Survey on Tasks and Approaches of Sentiment Analysis

[1]C. Muralidharan, [2]R. Anitha
[1,2]Department of Computer Science and Engineering, Sri Venkateswara College of Engineering
Sriperumbudur, Tamilnadu, India-602117
murali20infotech@gmail.com, ranitha@svce.ac.in

Abstract- **With the emergence of internet, people wholeheartedly involves insharing their experiences on the web ranging from the personal thoughts to the global issues. The world is now embellished by social media which paves way to share the user's experience of anything across the world. This enables the market to analyze the sentimentor opinion by excerpting and examining the feedbacks of the users. Sentiment analysis involves in understanding the user's opinions, believes, and feelings towards theindividuals, problems, business and other attributes. This intends to support for making the decision not only for other users but also to the firms. But the acquiescence here is, it is too hard to maintain the review sites and refining the data, as the propagation of sites is too high. This paper intends to present a literal survey on the sentiment analysis. It illustrates the ideas presented in the bunch of articles that are published in the last few years. It is concentrated to deliver the basic tasks, sub-tasks, techniques, methodologies and few applications of the review mining.**

Keywords- **sentiment survey, sentiment extraction, sentiment tasks, review evaluation**

I. INTRODUCTION

With the increase in the usage of internet, the online activities like e-commerce, video conferencing, surveillance, online transactions and reservations, chatting, blogging etc tends to increase which leads to fell up with large amount of data [1] (i.e) Bigdata. These can be examined by techniques such as Text, Web and Data Mining [1]. But analyzing this largedata related to user feedbacks is too clumsy and for this we need a vast variety of approaches to come up with the summary of the opinion. Many social network sites, blogs, repositories, e-commerce sites are available which acts as a platform through which we can observe the public opinions of social activities, politics, market strategy etc[2]. Sentiment analysis (SA) is thestudy about people's opinions and feelings that are expressed towards an individual by the means of texts [3]. Sentiment analysis is also called as review miningis the task to extract, classify and detect the sentiments and feelings of different aspects, which can be expressedas text[4]. It helps in accomplishingthe people's mood concerning politics, market strategy [5], the user satisfaction measurement, movie related predictions.

The most of the online customers rely on the feedback of the existing customer on e-commerce site. These opinions supports not only to the customer but also to the seller and the producer for improving the service and product's standards and quality. For example reviews that are provided on e-commerce sites like Flipkart, Amazon, Trip advisor etc will have impact over the customer opinionfor subscribing or buying the product[6]. Online social media has acquired the offline media rapidly, which emboldensthe people to engage in many discussions with theirindependent thoughts across many issues globally which seems to get the quickest response on all issues from the experts. Widest part is it also act as an intermediary platform for public discussions and ideas sharing with open minded. Thus, these can be used for exploiting the customer or people opinionsby understanding the attitudes of the consumer, market and the current trends [7]. Nowadays, Twitter plays its important role which has 255 million users active over every month and every day above 500 million tweets occurs. It acts as a platform for extracting variety of opinions that are provided by the people from different societies for various purposes such as quality improvements of products, expectation and demand prediction, etc. Usually online customers who wish to buy a product through e-commerce sites rely on other customer feedback before making the decision. The sentiments acquired may be positive or negative and it is not so easy to understand thesentiment by processing all the customer feels manually. To overcome this issue, sentiment analysisis performed by extracting the emotions and tones of the customers [8, 10].

II. RELATED WORK

Pang et al. [1] surveyed more than hundred articles which covers applications and issues of sentiment analysis. The tasks includes viz., extraction of opinions, polarity identification and sentiment classification. Tang et al. [12] confess about four major issues that are related to the review mining such as subjectivity identification, sentiment classification of words, files and beliefs extraction. They used few process as cut-based classifier dependent, Naïve Bayes and MultiNaïve Bayes classifier for getting the subjectivity. O'Leary et al. [89] concentrated on website mining, it discuss about website search, types, extraction of opinion units and its types from the websites, and some applicatrions. A. Montoyo et al. [4] tabulated few issues in the area of sentiment and subjectivity analysis. Tsytsarau et al. [1] come up with a detailed survey on

opinion mining which also elaborates about detection of spams and contradiction analysis. Liu [13] presented about various tasks of sentiment analysis. Major works such as opinion analysis based on aspects, subjectivity analysis and sentiment extraction, lexicon generation, summarization, comparative sentiments.

