



## A survey on sentiment classification algorithms, challenges and applications

Muhammad Rizwan Rashid RANA

University Institute of Information Technology  
Pir Mehr Ali Shah Arid Agriculture University  
Rawalpindi, Pakistan  
email: rizwanrana315@gmail.com

Asif NAWAZ

University Institute of Information  
Technology  
Pir Mehr Ali Shah Arid Agriculture  
University  
Rawalpindi, Pakistan  
email: asif.nawaz@uaar.edu.pk

Javed IQBAL

Department of Computer Science  
University of Engineering and  
Technology  
Taxila, Pakistan  
email: javed1797@hotmail.com

**Abstract.** Sentiment classification is the process of exploring sentiments, emotions, ideas and thoughts in the sentences which are expressed by the people. Sentiment classification allows us to judge the sentiments and feelings of the peoples by analyzing their reviews, social media comments etc. about all the aspects. Machine learning techniques and Lexicon based techniques are being mostly used in sentiment classification to predict sentiments from customers reviews and comments. Machine learning techniques includes several learning algorithms to judge the sentiments i.e Navie bayes, support vector machines etc whereas Lexicon Based techniques includes SentiWordnet, Wordnet etc. The main target of this survey is to give nearly full image of sentiment classification techniques. Survey paper provides the comprehensive overview of recent and past research on sentiment classification and provides excellent research queries and approaches for future aspects

**Computing Classification System 1998:** I.2.6, K.3.0, K.3.m

**Mathematics Subject Classification 2010:** 68R15

**Key words and phrases:** sentiment classification, supervised learning, opinion mining, machine learning, lexicon approaches, unsupervised learning

## 1 Introduction

The fast growth of World Wide Web (WWW) is constantly increasing the online communication. A huge number of customers reviews or suggestions on everything are present on the web and these customers reviews and suggestions are increasing day by day. Micro blogging like twitter, Facebook etc. are powerful channels for peoples sentiments, thoughts, ideas and feelings. It is now estimated that, in just 60 seconds, over 400,000 Twitter posts are being shared, about 300,000 Facebook statuses updates, about 25,000 items purchased from Amazon, over 5million YouTube videos viewed and about 2.7 million Google searches are being made among many other things [1]. Users rating and reviews, which have been found on many ecommerce and market websites is a good source that helps peoples to build opinion about specific products. As a result of this phenomenon, increasing numbers of opinions and thoughts are being spread and published over the internet

Before the rise of internet to answer the question of What do people think about any product or anything else, Surveys and polls are distributed in the form of paper in peoples. With the expeditious development of internet and the popularity of Micro-blogging sites like Facebook, Twitter etc enables an alternative option for getting opinions from large population [38]. Now a days Web become the necessity for people to share their ideas, experiences and opinions as well as seeking others experiences and opinions [47]. Millions of ideas and experiences are shared every day. It is impossible for peoples to read all ideas and experiences. About 2.7 million Google searches was made. A query Artificial Intelligence returns 98,400,000 results. This whole scenario demands fast, effective and accurate technique to track sentiments, opinions and ideas that are flowing on internet. Sentiment classification is the key component of such techniques [4, 41, 11].

Sentiments represent the viewpoint of customer such as like (positive), dislike (negative) and may be neutral viewpoint [7]. Sentiment classification also called opinion mining is the process to automatically determine the sentiment category to which the textual content belongs [46]. We can categorize the reviews, comments and document mainly in two types. These are numeric sentiment and categorical sentiment. Common example of a numeric sentiment is rating system in ecommerce sites. Using this rating system company judge the response of peoples. Categorical sentiment is a method to classify the comments or reviews in different categories. These categories are binary (positive and negative), ternary (positive, negative and neutral) and multiple categories (Anger, happy, sad etc.) [40]. Opinions are become the necessary

parts in all human activities. There are two types of reviewing sites, generic reviewing sites and specified reviewing sites. Generic reviewing sites include sites like amazon.com, epinions.com, rottentomatoes.com etc and Specified reviewing sites includes tripadvisor.com, yelp.com. Both of these reviewing sites have a significant effect in our decision meaning process. These decisions includes buying a camera, smart phone etc, making investments on any product etc, choosing schools, decision about any movie etc etc. Before the Internet, Other sources such as friends, relatives etc affect the human decision process. Positive review is showing in Figure 1 and negative review is showing in Figure 2.

