Dividing for Combination: A Bootstrapping Sentiment Classification Framework for Microblogs

Songxian Xie and Ting Wang

Department of Computer Science and technology, School of Computer,
National University of Defence Technology, Changsha, China
{xsongx,tingwang}@nudt.edu.cn
http://www.nudt.edu.cn

Abstract. Sentiment classification is an important area of Natural Language Processing. There are many challenges especially for sentiment classification of the social media such as microblogs. This paper focuses on context dependence, which has been the most challenging problem, and proposes a novel framework to solve the problem. By dividing the feature space of sentiment classification into two parts, a general classifier on general part of features is trained on off-the-shelf idiom resources, and a context classifier on context-dependent part of features is trained on random tweets retrieved from microblogs in distant supervision manner. Then a semi-supervised framework is developed to combine the general classifier and context classifier into a bootstrapping classifier. Experiments results show that both the general classifier and context classifier outperform baselines, the proposed semi-supervised framework is effective and achieves encouraging performance by outperforming supervised classifier upper bound.

Keywords: sentiment analysis; idioms; general classifier; context classifier;microblogs

1 Introduction

Sentiment analysis is the computational study of how opinions, attitudes, emotions, and perspectives are expressed in language. It can provide tools and techniques for extracting subjective information from large datasets and summarizing it for successive application such as Business Intelligence, Public Opinion Analysis, and Election Prediction, etc. [1] Sentiment classification, which deals with determining sentiment orientation of target text, is the classification form of sentiment analysis [2]. Although it can be viewed as a special text categorization problem, in fact sentiment classification is a more challenging task than text classification, because sentiment expressions critically depend on domains and context [3].

With the emergence and prosperity of microblogs, the amount of user-generated content(UGC) has risen exponentially over the last decade, and such content is now always at our fingertips. It has been well recognized that UGC of microblogs with rich sentiment information can trigger more attention, feedback or participation, and sentiment analysis researchers have begun to pay more and more attention to microblogs [4]. The distillation of subjective knowledge from such abundant language resources becomes an important part of applications in fields such as commerce, tourism, politics and health. But as a freely-publishing platform, users may use or create new words of abbreviations and acronyms that seldom appear in conventional text documents, for example, words like "cooooooool", OMG",":-(", are intuitive and popular in tweets. Although they may provide convenience for on-line communications of users, it is difficult for computer to accurately identify the semantic meanings of these words. To make situation worse, new words may arise and old words may change their meaning continuously in tweets. The noisy quantity, informal nature and explosive vocabulary make sentiment classification of tweets a very difficult and challenging task. The sentiment in tweets often depends on such particular expressions as emoticons, repeated letters and exclamation, etc., which resolve their semantic meanings only in the context of microblogs. Context dependence has been the main challenge sentiment classification of tweets must face. In this work, we have focused on the context-dependent problem of sentiment classification on microblogs.

For the problem, we propose a novel semi-supervised framework based on our two rational hypotheses. Firstly the problem is formulated as feature vector space model of text classification. And the feature space is divided into two parts consisted of general part and context part. Then a general classifier is trained on off-the-shelf labelled idiom resources by making use of general part of features, and a context classifier is trained on tweets randomly retrieved from microblogs in distant supervision manner by making use of context-dependent part of features. As a result, a semi-supervised framework is developed to combine the general classifier and context classifier into a bootstrapping classifier to harmonize with the two parts of features.

The rest of the paper is organized as follows: Related works are discussed in Section 2. The problem is formulated in Section 3, and our framework of semi-supervised sentiment classification is described in Section 4. The results and discussions of the experiment are presented in Section 5. Finally we conclude about our work in Section 6.

2 Related Works

Sentiment analysis has been a popular research area for years. Previous efforts mainly focuses on reviews and news comments. Generally, in terms of methodology, rule-based approaches and machine-learning based approaches are two major popular methods, and the machine-learning based approaches usually act as an upper bound for other methods to compare with because of the strong generalization ability of classifiers [1, 5].

