentation. The dictionary for Naïve Bayes removes stop words according to the Chinese stop word list of Zou et al. (2006), and retain tokens that span 90% of the accumulated frequency. In other words, the dictionary covers up to 90% of the tokens in the corpus. As for unseen events, we use Laplace smoothing in Naïve Bayes. Our system, denoted as *DKV*, is constructed as the following. At the outset of the training stage, word vectors are learned using *CBOW* with *HS*. The dimensionality is set to be 200, which is identical as that of *DM* and *DBOW*. Next, we calculate LLR values of every word in the corpus, and select top 150 distinctive ones for each sentiment. This threshold is found to be useful in (Chang et al., 2015). The weight of a keyword vector is set to be the log LLR value times a scaling factor of two. Afterwards, sum of the weighted keyword vectors are used as representations for each Weibo message, and support vector machine (SVM) (Chang and Lin, 2011) is trained to classify the sentiment and direction of the Weibo message. To ensure the fairness of this experiment, the keyword vectors are frozen at test time, so that new context words in the testing messages will not be learned by the model.

Task	***Accuracy (%)***					
	NB	**KW-SVM**	**DM**	**DKV**	**DKV_{+AI}**	**$DKV_{+AI+Sen}$**
Sentiment Classification	38.16	71.31	73.93	74.20	**74.79**	**-**
Direction Recognition	59.34	71.87	74.22	74.56	**77.83**	**78.9**

Table 2. System performance of sentiment classification and direction recognition

[5]http://crowdsdom.com/

[6] https://github.com/fxsjy/jieba

The system performances for sentiment classification and direction recognition are listed in Table 2. The baseline Naïve Bayes classifier is a keyword statistics-based model which only performs mediocrely. Since Naïve Bayes classifier only considers surface word weightings, it is difficult to represent inter-word relations using this method. The overall accuracy of the Naïve Bayes classifier is 38.16% for sentiment polarity and 59.34% for sentiment direction. As we can see, with accuracy at 71%, the KW-SVM substantially increases proficiency in recognizing polarity and direction of sentiment. The results indicate that polarity and direction of sentiment in Weibo messages can be recognized efficiently using the LLR scores of keywords, because the likelihood of a word contributing to a certain sentiment is not random. The words with large LLR values are closely associated with the corresponding emotions (Manning and Sch¨utze, 1999). It is worth noting that the neural network-based method (i.e. DM) can further improve upon the aforementioned methods for another 3% absolutely. They demonstrate the ability to overcome weaknesses of bag-of-word models due to the combinatorial nature of a vector-based model. On top of that, DKV can further surpass other methods. These results indicate that polarity and direction of sentiment can be sufficiently recognized by using only the information from keywords. It demonstrates that word embedding can encode the complex relations between keywords and their polarity and direction of sentiment into a dense vector. By means of an effective keyword weighting like LLR, we can give more discriminating power to those unique vectors. Paired with a robust vector-based classifier like SVM, our system can provide substantial performances. Moreover, such an approach has the advantage of requiring very little supervision and feature engineering. It automatically learns the importance or weights of various words inherently. Moreover, the performance of DKV on both tasks can be improved if we consider author information (i.e. +AI) of Weibo messages. This indicates polarity and direction of sentiment are associate with who post the Weibo message. Based on our further analysis, a large number of messages convey negative sentiment if author and target of message from different city (e.g. direction of sentiment is from local to non-local). Notably, DKV further obtain the best overall accuracy in the task of direction recognition through considering sentiment information (i.e. +Sen) which predicted in the first task.

Figure 2. The word cloud generated from sentiment keywords. Colors and their corresponding sentiment: Red - Negative, Green - Positive, and Gray - Neutral.