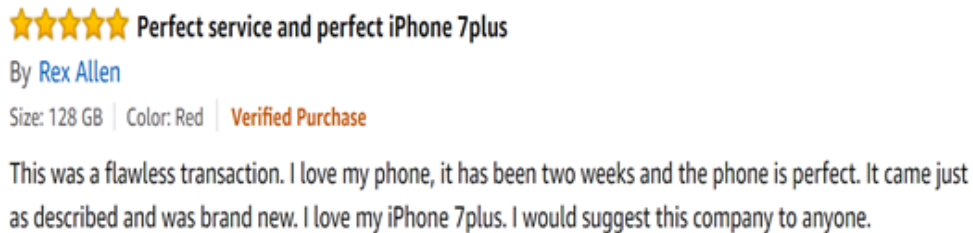


Figure 1: Positive review

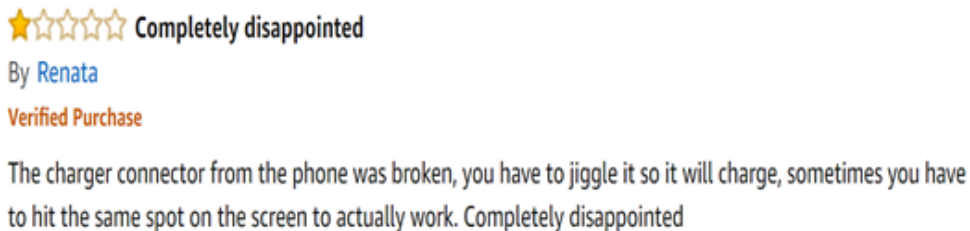


Figure 2: Negative review

The meaning of word sentiment itself is still very wide. Opinion mining mostly focuses on opinions which communicate or involve positive or negative sentiments from reviews, comments etc. These user reviews are very useful to organizations for making intelligent decisions about product purchasing and also helpful for merchants in knowing their products progress in the market

[29]. Nowadays everyone shares their views and experiences online. For example, if somebody wants to buy a mobile phone, such as the Samsung mobile, and he or she don't know about this mobile phone. He or she can use the internet, open mobile phone web site and read customer reviews about the product and then he or she can make a decision in the light of provided user reviews. This manual process is named as text mining, opinion mining or sentiment analysis. History calculation sentiment of document is the task of marketing team of any company. Humans have no trouble in reading the movie review, product review and any political comment and categorized it in positive, negative or neutral class. We humans use a technique of reading and understand the underlying meaning of a sentence but when there are a millions of reviews then its a time consuming task to read all reviews and categorized it in positive, negative or neutral class.

In the wake of digital age, thousands of movies are directed per year and peoples share millions of reviews and comments about these movies on the internet. Usually a movie reviews peoples share their comments about the movie. As it is the time consuming task so it is very hard for humans to judge the tone of these reviews and classify it in positive or negative category. There is a strong demand for automatically analyzing and summarizing the opinions expressed in natural language text. We can automatically analyze and classify the reviews using machine learning techniques. Sentiment classification can be helpful for customers who need to research the sentiment of product before purchase or companies that need to watch the general public sentiment of their brands.

Sentiment classification or opinion mining has been studied at three different levels of classification [28] these include sentiment classification at document level, sentence level and at aspect level. Classification of whole document in a positive or negative is called document level classification. Sentence-level sentiment classification techniques read document sentence by sentence and decide whether each sentence gives a positive, negative or neutral opinion for a service, product etc. Aspect level or entity level sentiment classification is the most modern technique which classifies the reviews or comments on the basis of aspects or entities

## **2 Literature review**

In the last decade, lot of research work has been carried out in sentiment classification. Techniques of sentiment classification (i.e judging tone of the

text) have been performed for a variety of applications over a wide range of classification algorithms. A Review of some existing techniques from literature is provided in the following section. There exists a four different techniques for sentiment classification as shown in Figure 3.

