
Text-based sentiment analysis: review

V.P. Lijo* and Hari Seetha

School of Computer Science and Engineering,
VIT University,

Vellore-632014, Tamil Nadu, India

Email: lijo.vp@vit.ac.in

Email: hariseetha@vit.ac.in

*Corresponding author

Abstract: The impact of the social networks-based sentiment analysis (SA) and opinion mining has increased in recent times. Decision-makers consider the opinions of the thought leaders and laymen, and plenty of opinions are available in social networks. When a user wants to get a service or buy a product he or she will check for the reviews and opinions provided by other people about various offerings. Opinion rich data sources are available in digital form; this attracts many researchers to focus research on SA. The ‘sentiments’ available in social networks and review pages are highly valuable for industries and individuals who want to closely monitor their reputation and live feedback about their services and products. This paper presents a review covering techniques, tools, data resources and applications in the area of text-based SA.

Keywords: sentiment analysis; SA; feature selection; text mining; semantic orientation; SO; text classification; opinion mining; lexicon.

Reference to this paper should be made as follows: Lijo, V.P. and Seetha, H. (2017) ‘Text-based sentiment analysis: review’, *Int. J. Knowledge and Learning*, Vol. 12, No. 1, pp.1–26.

Biographical notes: V.P. Lijo obtained his Master’s degree in Computer Science and Engineering from National Institute of Technology (formerly R.E.C.) Calicut and he is doing PhD at V.I.T University. He has research interests in the fields of big data, data mining, text mining and machine learning. He has published a few research papers in national and international journals and conferences. He is currently working as Assistant Professor in the School of Computer Science and Engineering at VIT University, Vellore, India.

Hari Seetha obtained her Master’s degree in Engineering Physics with specialisation in electronics from National Institute of Technology (formerly R.E.C.) Warangal and obtained MPhil as well as PhD from V.I.T University. She has research interests in the fields of pattern recognition, data mining, text mining and machine learning. She has published a few research papers in national and international journals and conferences. She is a member of editorial board for various international journals. She is presently guiding six PhD students. She is listed in the 2014 edition of Who’s who in the world published by Marquis Who’s Who, as the biographical reference representing the world’s most accomplished individuals. She had been a co-investigator to a major research project sponsored by the Department of Science and

Technology, Government of India. She is working as an Associate Professor in School of Computer Science and Engineering and as Assistant Director, Ranking and Accreditation, VIT University Vellore.

1 Introduction

The social networks and other online applications provide a space for their users to present their opinion about a product, an incident or all tangibles or intangibles in the world. The users focus on reviews for buying a product, or before going for a movie. Social networks have a high impact in current political scenario as well. Manual evaluation of reviews is impossible as the number of users increase and the corpus of entries is also very huge. An automated evaluation system is desirable to classify the big corpus effectively. Sentiment analysis (SA) deals with automatic detection of the opinion's polarity and classifies them based on its contrariety. The data sources are blogs, review sites, datasets or micro-blogging sites (Vinodhini and Chandrasekaran, 2012). The SA is conducted either at document-level, sentence-level or feature (aspect) level (Feldman, 2013). In document-level, SA assumes single opinion for the entire document, but in sentence-level SA expects multiple opinions in the document. Sometimes, a sentence itself contains multiple opinions that will lead to aspect-level SA. Jagtap and Pawar (2013) have discussed challenges and solutions of sentence-level SA. In addition to this, they have compared various methods used for sentence-level SA.

The generic SA progresses through the following process.

- Document processing: the documents in the corpus (may be in HTML, XML, Word, PDF format) are converted into text and are pre-processed using linguistic tools. The SA system may get the support of a set of lexicons and linguistic resources.
- The document analysis is performed, which will annotate the pre-processed documents with sentiments. These sentiment annotations are presented to the user by using various presentation or visualisation tools.

The SA may use supervised or unsupervised approaches to classify the input documents. In supervised approaches, it is assumed that there are a fixed number of classes, for instance, neutral, negative and positive, and sufficient training sets are available for those classes. The system learns classification model from the given training data using one of the classifier algorithms such as Naïve Bayes (NB), K-nearest neighbour (KNN), support vector machine (SVM), etc. (Feldman, 2013). Unsupervised approaches use the semantic orientation (SO) of words in the documents to classify the documents. SO can be identified using either parts-of-speech patterns or lexicon of sentiment phrases. The characteristics, similarity and contrast among affect, feeling, emotion, sentiment and opinion detection are discussed in Munezero's paper (Munezero et al., 2014). The SA can be done through various approaches based on three aspects such as

- *Based on techniques:* Machine-learning-based, lexicon-based or statistical.
- *Based on text-view:* Document, sentence, word.
- *Based on rating level:* Aspect and global rating (Collomb et al., 2014).

The SA applications, word sentiment classification, document-level sentiment classification, opinion extraction strategies and evaluation of classification are listed systematically in Pang and Lee's paper (Pang and Lee, 2008).

There are plenty of papers presented every year in the field of SA. It is desirable to have survey on papers to summarise recent trends in sentimental analysis. The readers can find enlightened surveys which discuss various applications of SA (Cambria et al., 2013; Liu, 2012; Montoyo et al., 2012; Tsytarau and Palpanas, 2012). A detailed survey on application, challenges and solution to each of the challenges is also available in literature (Pang and Lee, 2008; Liu, 2012). Hussein (2016) presented the SA challenges and categorised them as either theoretical or technical. A review and comparative analysis of web services for SA is presented by Serrano-Guerrero et al. (2015). Tsytarau and Palpanas (2012) reviewed the most prominent approaches for the problems of opinion mining and opinion aggregation, as well as contradiction analysis. Feldman (2013) presented a short survey on trends in SA. SA on reviews and its various aspects are surveyed and presented by Tang et al., (2009). A detailed survey on SA techniques and categories of sentiment classification are reported by Medhat et al. (2014). A list of publicly available datasets, a list of papers of various SA approaches, a list of text pre-processing tools, a list of SA applications and a comparison on accuracy of various SA approaches on same dataset are presented in a survey paper by Ravi and Ravi (2015). An analytical mapping of activities in SA research is available (Piryani et al., 2017). A series of ongoing evaluation tasks on computational semantic systems, known as SemEval, which consider SA on twitter as a subtask. As a part of this, a lot of datasets and tools for message classification based on context/topic have been developed (Rosenthal et al., 2014, 2015; Nakov et al., 2016). A differential analysis on SA of formal and informal text is presented by Kaur and Saini (2014). A large number of the formal and informal corpuses are listed and they identified some of the common and unique techniques used in formal and informal text SA.

Figure 1 illustrates the SA process on tweets. The SA consists of two main tasks sentiment detection and sentiment classification. The sentiment detection phase finds the words, phrases or sentences bearing opinion or sentiment in documents. The sentiment classification assigns the sentences or documents to various classes based on their sentiment polarity.

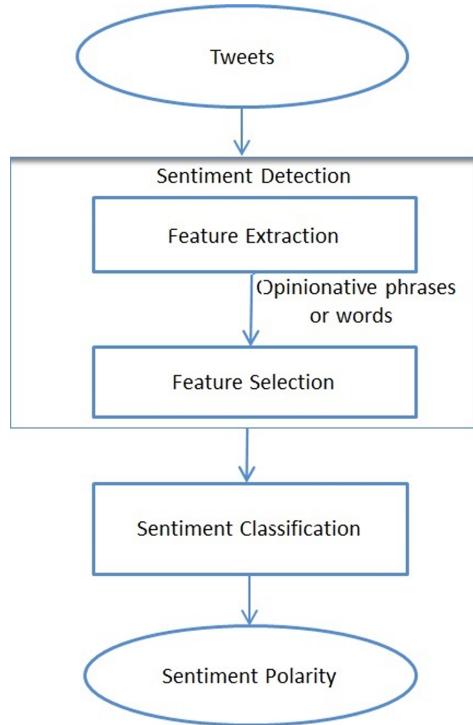
The following sections are arranged as follows: Section 2 presents various approaches to the sentiment detection. Section 3 includes main classes of sentiment classification. Important applications are discussed in Section 4 and various lexical resources which are publicly available are discussed in Section 5. Section 6 presents conclusion and future work.

2 Sentiment detection

SA uses different approaches based on data sources, tools and its application. The sentiment detection and sentiment classification are the two main tasks in SA. In sentiment detection, the system will identify the sentences with opinions and sentiments in the documents. Various strategies such as lexicon-based SO, word cooccurrence and probability-based models are used to detect sentiments in the documents. The sentiment detection consists of two phases, namely feature extraction and feature selection. In the first phase, the system identifies the features in one of forms

such as n -grams, parts-of-speech, SO and opinionative words or phrases from the sentence in the document. The second phase selects a subset of relevant features from the set of features extracted in first phase. SA relies on text interpretation and a direction-based text interpretation is proposed by Hearst (1992). He discussed different conceptual models (the force dynamic model and the path model) and their uses in addition to the role of syntax and role of general metaphor in text interpretation.

Figure 1 SA process on tweets (see online version for colours)



2.1 Feature extraction

Feature extraction is an important activity in text classification task. The efforts for feature selection and extraction started in 1990. The researchers derived many methods based on syntactic phrases to extract features (Lewis, 1992a, 1992b, 1992c; Apte et al., 1994). The set of words in the document which give some grammatical sense are called syntactical phrase. The performance of the classifiers using syntactical phrases is not good as the performance of classifiers based on bag-of-words. Lewis has identified and presented the reasons for such failures based on desirable characteristic of text classification. Many of the later works also failed to gain good performance using syntactical phrases. A rule-based approach is presented by Scott and Matwin (1999). They have concluded that the rule-based method could not give better performance than bag-of-words approach. The following subsections present some of the popular features and feature selection techniques used in text-based SA.

2.1.1 Word n-grams

Some of the sentiment classification algorithms use n -gram as a feature, where n -gram is a contiguous sequence of n words from a given sequence of text. The researchers investigated the use of word n -grams (string with various lengths) for text classification. Mladenic and Grobelnik (1998) proposed the n -gram strategy and experimented with features with sizes up to five. The classifier which uses 3-gram feature improves the performance and accuracy, and the larger n -grams decrease the accuracy (Mladenic and Grobelnik, 1998). Furnkranz (1998) claims that the works which consider the n -grams and their frequency in documents do not gain acceptable results.