Recently, sentiment analysis researchers have begun to pay more and more attention to the massive user-generated content of social networks such as Twitter[1]. Many studies showed that the unique characteristics of Twitter can be incorporated into sentiment analysis techniques. Barbosa and Feng [6] first investigated to use a two-stage Support Vector Machine (SVM) classifier for tweets sentiment classification which proved to be robust regarding biased and noisy data. Hu et al. [7] interpreted emotional signals available in social media data for unsupervised sentiment analysis by providing a unified way to model two main categories of emotional signals: emotion indication and emotion correlation. Jiang et al. [3] focused on target-dependent Twitter sentiment classification, they proposed to improve target-dependent Twitter sentiment classification by taking both target-dependent features and related tweets into consideration. Wang et al. [8] focused their study on hashtag-level sentiment classification, they proposed a novel graph model and further improved the model using an enhanced boosting classification setting. Amir Asiaee T et al. [9] presented a cascaded classifier framework for per-tweet sentiment analysis by extracting tweets about a desired target subject, separating tweets with sentiment, and setting apart positive from negative tweets. Hu et al. [10] extracted sentiment relations between tweets based on social theories, and proposed a novel sociological approach to utilize sentiment relations between messages to facilitate sentiment classification and effectively handle noisy Twitter data. Motivated by sociological theories arguing that humans tend to have consistently biased opinions, Guerra et al. [11] addressed challenges of topic-based real-time sentiment analysis by proposing a novel transfer learning approach with a suitable source task of opinion holder bias prediction. Thelwall et al. [12, 13] designed SentiStrength, an algorithm for extracting sentiment strength from informal English text. The algorithm built on human-evaluated dictionaries for words connotated with positive or negative sentiments and exploited the grammar and spelling styles in typical microblogs.

All works above have tried to adapted their methods to microblogs by making use of the network and language characteristics, no matter what approaches they have taken. However, in this paper, we solve the context-dependent problem from a novel perspective by separating context-independent part of features from context-dependent part of features in the feature space of sentiment classification.

3 Problem Formulation and Hypotheses

3.1 Problem Formulation

Simply speaking, sentiment classification aims to classify text as predefined sentiment polarity classes (negative or positive). Formally, Given document corpus $D = \{d_1, \ldots, d_n\}$, and predefined sentiment classes $C = \{1, -1 \mid \text{positive} = 1, \text{negative} = -1\}$, the task of sentiment classification is to predict each d_i with

[1] `twitter.com`

a label c_i. To be along with text categorization, each document can be represented as a vector of features $x = \mathbb{R}^n$, where n is the size of a pre-specified feature volume V. For sentiment classification, the weight of each entry in the vector usually is often specified as binary, with weight equals to 1 for feature present in the vector and 0 for absent. Given a training dataset $X = \{x_1, \ldots, x_m\}$, a classifier can be build:

$$f : X \longrightarrow Y, Y = \{1, -1\} \ . \tag{1}$$

and it is employed to predict label for an unseen instance x by computing $f(x)$, with each instance represent as a vector $x = (w_1, \ldots, w_v)$, in which w_i is the ith features weight.

3.2 Feature Space Division

In previous sentiment classification researches, there is an underlying hypothesis, which implies all features in the text vector represent the document's sentiment polarity equally. In fact, some features may only appear in specific context, while others appear across any context. As an example, an English tweet is listed as following:

@Kid_Cloudz: Happy birthday to Yessicaaaa! :D lovee you feggit wish you the best day everrrrr!!!!! @030268.

If simple "bag-of-words" features are considered, all words would be extracted as equal features to model the sentiment when computationally classifying sentiment polarity of this tweet. However with thorough considerations, we can find that such words as "@Kid_Cloudz, :D, lovee, everrrrr,!!!!!" are mainly used in the language context of microblogs, while "Happy, birthday, wish, best, thanks" are sentiment indicators of text almost in any language context. With such intuition, we proposed a feature space division hypothesis as:

Hypothesis 1. In the feature space of sentiment classification, features can be divided into two different parts:

- context-independent part, i.e. general features, which are indicators of sentiment independent of any language context;
- context-dependent part, i.e. context features, which indicate sentiment in specific language context.