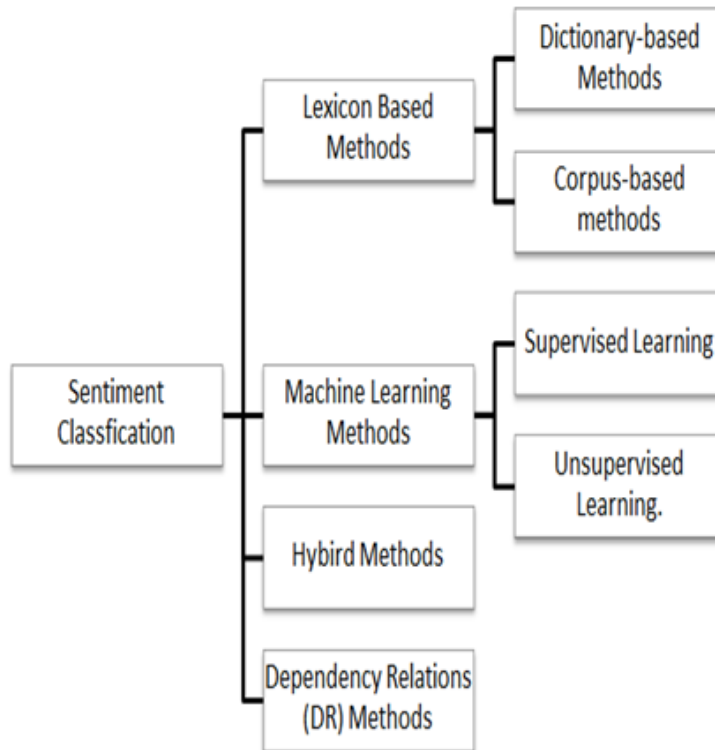


Figure 3: Sentiment classification techniques

## 2.1 Lexicon based methods

Lexicon based methods adopt a lexicon to perform aspect based sentiment analysis. These methods can work by counting, analyzing and weighting opinion words. These methods are further divided in two broad classes. They are dictionary-based methods and corpus-based methods

### **2.1.1 Dictionary based methods**

Dictionary-based methods used the lexicon database to judge the tone of text. Popular lexicon databases are WordNet, Sentiwordnet etc. WordNet groups the words into synsets (synonym sets) and the semantic relation between synsets [32]. WordNet is used on adjectives in order to find their semantic orientation [20]. Process first count the number of synonym links for adjectives such as bad,good etc. Another paper used the WordNet to create semantic lexicon. They used the antonym relation of adjective and WordNet synonym [10]. This idea is used for constructing another lexicon named as SentiWordNet. SentiWordNet provides the three types of sentiment scores of each words. These types are positive score, negative score and objective score. Another paper uses the three different lexical relations in WordNet [2]. These lexical relations consists on antonymy, hyponymy and synonymy). It takes the adjective from epinions.com reviews and mapped to the star rating. Paper proposed method uses the breadth first search adjective on WordNet synonymy graph with unknown sentiment and then distance-weighted nearest neighbor algorithm is to calculate the weights of two average rating of two nearest neighbor as related adjective. Different bootstrapping method using WordNet is proposed in [43]. Algorithms take the known sentiment orientations as input and generate the set of synonyms (Synsets) as output. The new generated synsets are then used to calculate the polarities of words. Constrained symmetric nonnegative matrix factorization (CSNMF) technique with sentiment lexicon generation is used in [14]. Proposed method words on two steps. In first step dictionary is used to find the candidate sentiment and in second step corpus is used to assign the polarity score to each word.