In Colace et al.'s (2014) proposal; a feature selection model based on the probabilistic topic model is given; which find the pairs of words that are most discriminative. A graph-based classifier is used in this proposal; where graph is made of several clusters and each contains a set of words and the Keywords: which relate to these words. *Reuters-21578* repository is used as the dataset (Colace et al., 2014).

2.1.2 Semantic orientation

Taboada et al. (2011) proposed a vocabulary-based SA, in which the documents are classified based on its SO. Automatic or semiautomatic adjective dictionary creation helps to include many new words in dictionary and this will help to predict overall SO of the document effectively. The SO is measured by considering different components such as adjectives, verb-noun-adverb, negation, intensification, irrealis and text-level features. The negation has an important role in SA (Wiegand et al., 2010). To experiment the validity of the dictionary Taboada has used Mechanical Turk service of Amazon. Pang and Lee (2008) described the sentiment aware applications and they explored the inherent problems of opinion and SA.

2.1.3 Word cooccurrence (non-adjacent)

The relationship among words consider as potential feature for SA. The cooccurrence of the words reveals the real sentiment polarity of the sentences. Figueiredo et al. (2011) have proposed a text classification based on cooccurrence of the words (c-features: compound features) in the document. They have used c-features with various lengths in the experiments. The process started from finding all s-features (single term) and combined them for larger length c-features. The ranking process gives the facility to identify c-features with most discriminative power.

Similarity measure for text processing (SMTP) (Lin et al., 2014) is a new similarity measure which considers the presence and absence of a feature in a set of document. Single label-KNN and Multi-label-KNN are used with different similarity measures such as Euclidean Cosine, Pairwise and IT-Sim. The results showed that the SMTP gave better accuracy than KNN which is used with other similarity measures.

The features discussed above are common for formal (example: Blogs) and informal (example: Microblogs). The informal texts' special characteristics such as short in length and span in one or less sentences bring new challenges in SA. Informal text tends to have many misspellings, informal intensifiers, slang terms and abbreviations (Kiritchenko et al., 2014). They also have special features such as emoticons that are used to indicate sentiment, hashtags that are used to facilitate search and to indicate a topic. The features such as emoticons, hashtags and intensifiers are useful in SA (Kouloumpis et al., 2011).

2.2 Feature selection

The feature selection is an important task in SA of high-dimensional data. Large set of features demands large amount of memory and processing power. The classification algorithm may overfit to the noisy training samples and it may be failed to generalise new data. The low dimensionality improves the performance of the classifier and saves from the risk of overfitting. Some of the well-known feature selection methods are Chi square (CHI) (Tallarida and Murray, 1987), Gini index (Singh et al., 2010; Park et al., 2010), correlation coefficient (CC) (Hsu and Hsieh, 2010), latent semantic analysis (LSA) (Landauer et al., 2013) and mutual information (MI) (Doquire and Verleysen, 2013). The feature selection can be divided in to statistical (automatic model and it is used more frequently) and lexicon-based where human intervention needed to annotate the samples. In the following subsections, we present two of the statistical methods.

2.2.1 Chi square

The chi-square or χ^2 distribution is very popular feature selection technique and it is widely used in text classification (Mesleh, 2007). The χ^2 distribution is used in the common χ^2 tests for the independence of two criteria of classification of qualitative data and goodness of fit of an observed distribution to a theoretical one. CHI measures dependence between a term and a category (Zheng et al., 2004). The CHI and various feature selection techniques are presented discussed by Forman (2003). They claimed that the CHI plays an important role in text classification.

The CHI is defined as follows:

$$\text{CHI}^2 = \frac{n \times F(W)^2 \times (p_i(W) - P_i^2)}{F(W) \times (1 - F(W)) \cdot P_i \times (1 - P_i)}. \quad (1)$$

For n number of documents, $p_i(W)$ the conditional probability of class i of documents which contain w , $F(W)$ the global fraction of the documents which contains the word and P_i the global fraction of the document which contain the class i (Medhat et al., 2014).

2.2.2 Latent semantic analysis

Latent semantic analysis (LSA) is a statistical approach in natural language processing to analyses relationships between a set of documents and the terms they contain by producing a set of concepts related to the documents and terms. LSA assumes that words that are close in meaning will occur in similar pieces of text (Landauer et al., 1998). The term ‘LSA’ used to represent the theory as well as the method for extracting and representing the contextual usage meaning of words by statistical computations applied to a large corpus of text (Landauer and Dumais, 1997). The latent semantic indexing (LSI) is an automatic indexing where LSA is the key component (Dumais, 1995). The LSI employs the singular value decomposition (SVD) to find the lower-dimensional subspace that consider the terms and document relationship in the form of term document matrix. LSI along with principal component analysis (PCA) is used to create small set of features as a function of original set (Jolliffe, 2002). The LSI is an unsupervised technique and it follows the underlying class distribution blindly. So the features selected by the LSI may not be best separated the classes (Aggarwal, 2015).

3 Sentiment classification

Sentiment classification is the process of classifying a text according to the sentimental polarities of opinions it contains (Pang et al., 2002). It can be divided into lexicon-based approach, machine-learning approach and hybrid approach (Maynard and Funk, 2011). The lexicon is a collection of known and precompiled collection of sentiment terms. The performance of lexicon-based approach is purely based on the soundness and completeness of the sentiment lexicon. Lexicon-based approach is divided into dictionary-based approach and corpus-based approach, and the corpus-based approaches use either statistical or semantic method to determine sentiment polarity. The machine-learning approach follows either supervised learning or unsupervised learning strategies. Supervised learning applies popular machine learning classification algorithms such as SVM, NB, Decision Tree and Rule-based classifiers (RBCs) along with linguistic features (Medhat et al., 2014). Table 1 gives a summary of 20 research papers in the field of SA. The algorithms used in the papers are listed in the second column and the data scope and data source are in column number three and five, respectively. The sentiment polarity considered as positive (P), negative (N), neutral (NL) and some of them treated sentiment in differently as general (G). The polarity is listed in fourth column of the table. To support the SA process some components (lexicons, tools, datasets, etc.) are used along with the classification algorithms. These supporting components are listed in sixth column. The popular evaluation metrics used are Precision (P), Accuracy (A), Recall (R), F-measure (F), Area Under ROC (AUC), etc. (Davis and Goadrich, 2006; Powers, 2011). Seventh column consists of the evaluation measures considered in various papers. The combination of machine-learning and lexicon-based approach is considered as hybrid approach and in most of these methods the lexicon is used as a key component. Figure 2 shows sentiment classification methods as discussed above. The sentiment classification task performs binary sentiment classification or multi-class sentiment classification. In binary class, the sentiments classify as either positive or negative, but multi-class classifier considers finite number of classes.

Table 1 List of papers in sentiment analysis

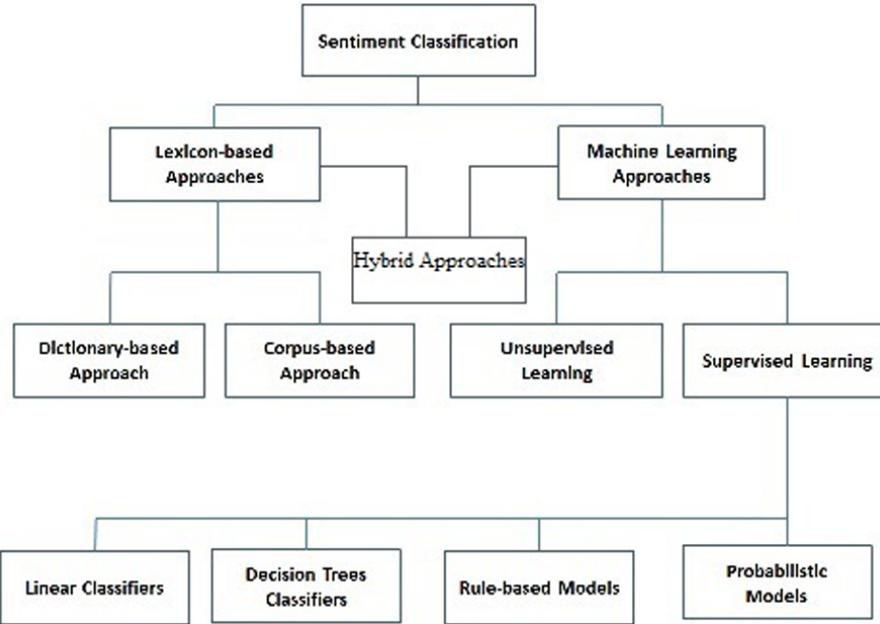
| S. No. | Reference | Algorithm used | Data scope | Polarity | Dataset/ source | Supporting components | Evaluation metrics |
|--------|-----------------------|--|-----------------------|----------|--|---|-------------------------------|
| 1 | Mamgain et al. (2016) | NB, SVM, NN-multilayer Perceptron | Tweets | P, N | Twitter | Twitter API | A- 92.6% |
| 2 | Devi et al. (2016) | SVM | E-commerce Reviews | P, N | Customised data e-commerce reviews | SentiWordNet (Esuli and Sebastiani, 2006) | P- 87.15% R- 89.76% A- 88.13% |
| 3 | Ashok et al. (2016) | NB, SVM, Random Forest, ME, Rule-based, Aspect-based | Cornell movie reviews | P, N | Cornell movie reviews (Pang and Lee, 2004) | Word2Vec-gensim- python, sklearn, nltk | A- 77.7% |
| 4 | Apoorva et al. (2016) | Rule-based | Tweets | P, N | Twitter | Twitter API | N/A |