Formally, for the text feature vector of a tweet $x = (w_1, \ldots, w_l, w_{l+1}, \ldots, w_v)$, it could be divided into two parts:

$$x = \begin{cases} x_g & : \text{general features} \\ x_c & : \text{context features} \end{cases} \ . \tag{2}$$

where $x_g = (w_1, \ldots, w_l)$ denotes the general part of feature space, while $x_c = (w_{l+1}, \ldots, w_v)$ denotes the context dependent part.

4 Sentiment Classification Framework

Besides traditional expressions, some languages of microblogs are often considered "Mars Language" because of their obscurity. However users could still distinguish the sentiment polarity of the tweet even if some words seem strange to them. Intuitively, this kind of phenomenon may be explained by the general part of the tweet which is used to express the holistic sentiment polarity across any context, and the polarity of general sentiment words are prone to be recognized by anyone independent of context knowledge. Comparably, for sentiment classification, we put forward that, the sentiment polarity of a document could sometimes still be recognized with only general part of feature space (x_g in Equation 2). That is to say, if general sentiment knowledge could be modeled, what sentimental polarity a tweet prefers for could still be classified based on such general sentiment models.

In fact, many researchers have tried to establish all kinds of sentimental ontology lexicons to represent general knowledge of humans sentiment, such as SentiWordNet [14] and General Inquiry [15] in English, HowNet[2] [16] and NTUSD [17] in Chinese, etc. However, some entries of these lexical resources have multiple senses with different sense representing different sentimental polarity, and the exact sentiment of each sense unavoidably depends on the context. It seems a rather difficult task to find a general resources to model the context independent knowledge. Actually such knowledge exists in many situations, and many regular language expressions in the form of one general phrase or combination of a few words could identify exact sentimental polarity independent of language context, such as idioms and proverbs.

4.1 General Sentiment Classifier

There are many linguistic resources highly valuable for sentiment classification, of which idiom resources attract interests of this research. Idioms are common phenomena of many languages beside Chinese, such as "castles in the air", "a bed of thorns", and "bring down the house" in English. As is convinced, the sentimental polarity of idioms is independent and unchangeable under any context. There are many off-the-shelf idiom lexicons in all kinds of languages with entries taking the example form as:

castles in the air: a derogatory term, indicate the illusive things or impractical fanciness metaphorically.

In this example, the entry is composed of three parts: the idiom "castles in the air", the semantic orientation "a derogatory term" representing negative sentimental polarity and a short paraphrase with three general negative words ("illusive, impractical and fanciness"). The example provides us with a labelled sentimental instance with general sentiment features and a negative label. Most importantly, the sentimental polarity of such instance is independent of any context just as the idiom it explains. Based on such observation, another hypothesis is proposed as follows:

[2] http://www.keenage.com

Hypothesis 2. The sentimental polarity of the idiom paraphrase is independent of language context just as the idiom it describes.

With Hypothesis 2 admitted, we could constructed a training dataset with the features extracted from paraphrases of idioms as general feature vector, the sentimental polarity of idioms as sentimental labels. Then a context-independent classifier could be trained to model the general sentiment knowledge.

4.2 Context Sentiment Classifier

As the general sentiment features are only one part of all features in the whole feature space, the other context-dependent part of features should be considered in order to capture the subtle clues embedded in the specific sentiment expressions in language context of microblogs.

To model the context-dependent part of tweets, there are two questions must be solved. The first is to identify the context-dependent part of features extracted from tweets. In fact, new expressions appears on microblogs with the explosively increasing of UGC, which makes it rather difficult to clearly tell whether each word is context dependent or not. However tweets are conventionally short with the limitation of 140 characters, and users often express one particular sentiment in one tweet with a few words. Based on the particular characteristics of tweets, we make an assumption that if a tweet contains idioms, sentimental polarity of words in a tweet are often context-dependent except for the idioms. The second is how to find labelled instances to train the context-dependent classifier. Some researchers have proposed distant supervision to solve the training data shortage of Twitter [18, 19]. In this paper we establish our training dataset in a similar way. We retrieve tweets on microblogs and try to get as many tweets as possible that contains idioms. By stripping off idioms, we extracted context-dependent features from left words, and took the sentimental polarity of idioms as labels. So we get our noisy labelled dataset and a context-dependent classifier is trained to model the context-dependent knowledge.