### **2.1.2 Corpus based methods**

One of the earliest ideas that use the Corpus-based method was presented by Hazivassiloglou and McKeown [13]. Proposed idea used the seed adjective and corpus to find other sentiment adjectives in corpus. This technique also used some linguistic rules. One of the rule is about conjunction AND. According to this rule where ever conjoined adjective comes, they have same orientation. For example if there is a sentence Today I am happy and delighted Here if happy is a positive then definitely delighted is also a positive. Basic idea behind conjunction AND rule is peoples always express same sentiments on both side of AND. Other rules are OR, BUT, NEITHERNOR etc.

Above approach was extended by introducing call coherency [21]. Call coherency includes intra-sentential consistency and inter-sentential consistency. Intra-sentential consistency exists within sentence and inter-sentential consistencies exist between neighboring sentences. This technique is used to find the domain dependent sentiment words. Later this technique was also used in [19]. There are many words in same domain have different orientations in different way of writing [9]. Same word is used as positive in one context and in second context it will be used as negative. For example in mobile domain lets take two reviews. First review is Mobile have long battery life and second review It takes long time to open contacts. In first review long is used in positive sense and in second review long is used in negative sense. Author presents the solution to solve this problem. First find the aspects and sentiment words or opinion words from text then use both aspects and sentiment word in pair like (aspect, sentiment word). For example (contacts, long). To predict which pair is positive and which is negative, the call coherency will also be used.

In 2011, authors argue the technique to study the lengthening of words (e.g. thanksssssssss) in social media sites [5]. Usually in comments and tweets are many lengthy words present. According to authors these lengthy words show the high sentiments in comments and proposed an automatic technique for finding the sentiments. Connotation lexicon is very much changed from simple sentiment lexicon [12]. Using Connotation lexicon paper achieved the better results as compared to other lexicons.

## **2.2 Machine learning**

Machine learning is further divided into Supervised Learning and Unsupervised Learning. In supervised learning, output dataset is necessary, we train algorithm on output dataset and get the desired outputs whereas in unsupervised learning we don't have any output datasets, instead the data is clustered into different classes [45].

### **2.2.1 Supervised learning**

Pang et al. applied Support Vector Machine, Naive Bayes and Maximum entropy with different feature extracting techniques on movie reviews [34]. Experimental results show that SVM has the best performance with unigram text representation. It has been noted that without POS tagging information accuracy of naive Bayes and maximum entropy increases but it decreases the performance of SVM. Liu et al. argued the sentiment classification system

that uses Nave Bayes Classifier and Map Reduce framework [26]. Paper uses machine leaning algorithm Nave Bayes Classifier to classify the sentiments in two classes positive and negative. Paper also uses Map Reduce framework with Naive Bayes Classifier to get better results. Map Reduce framework usually use to analyze extremely large datasets such as tweets collections, movie reviews etc. Experimental results show the accuracy of Naive Bayes classifier on large data sets is 82%.

Dhande and Patnaik uses neural network with Naive Bayes classifier because in many complex real world situations Naive Bayes cannot work well [8]. An experimental result shows that, when we are using Naive Bayes Classifier then Accuracy will be 62.35. So for getting better results Paper uses the Neural Network with Naive Bayes Classifier. Results show that accuracy of sentimental analysis increased up to 80.65 by combining Neural Network with Naive Bayes Classifier. Shaziya et al. takes the dataset of movie reviews and apply two well knows classifier on this dataset. These classifiers are Naive Bayes classifier (NB) and Support Vector Machine (SVM) [37]. The dataset is preprocessed and various filters have been applied to reduce the feature set. Papers use the Feature selection method for getting most valuable words and use Information Gain, and Gain Ratio methods for getting distinctive word. Using these methods we get the most value data from dataset. An experimental result shows that Navies Bayes results are better than SVM results. Accuracy of Naive Bayes classifier is 86.1% for positive reviews and 84.1% for negative reviews and accuracy for SVM classifier is 84.7% for positive reviews and 84.4% for negative reviews.Support Vector Machine (SVM) is more efficient classifier than Naive Bayes in many cases.