Table 1 List of papers in sentiment analysis (continued)

| <i>S. No.</i> | <i>Reference</i> | <i>Algorithm used</i> | <i>Data scope</i> | <i>Polarity</i> | <i>Dataset/source</i> | <i>Supporting components</i> | <i>Evaluation metrics</i> |
|---------------|---------------------------|---|----------------------------------|-----------------|--|----------------------------------|---|
| 5 | Gao et al. (2016) | Dictionary-based, PAD model (Mehrabian and Russell, 1974) | Micro-Blogs | P, N | Chinese Micro-blog- Sina microblog OAuth2.0 Interface | Google word2vec model tools | P- 81.56%, R- 89.32% F- 80.10% |
| 6 | Rabab'ah et al. (2016) | Statistical-based | Reviews, posts, comments, tweets | P, N | OCA, AWATIF, SentiStrength Yahoo!- Maktoob and Gazza Attacks | (Thelwall, 2013) | A- 62.0% P- 83.7% R- 64.0% F- 68.0% |
| 7 | Yang et al. (2016) | Treebank Convolutional NN (Tb-CNN) | Sentences, phrases | P, N | Stanford Sentiment Treebank (SSTb) | word2vec-Python | N/A |
| 8 | Qaisi and Aljarrah (2016) | NB | Tweets | P, N, NI | AWS, ec2cloud, Amazon web services Microsoft Azure | Twitter API | Pn, Nn |
| 9 | Khan et al. (2016) | Hybrid approach-(SWIMS) | Reviews | N, P | Large movie review dataset (Maas et al. (2011), Cornell movie review dataset (Pang and Lee, 2004), Multi-domain sentiment dataset (Blitzer et al., 2007) | SentiWordNet | A- 84.50% P- 81.08% R- 90.00% F- 85.31% |
| 10 | Yuan et al. (2016) | Hybrid approach, NB | Micro blogs | P, N, NI | Chinese micro-blog- sina microblog | CLIWC, Emoticons | F- 33.33% A- 97.83% |
| 11 | Xiao et al. (2016) | Deep NN | Micro blogs | N, P, NI | Twitter dataset, Sina micro- blog sentiment corpus (MSMC) | RBM-Restricted Boltzmann Machine | P- 86.2% R- 73.0% F- 79.1% |
| | | | | | The Fifth Chinese opinion analysis evaluation- corpus (COAE2014) | | |
| 12 | Sahu and Ahuja (2016) | Bagging, random forest, decision tree, NB, KNN | Movie reviews | P, N | IMDB movie review database | SentiWordNet | P- 89.2% R- 89.0% F- 89.0% A- 88.95% |

Table 1 List of papers in sentiment analysis (continued)

| <i>S. No.</i> | <i>Reference</i> | <i>Algorithm used</i> | <i>Data scope</i> | <i>Polarity</i> | <i>Dataset/source</i> | <i>Supporting components</i> | <i>Evaluation metrics</i> |
|---------------|--|------------------------|-------------------|-----------------|---|--|--|
| 13 | Onan et al. (2016b) | Ensemble classifier | Service reviews | G | Amazon.com, CampRatingz.com, RateMDs.com, DrugRatingz.com, etc | AR1 and PC1 datasets (Tosun and Bener, 2009), Spambase dataset (Lichman, 2013). | A- 98.86% |
| 14 | Fernández-Gavilanes et al. (2016) | Unsupervised | Online texts | | Cornell movie review (Pang and Lee, 2004), Obama-McCain Debate (Shamma et al., 2009) | PolarityRank (Cruz et al., 2011) | P- 76.38% |
| | | | | | SemEval-2015 (http://alt.qcri.org/semeval2015/) | R- 71.80% F- 74.02% A- 74.80% | |
| 15 | Jang et al. (2016) | NB | Unstructured data | P, N | http://www.sejongcity.com . | SQL, KKM (Kind Korean Morpheme) Analyser | A- 75.0% |
| 16 | Moreno-Ortiz and Fernández-Cruz (2015) | Corpus-based | Finance texts | P, N | Contemporary American English (http://corpus.byu.edu/coca/) | Context rules database | P- 84.69% R- 77.57% |
| | | | | | Global web-based English (http://corpus.byu.edu/glowbe/) | | |
| 17 | Chikarsal et al. (2015) | Rule-based, SVM | Tweets | P, N, NI | Twitter | Bing Liu Lexicon, NRC Emotion Lexicon, SentiWordNet | P- 82.4% R- 62.9% F- 66.2% |
| 18 | Katz et al. (2015) | Context-based (ConSen) | Reviews | P, N | TripAdvisor, The internet | Word Sense Disambiguation (WSD)(Ide and Véronis, 1998) | A-76.9% AUC- 85.77% F- 77.95% |
| | | | | | Movie database (IMDB) | | |
| 19 | Wang et al. (2014) | Rule-based | Micro blog | G | Sina Weibo | CBoO, OVD Visualisation tools | P- 77.6% R- 81.2% F- 70.2% |
| 20 | Da Silva et al. (2014) | Ensemble classifier | Tweets | P, N | Sanders, Stanford, Obama-McCain Debate | WEKA (http://www.cs.waikato.ac.nz/ml/weka/) LibSVM | A- 84.89% P- 82.10% R- 86.30% F- 84.20% |

Figure 2 Sentiment classification methods

3.1 Machine-learning approach

The machine-learning algorithms are trained on existing data and capable to learn and identify patterns when exposed to new data. The following subsections present some of the supervised and unsupervised learning approaches.

3.1.1 Supervised learning

The supervised learning methods depend on the availability of labelled training documents. It is the machine learning task which infers a function from the labelled set of training examples. In supervised learning, each example is a pair of input vector and an expected output. In the following subsections, we present some of the most commonly used classifiers in sentiment classification.

3.1.1.1 Linear classifiers

The linear classifier constructs a hyperplane that separates data between different classes. A linear predictor p is defined as

$$p = \bar{A}\bar{W} + s, \quad (2)$$

where $\bar{A} = \{a_1, a_2, a_3, \dots, a_n\}$ is a vector of linear coefficients, $\bar{W} = \{w_1, w_2, w_3, \dots, w_n\}$ the normalised document word frequency and s is a scalar. The predictor p is a separating hyperplane between different classes, which is the output of the linear classifier (Medhat et al., 2014). There are a number of linear classifiers, such as SVM (Burges, 1998) and neural network (NN) (Ruiz and Srinivasan, 1999).

SVM works on the principle of determining linear separators in the search space which can best separate the different classes. Mullen and Collier (2004) introduced an approach to SA in which they used SVM classifier. This strategy uses SO (positive or negative) values from various data sources and uses them to create feature space which is classified using SVM. They used 3-fold and 10-fold validation with Hybrid SVM models, and achieved the accuracy of 84.6% and 86.0%, respectively.

Li and Xu (2014) propose an emotion classification using the emotion cause detection technique, where they experimented on Weibo dataset. They performed the experiments in two steps as

- extracting emotion cause events
- training and testing classifier.

SVM is used for classification and χ^2 test is used to generate final feature set.

The NN is another frequently used linear classifier. The neuron is the basic unit in neural network and NN works on set of neurons. The input of the NN is the vector W_i , the frequency vector of the document i . A set of weight G associated with each neuron which is used to compute input functions. The linear predictor function of NN is

$$P_i = G \cdot \overline{W}_i. \quad (3)$$

The sign of the P_i yields the class label y_i . The NN implementation can be either single or multi-layer. The multi-layer implementation is applicable for non-linear boundaries. The NN implementation is available for text data (Ruiz and Srinivasan, 1999). The NNs are useful in sentiment classification (Chen et al., 2011).

3.1.1.2 Rule-based model

In RBCs, a set of rules are used to model the data space. The left-hand side of the rule represents a condition on the feature set expressed in disjunctive normal form while the right-hand side is the class label. The term presence considers for making conditions. Term absence is rarely used because it is not informative in sparse data. The *support* and *confidence* are the important criteria using for generating rules (Ma and Liu, 1998). The *support* refers to the absolute number of relevant instances which follows rule in the training data. The *confidence* is the conditional probability that the right hand side of the rule is satisfied if the left-hand side is satisfied. Shaheen et al. (2014) had presented a rule-based emotion recognition model. They consider six emotions: sadness, happiness, anger, fear, disgust and surprise. The emotion recognition is performed based on emotion recognition rules (ERRs). They have used KNN and point mutual information (PMI) classifiers to classify the sentences based on the semantic similarity and keyword similarity. They have considered two annotated datasets in which each sentences annotated with one of the six Ekman emotions. The Aman 2007 (Aman and Szpakowicz, 2007) dataset is composed of emotion-rich sentences and the other Tweets collected from Twitter. Out of six, Ekman emotions Tweets are annotated with five emotions and disgust emotion is left out. The KNN classifier classifies input ERRs using annotated ERR and dictionaries such as WordNet (Esuli and Sebastiani, 2006) and ConceptNet. The dictionaries are used to calculate similarity score based on word and concept similarity. The experiments rely on short sentences with lengths varying from 0 to 10 words. According to Shaheen et al. (2014), the short sentences provide more accurate emotion

indication than large sentences. The rule-based approach can be combined with other supervised classifiers such as SVM for sentiment classification. Rule-based approach with SVM for Twitter SA is presented by Chikersal et al. (2015).

3.1.1.3 Probabilistic models

The probabilistic model assumes that each class is a component of a mixture and each mixture provides the probability of sampling a particular term for that component. There are many probabilistic (Generative) classifiers, e.g., NB, maximum entropy classifier (ME) and Bayesian network (BN).

NB classifier is a very simple and powerful classifier and it assumes independent features in the dataset and works based on NB' theorem as follows:

$$p(c/f) = p(c) \times p(f/c) / p(f), \quad (4)$$

where $p(c/f)$ is the probability of the given feature set which belongs to c class, $p(f/c)$ is the prior probability that a given feature set is being labelled as a class c , $p(c)$ is prior probability of a class and $p(f)$ is the prior probability that a given feature set is occurred. The NB classifier is useful and effective for text-based SA (Chen et al., 2009; Qaisi and Aljarah, 2016). The comparative level SA does not expect explicit opinions, but instead formulate an opinion by comparing with other products. The Chen's study reveals that relatively a small set of words can cover 98% of the comparative opinions. Hence the recall of these words is very high and precision is low. In this scenario, the NB classifier can be used to filter out the sentences those not supporting comparative opinions.