4.3 Combination of Two Classifiers

Although theoretically the general classifier and context classifier could be able to model different sentiment knowledge separately and classify the sentiment of a tweet accurately to some extent, the coverage and efficiency of them are limited by the quality and quantity of training datasets. Besides, it is obvious that the paraphrase of idiom and tweet segments (lefts by stripping off idioms) is usually very short, so the feature vectors of datasets must be very sparse, which degrades the performance of the classifiers. For above reasons, a consistent bootstrapping machine learning framework is chosen to combine the two classifiers together. The framework is illustrated in Figure 1. As illustrated in the framework, in an iterating manner, the general classifier P_g and context classifier P_c are applied to test dataset so that the every test instance x_i is predicted labels $c_i = \{c_g, c_c\}$

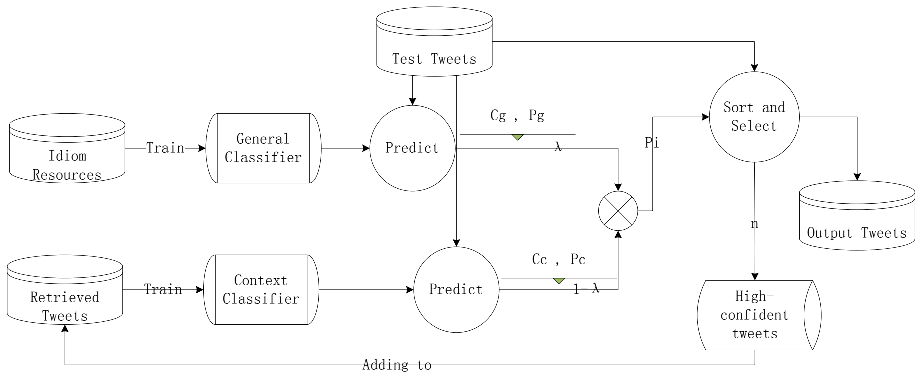

Fig. 1. The Bootstrapping Sentiment classification Framework

initially, with confidence $p_i = \{p_g, p_c\}$. Then a combined confidence score is calculated with Equation 3:

$$p_i = \begin{cases} \lambda * p_g + (1-\lambda) * p_c & \text{if} \quad c_g = c_c; \\ 0 & \text{if} \quad c_g \neq c_c; \end{cases} \quad . \tag{3}$$

where λ is the coefficient to control impacting weights of different part of features. We initialize $\lambda = 0.5$ with equal weights of general part and context part of features, and to make combined classifier more adaptable for microblogs, we increase the weight of context part step by step with the bootstrapping iteration progressing. The test dataset initially labelled as c_i $(c_i \in \{1, -1\})$ was sorted with descending confidence p_i $(p_i \neq 0)$ for two sentimental categories $C = \{1, -1\}$ separately. The n positive and negative instances with top high confidence are selected as new instances to add into the training dataset and improve the context classifier to a more context-aware classifier. Such a procedure iterates until convergence. The output of such semi-supervised sentiment classification framework is the predicted results of test dataset. Above all, the whole framework can combine the two classifiers which constructed on divided feature space into a stronger classifier.

4.4 Classifier Description

We adopt the same methods as Pang et al. [5], in that they have applied Naïve Bayes, Maximum Entropy and Support Vector Machine classifiers to identify the effectiveness of machine learning techniques on sentiment classification, and they got satisfying result (accuracy of 82.9%).

Naïve Bayes Classifier. Nave Bayesian classifier is the most popular text classification techniques. For sentiment classification problem as formulated in

Section 3.1, to determine which sentimental category c_j a document d_i belongs to, it is needed to compute the posterior probability $P(c_j \mid d_i)$. According to the Bayesian probability and the multinomial model, based on the hypothesis that the probabilities of features $w_{d_i,k}$ are independent given document class, Equation 4 is got:

$$P(c_j \mid d_i) = \frac{P(c_j) \prod_{k=1}^{|d_i|} P(w_{d_i,k} \mid c_j)}{\sum_{r=1}^{|C|} P(c_r) \prod_{k=1}^{|d_i|} P(w_{d_i,k} \mid c_r)} \quad . \tag{4}$$

The class with the highest probability is assigned as the sentimental category of the document d_i.