Manek et al. adopted SVM classifier with Gini Index feature selection method for sentiment classification for large movie review data set [28]. Experimental results show that Gini Index feature selection method is better in terms of accuracy and performance. Paper achieves the accuracy 78% using SVM classifier on large dataset of movie reviews. Paper [15] implement the three sentiment analysis algorithms for identifying the sentiments (positive or negative) from reviews. Experiential results are then compared with the numerical ratings of hotels. Dataset of One million reviews with numerical rating is collected from Tripadvisor. Results shows that predicted rating from sentiment analysis algorithms are very close to actual ratings of the hostel. Sentiment classification using Bayesian Classifier was implemented in [36]. Experimental results show that bayesian classifier works well on large dataset as well as small dataset.



### 2.2.2 Unsupervised learning

Unsupervised learning is the second type of Machine Learning. They proposed a technique to find the potential sentiment phrase by explicit aspects in its surrounding. Surrounding of any aspect is measured using syntactic dependencies. All potential sentiment phrases are examined and the phrase which shows positive or negative sentiments is retained. Semantic orientation and polarity is calculated by unsupervised technique. Unsupervised technique that is used are named as relaxation labeling. In paper [30], the probabilistic latent semantic indexing (pLSI) [17] was used to develop Topic-Sentiment Mixture (TSM) model which reveal latent topics including sentiment classes as additional topics. The dynamic nature of social media data whereby sentiments and topics constantly change means that sentiment/topic models also need to be updated over time. This is addressed by the dynamic JST [24] which captures both topic and sentiment dynamics by assuming that the current sentiment-topic specific word distributions are generated according to the word distributions.

Consequently the Dependency Sentiment-LDA model, which relaxes the sentiment independent assumption, was introduced by Li et al [23]. In this model the sentiments of the words in a document are viewed to form a Markov chain, where the sentiment of a word is dependent on the previous one. Although topic modeling approaches to sentiment classification do not require labeled data, they still rely on sentiment lexicons as the source of prior sentiment knowledge. Like with purely lexicon-based methods, their performance was shown to be dependent on both the coverage and quality of the lexicons used by Lin and He [25]. However, the lexicon-based methods offer greater flexibility to incorporate linguistically derived contextual knowledge making for a more transparent and accessible approach to sentiment classification.

### 2.3 Hybrid methods

Sentiment classification was also observed to improve when multiple classifiers, formed from machine learning and lexicon-based methods, are used to classify a document [35]. The hybrid method also helps overcome certain limitations of the combined methods. For instance, in a system called P Senti lexicon knowledge was used to filter out non sentiment-bearing words from the feature set subsequently used for machine learning [22].

Evaluation of P Senti shows the hybrid approach achieved better performance compared to pure lexicon-based and better cross-domain portability compared to pure machine learning. In another work, a small amount of train-

ing data for machine learning was compensated with lexicon knowledge [31]. In some other work, machine learning was applied to optimize sentiment scores prior to lexicon based sentiment classification [42]. This approach has the tendency to produce domain adapted lexicons which in turn improve sentiment classification. It is noteworthy, however, that although the hybrid approach can help overcome certain limitations of either of the combined methods (lexicon-based or machine learning) alone, it can also combine challenges from both methods. For instance, it often requires both labeled data, which can be difficult to obtain, as well as a sentiment lexicon.

## 2.4 Dependency relationship (DR) techniques

DR can be used to generalize the changing relationship of opinion words and aspects. Paper [3] enlisted the DR to get paired aspect-opinion by using movie reviews. By using dependency relationship parser, the parsed words in a sentence are joined by definite dependency relationship. By using dependency sequence, encouraging results in various research fields by employing distinct approaches to point product features and their kindred point of view from various language reviews. Different feature selection techniques have been used besides with machine learning approaches like bigrams and unigrams [34]. Paper [16] deployed syntactic relations between words in different sentences for the organization of document sentiments. Agarwal et al. (2015) used ConceptNet ontologybased dependency relations to extract features from text. They also used a method called mRMR which is basically a feature selection scheme to remove redundant information. Paper [39] proposed a technique that retrieves product aspects and opinions by taking signifies and linguistics information based on dependency relationship.