The Bayesian network (BN) assumes that all the features are fully dependent. In BN, the main component is a directed acyclic graph whose nodes represent random variables, and edges represent conditional dependencies. The BN computation is very complex for text mining, so it is not frequently used (Aggarwal, 2015).

Maximum entropy (ME) does not assume that the features are conditionally independent of each other. The ME classifier is based on the *Principle of Maximum Entropy* (Aggarwal, 2015) and selects the one model which has the largest entropy, from all the models that best fit the training data. ME outperforms NB sometimes for text classification (Nigam et al., 1999). SA on online news text using ME and supporting components FrameNet (Baker et al., 1998) and clustering by committee (CBC) (Pantel and Lin, 2002) has done in Kim's work (Kim and Hovy, 2006). The importance of the ME is emphasised by El-Halees (2015).

A filter-based feature selection method is proposed in Uysal and Gunal's (2012) works, where the filter is a probabilistic-based filter. This filter helps assign high score to discriminative features. The distinguishing feature selector (DFS) filter does not depend on the learning model. They used three different classifiers such as SVM, decision tree and NN, to investigate contributions of the selected features to the classification accuracy. Reuters-21578 ModApt split, 20 NewsGroups (Asuncion and Newman, 2007) and short message service (SMS) (Almeida et al., 2011) are used as datasets. *Macro-F1* and *Micro-F1* are calculated, 96.97 and 96.68, respectively, for NewsGroups datasets. They performed term similarity analysis for checking the effectiveness of the algorithm based on the profile of features selected. The classification accuracy is higher if the distinctive features get high score. The observation reveals that the 71% of the selected

features from top 500 features in Reuters dataset is common for DFS and CHI2 (Uysal and Gunal, 2012) and 29% specific to the DFS. The DFS is superior or giving good results (slightly lower) in when compared with CHI2. They have included dimension reduction analysis for measuring the effectiveness of the DFS in terms of dimension reduction rate and compared algorithmic complexities of the DFS with other methods.

3.1.2 Unsupervised learning

The unsupervised learning is inferring a function to learn hidden structure of the unlabelled data. The supervised strategy needs large set of training data to achieve high accuracy. This is a difficult task to prepare training set manually by assigning labels to the data elements. So many researchers are focusing their works on finding solution to classify unlabelled documents. Ko and Seo (2002) introduced an unsupervised or semi-supervised text classification technique for unlabelled documents using bootstrapping and text categorisation using feature projections (TCFP) classifier. TCFP is a classifier using feature projections as key component. Feature projection is used to assign high weights to the features with discriminating characteristics. The text categorisation is progressed as follows: the collected documents are pre-processed to get a pool of contexts. Then, the keywords for each category are created automatically using the cooccurrence features and the title and keywords are used to extract basic contexts which are used as centroids of context clusters. The similarity measure is used to extract keywords for each category. Degree of similarity between the title words of each category with other words will help to extract keywords. The TF-IDF gives the term weight and they calculate the vote for each class and assign the document to the class with majority votes.

They compared the effectiveness of the TCFP with KNN and KNNFP (KNN with feature projection). The results show that TCFP is superior to other classifiers such as KNN and KNNFP for noisy training data. The TCFP's performance depends on its title and keyword quality (Ko and Seo, 2002). The effectiveness of the strategy for the dataset WebKb is worst due to the high frequency of the keywords of one category in other categories. They suggest providing keywords and title words for each category manually by human developers for improving performance. The robustness of TCFP in noisy data is appreciable. This method can be used as assisting tool, with less expense, to create training data for text classification. Hofmann (2001) proposed an unsupervised learning and used LSA as key component. Xianghua et al. (2013) used an unsupervised learning approach with latent Dirichlet allocation (LDA) (Blei et al., 2003) to detect aspects in social reviews. The unlabelled data (unsupervised information) can be used along with a classifier for improving semi-supervised sentiment classification and the similarity matrix is used as the powerful knowledge recovery tool (Da Silva et al., 2016).

3.2 Lexicon-based approaches

Lexicon-based approaches are either dictionary-based or corpus-based. Publicly available data resources (lexicons) are supporting this classification tasks. Lexicon-based SA consider the word level SOs. The lexicon considers verbs, adjectives and nouns in the given sentences. Verb description can be used to perform deep SA (Maks and Vossen, 2011). Maks and Vossen (2011) used the database for the Dutch, which combines two lexicons:

- the Dutch Wordnet (DWN)
- the Dutch Reference Lexicon (RBN).

The verbs are categorised in different types based on how they give reference to the emotions. They investigated that the combinations of the types of verbs are capable to identify sentimental oriented subjectivity.

Hatzivassiloglou and Wiebe (2000) have analysed the adjective orientation-based approaches. The adjective orientation and gradability are useful to identify the sentence level subjectivity. An automatic reliable method for extracting gradability values is included in their work. Godbole et al. (2007) have presented a lexicon-based SA with path analysis for large set. The path analysis is essential to overcome pitfalls of selecting spurious features. The synonyms inherit the polarity from its parent feature and antonyms get opposite polarity. So the paths which connect between synonyms and antonyms are most probably spurious.

An appraisal groups-based sentimental analysis is presented by Whitelaw et al. (2005). A semi-automated method is used to build lexicon of appraisal and their variations. They claimed the accuracy of their method is more than 90%. A value is assigned for each appraisal adjective based on its attitude, orientation, force, focus and polarity. They have measured the effectiveness of the appraisal groups for SA using publicly available movie reviews. They have considered the reviews which are annotated as only positive and negative. The adjectival appraisal groups are giving good classification results than other appraisal groups. Whitelaw et al. (2005) have not considered nominal and verbal appraisal groups. The hash-tagged dataset (HASH) and emotion dataset (EMOT) are useful for Twitter SA (Kouloumpis et al., 2011). Kouloumpis et al. (2011) have conducted experiments to measure the classification accuracy over different combinations of HASH and EMOT along with n -gram, lexicon features, part-of-speech (POS) and micro-blogging features (Li et al., 2010). The best performance got from using the n -grams together with the lexicon features and the micro-blogging features. The POS inclusion adversely affects the performance. The best result got from the n -grams, lexical and Twitter features trained on the hash-tagged data alone.

The dictionary-based analysis (Hu and Liu, 2004) is proceeding as follows: a small set of terms with known SO is collected manually. Then make this set grown by searching available corpora for their antonyms and synonyms. The newly collected terms is added to the seed set and start next iteration. The iteration is over when no new terms found. A manual investigation is needed to correct errors in the collection of terms. The demerit of this method is that there is no provision to identify the terms based on its domain and context orientation. Qiu et al. (2010) have presented a method to perform context-based SA and the results of their works reveal the effectiveness of this approach.

Scalable SA based on large dictionary of words and its polarity is presented by Kaushik and Mishra (2014). This approach is suitable for domain specific SA. They considered the positive, negative and neutral polarities of the words. Dictionary contains all possible verbal forms of words to eliminate stemming and achieving fast process. The role of negation and blind negation is also considered and add negations in dictionary. Flume (Kaushik and Mishra, 2014) is used to collect large amount of data from Twitter and transfer to Hadoop distributed file system (HDFS). This method achieved fast analysis but it compromise the accuracy as 73%.

The Corpus-based approach helps to solve the problem of finding context specific opinion words. This approach can be either statistical or semantic approach. These methods use either syntactic patterns or patterns that occur together along with a seed list of opinion words to find other opinion words in a large corpus (Medhat et al., 2014). SmartSA is a lexicon-based sentiment classification system for social media genres and it provides a contextual polarity detection by integrating local context detection and global context detection (Muhammad et al., 2016).

3.3 Hybrid approach

The combination of machine-learning and lexicon-based approaches outperforms sometimes the individual performance of these approaches. The large computational complexity of the Hybrid model made them not frequent. Prabowo and Thelwall (2009) proposed a hybrid model of classifiers for sentimental analysis. They considered various combinations of general inquirer-based classifier (GIBC), RBC, statistics-based classifier (SBC), induction RBC (IRBC) and SVM. The experimental results show that the GIBC reduce the effectiveness of the hybrid classification. The combination of SBC and SVM improved the effectiveness of the classification. They used four different datasets such as movie reviews with different number of positives and negatives, product reviews and MySpace comments and calculated the *MicroF1* and *MacroF1*.

3.4 Other approaches

The researchers developed new approaches for SA to overcome the shortcoming of the existing approaches. The ensemble methods such as AdaBoost, Bagging, Dagging, Random Subspace and Majority Voting combine various classifiers' output to get better classification performance (Onan et al., 2016a) and they provide better performance in SA (Onan et al., 2016b). Agarwal and Sabharwal (2012) proposed an end-to-end pipeline for classifying tweets in Twitter into different classes. They give four classes: objective, neutral, positive and negative. A 4-way classifier and Cascaded design of three classifiers stacking on top of each other are proposed in this paper. The trade-off between Rejection rate and F1-measure is tested and illustrated. It shows that rejection rate always increased faster than F1-measure.

The Readme (Hopkins and King, 2010) and iSA (Ceron et al., 2016) are two alternatives for contextual SA. Instead of aggregating estimates of classifiers, they consider an entire corpus of texts and directly estimate the aggregated distribution of sentiment. A context-based learning and classification (ConSen) is presented by Katz et al. (2015). A fuzzy-based approach is used to handle unclear and uncertain texts, where main component is Formal Concept Analysis-FCA (Li and Tsai, 2013).

4 Applications

SA of reviews of services, actions and products is a general application. There are many startups to provide automatic review summery. Social networks are providing good impact on reputation of brands and events. Sentimental analysis on Twitter, Facebook, Weibo, etc. provides live feedback on actions, events, products and individuals. Twitter

provides real time analysis of tweets to help the user to get a set of tweets (with polarity assigned) about a given term.