Maximum Entropy Classifier. The Maximum Entropy classifier assigns the class with the higher conditional probability to the sentimental category of document d_i as follows:

$$P\left(c_j \mid d_i, \overrightarrow{\theta}\right) = \frac{1}{Z} exp\left(\overrightarrow{\theta}, \overrightarrow{f}(d_i, c_j)\right) \quad . \tag{5}$$

where $\overrightarrow{\theta}$ denotes the vector of feature weights, $\overrightarrow{f}(d_i, c_j)$ denotes feature function that maps pair (d_i, c_j) to a feature vector, and Z denotes normalization factor. With labelled dataset D, the training procedure is trying to solve such an optimization problem as:

$$\overrightarrow{\theta^*} = argmax_{\overrightarrow{\theta}} \prod_{i=1}^{|D|} P\left(c_j \mid d_i, \overrightarrow{\theta}\right) \quad . \tag{6}$$

Support Vector Machine Classifier. Support Vector Machines classifier (SVM), is a kind of discriminative method of machine learning techniques. SVM tries to find a decision surface to separate the training data into two classes and makes decisions based on support vectors. In this research, linear SVM has been adopted due to its popularity and sound performance in sentiment classification task. The training of SVM is trying to minimize an constraint optimization problem:

$$\begin{gathered} \overrightarrow{\alpha^*} = argmin\left(-\sum_{i=1}^{n} \alpha_i + \sum_{i=1}^{n}\sum_{j=1}^{n} \alpha_i \alpha j x_i x_j < \overrightarrow{x_i}, \overrightarrow{x_j} >\right) \\ \text{Subject to:} \sum_{i=1}^{n} \alpha_i y_i = 0, 0 \leqslant \alpha_i \leqslant 1 \end{gathered} \quad . \tag{7}$$

5 Experiment

In this section, the proposed semi-supervised sentiment classification framework is verified in an empirical test. The test has been carried out with Chinese dataset constructed from an off-the-shelf on-line idiom dictionary and microblog platform Tencent[3].

5.1 Experiment Description

Dataset: We crawled the on-line idiom dictionary from China Education Network[4] and got an idiom dataset of 8,160 instances labelled with positive (3,648 instances) and negative (4,512 instances) sentiment, which was used to train the general classifier. From Apr.15th,2013 to May 1st,2013, we monitored Tencent time-line tweet streams, retrieved and extracted the tweets that contained at least one idiom in our idiom dataset, resulting in about 120,346 tweets. After stripping off idioms from all tweets and removing tweets with characters less than 4, we got a dataset of 91,268 instances which was used to train context classifier. The dataset from the First Chinese tweet Sentiment Analysis and Semantic Relationship Extraction Evaluation of CCF Natural Language Processing and Chinese Computing[5] was used to evaluate performance of our framework.

Classifiers and performance measurement: There are various complicated measurements to evaluate the performance of computational algorithm, of which the simplest accuracy index was chosen to evaluate the performance of our framework, because the comparison between measurements was not the important points of our research. As for classifiers, Naïve Bayes classifier and Maximum Entropy classifier of NLTK (Natural Language ToolKits) [20] package and Support Vector Machine classifier of Libsvm [21] package were used for classification. All the parameters and settings were optimized by cross-validation.

Baseline and upper bound: Two baselines were used to compare with the proposed framework, the first one was naïve 50% baseline since the test dataset was balanced with respect to the sentiment classes, the other one was the traditional lexicon-based classifier by comparing the number of positive words and negative words in HowNet sentiment lexicon for a tweet to determine its sentiment class. As mentioned in Section 2, supervised machine learning methods are often used as upper bound to be compared by other methods. In the experiments, an upper bound was also set up with test dataset in the five-folded cross-validation manner.

Preprocessing: Text written in Chinese are not well formatted in that words in a sentence are not separated by space as English. All the text in Chinese must be segmented before bag-of-words features being extracted. In the experiment, all text of training and testing dataset were segmented with Chinese segmentation software ICTCLAS[6].