## 3 Applications

Sentiment classification is a large field that contains the vast range of applications being discussed in past research. In last decade, a lot of research has been done to examine the influence of media on the business world. The Internet has turned into a vast source of all kind of knowledge for everybody. Using internet there is an opportunity to discover the perspectives of people in general about organization methodologies, political developments, business world etc. In short, numerous applications of sentiment classification have emerged in domains of daily life like sentiment classification for the business world, political reviews, movie reviews etc. Some of these applications are OpinionMiner [27], Opinion observer [18] and OpinionFinder [44].

There are hundreds of companies that develop sentiment classification tools for their themselves and for their clients. These companies include Oracle, Azamon, Google, Hawlett-Packart, SAS and Facebook. Some small companies also build sentiment classification tool for their clients. They are Lexalytics, Semantria, Synapsify, ThriveMetrics, Etuma and MeshLabs. Facebook's Gross National Happiness application developed by facebook to predict the happiness of peoples on facebook, by countries. This application works by checking positive and negative words from peoples statuses [33]. For political parties tracking sentiments of peoples are very important. A site named as Sentex.com tracks sentiments of political parties and political topics and provides sentiment classification in ternary category (positive, negative and neutral) [6].

## 4 Challenges

There are several challenges in judging sentiments from reviews, comments etc. Usually in reviews there is inconsistent and erratic data. People have various ways of expressing sentiments; sometime they use shorthand and lots of abbreviations. Usually they cannot use proper grammar in reviews. We judge positive or negative opinions from reviews using opinion words and phrases are usually used to express. These phrases and opinion words may be used in positive and negative situations. For example good is for positive and not good is for negative. Judgment of positive and negative sentiments from review depends on context of what is around it. There are very less words that will always attach a positive or negative sentiment to an expression. Comments and reviews also contain irony and hidden emotions. The task of judging sentiments is also a challenging task, due subjective sentences and also ambiguity naturally found in opinionated text. Ambiguity words are the same meaning words which come in more than one time in same sentence. Ambiguity becomes a serious problem when it come with irony and convey words. Take for example the sentence A great mobile, yeah right!this may look like a positive review, but it may be taken as a negative review. One of the major issue in Lexicon based approaches is need of lexicons for other languages. Lexicons are only available for some popular languages like English, Arabic, Chinese etc but for unpopular languages there is no lexicons available. Also lexicons of Arabic, Chinese languages are limited in term of words, they cannot cover all words of these languages