Sentiment analysis is used to measure the acceptance and reputation of the candidate of political elections. Social networks provide corpus of opinion oriented data. Some systems are developed to detect emerging political topics on twitter earlier than Google trend and collect those topics to extend the knowledge bases for concept-level SA (Rill et al., 2014). The SA can be used to check whether the controversial topics arise quicker than other topics. Spatial distribution of the tweets can be analysed to know the impact of emerging topics in different Geographical regions. The scope of the SA is growing day by day and it is being applied in most of the day to day life activities. SA is used for Ranking among services, for example ranking educational institutions in a country based on opinions in tweets (Mamgain et al., 2016) and teachers performance evaluation based on students' feedback (Balahadia et al., 2016). The SA is also used to identify society's pending issues (Jang et al., 2016) and to improve government services (Seki, 2016). Nowadays, SA on social networks is used for predicting the wins and spread of mega sports events such as premier league, FIFA World Cup, etc. (Yu and Wang, 2015; Schumaker et al., 2016).

Annotating large datasets based on emotions identify emotions automatically. Annotated large dataset for emotions: Anger, Disgust, Fear, Joy, Sadness and Surprise are available (Strapparava and Mihalcea, 2008).

Another important application of SA is to measure the social impact on a natural calamity or a big event such as world cups. Takahashi et al. (2015) conducted an analysis on tweets during Typhoon Haiyan and they presented various uses of social media during disaster. They classified the tweets as requesting help, co-coordinating relief efforts, providing mental counselling, criticising the government, expressing well wishes and memorialising, discussing causes, etc. The use of emotion expressions vary from the places close to the affected area to farthest places. The SA helps predict the people's reactions after the natural disaster and based on these results the government can take corrective actions to reduce the negative effects. Affective learning is an emerging application of SA, opinion mining or emotion detection. Automatic detection of emotion feedback is useful to improve learning experience in case of large pervasive environment. An affective e-learning model is proposed by Shen et al. (2009) which combined the learner's emotions with Shanghai e-learning platform and analysed the methods which are utilising affective learning to improve learning. The SA can be used for detecting spammers who unfairly overwhelm normal users with unwanted or fake content (Hu et al., 2014).

5 Lexical resources

Opinion Mining or SA is relying on some tasks such as determining SO, determining texts' negative or positive polarity and determining the strength of the PN-polarity (Esuli and Sebastiani, 2006). To aid these tasks the researchers are using many lexical resources to annotate words automatically with its numeric scores. Most of the Lexicons are developed from online resources. The acquisition of lexical components from online resources and various aspects to build lexicon are discussed and illustrated by Uri (1991). Ontological principles and different aspects are discussed by Guarino (1998). He discussed the main problems in upper level construction of lexical resources, such as ISA

overloading, confusion in senses, sense reduction, generalisation issues and confusion in roles. Dependence theory, parts theory, theory of whole and identity theory are considered under formal ontology. FrameNet Database and associated software tools are developed by Baker et al. (1998). FrameNet consists of lexicon, frame database and annotated example sentences. An example of free available lexical resource is SentiWordNet (Esuli and Sebastiani, 2006), in which all synsets are annotated with positive, negative and objective scores. This resource is based on the WordNet 2.0 synsets where WordNet 2.0 maintains a structure of lexical components to support fast access and effective linkage among components. The words synsets are annotated with three different polarities such as Positive, Negative and Objective. A committee of classifiers is formed to improve accuracy. The experiments show that the SentiWordNet 1.0 is very helpful for sentiment classification more accurately. SentiWordNet 1.1 and SentiWordNet 2.0 (Esuli and Sebastiani, 2007; Esuli, 2008) are presented by Esuli in 2007 and 2008, respectively, but they are not publicly available.

SentiWordNet 3.0 is an enhanced lexical resource based on WordNet 3.0 (Baccianella et al., 2010). The SentiWordNet 1.0 and SentiWordNet 3.0 differ in the algorithms used for annotation. SentiWordNet 3.0 uses an additional random-walk (Esuli and Sebastiani, 2007) for refining scores and in the version of WordNet which they annotated. There are two steps in automatic annotation, the semi-supervised learning and the Random-walk. Semi-supervised learning step consists of seed set expansion, classifier training, synset classification and classifier combination. The classifier combination, a committee of ternary classifiers instead of single classifier, helps improve accuracy (Esuli and Sebastiani, 2006). The random-walk is an iterative scanning through the node of graph to improve the positive and negative score possibly. It assumes a direct linkage between various synsets if the synonyms in the synsets come in the gloss of each other. The mapping of synset is based on three different strategies: the WordNet sense mapping which is limited for noun and verbs only, the synset term matching which identify the terms to be selected from various synsets and the gloss similarity. The gloss similarity is calculating using Dice coefficient (https://en.wikipedia.org/wiki/Dice_coefficient). The high value of Dice coefficient indicates high similarity. Princeton WordNet Gloss corpus is giving manually disambiguated glosses and it can be used to generate SentiWordNet. But the performance comparison with automatically disambiguated one from eXtendedWordNet is not possible as eXtendedWordNet is available only for WordNet 2.0.

The linguistic resource named WordNet-Affect (Valitutti et al., 2004) is useful for affective learning applications. WordNet-Affect focused on the affective meanings instead of nature of emotions. WordNet is used as a model for the affective concepts. They showed that new affective synsets can be obtained by applying WordNet relations to the synsets of WordNet-Affect. In this study, they have included only WordNet relations. It is desirable to validate its efficiency and performance using machine learning techniques applied to large corpora.

The lexicon resource called STS Gold (Saif et al., 2013) provides lexicon for micro-blogs such as Twitter. STS Gold composed of lexicon for the unstructured sentences such as tweets. Hence STS Gold is not suitable for the analysis of reviews or other document level SA.

There are a variety of lexicon resources such as NRC emotion lexicon, Bing Liu's lexicon, MPQA Subjectivity Lexicon, Hashtag sentiment lexicons (HS), Sentiment140 lexicons available (Kiritchenko et al., 2014), but they are not framed in common

platform. Lack of common framework increases the complexity of using lexicon resources. A lexical markup framework (LMF) (Francopoulo et al., 2006) which provides a common platform for the construction of lexicons is developed in 2006. The structure of the LMF consists of the core package and Extensions. This framework is suitable for coordinating creation of NLP lexicons. Various classes of tasks are described in detail, which is helpful to formulate the structure of the lexicons with fewer efforts. Lexical and linguistic patterns play major role in text-based sentiment and emotion cause detection (Hunston and Francis, 2000; Li and Xu, 2014).

6 Conclusions and future work

Ideally, the text-based intelligent system would interpret its input corpus and allow the user to get results according to his preferences. The state of the art is very far from ideal goal. This review paper presented a review on progress of SA techniques and applications. The review reveals that sentiment detection and sentiment classification are still open for research. NB, KNN and SVM are most frequently used in sentiment classification and they are used as reference to compare with other proposed algorithms. Even though the Hybrid and ensemble approaches are using to improve the SA performance, the short comings of existing classification and feature selection techniques limit the application of SA. So developments of new and refined techniques are needed and these are open research challenges.

The researches on SA in many natural languages other than English are growing as building lexicons. The publicly available lexical resources such as SentiWordNet and its various versions, WordNet and its successors are available for many languages other than English. Building resources for natural languages is still an open challenge. Automatic annotation of sentiment phrases in corpora gives more coverage but credible annotation is manual.

The text-based SA for information from micro-blogs, blogs, forums and news sources have gained high attention in recent years. The short and unstructured presentation of sentences in micro-blogs and social media raise challenges to detect and classify sentiments with high accuracy. The lexical resource STS Gold is developed exclusively for the micro-blogs such as Twitter.

In many applications, context-based SA is necessary, and they may consider the user preferences as well. So context-based SA is an open research area. The forums and news sources may keep more than one sentiment in a single sentences and it leads to aspect level analysis. But Aspect level SA is still not well addressed.

References

- Agarwal, A. and Sabharwal, J.S. (2012) ‘End-to-end sentiment analysis of twitter data’, *Proceeding of the Workshop on Information Extraction and Entity Analytics on Social Media Data, COLING 2012*, Mumbai, India, pp.39–44.
- Aggarwal, C.C. (2015) ‘Mining text data’, *Data Mining*, Springer International Publishing, Switzerland, pp.429–455.
- Almeida, T.A., Hidalgo, J.M.G. and Yamakami, A. (2011) ‘Contributions to the study of SMS spam filtering: new collection and results’, *Proceedings of the 11th ACM Symposium on Document Engineering*, September, ACM, CA, USA, pp.259–262.