To visualize the keyword selection method, we present the keywords as word clouds in Figure 2. Each keyword is color according to its corresponding sentiment, and scaled according to its LLR score. In this way, we can easily identify the features tied to each sentiment. For example, we observed that keywords related to "Positive" (in green) are mostly about food such as "下午茶 $_{\text{afternoon tea}}$" and "特色小吃 $_{\text{special snack}}$". On the contrary, "Negative"-related keywords (in red) consist largely of targeted terms. For instance, the most noticeable word "你们 $_{\text{you}}$" indicates deep resentment toward migrants in the Shanghai stampede. Consequently, numerous targeted terms such as "警察 $_{\text{policemen}}$", "上海 $_{\text{Shanghai}}$", and "侵佔 $_{\text{encroach}}$" are also keywords that express negative emotion. The figure highlights the fact that the extracted sentiment keywords are highly correlated with polarity of Weibo messages, and including them in the distributed vector representations helps DKV achieve more accurate sentiment classification.

Weibo messages are very short, which imposes challenges for learning users' background information. Among all the messages examined, 42 percent concentrates on either migrants or native urbanites, but the subject of the rest messages is ambiguous. Based on the messages with identified subjects, 16 percent mentions migrants, 18 percent native residents, and 14 percent both social groups. These results suggest that migrants and native residents receive roughly equal attention in online discussions on the issue of migration. Furthermore, our sentiment analysis of regional prejudice demonstrates deep mistrust between insiders and outsiders in Chinese cities. Overall, only 6 percent of the Weibo messages about migrants show positive sentiment. In contrast, 51 percent of the messages have a negative tone and 44 percent are neutral. Our research corroborates previous findings about interpersonal trust in China. By analysing the fifth wave of the World Value Survey (2005-2009), Delhey et al. (2011) compared 51 countries and found that interpersonal trust in China has a short radius. Chinese people have a high level of trust in their family members, friends, colleagues, and fellow townsmen, but lack trust in strangers and people from a different town (Tang, 2016).

5 CONCLUSION

Regional prejudice is a serious urban issue in China. In Chinese cities, native residents and migrants lack a basic level of trust in the other group. In this research, we collected 4.6 million migration-relevant messages from Sina Weibo. We propose a neural network-based vector representation approach to measure individuals' regional prejudice based on social media data. The contributions of this work are three-fold. First, we demonstrate that our method can yield substantial improvement over other models on the Weibo dataset. Next, we also find that the amount of keywords is positively related to the performance of our system, though it stops to be so effective beyond a certain limit. Finally, the findings from our sentiment analysis confirm previous concern with regional prejudice: in our Weibo data sample, 51 percent of messages contain negative sentiment, 44 percent neutral, and merely 6 percent positive. Regional prejudice in China is widespread and severe. In the future, we want to investigate algorithms for combining keyword vectors that can lead to a better representation of a message. Also, exploring how to project semantic knowledge into the vector space is another interesting field of research. Lastly, more research is needed for finding the causes and consequences of regional prejudice. For example, is regional prejudice stable over time or easily stirred up by incidents and opinion leaders? Local residents are hostile toward migrants partly because of realistic conflicts and partly because of cultural differences, but which type of causes play a more important role in Chinese cities? And does the household registration system not only prevent migrants from settling in destination cities but also create psychological distance between migrants and local residents? We acknowledge this as an important issue for future research.

References

Bengio, Y., Ducharme, R., Vincent, P., and Janvin, C. (2003) A neural probabilistic language model. *The Journal of Machine Learning Research*, vol. 3, pp. 1137-1155.

Chang, C.C. and Lin C.J. (2011) LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*, vol. 2, no. 3, pp. 1-27.

Chang, Y.C., Hsieh, Y.L., Chen, C.C., and Hsu, W.L. (2015). A semantic frame-based intelligent agent for topic detection. *Soft Computing*, pp. 1-11.

Chen Y., Lee, S., Li, S., and Huang, C. (2010). Emotion cause detection with linguistic constructions. In *Proceedings of the 21th International Conference on Computational Linguistics*, pp. 179-187.

Cortes, C. and Vapnik, V. (1995). Support vector networks. *Machine Learning*, vol. 20, pp. 1-25.

Delhey J., Newton K. and Welzel C. (2011). How general is trust in "most people"? Solving the radius of trust problem. *American Sociological Review* 76(5): 786-807.

Fan, R., Zhao, J., Chen, Y., & Xu, K. (2014). Anger Is More Influential than Joy: Sentiment Correlation in Weibo. PLoS ONE, 9(10), e110184.