- 
- [8] L. Dey, S. Chakraborty, A. Beepa, S. Tiwari, Sentiment analysis of review datasets Using naive bayes and k-nn classifier, *Information Engineering and Electronic Business* (2016) 54–62. ⇒65
  - [9] X. Ding, B. Liu, P. S. Yu, A holistic lexicon-based approach to opinion mining, *Proc. of the 2008 international conference on web search and data mining, ACM*, 2008, pp. 231–240. ⇒64
  - [10] A. Esuli, F. Sebastiani, Determining term subjectivity and term orientation for opinion mining, *11th Conference of the European Chapter of the Association for Computational Linguistics*, 2006. ⇒63
  - [11] Y. Fei, Simultaneous support vector selection and parameter optimization using support vector machines for sentiment classification, *Software Engineering and Service Science (ICSESS), 7th IEEE Int. Conference*, 2016, pp. 59–62. ⇒59
  - [12] S. Feng, R. Bose, Y. Choi, Connotation lexicon: A dash of sentiment beneath the surface meaning, *Proc. 51st Annual Meeting of the Association for Computational Linguistics*, 2013, pp- 1774–1784. ⇒64
  - [13] V. Hatzivassiloglou, K. R. McKeown, Predicting the semantic orientation of adjectives, *Proc. 35th Annual Meeting of the Association for Computational Linguistics*, 1997, pp. 174–181. ⇒63
  - [14] W. Haywood, J. Ricky, J. B. Holcomb, E. A. Gonzalez, Z. Peng, S. Pati, P. W. Park, W. Wang, A. M. Zaske, T. Menge, R. A. Kozar, Modulation of syndecan-1 shedding after hemorrhagic shock and resuscitation, *PloS*, 2011. ⇒63
  - [15] W. He, X. Tian, R. Tao, W. Zhang, G. Yan, V. Akula, Application of social media analytics: a case of analyzing online hotel reviews, *Online Information Review* (2017) 921–935. ⇒65
  - [16] D. T. Hess, Akio Matsumoto, Sung-Oog Kim, H. E. Marshall, J. S. Stamler, Protein s-nitrosylation: purview and parameters, *Nature Reviews Molecular Cell Biology* (2005). ⇒67
  - [17] T. Hofmann, Probabilistic latent semantic indexing, *Proc. 16th Int. Conference on World Wide Web*, 1999, pp. 50–57. ⇒66
  - [18] W. Jin, H. H. Ho, R. K. Srihari, Opinionminer: a novel machine learning system for web opinion mining and extraction, *Proc. 15th ACM SIGKDD Int. Conference on Knowledge Discovery and Data Mining*, 2009, pp. 1195–1204. ⇒67
  - [19] N. Kaji, M. Kitsuregawa, Building lexicon for sentiment analysis from massive collection of html documents, *Proc. of the 2007 Joint Conference on Empirical Methods in Natural Language Processing and Computational Natural Language Learning (EMNLP-CoNLL)*, 2007. ⇒64
  - [20] J. Kamps, M. Marx, R. J. Mokken, M. D. Rijkkel, Using wordnet to measure semantic orientations of adjectives, *LREC*, 2004, pp. 1115–1118. ⇒63
  - [21] H. Kanayama, T. Nasukawa, Fully automatic lexicon expansion for domain-oriented sentiment analysis, *Proc. 2006 Conference on Empirical Methods in Natural Language Processing. Association for Computational Linguistics*, 2006, pp. 355–363. ⇒64

- 
- [22] A. Mudinas, D. Zhang, M. Levene, Combining lexicon and learning based approaches for concept-level sentiment analysis, *Proc. of the First Int. Workshop on Issues of Sentiment Discovery and Opinion Mining*, 2012, pp. 51–58. ⇒66
- [23] S. Li, C. R. Huang, G. Zhou, S. Y. M. Lee, Employing personal/impersonal views in supervised and semi-supervised sentiment classification, *Proc. 48th Annual Meeting of the Association for Computational Linguistics*, 2010, pp. 414–423. ⇒66
- [24] F. Li, M. Huang, X. Zhu, Sentiment analysis with global topics and local dependency, *Association for the Advancement of Artificial Intelligence* **10** (2010) 1371–1376. ⇒66
- [25] C. Lin, Y. H. Lee, Joint sentiment/topic model for sentiment analysis, *Proc. 18th ACM Conference on Information and Knowledge Management*, 2009, pp. 375–384. ⇒66
- [26] B. Liu, E. Blasch, Y. Chen, D. Shen, G. Chen, Scalable sentiment classification for big data analysis using naive bayes classifier, *IEEE Int. Conference on Big Data*, 2013, pp. 99–104. ⇒65
- [27] B. Liu, H. Minqing, C. Junsheng, Opinion observer: analyzing and comparing opinions on the Web, *Proc. 14th Int. Conference on World Wide Web*, 2005, pp. 342–351. ⇒67
- [28] A. Manek, P. Deepa, C. Mohan, K. Venugopal, Aspect term extraction for sentiment analysis in large movie reviews using gini index feature selection method and svm classifier, *World wide web* **20** (2016). ⇒61, 65
- [29] C. Mate, Product aspect ranking using sentiment analysis: a survey, , *Int. Research Journal of Engineering and Technology* (2014). ⇒61
- [30] Q. Mei, X. Ling, M. Wondra, H. Su, C. Zhai, Topic sentiment mixture: modeling facts and opinions in weblogs, *Proc. 16th Int. Conference on World Wide Web*, 2007, pp. 171–180. ⇒66
- [31] P. Melville, W. Gryc, R. D. Lawrence, Sentiment analysis of blogs by combining lexical knowledge with text classification, *Proc. 15th ACM SIGKDD Int. Conference on Knowledge Discovery and Data Mining*, 2012, pp. 163–173. ⇒67
- [32] G. M. Miller, Wordnet: a lexical database for english, *Communications of the Association for Computing Machinery* (1995) 39–41. ⇒63
- [33] A. Ortigosa, J. Martin, R. Carro, Sentiment analysis in facebook and its application to e-learning, *Computers in Human Behavior* **31** (2014) 527–541. ⇒68
- [34] B. Pang, L. Lee, S. Vaithyanathan, Thumbs up?: sentiment classification using machine learning techniques, *Proc. ACL-02 Conference on Empirical Methods in Natural Language Processing*, 2002, pp. 79–86. ⇒64, 67
- [35] R. Prabowo, M. Thelwal, Sentiment analysis: a combined approach, *Journal of Informetrics* **3** (2009) 143–157. ⇒66
- [36] M. R. R. Rana, M. A. Akbar, T. Ahmad, Sentiment classification of customer reviews using bayesian classifier, *Asian Journal of Engineering, Sciences & Technology* **7** (2017). ⇒65