Cambria et al. [11] showedintricacy of Sentiment Analysisregarding to the current trends. Further, Feldman [9] spotlighted five important problems with sentiment analysis which includes sentence level, document level, and aspect level, comparative and, lexicon level sentiment analysis. Few issues like same review by multiple users, subjectivity and sarcasm are concentrated. Recently, Medhat et al. [3] surveyed about feature selection and discussed about methods of sentiment classification. They outlined more than fifty articles and listed out the tasks, algorithm used, data source and polarity. Three different problems such as Time, Domain and Topic is coined and experimented the polarity score which varies with time [14]. Mohammad et al. developed Amazon Mechanical Turk, it as an Amazon online service, for gaining the common annotations of emotion lexicon. Erroneous annotations are rejected and re-annotated through various validations[15]. The recent examples ofsuchkindofwork to extract review from social media and other online platforms are discussed in [7][8][9].

This survey intends to concentrate on (i) the existing studies of classification of opinion mining approaches and its applications, (ii) To the core it concentrates more onvarious tasks and issues that are notably concentrated by other papers as sentiment extraction from unrefined to refined level, measurement of efficiency of feedbacks, unrelated message detection, formation of lexicons, and aspect extraction (iii) Few feature selection techniques are expressed for sentiment analysis. The organization of the tasks and approaches of the sentiment analysis is shown in Fig. 1.
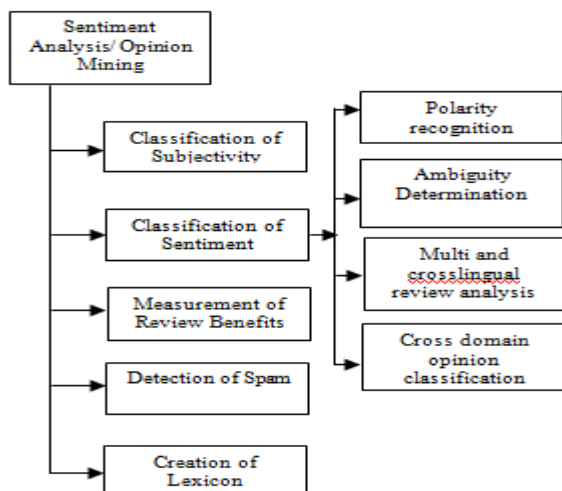

Figure 1. Organization of the Survey

III. PREPARATORY STEPS

Most of the available data such as text, audio, video etc. are being unstructured which imposes greatchallenges in research. The traditional methods of Natural Language Processing i.e. extraction and retrieval of information came into practice [4]. To retrieve the feel from the text, many studies are going on which leads to an automated sentiment analysis. Sentiment extraction is a multi-faceted issue which includes various steps to abstract the sentiment. Acquisition of data and preprocessing of data are the frequent subtasks of analysis, these are concentrated in the below section.

A. Procurement of Data

Data procurement is intensiveto the data typethat are assisted by online publicmedia and the different types of analysis, as the information is from online medias. Twitter like social medias have application program interface to collect the data from their sites. Twitter has an API named Rest for retrieving the data likedetails of twitter user profile, and similarly streaming interface and twitter4J to get tweets that are streamed [17, 8]. Similarly, Facebook Graph API4 is used for Facebook and Tancent API5 for Sina Wiebo . These APIs supports us to getthe information such as posts from their sites [3,19]. Xu et al. [20] collected data about 23,507 friendships and 5,012 members from a commodity feedback website called UrCosme and measures the impacts.