[3] `http://t.qq.com/`
[4] `http://chengyu.teachercn.com`
[5] `http://tcci.ccf.org.cn/conference/2012/pages/page04_eva.html`
[6] `http://ictclas.nlpir.org/`

5.2 Result and Analysis

To determine the values of λ in Equation 3, we carried out traversal experiment by varying the value of λ from 0 to 1 with step of 0.1. The result is illustrated in Figure 2. The values of λ for each combined bootstrapping classifier based on

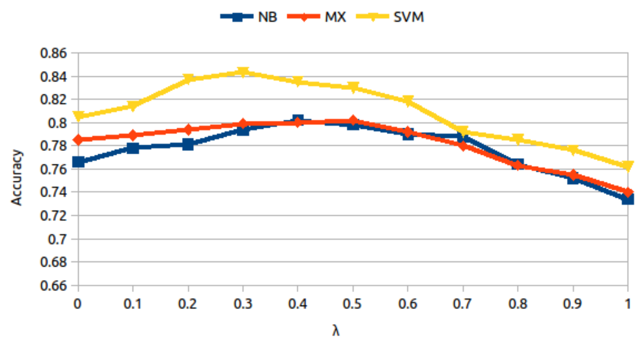

Fig. 2. The Impact of λ on Different Classifiers: NB denotes Naïve Bayes, MX denotes Maximum Entropy and SVM denotes Support Vector Machine.

three different classifiers are fixed as the accuracy reaches its peak, which are: 0.4 for Naïve Bayes classifier, 0.5 for Maximum Entropy classifier, and 0.3 for Support Vector Machine classifier.

The final sentiment classification results are shown in Table 1.

Table 1. Results for Different Classifiers

Classifier	NB	MX	SVM
Lexicon Classifier	0.725	0.725	0.725
Supervised Classifier	0.785	**0.806**	0.826
General Classifier	0.734	0.740	0.762
Context Classifier	0.766	0.785	0.805
Combined Classifier	**0.802**	0.802	**0.843**

Firstly, the accuracies of general classifier and context classifier all surpass the naïve baseline (50%), which proves that they are superior to random selection and may be better choice when there are no labelled dataset available for supervised or semi-supervised machine-learning sentiment classification.

Secondly, the accuracy of general classifier is a little higher than the traditional lexicon-based classifier, in that although they can both model the general sentiment knowledge, the general classifier is trained on context-independent part of features, so it can better catch the holistic sentimental polarity of tweets, while the sentiment lexicon often contains ambiguous entries. As for the context classifier, the performance outperforms the traditional lexicon-based classifier and general classifier, because it has been trained on context-aware part of features, and the users are more willing to express their sentiment with “Mars Language” of microblogs, so the context classifier is more adaptable for microblog context.

Finally, the combined classifier shows the best performance by combining general classifier and context classifier. It even outperforms the upper bound supervised classifier, which proves the effectiveness our proposed framework because it can not only catch the holistic sentimental polarity of tweets by modelling general sentiment knowledge but also adapt to the microblog language context by considering the subtle sentiment expressions articulated in tweets.

6 Conclusion

Context-dependent problem has always been a main challenge of sentiment classification. In this paper, we have proposed a novel semi-supervised framework to get it solved in the microblog language context. From a different perspective of feature space, we put forward the assumption that feature space can be divided into the general part and the context part. To make use of two parts of features, two classifiers are trained on dataset constructed from idiom resources and tweets separately. Our framework combines the classifiers with a semi-supervised bootstrapping learning algorithm. The experiment results show that the proposed framework could outperform the state-of-art supervised classifier. In future, we will try to improve the sentiment classification performance by enlarging the context-independent resources and extracting richer features besides bag-of-words feature.