- [37] H. Shaziya, G. Kavitha, R. Zaheer, Text categorization of movie reviews for sentiment analysis, *Int. Journal of Innovative Research in Science, Engineering and Technology* (2015) 11255–11262. ⇒65
- [38] S. P. Sivasubramanian, N. Suganya, Sentiment analysis on micro-blogs, *Int. Innovative Research Journal of Engineering and Technology* **2** (2017) 46–51. ⇒59
- [39] G. Somprasertsri, P. Lalitrojwong, Mining feature-opinion in online customer reviews for opinion summarization, *J. UCS* **16** (2010) 938–955. ⇒67
- [40] J. Steinberger, T. Brychein, M. Konkol, Sentiment and social media analysis, *Proc. 5th Workshop on Computational Approaches to Subjectivity*, 2014. ⇒59
- [41] B. N. Supriya, V. Kallimani, S. Prakash, C. B. Akki, Twitter sentiment analysis using binary classification technique, *Int. Conference on Nature of Computation and Communication*, 2016, pp. 391–396. ⇒59
- [42] M. Thelwall, K. Buckley, G. Paltoglou, Sentiment strength detection for the social web, *Journal of the Association for Information Science and Technology* **63** (2012) 163–173. ⇒67
- [43] P. D. Turney, M. L. Littman, Measuring praise and criticism: inference of semantic orientation from association, *ACM Transactions on Information Systems (TOIS)* (2003) 315–346. ⇒63
- [44] T. Wilson, P. Hoffmann, S. Somasundaran, J. Kessler, Opinionfinder: a system for subjectivity analysis, *HLT-Demo '05 Proc. of HLT/EMNLP on Interactive Demonstrations*, 2005, pp. 34–35. ⇒67
- [45] Q. Ye, Z. Zhang, R. Law, Sentiment classification of online reviews to travel destinations by supervised machine learning approaches, *Expert systems with applications* **36** (2009) 6527–6535. ⇒64
- [46] N. Zainuddin, A. Selamat, V. Kekan, Sentiment analysis using support vector machine, *Computer, Communications, and Control Technology (I4CT), 2014 Int. Conference*, 2014, pp. 333–337. ⇒59
- [47] A. Zubiaga, I. S. Vicente, P. Gamallo, J. R. P. Campos, I. A. Loinaz, N. Aranberri, A. Ezeiza, V. F. Fernandez, Overview of tweetlid: tweet language identification, *TweetLID SEPLN*, 2014, pp. 1–11. ⇒59

*Received: May 30, 2018 • Revised: July 15, 2018*