B. Preprocessing of Data

The data collected from different sources will be crude hence it has to be preprocessed before conducting analysis. Few popular preprocessing steps are: tokenization, removal of stop words, stemming, tagging, and extraction of features. A sentence can be broken into symbols, words, phrases, other tokens by removing punctuation marks which is said to be the tokenization. Stemming is the method which finds the root word from the word by neglecting parts of speech from the word. Tagging is another method which tend to find the parts of speech that are available with the words and isthe fundamental step for NLP. The data will be sparse and noise prevalent in it which requires feature extraction. A detailed study of sentiment analyses is highlighted in [3] that explains manyselection techniques for extracting features like chi square, Point wise information and latent semantic indexing. Wang et al. [21] used improved discriminant ratio based on fischer for selecting the feature. Abbasi et al. [22] used multi-lingual sentiment classification viz. Arabic and English. It was performed based on the syntactic features such as word length, measures of vocabulary, characters, special-character and word level features of lexica. Further feature selection based studies surveyed which are implemented using syntactic[1, 23, 24, 25] semantic [23, 25, 27] based methods.

IV. REVIEW OF METHODOLOGY

In sentiment analysis the text is categorizedat two levelsas subjective and objective nature. Subjective nature means that the text contains the decision support opinion and objective

nature means that the text contains some of the facts. In particular, Subjectivity is the tropical based opinions [28].Genetic Algorithm by Das et.al. (2011)got a great success with multiple objective for subjectivity detection. Example- Subjective- The train is comfortable. (It expresses feel which seems to be the opinion). Objective- Government launched bullet train. (It contains a fact) The analysis is done in five ways viz. classification of subjectivity, classification of sentiment, assessment of review benefits, detection of spam and creation of lexicons.

A. *Classification of Subjectivity*

The subjectivity deals with the state of the review or opinion that encloses the sentiments, emotions, believes and speculations. Maks et. al. [24] expressed the subjectivity in different ways.It is a feature which is correlated with the sense of the word. Other sentiment expressing texts are reports, the records related to politics, and online public arguments etc.The problem of differentiating subjective and objective instances is more difficult than classifying the polarity. Hence the subjectivity classification has to be considered precisely for sentiment analysis.

B. *Classification of sentiment*

Sentiment classificationis the assimilation of sentiments from the text in two or more number of classes. It is further classified into four: polarity recognition, ambiguity determination, multi and cross lingual review analysis, Cross-domain opinion classification and Lexicon based polarity identification.

1) Polarity recognition: Classification of sentiment has its major concern which findsthe polarity of the sentences. A sentence may express positive, negative or neutral sentimentsregarding the subjects. Hence, it is coined as polarity recognition. Polarity recognition is accomplished for feedbacks of the product, websites, micro-blogs and articles. In micro blogs the number of words limited to 140 hence it does not contain full sentence and occurs with abbreviated words which produces more noise hence preprocessing has to be done to rectify it [6, 19, 29]. Pang et. al. [30] applied algorithms of ML viz. Naive bayes, Support vector machine and maximum entropy for classifying the sentiments of movie related reviews by means of binary values. They have collected movie reviews from IMDb.com.Zhang et al. [31] used both Naïve bayes and support vector machine for restaurant reviews which is written in Cantonese. Experiments were done using 1500 positive and 1500 negative reviews. They have considered many feature representations like uni-gram, uni-gram frequent, bi-gram, bi-gram frequent, tri-gram, and tri-gram frequent. Wang et al. [32] made a comparison on three methods. It is based on five learners namely naive bayes, decision tree, maximum Entropy, support vector machine and k nearest neighbor and Support Vector Machine for analyzing the sentiments.

2) Ambiguity determination: The user reviews are not written in standard form hence the obscurity and vagueness

occurs which seems to be the major issues. Public discussion will includesarcastic words, detecting such words is too hard and more challenging in the area of opinion mining. The fuzzy methods are useful for representing those expressions. Li et. al. [34] developed a framework using fuzzy based analysis method. For reducing the dimension and feature extraction, uniformity and inverted frequency is used. Reyes et.al. [35] identified six different features for dealing with irony. The six features are ngrams, Parts of Speech n grams, positive, negative, affective and pleasant profiling. Justo et al. [36] tried to remove the nasty words andclassify the sarcasm that exists in the dialogic language. They have combined semantic, linguistic and emotional features in the review for trainingRandom Naïve Bayes andNaive Bayes classifier. Linguistic and sensitive features are captured by Senticnet-3. The prominent features are employed using Chi-square feature selection.