References

1. Liu, B.: Sentiment Analysis and Opinion Mining. Synthesis Lectures on Human Language Technologies. Morgan & Claypool Publishers (2012)
2. Pang, B., Lee, L.: Opinion mining and sentiment analysis. Foundations and Trends in Information Retrieval **2**(1-2) (2007) 1–135
3. Jiang, L., Yu, M., Zhou, M., Liu, X., Zhao, T.: Target-dependent Twitter sentiment classification. In: Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies - Volume 1. HLT ’11, Stroudsburg, PA, USA, Association for Computational Linguistics (2011) 151–160
4. Stieglitz, S., Dang-Xuan, L.: Political communication and influence through microblogging-an empirical analysis of sentiment in twitter messages and retweet behavior. In: HICSS, IEEE Computer Society (2012) 3500–3509

5. Pang, B., Lee, L., Vaithyanathan, S.: Thumbs up?: sentiment classification using machine learning techniques. In: Proceedings of the ACL-02 conference on Empirical methods in natural language processing - Volume 10. EMNLP '02, Stroudsburg, PA, USA, Association for Computational Linguistics (2002) 79–86
6. Barbosa, L., Feng, J.: Robust sentiment detection on twitter from biased and noisy data. In: Proceedings of the 23rd International Conference on Computational Linguistics: Posters. COLING '10, Stroudsburg, PA, USA, Association for Computational Linguistics (2010) 36–44
7. Hu, X., Tang, J., Gao, H., Liu, H.: Unsupervised sentiment analysis with emotional signals. In: Proceedings of the 22nd international conference on World Wide Web. WWW '13, Republic and Canton of Geneva, Switzerland, International World Wide Web Conferences Steering Committee (2013) 607–618
8. Wang, X., Wei, F., Liu, X., Zhou, M., Zhang, M.: Topic sentiment analysis in twitter: a graph-based hashtag sentiment classification approach. In: Proceedings of the 20th ACM international conference on Information and knowledge management. CIKM '11, New York, NY, USA, ACM (2011) 1031–1040
9. Asiaee T., A., Tepper, M., Banerjee, A., Sapiro, G.: If you are happy and you know it... tweet. In: Proceedings of the 21st ACM international conference on Information and knowledge management. CIKM '12, New York, NY, USA, ACM (2012) 1602–1606
10. Hu, X., Tang, L., Tang, J., Liu, H.: Exploiting social relations for sentiment analysis in microblogging. In: Proceedings of the sixth ACM international conference on Web search and data mining. WSDM '13, New York, NY, USA, ACM (2013) 537–546
11. Calais Guerra, P.H., Veloso, A., Meira, Jr., W., Almeida, V.: From bias to opinion: a transfer-learning approach to real-time sentiment analysis. In: Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining. KDD '11, New York, NY, USA, ACM (2011) 150–158
12. Thelwall, M., Buckley, K., Paltoglou, G., Cai, D., Kappas, A.: Sentiment in short strength detection informal text. J. Am. Soc. Inf. Sci. Technol. **61**(12) (December 2010) 2544–2558
13. Thelwall, M., Buckley, K., Paltoglou, G.: Sentiment strength detection for the social web. J. Am. Soc. Inf. Sci. Technol. **63**(1) (January 2012) 163–173
14. Baccianella, A.E.S., Sebastiani, F.: Sentiwordnet 3.0: An enhanced lexical resource for sentiment analysis and opinion mining. In: Proceedings of the Seventh conference on International Language Resources and Evaluation (LREC'10), Valletta, Malta, European Language Resources Association (ELRA) (May 2010)
15. Stone, P.J., Dunphy, D.C., Smith, M.S., Ogilvie, D.M.: The General Inquirer: A Computer Approach to Content Analysis. MIT Press (1966)
16. Zhendong, D.: Hownet. (1999) `http://www.keenage.com.`
17. Ku, L.W., Chen, H.H.: Mining opinions from the web: Beyond relevance retrieval. J. Am. Soc. Inf. Sci. Technol. **58**(12) (October 2007) 1838–1850
18. Go, A., Bhayani, R., Huang, L.: Twitter Sentiment Classification using Distant Supervision. Technical report, Stanford University
19. Marchetti-Bowick, M., Chambers, N.: Learning for microblogs with distant supervision: Political forecasting with twitter. In Daelemans, W., Lapata, M., Mrquez, L., eds.: EACL, The Association for Computer Linguistics (2012) 603–612
20. Loper, E., Bird, S.: NLTK: The Natural Language Toolkit (May 2002)
21. Chang, C.C., Lin, C.J.: LIBSVM: A library for support vector machines. ACM Transactions on Intelligent Systems and Technology **2** (2011) 27:1–27:27 Software available at `http://www.csie.ntu.edu.tw/~cjlin/libsvm`.