- Aman, S. and Szpakowicz, S. (2007) ‘Identifying expressions of emotion in text’, *Text, Speech and Dialogue*, Springer Berlin Heidelberg, pp.196–205.
- Apoorva, G., Vaishnav, N.R., Chowdary, E.D. and Uddagiri, C. (2016) ‘An approach to sentiment analysis in Twitter using expert tweets and retweeting hierarchy’, *International Conference on Microelectronics, Computing and Communications (MicroCom)*, IEEE, January, Durgapur, India, pp.1–8.
- Apte, C., Damerau, F. and Weiss, S.M. (1994) ‘Automated learning of decision rules for text categorization’, *ACM Transactions on Information Systems* Vo. 12, No. 3, pp.233–251.
- Ashok, M., Rajanna, S., Joshi, P.V. and Kamath, S. (2016) ‘A personalized recommender system using machine learning based sentiment analysis over social data’, *IEEE Students’ Conference on Electrical, Electronics and Computer Science (SCEECS)*, IEEE, Bhopal, India, pp.1–6.
- Asuncion, A. and Newman, D.J. (2007) *UCI Machine Learning Repository*, School of Information and Computer Science, University of California, Irvine, CA.
- Baccianella, S., Esuli, A. and Sebastiani, F. (2010) ‘Sentiwordnet 3.0: an enhanced lexical resource for sentiment analysis and opinion mining’, *LREC*, Vol. 10, pp.2200–2204.
- Baker, C.F., Fillmore, C.J. and Lowe, J.B. (1998) ‘The Berkeley frameNet project’, *Proceedings of the 36th Annual Meeting of the Association for Computational Linguistics and 17th International Conference on Computational Linguistics*, Vol. 1, Association for Computational Linguistics, Montreal, Quebec, Canada, August, pp.86–90.
- Baker, C.F., Fillmore, C.J. and Lowe, J.B. (1998) ‘The Berkeley frameNet project’, *Proceedings of the 17th International Conference on Computational Linguistics*, Association for Computational Linguistics, Vol. 1, pp.86–90.
- Balahadia, F.F., Fernando, M.C.G. and Juanatas, I.C. (2016) ‘Teacher’s performance evaluation tool using opinion mining with sentiment analysis’, *Region 10 Symposium (TENSYMP)*, IEEE, IEEE, May, pp.95–98.
- Blei, D.M., Ng, A.Y. and Jordan, M.I. (2003) ‘Latent Dirichlet allocation’, *Journal of Machine Learning Research*, Vol. 3, January, pp.993–1022.
- Blitzer, J., Dredze, M. and Pereira, F. (2007) ‘Biographies, bollywood, boom-boxes and blenders: Domain adaptation for sentiment classification’, *ACL*, Vol. 7, June, pp.440–447.
- Burges, C.J.C. (1998) ‘A tutorial on support vector machines for pattern recognition’, *Data Mining and Knowledge Discovery*, Vol. 2, No. 2, pp.121–167.
- Cambria, E., Schuller, B., Xia, Y. and Havasi, C. (2013) ‘New avenues in opinion mining and sentiment analysis’, *IEEE Intelligent Systems*, Vol. 28, No. 2, pp.15–21.
- Ceron, A., Curini, L. and Iacus, S.M. (2016) ‘iSA: a fast, scalable and accurate algorithm for sentiment analysis of social media content’, *Information Sciences*, Vol. 367, pp.105–124.
- Chen, J., Huang, H., Tian, S. and Qu, Y. (2009) ‘Feature selection for text classification with Naïve Bayes’, *Expert Systems with Applications*, Elsevier, Vol. 36, No. 3, pp.5432–5435.
- Chen, L.S., Liu, C.H. and Chiu, H.J. (2011) ‘A neural network based approach for sentiment classification in the blogosphere’, *Journal of Informetrics*, Vol. 5, No. 2, pp.313–322.
- Chikarsal, P., Poria, S. and Cambria, E. (2015) ‘SeNTU: sentiment analysis of tweets by combining a rule-based classifier with supervised learning’, *Proceedings of the International Workshop on Semantic Evaluation*, SemEval, Denver, Colorado, pp.647–651.
- Colace, F., De Santo, M., Greco, L. and Napoletano, P. (2014) ‘Text classification using a few labeled examples’, *Computers in Human Behavior*, Vol. 30, pp.689–697.
- Collomb, A., Costea, C., Joyeux, D., Hasan, O. and Brunie, L. (2014) *A Study and Comparison of Sentiment Analysis Methods for Reputation Evaluation*, Rapport de recherche RR-LIRIS-2014-002, Retrieved from <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.646.9070&rep=rep1&type=pdf>

- Cruz, F.L., Troyano, J.A., Ortega, F.J. and Enríquez, F. (2011) ‘Automatic expansion of feature level opinion lexicons’, *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis, Association for Computational Linguistics*, June, Portland, Oregon, pp.125–131.
- Da Silva, N.F., Hruschka, E.R. and Hruschka, E.R. (2014) ‘Tweet sentiment analysis with classifier ensembles’, *Decision Support Systems*, Vol. 66, pp.170–179.
- Da Silva, N.F.F., Coletta, L.F., Hruschka, E.R. and Hruschka Jr., E.R. (2016) ‘Using unsupervised information to improve semi-supervised tweet sentiment classification’, *Information Sciences*, Vol. 355, pp.348–365.
- Davis, J. and Goadrich, M. (2006) ‘The relationship between precision-recall and ROC curves’, *Proceedings of the 23rd International Conference on Machine Learning*, ACM, June, Pittsburgh, Pennsylvania, USA, pp.233–240.
- Devi, D.N., Kumar, C.K. and Prasad, S. (2016) ‘A feature based approach for sentiment analysis by using support vector machine’, *IEEE 6th International Conference on Advanced Computing (IACC)*, IEEE, February, Bhimavaram, India, pp.3–8.
- Doquire, G. and Verleysen, M. (2013) ‘Mutual information-based feature selection for multilabel classification’, *Neurocomputing*, Vol. 122, pp.148–155.
- Dumais, S.T. (1995) *Latent Semantic Indexing (LSI): TREC-3 Report*, NIST Special Publication SP, Gaithersburg, Maryland, USA, pp.219–219.
- El-Halees, A.M. (2015) ‘Arabic text classification using maximum entropy’, *IUG Journal of Natural Studies*, Vol. 15, No. 1, pp.157–167.
- Esuli, A. (2008) ‘Automatic generation of lexical resources for opinion mining: models, algorithms and applications’, *ACM SIGIR Forum*, ACM, Vol. 42, No. 2, pp.105–106.
- Esuli, A. and Sebastiani, F. (2006) ‘Sentiwordnet: a publicly available lexical resource for opinion mining’, *Proceedings of LREC*, Vol. 6, Citeseer, pp.417–422.
- Esuli, A. and Sebastiani, F. (2007) ‘Random-walk models of term semantics: an application to opinion-related properties’, *Proceedings of LTC 2007*, Poznań, pp.221–225.
- Feldman, R. (2013) ‘Techniques and application for sentiment analysis’, *Communications of the ACM*, ACM, Vol. 56, No. 4, pp.82–89.
- Fernández-Gavilanes, M., Álvarez-López, T., Juncal-Martínez, J., Costa-Montenegro, E. and González-Castaño, F.J. (2016) ‘Unsupervised method for sentiment analysis in online texts’, *Expert Systems with Applications*, Vol. 58, pp.57–75.
- Figueiredo, F., Rocha, L., Couto, T., Salles, T., Gonçalves, M.A. and Meira Jr., W. (2011) ‘Word co-occurrence features for text classification’, *Information Systems*, Vol. 36, No. 5, pp.843–858.
- Fischer, A. and Igel, C. (2012) ‘An introduction to restricted Boltzmann machines’, *Iberoamerican Congress on Pattern Recognition*, Springer Berlin Heidelberg, September, pp.14–36.
- Forman, G. (2003) ‘An extensive empirical study of feature selection metrics for text classification’, *Journal of Machine Learning Research*, Vol. 3, March, pp.1289–1305.
- Francopoulo, G., George, M., Calzolari, N., Monachini, M., Bel, N., Pet, M. and Soria, C. (2006) ‘Lexical markup framework (LMF)’, *International Conference on Language Resources and Evaluation-LREC 2006*, May, Genoa, Italy, Vol. 6, pp.233–236.
- Furnkranz, J. (1998) *A Study Using n-Gram Features for Text Categorization*, Technical Report, OEFAI-TR-9830, Austrian Institute for Artificial Intelligence.
- Gao, F., Sun, X., Wang, K. and Ren, F. (2016) ‘Chinese micro-blog sentiment analysis based on semantic features and PAD model’, *IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS), 2016*, IEEE, Okayama, Japan, pp.1–5.
- Godbole, N., Srinivasaiah, M. and Skiena, S. (2007) ‘Large-scale sentiment analysis for news and blogs’, *ICWSM’2007*, Boulder, Colorado, USA, Vol. 7, pp.21–24.

- Guarino, N. (1998) ‘Some ontological principles for designing upper level lexical resources’, *First International Conference on Language Resources and Evaluation Granada*, Spain, pp.28–30, ArXiv, Preprint Cmp-Lg/9809002.
- Hatzivassiloglou, V. and Wiebe, J.M. (2000) ‘Effects of adjective orientation and gradability on sentence subjectivity’, *Proceedings of the 18th Conference on Computational Linguistics*, Association for Computational Linguistics, Saarbrücken, Germany, Vol. 1, pp.299–305.
- Hearst, M.A. (1992) ‘Direction-based text interpretation as an information access refinement’, *Text-Based Intelligent Systems: Current Research and Practice in Information Extraction and Retrieval*, Lawrence Erlbaum Associates, Mahwah, NJ, pp.257–274.
- Hofmann, T. (2001) ‘Unsupervised learning by probabilistic latent semantic analysis’, *Machine Learning*, Vol. 42, Nos. 1–2, pp.177–196.
- Hopkins, D.J. and King, G. (2010) ‘A method of automated nonparametric content analysis for social science’, *American Journal of Political Science*, Vol. 54, No. 1, pp.229–247.
- Hsu, H.H. and Hsieh, C.W. (2010) ‘Feature selection via correlation coefficient clustering’, *Journal of Software*, Vol. 5, No. 12, pp.1371–1377.
- Hu, M. and Liu, B. (2004) ‘Mining and summarizing customer reviews’, *Proceedings of the Tenth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, August, Washington, USA, pp.168–177.
- Hu, X., Tang, J., Gao, H. and Liu, H. (2014) ‘Social spammer detection with sentiment information’, *2014 IEEE International Conference on Data Mining*, IEEE, December, Shenzhen, China, pp.180–189.
- Hunston, S. and Francis, G. (2000) *A Corpus-Driven Approach to the Lexical Grammar of English*, MIT Press, Amsterdam/Philadelphia, Benjamins.
- Hussein, D.-M.E.D.M. (2016) ‘A survey on sentiment analysis challenges’, *Journal of King Saud University – Engineering Sciences*, <http://dx.doi.org/10.1016/j.jksues.2016.04.002>
- Ide, N. and Véronis, J. (1998) ‘Introduction to the special issue on word sense disambiguation: the state of the art’, *Computational Linguistics*, Vol. 24, No. 1, pp.2–40.
- Jagtap, V.S. and Pawar, K. (2013) ‘Analysis of different approaches to sentence-level sentiment classification’, *International Journal of Scientific Engineering and Technology*, Vol. 2, No. 3, pp.164–170.
- Jang, J.S., Lee, B.I., Choi, C.H., Kim, J.H., Seo, D.M. and Cho, W.S. (2016) ‘Understanding pending issue of society and sentiment analysis using social media’, *Eighth International Conference on Ubiquitous and Future Networks (ICUFN), 2016*, IEEE, July, Vienna, Austria, pp.981–986.
- Jolliffe, I. (2002) *Principal Component Analysis*, John Wiley & Sons, Ltd. Hoboken, New Jersey, USA.
- Katz, G., Ofek, N. and Shapira, B. (2015) ‘ConSent: context-based sentiment analysis’, *Knowledge-Based Systems*, Vol. 84, pp.162–178.
- Kaur, J. and Saini, J.R. (2014) ‘Emotion detection and sentiment analysis in text corpus: a differential study with informal and formal writing styles’, *International Journal of Computer Applications*, Vol. 101, No. 9, pp.1–9.
- Kaushik, C. and Mishra, A. (2014) ‘A scalable, lexicon based technique for sentiment analysis’, *International Journal in Foundations of Computer Science and Technology (IJFCST)*, Vol. 4, No. 5, pp.35–43.
- Khan, F.H., Qamar, U. and Bashir, S. (2016) ‘Swims: semi-supervised subjective feature weighting and intelligent model selection for sentiment analysis’, *Knowledge-Based Systems*, Vol. 100, pp.97–111.
- Kim, S.M. and Hovy, E. (2006) ‘Extracting opinions, opinion holders and topics expressed in online news media text’, *Proceedings of the Workshop on Sentiment and Subjectivity in Text*, July, Association for Computational Linguistics, Sydney, Australia, pp.1–8.