3) Multi and cross lingual review analysis: The expressions may differ for different languages around the world regarding sentiments. This multi and cross lingual analysis has addressed this issue. The accuracy is not high but have explored several approaches that are necessary. Cross-lingual analysis is performed in two ways such as (a) Lexicon-based approach (b) Corpus-based approach. K. Hiroshi et al. [37] developed reviewanalysis methodto japan for understandingthe opinions from the corpus that exists in Japanese language. It doestranslation to English from the Japanese language using the machine language engine. Seki et al. [38] dealt with multi-lingual based opinion identification using the view of author.

*4) Crossdomain opinion classification:*Cross domain analysis includes two domains: source and target domain where the classifier is trained in source and testing is done in target domain [1]. H. Tang et al. [12] usedclassificationmethods for cross domain based opinion analysis. They proposed Frequently Co-occurring Entropy, for selectingthe general characteristics which exists persistently.H. Cho et al. [39] used cross domain opinion analysis by using different dictionaries such as Word net affect [40], Senti word net [41], Word net [49], Opinion-lexicon [3], Sentic net [42], Senti sense [50] together.

5) Lexicon based polarity identification: Turney [33] have implemented two word phrase extraction using two pattern of tags. They developed PMI-Information Retrieval for determining the orientation of semantics in the reviews, where "fair" and "tidy" can be expressed as positive and negative reference words. Whitelaw et al. [27] partially created a lexicon of 1,329 adjectives. They have concentrated on speaker, belief targetand the orientation. Benamara et al. [24] deployeda scoring method using three Adverb-Adjective Combinations i.e. adjective based variable scoring, scoring based on priority of adjective relevant to the adverb, and countingadverb related with the adjective. They have used BBC news articles for the experiment. Popescu et. al. [7] used Google N-gram based corpora for studying the diachronic

phenomenon of two domains such as sports and socio-political reviews.

C. Measurement of Review benefits

The seller or producer may tend to create fake reviews by hiring fake for promoting their product [13]. This is said as spam which can be detected and the benefits of the reviewscan be measured [13]. A spammed review usually concerns only the good quality comments and will not concentrate on negative comments. It is written either to boost their product or to discredit other's product. Ghose et al. [44] determined few features for measuring review helpfulness such assubjectivity, readability measures and spelling errors for identifyingtext features. Multiple reviewer features such as average usefulness of opinions, identity measure ofopinionates are considered. Racherla et al. [45] concentrated on boththe review and the reviewer'scharacteristics for measuring the usefulness. Linear relationship between reviewers, expert, and reputation is used for credibility measure. Mudambi et. al. [46] used regression model for measuring the helpfulness of reviews.

D. Opinion spam detection

The professionals may be engaged for writing false reviews for promoting the product which will increase their sale but will acquire the new user with false attention. These types of reviews are said to be fake or spam or fraudulent review.The professional who writes these types of reviews are known as spammer, if it is done by a group it may be called as group spammers. Most spam detection techniques depends on three features as review content, meta-data, and real-life knowledge of the product [13]. Ott et al. [47] done a spam detection usinga spam dataset.For detecting the unrelated texts they haveused the genre based parts of speech tag, psycho-linguistic detection method and naive bayesand support vector machine using n gram based representation. Fusilier [48] developed an altered non negative and clustered sample related learning methodfor detecting both positive and negative sentiments.

E. Lexica and corpus creation

A collection of words with its stability and polarity value is called as lexicon[1]. It starts with the seed words, this can be further extended using the antonyms and synonyms of the seed words. These are captured with the help of the dictionary word net [49]. This can be repeated for developing the sequence. It include two categories such as ontology and non-ontology based approaches.

*1) Non-ontology based approach:*M. Thelwall et al. [51] developed a lexicon calledsenti strength. The lexica is created based on myspace which recognize ungrammatical spelling andother textual methods. Thelwall et. al. [52] extended the senti strength i.e. senti strength-2. Senti word net-3is an extendedresource for scoring with the sentimental words [41]. It defines three sentiment scores Objective, positive, and negative score as in the WordNet [49].