Kim S.M. and Hovy, E. (2004). Determining the Sentiment of Opinions. In *Proceedings of the 20th international conference on Computational Linguistics*, pp. 1367-1374.

Kuklinski, James H., Michael D. Cobb, and Martin Gilens. "Racial attitudes and the "New South"." The Journal of Politics 59, no. 02 (1997): 323-349.

Kuklinski, James H., Paul M. Sniderman, Kathleen Knight, Thomas Piazza, Philip E. Tetlock, Gordon R. Lawrence, and Barbara Mellers. "Racial prejudice and attitudes toward affirmative action." American Journal of Political Science (1997): 402-419.

Le Q. and Mikolov T. (2014) Distributed representations of sentences and documents. In *Proceedings of the 31st International Conference on Machine Learning*, pp. 1188-1196.

Manning, C.D. and Sch¨utze, H. (1999) *Foundations of Statistical Natural Language Processing*.MIT Press.

McCallum, A. and Nigam, K. (1998). A comparison of event models for Naïve Bayes text classification. In *Proceedings of AAAI/ICML-98 Workshop on Learning for Text Categorization*, 41-48.

Mikolov, T., Chen, K., Corrado, G., and Dean, J. (2013) Efficient estimation of word representations in vector space. In *Proceedings of Workshop at ICLR*, pp. 1-12.

Mikolov, T., Sutskever, I., Chen, K., Corrado, G.S., and Dean, J. (2013) Distributed representations of words and phrases and their compositionality. In *Proceedings of Advances in Neural Information Processing Systems*, pp. 3111-3119.

Miller G.A. and Charles, W.G. (1991) Contextual correlates of semantic similarity. *Language and Cognitive Processes*, vol. 6, no. 1, pp. 1-28.

Mishne, G. (2005). Experiments with mood classification in blog posts. In *Proceedings of the 1st Workshop on Stylistic Analysis of Text for Information Access*.

Mnih, A. and Kavukcuoglu, K. (2013). Learning word embeddings efficiently with noise-contrastive estimation. In *Proceedings of Advances in Neural Information Processing Systems*, pp. 2265-2273.

Morin, F. and Bengio, Y. (2005). Hierarchical probabilistic neural network language model. In *Proceedings of the International Workshop on Artificial Intelligence and Statistics*, pp. 246-252.

Pang, B., Lee, L., and Vaithyanathan, S. (2002). Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing*, 79-86.

Purver M. and S. Battersby. 2012. Experimenting with distant supervision for emotion classification. In *Proceedings of the 12th Conference of the European Chapter of the Association for Computational Linguistics*, 482-491.

Qiu, L., Cao, Y., Nie, Z., Yu, Y., and Rui, Y. (2014) Learning word representation considering proximity and ambiguity. In *Proceedings of the 28th Association for the Advancement of Artificial Intelligence*, pp. 1-7.

Tang W. (2016). *Populist Authoritarianism: Chinese Political Culture and Regime Sustainability*, Chapter 4. Oxford University Press.

Thayer ,R.E. (1989). *The Biopsychology of Mood and Arousal*, Oxford University Press.

Wu, C.H., Chuang, Z.J., and Lin, Y.C. (2006). Emotion recognition from text using semantic labels and separable mixture models. *ACM Transactions on Asian Language Information Processing*, 5(2), 165–183.

Yang, C.H., Lin, K.H.Y. and Chen, H.H. (2007).Building emotion lexicon from Weblog corpora. In *Proceedings of 2007 IEEE/WIC/ACM International Conference on Web Intelligence*, 275-278.

Yang, Y.H., Liu, C. C., Chen, H. H. (2006). Music emotion classification: A fuzzy approach. In *Proceedings of the 14th Annual ACM International Conference on Multimedia*, pp. 81-84.

Yu, H. and Hatzivassiloglou, V. (2003). Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pp. 129-136.

Zou, F., Wang, F.L., Deng, X., Han, S., and Wang, L.S. (2006). Automatic construction of Chinese stop word list. In *Proceedings of the 5th WSEAS International Conference on Applied Computer Science*, 1010-1015.