- Kiritchenko, S., Zhu, X. and Mohammad, S.M. (2014) ‘Sentiment analysis of short informal texts’, *Journal of Artificial Intelligence Research*, Vol. 50, pp.723–762.
- Ko, Y. and Seo, J. (2002) ‘Text categorization using feature projections’, *Proceeding of the 19th International Conference on Computational Linguistics, COLING2002*, Taipei, Taiwan, pp.467–473.
- Kouloumpis, E., Wilson, T. and Moore, J.D. (2011) ‘Twitter sentiment analysis: the good the bad and the omg!’, *Icwsmt*, Vol. 11, pp.538–541.
- Landauer, T.K. and Dumais, S.T. (1997) ‘A solution to Plato’s problem: the latent semantic analysis theory of acquisition, induction and representation of knowledge’, *Psychological Review*, Vol. 104, No. 2, pp.211–240.
- Landauer, T.K. and Dumais, S.T. (1997) ‘A solution to Plato’s problem: the latent semantic analysis theory of acquisition, induction and representation of knowledge’, *Psychological Review*, Vol. 104, No. 2, p.211.
- Landauer, T.K., Foltz, P.W. and Laham, D. (1998) ‘An introduction to latent semantic analysis’, *Discourse Processes*, Vol. 25, Nos. 2–3, pp.259–284.
- Landauer, T.K., McNamara, D.S., Dennis, S. and Kintsch, W. (Eds.) (2013) *Handbook of Latent Semantic Analysis*, Psychology Press, UK.
- Lewis, D.D. (1992a) ‘An evaluation of phrasal and clustered representations on a text categorization task’, *SIGIR ’92: Proceedings of the 15th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, New York, NY, USA, pp.37–50.
- Lewis, D.D. (1992b) ‘Feature selection and feature extraction for text categorization’, *HLT ’91: Proceedings of the Workshop on Speech and Natural Language*, Association for Computational Linguistics, Morristown, NJ, USA, pp.212–217.
- Lewis, D.D. (1992c) *Representation and Learning in Information Retrieval*, PhD Thesis, Amherst, MA, USA.
- Li, G., Hoi, S.C., Chang, K. and Jain, R. (2010) ‘Micro-blogging sentiment detection by collaborative online learning’, *2010 IEEE International Conference on Data Mining*, IEEE, Sydney, Australia, pp.893–898.
- Li, S.T. and Tsai, F.C. (2013) ‘A fuzzy conceptualization model for text mining with application in opinion polarity classification’, *Knowledge-Based Systems*, Vol. 39, pp.23–33.
- Li, W. and Xu, H. (2014) ‘Text-based emotion classification using emotion cause extraction’, *Expert System with Applications*, Vol. 41, No. 4, Elsevier, pp.1742–1749.
- Lichman, M. (2013) *UCI Machine Learning Repository*, University of California, <http://archive.ics.uci.edu/ml>
- Lin, Y.S., Jiang, J.Y. and Lee, S.J. (2014) ‘A similarity measure for text classification and clustering’, *IEEE Transactions on Knowledge and Data Engineering*, Vol. 26, No. 7, pp.1575–1590.
- Liu, B. (2012) ‘Sentiment analysis and opinion mining’, *Synthesis Lectures on Human Language Technologies*, Vol. 5, No. 1, pp.1–167.
- Ma, B.L.W.H.Y. and Liu, B. (1998) ‘Integrating classification and association rule mining’, *Proceedings of the Fourth International Conference on Knowledge Discovery and Data Mining*, August, New York, pp.1–7.
- Maas, A.L., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y. and Potts, C. (2011) ‘Learning word vectors for sentiment analysis’, *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Vol. 1, June, Association for Computational Linguistics, Portland, Oregon, USA, pp.142–150.
- Maks, I. and Vossen, P. (2011) ‘A verb lexicon model for deep sentiment analysis and opinion mining applications’, *Proceedings of the 2nd Workshop on Computational Approaches to Subjectivity and Sentiment Analysis*, Association for Computational Linguistics, Oregon, USA, pp.10–18.

- Mamgain, N., Mehta, E., Mittal, A. and Bhatt, G. (2016) ‘Sentiment analysis of top colleges in India using Twitter data’, *International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT)*, IEEE, New Delhi, India, March, pp.525–530.
- Maynard, D. and Funk, A. (2011) ‘Automatic detection of political opinions in tweets’, *Extended Semantic Web Conference*, Springer Berlin Heidelberg, May, pp.88–99.
- Medhat, W., Hassan, A. and Korashy, H. (2014) ‘Sentiment analysis algorithms and applications: a survey’, *Ain Shams Engineering Journal*, Vol. 5, No. 4, pp.1093–1113.
- Mehrabian, A. and Russell, J.A. (1974) *An Approach to Environmental Psychology*, Mehrabian, A. and Russell, J.A. (Eds.), MIT Press, Cambridge, Massachusetts, USA.
- Mesleh, A.M.A. (2007) ‘Chi square feature extraction based SVMs Arabic language text categorization system’, *Journal of Computer Science*, Vol. 3, No. 6, pp.430–435.
- Mladenic, D. and Grobelnik, M. (1998) ‘Words sequences as features in text-learning’, *Proceedings of ERK-98, the 7th Electro Technical and Computer Science Conference*, Ljubljana, SL, pp.145–148.
- Montoyo, A., Martínez-Barco, P. and Balahur, A. (2012) ‘Subjectivity and sentiment analysis: an overview of the current state of the area and envisaged developments’, *Decision Support Systems*, Vol. 53, No. 4, pp.675–679.
- Moreno-Ortiz, A. and Fernández-Cruz, J. (2015) ‘Identifying polarity in financial texts for sentiment analysis: a corpus-based approach’, *Procedia-Social and Behavioral Sciences*, Vol. 198, pp.330–338.
- Muhammad, A., Wiratunga, N. and Lothian, R. (2016) ‘Contextual sentiment analysis for social media genres’, *Knowledge-Based Systems*, Vol. 108, pp.92–101.
- Mullen, T. and Collier, N. (2004) ‘Sentiment analysis using support vector machine with diverse information sources’, *EMNLP*, Vol. 4, pp.412–418.
- Munezero, M.D., Montero, C.S., Sutinen, E. and Pajunen, J. (2014) ‘Are they different? affect, feeling, emotion, sentiment and opinion detection in text’, *IEEE Transactions on Affective Computing*, Vol. 5, No. 2, pp.101–111.
- Nakov, P., Ritter, A., Rosenthal, S., Sebastiani, F. and Stoyanov, V. (2016) ‘SemEval-2016 task 4: sentiment analysis in twitter’, *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016)*, San Diego, USA (Forthcoming).
- Nigam, K., Lafferty, J. and McCallum, A. (1999) ‘Using maximum entropy for text classification’, *IJCAI-99 Workshop on Machine Learning for Information Filtering*, Vol. 1, August, pp.61–67.
- Onan, A., Korukoğlu, S. and Bulut, H. (2016a) ‘A multiobjective weighted voting ensemble classifier based on differential evolution algorithm for text sentiment classification’, *Expert Systems with Applications*, Vol. 62, pp.1–16.
- Onan, A., Korukoğlu, S. and Bulut, H. (2016b) ‘Ensemble of keyword extraction methods and classifiers in text classification’, *Expert Systems with Applications*, Vol. 57, pp.232–247.
- Pang, B. and Lee, L. (2004) ‘A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts’, *Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics*, July, Association for Computational Linguistics, pp.271–278.
- Pang, B. and Lee, L. (2008) ‘Opinion mining and sentiment analysis’, *Foundations and Trends in Information Retrieval*, Vol. 2 Nos. 1–2, pp.1–135.
- Pang, B., Lee, L. and Vaithyanathan, S. (2002) ‘Thumbs up?: sentiment classification using machine learning techniques’, *Proceedings of the ACL-02 Conference on Empirical Methods in Natural Language Processing-Volume 10*, Association for Computational Linguistics, July, pp.79–86.