*2) Ontology based approaches:*Ontology specifies the concepts explicitly. Since it is based on semantics it is considered to be better than other. Hence Sentiment analysis moved drastically towards ontological based approaches. A. C. R. Tsai et al. [53] proposed and developed a dictionary called concept net which uses the regression method. Itis erected from the concept opinion value andpolarity. Word Net-3 isused for induction of lexica to the sentic net-1.E. Cambria et al. [54] proposed and developed version 2 of sentic net, which is based on the computer and social sciences and sentic computing. Poria et al. [55] developed a hybrid model that includes linguistics, concept and common-sense computing for analyzing the opinion Balahur et al. [56] developed a EmotiNet which is based on knowledge to represent and storeemotional reaction in context of real-life. They accomplished few other existing corpora and dictionaries like Senti word net [41], Concept net, Verb ocean and ISEAR for extracting emotional words.

V. FEW APPLICATIONS OF OPINION MINING

A. FOREX market prediction

Bollen et al. [6] analyzed the market prediction by processing thesentiments from twitter to financial data. To analyze the twitter opinions, two tools were used for tracking the feel such as Opinion finder [39] and Mood States for the google profiles which measuresthe feelings by means of six factors.L. Qiu et al. [57] crowd accessfor predicting the market trend. Yu et al. [59] examined the public broadcasts and other medias and interrelations about the firm's performances.

B. Box office prediction

Rui et al. [61] evaluated the twitter sentiment for predicting the sales of the movies. Instead using the post consumption concept they developed the pre consumption approach.Experiments are executedbased on the sales data of the movies from BoxOfficeMojo.com, they have took 4166623 tweets on 63 different movies. SVM is used for the classification of subjectivity and NB was used for classifying the sentiments of the tweets.

C. Business analytics

Coussement et. al. [60] took the emotions that are used in the customer mails which thencombined in the prediction model named churn. It is implemented for subscription of newspaper business. 18,331 electronic mailswas purchased from a Belgium news. 690 positive words, 1,327 negative words are acquired by Linguistic Inquiry and Word Count Logistic Regression (LR).

D. Marketing Intelligence

Market intelligence is an important information regarding the company's trend in the market. It is used for decision making in different aspects such as determining the threats,contestants, acquiringcontestant's actions and other marketingrelated judgments. Li et. al. [5] proposeda method for market perceptions and important tasks like trend detection, classifying opinion, assessing the credibility, and

summarizations. The initial workis implemented by assuming a score, frequency of documents, and patterns ofthe terms. The next workis done at two phases such as classification of subjectsand opinions using support vector machine and with radial basis function kernel and naïve bayes .

VI. CONCLUSION

This paper presents an extensiveliterature review about the previous work that is done withthe different facts of opinion mining. The review is made in six different dimensions i.e. Classification of subjectivity, sentiment, measurement of review benefit, creation of lexicon, aspect based extraction, and various applications of sentiment analysis. These dimensions refers to the tasks that has to be done for analyzing the sentiments. All the sub-tasks are reviewed such as addressed problem, exploitation of dataset, feature selection, several methods and techniques. It provides a clear cut view on the process of sentiment analysis. The approaches such as support vector machine, naïve bayes and lexicon based practices have been exploited comprehensively.

REFERENCES

[1] K. Ravi,V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications",*Knowledge-Based Systems*Vol. 89, pp. 14- 46, 2015.

[2] M.R. Saleh, M.T. Martín Valdivia, A. MontejoRáez, L. A. Ureña López, "Experiments with SVM to classify opinions in different domains", *Expert Systems with Applications*, Vol. 38, pp. 14799–14804, 2011.

[3] W. Medhat, A. Hassan, H. korashy,"Sentiment analysis algorithms and applications: A survey", *Ain Shams Engineering Journal,* Vol. 5, pp. 1093-1113, 2014.

[4] A. Montoyo, P. Martínez Barco, A. Balahur, "Subjectivity and sentiment analysis: An overview of the current state of the area and envisaged developments", *Decision Support Systems*, Vol. 53, pp. 675–679, 2012.

[5] Y. M. Li, T. Y. Li, "Deriving market intelligence from microblogs", *Decision Support Systems*Vol. 55,pp. 206–217, 2013.

[6] J. Bollen, H. Mao, X. Zeng, "Twitter mood predicts the stock market", *Journal of Computational Science*, Vol. 2, pp. 1-8, 2012.

[7] O. Popescu, and C. Strapparava, "Time corpora: Epochs, opinions and changes", *Knowledge-Based Systems*, Vol. 69, pp. 3-13, 2014.

[8] X. Fang and J. Zhan, "Sentiment analysis using product review data," *Journal of Big Data*, Vol. 2, No. 1, pp. 1, 2015.

[9] R. Feldman, "Techniques and Applications for Sentiment Analysis", *Communications of the ACM*, Vol. 56 No. 4, pp. 82-89, 2013.

[10] A. M. Alkalbani, L. Gadhvi, B. Patel, F. K. Hussain, A. M. Ghamry, O. K. Hussain, "Analysing cloud services reviews using opining mining", in *Proc* of *IEEE 31st Int. Conf. on Advanced Information Networking and Applications*, 2017.

[11] E. Cambria, B. Schuller, and Y.-Q. Xia,"New avenues in opinion mining and sentiment analysis", in *Proc* of IJCAI, Buenos Aires, 2015.

[12] H. Tang, S. Tan, X. Cheng, "A survey on sentiment detection of reviews", *Expert Systems with Applications* Vol.36,pp. 10760–10773, 2009.

[13] B. Liu, "Sentiment analysis and opinion mining", *Morgan and Claypool publishers*, 2012.

[14] J. Read, "Using Emotions to Reduce Dependency in Machine Learning Techniques for Sentiment Classification", in *Proc. of the Student Research Workshop*, pp. 43-48, 2005.

[15] B. Pang, L. Lee, and S. Vaithyanathan, "Thumbs up? Sentiment Classification using Machine Learning Techniques", in *Proc. of the Empirical Methods on Natural Language Processing*, pp. 79-86, 2002.

[16] A. Minanovic, H. Gabelica, and ˘Z. Krsti´c, "Big data and sentiment analysis using knime: Online reviews vs. social media","in *proc. of 37th International Convention on. IEEE*", pp. 1464–1468, 2014.

[17] S. Kumar, F. Morstatter, H. Liu, "*Twitter Data Analytics*", August 19, 2013.

[18] F.H. Khan et al., "TOM: Twitter opinion mining framework using hybrid classification scheme", *Decision Support Systems*, Vol. 47, pp. 245- 257,2013.

[19] H. Bao, Q. Li, S. S. Liao, S. Song, H. Gao, "A new temporal and social PMF-based method to predict users' interests in micro-blogging", *Decision Support Systems*, Vol. 55, pp. 698–709, 2013.

[20] Y.C. Xu, C. Zhang, L. Xue, "Measuring product susceptibility in online product review social network", *Decision Support Systems,* 2013.

[21] S. Wang, D. Li, X. Song, Y. Wei, H. Li, "A feature selection method based on improved Fisher's discriminant ratio for text sentiment classification", *Expert Systems with Applications,* Vol. 38, pp. 8696–8702, 2011.

[22] A. Abbasi, H. Chen, A. Salem, "Sentiment Analysis in Multiple Languages: Feature Selection for Opinion Classification in Web Forums", *ACM Transactions on Information Systems*, Vol. 26, No. 3, pp. 12.1- 12.34, 2008.

[23] A. Duric, F. Song, "Feature selection for sentiment analysis based on content and syntax models", *Decision Support Systems,* Vol. 53, pp. 704–711, 2011.

[24] F. Benamara, C. Cesarano, A. Picariello, D.R. Recupero, V.S. Subrahmanian, "Sentiment Analysis: Adjectives and Adverbs are Better than Adjectives Alone", in *Proc. ICWSM*, 2007.