- Pantel, P. and Lin, D. (2002) ‘Discovering word senses from text’, *Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, ACM, July, Edmonton, AB, Canada, pp.613–619.
- Park, H., Kwon, S. and Kwon, H.C. (2010) ‘Complete Gini-index text (GIT) feature-selection algorithm for text classification’, *2nd International Conference on Software Engineering and Data Mining (SEDM), 2010*, IEEE, June, Chengdu, China, pp.366–371.
- Piryani, R., Madhavi, D. and Singh, V.K. (2017) ‘Analytical mapping of opinion mining and sentiment analysis research during 2000–2015’, *Information Processing & Management*, Vol. 53, No. 1, pp.122–150.
- Powers, D.M. (2011) ‘Evaluation: from precision recall and f-measure to roc informedness markedness and correlation’, *J. Mach. Learn. Technol.*, Vol. 2, No. 1, pp.37–63.
- Prabowo, R. and Thelwall, M. (2009) ‘Sentiment analysis: a combined approach’, *Journal of Informatics*, Vol. 3, No. 2, pp.143–157.
- Qaisi, L.M. and Aljarah, I. (2016) ‘A twitter sentiment analysis for cloud providers: a case study of Azure vs. AWS’, *7th International Conference on Computer Science and Information Technology (CSIT)*, IEEE, July, Amman, Jordan pp.1–6.
- Qiu, G., He, X., Zhang, F., Shi, Y., Bu, J. and Chen, C. (2010) ‘DASA: dissatisfaction-oriented advertising based on sentiment analysis’, *Expert Systems with Applications*, Vol. 37, No. 9, pp.6182–6191.
- Rabab’ah, A.M., Al-Ayyoub, M., Jararweh, Y. and Al-Kabi, M.N. (2016) ‘Evaluating SentiStrength for Arabic sentiment analysis’, *7th International Conference on Computer Science and Information Technology (CSIT), 2016*, IEEE, July, Amman, Jordan, pp.1–6.
- Ravi, K. and Ravi, V. (2015) ‘A survey on opinion mining and sentiment analysis: tasks, approaches and applications’, *Knowledge-Based Systems*, Vol. 89, pp.14–46.
- Rill, S., Reinel, D., Scheidt, J. and Zicari, R.V. (2014) ‘Politwi: early detection of emerging political topics on twitter and the impact on concept-level sentiment analysis’, *Knowledge-Based Systems*, Elsevier, Vol. 69, pp.24–33.
- Rosenthal, S., Nakov, P., Kiritchenko, S., Mohammad, S.M., Ritter, A. and Stoyanov, V. (2015) ‘Semeval-2015 task 10: sentiment analysis in twitter’, *Proceedings of SemEval-2015*, Denver, Colorado, pp.451–463.
- Rosenthal, S., Ritter, A., Nakov, P. and Stoyanov, V. (2014) ‘Semeval-2014 task 9: sentiment analysis in twitter’, *Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014)*, Dublin, Ireland, pp.73–80.
- Ruiz, M.E. and Srinivasan, P. (1999) ‘Hierarchical neural networks for text categorization (poster abstract)’, *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, ACM, August, USA, pp.281, 282.
- Sahu, T.P. and Ahuja, S. (2016) ‘Sentiment analysis of movie reviews: a study on feature selection and classification algorithms’, *International Conference on Microelectronics, Computing and Communications (MicroCom), 2016*, IEEE, January, Durgapur, India, pp.1–6.
- Saif, H., Fernandez, M., He, Y. and Alani, H. (2013) ‘Evaluation datasets for twitter sentiment analysis: a survey and a new dataset, the sts-gold’, *1st International Workshop on Emotion and Sentiment in Social and Expressive Media: Approaches and Perspectives from AI (ESSEM 2013)*, Turin, Italy.
- Schumaker, R.P., Jarmoszko, A.T. and Labedz, C.S. (2016) ‘Predicting wins and spread in the premier league using a sentiment analysis of twitter’, *Decision Support Systems*, Vol. 88, pp.76–84.
- Scott, S. and Matwin, S. (1999) ‘Feature engineering for text classification’, *ICML ’99: Proceedings of the 16th International Conference on Machine Learning*, Morgan Kaufmann Publishers, Inc., San Francisco, CA, USA, pp.379–388.
- Seki, Y. (2016) ‘Use of twitter for analysis of public sentiment for improvement of local government service’, *IEEE International Conference on Smart Computing (SMARTCOMP)*, IEEE, May, Missouri, USA, pp.1–3.

- Serrano-Guerrero, J., Olivas, J.A., Romero, F.P. and Herrera-Viedma, E. (2015) ‘Sentiment analysis: a review and comparative analysis of web services’, *Information Sciences*, Vol. 311, pp.18–38.
- Shaheen, S., El-Hajj, W., Hajj, H. and Elbassuoni, S. (2014) ‘Emotion recognition from text based on automatically generated rules’, *International Conference on Data Mining Workshop (ICDMW)*, IEEE, December, Shenzhen, China, pp.383–392.
- Shamma, D.A., Kennedy, L. and Churchill, E.F. (2009) ‘Tweet the debates: understanding community annotation of uncollected sources’, *Proceedings of the First SIGMM Workshop on Social Media*, ACM, October, Beijing China, pp.3–10.
- Shen, L., Wang, M. and Shen, R. (2009) ‘Affective e-learning: using ‘emotional’ data to improve learning in pervasive learning environment’, *Educational Technology & Society*, Vol. 12, No. 2, pp.176–189.
- Singh, S.R., Murthy, H.A. and Gonsalves, T.A. (2010) ‘Feature selection for text classification based on Gini coefficient of inequality’, *FSDM*, Vol. 10, pp.76–85.
- Strapparava, C. and Mihalcea, R. (2008) ‘Learning to identify emotions in text’, *Proceedings of the 2008 ACM Symposium on Applied Computing*, ACM, Fortaleza, Ceara, Brazil, pp.1556–1560.
- Taboada, M., Brooke, J., Tofiloski, M., Voll, K. and Stede, M. (2011) ‘Lexicon-Based methods for sentiment analysis’, *Computational Linguistics*, MIT Press, Cambridge, Massachusetts, USA, Vol. 37, No. 2, pp.267–307.
- Takahashi, B., Tandoc, E.C. and Carmichael, C. (2015) ‘Communicating on Twitter during a disaster: an analysis of tweets during Typhoon Haiyan in the Philippines’, *Computers in Human Behavior*, Vol. 50, pp.392–398.
- Tallarida, R.J. and Murray, R.B. (1987) ‘Chi-square test’, *Manual of Pharmacologic Calculations*, Springer, New York, pp.140–142.
- Tang, H., Tan, S. and Cheng, X. (2009) ‘Survey on sentiment detection on reviews’, *Expert Systems with Applications*, Vol. 36, No. 7, Elsevier, pp.10760–10773.
- Thelwall, M. (2013) ‘Heart and soul: sentiment strength detection in the social web with sentistrength’, *Proceedings of the CyberEmotions*, Warsaw, Poland, pp.1–14.
- Tosun, A. and Bener, A. (2009) ‘Reducing false alarms in software defect prediction by decision threshold optimization’, *Proceedings of the 2009 3rd International Symposium on Empirical Software Engineering and Measurement*, IEEE Computer Society, October, Florida, USA, pp.477–480.
- Tsytsarau, M. and Palpanas, T. (2012) ‘Survey on mining subjective data on the web’, *Data Mining and Knowledge Discovery*, Vol. 24, No. 3, pp.478–514.
- Uri, Z. (1991) *Lexical Acquisition: Exploiting On-Line Resources to Build Lexicon*, Psychology Press, UK.
- Uysal, A.K. and Gunal, S. (2012) ‘A novel probabilistic feature selection method for text classification’, *Knowledge Based Systems*, Vol. 36, Elsevier, pp.226–235.
- Valitutti, A., Strapparava, C. and Stock, O. (2004) ‘Developing affective lexical resources’, *Psychology Journal*, Vol. 2, No. 1, pp.61–83.
- Vinodhini, G. and Chandrasekaran, R.M. (2012) ‘Sentiment analysis and opinion mining: a survey’, *International Journal of Advanced Research in Computer Science and Software Engineering*, Vol. 2, No. 6, pp.282–292.
- Wang, Z., Yu, Z., Chen, L. and Guo, B. (2014) ‘Sentiment detection and visualization of Chinese micro-blog’, *International Conference on Data Science and Advanced Analytics (DSAA), 2014*, IEEE, October, Shanghai, China, pp.251–257.
- Whitelaw, C., Garg, N. and Argamon, S. (2005) ‘Using appraisal groups for sentiment analysis’, *Proceedings of the 14th ACM Conference on Information and Knowledge Management*, (2005), Bremen, Germany, pp.625–631.

- Wiegand, M., Balahur, A., Roth, B., Klakow, D. and Montoyo, A. (2010) ‘A survey on the role of negation in sentiment analysis’, *Proceedings of the Workshop on Negation and Speculation in Natural Language Processing*, Uppasala, pp.60–68.
- Xianghua, F., Guo, L., Yanyan, G. and Zhiqiang, W. (2013) ‘Multi-aspect sentiment analysis for Chinese online social reviews based on topic modeling and howNet lexicon’, *Knowledge-Based Systems*, Vol. 37, pp.186–195.
- Xiao, S.U.N., Chengcheng, L.I. and Fuji, R.E.N. (2016) ‘Sentiment analysis for chinese microblog based on deep neural networks with convolutional extension features’, *Neuro Computing*, Vol. 210, pp.227–236.
- Yang, T., Li, Y., Pan, Q. and Guo, L. (2016) ‘Tb-CNN: joint tree-bank information for sentiment analysis using CNN’, *Control Conference (CCC), 2016 35th Chinese*, IEEE, July, Chengdu, China, pp.7042–7044.
- Yu, Y. and Wang, X. (2015) ‘World cup.2014 in the twitter world: a big data analysis of sentiments in US sports fans’ tweets’, *Computers in Human Behavior*, Vol. 48, pp.392–400.
- Yuan, S., Wu, J., Wang, L. and Wang, Q. (2016) ‘A hybrid method for multi-class sentiment analysis of micro-blogs’, *13th International Conference on Service Systems and Service Management (ICSSSM), 2016 IEEE*, June, Kunming, China, pp.1–6.
- Zheng, Z., Wu, X. and Srihari, R. (2004) ‘Feature selection for text categorization on imbalanced data’, *ACM SIGKDD Explorations Newsletter*, Vol. 6, No. 1, pp.80–89.

Website

Dice coefficient, https://en.wikipedia.org/wiki/Dice_coefficient (Accessed 7 January, 2016).