[25] K. Zhang, Y. Xie,Y. Yang, A. Sun, H. Liu, A. Choudhary,"Incorporating conditional random fields and active learning to improve sentiment identification", *Neural Networks*, Vol. 58, pp. 60-67, 2014.

[26] J. Zhang, S. O. Williams, H. Wang, "Intelligent Computing System based on pattern recognition and data mining algorithms", *Sustainable Computing : Informatics and Systems*, 2017

[27] C. Whitelaw, N. Garg, S. Argamon, "Using appraisal groups for sentiment analysis", in *Proc. of the 14th ACM Int. Conf. on Information and Knowledge Management*, pp. 625-631, 2005.

[28] A. Kaur, V. Gupta, "A Survey on Sentiment Analysis and Opinion Mining Techniques", *Journal of Emerging Technologies in Web Intelligence*, Vol. 5, pp. 367-371, 2013.

[29] M. Saif, D. Cody, D. Bonnie, "Generating high-coverage semantic orientation lexicons from overtly marked words and a thesaurus", in *Proc. of the 2009 Conference on EMNLP*, Association for Computational Linguistics, Morristown, NJ, USA, pp. 599– 608,2009.

[30] B. Pang, L. Lee, S. Vaithyanathan, "Thumbs up? sentiment classification using machine learning techniques", in *Proc. of the ACL02 conference on empirical methods in natural language processing,* Association for Computational Linguistics,Vol. 10, pp. 79–86, 2002.

[31] Z. Zhang, Q. Ye, Z. Zhang, Y. Li, "Sentiment classification of Internet restaurant reviews written in Cantonese", *Expert Systems with Applications*,Vol. 38, pp. 7674–7682, 2011.

[32] G. Wang, J. Sun, J. Ma, K. Xu, J. Gu, "Sentiment classification: The contribution of ensemble learning"*, Decision Support Systems*, Vol. 57, pp. 77- 93, 2013.

[33] P. Turney, "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews", in*Proc. of the 40th annual meeting on association for computational linguistics ACL'02*, Stroudsburg, PA, USA: Association for Computational Linguistics, pp. 417–424, 2002.

[34] S.-T. Li, F.-C. Tsai, "A fuzzy conceptualization model for text mining with application in opinion polarity classification", *KnowledgeBased Systems,*Vol. 39, pp. 23–33, 2013.

[35] A. Reyes, P. Rosso, "Making objective decisions from subjective data: Detecting irony in customer reviews", *Decision Support Systems,*Vol. 53, pp. 754–760, 2012.

[36] R. Justo, T. Corcoran, S.M. Lukin, M. Walker, M. Ins Torres, "Extracting relevant knowledge for the detection of sarcasm and nastiness in the social web", *Knowledge-Based Systems,*Vol. 69, pp. 124–133, 2013.

[37] K. Hiroshi, N. Tetsuya, W. Hideo, "Deeper sentiment analysis using machine translation technology", in *Proc. of the 20th Int. conf. on Computational Linguistics*, Association for Computational Linguistics, 2004.

[38] Y. Seki, N. Kando, M. Aono, "Multilingual opinion holder identification using author and authority viewpoints", *Information Processing and Management,*Vol. 45, pp. 189–199, 2009.

[39] H. Cho, S. Kim, J. Lee, J. S. Lee, "Data-driven integration of multiple sentiment dictionaries for lexicon-based sentiment classification of product reviews", *Knowledge Based System*, Vol. 71, pp. 60-71, 2014.

[40] C. Strapparava, A. Valitutti, "WordNet-Affect: an affective extension of WordNet", in *Proc. of LREC*, Vol. 4, pp. 1083–1086, 2004.

[41] S. Baccianella, A. Esuli, F. Sebastiani, "SENTIWORDNET 3.0: An Enhanced Lexical Resource for Sentiment Analysis and Opinion Mining", in *Proc. of LREC-10*, Malta, pp. 2200–2204, 2010.

[42] E. Cambria, R. Speer, C. Havasi, A. Hussain, "SenticNet: A Publicly Available Semantic Resource for Opinion Mining", in *AAAI Fall Symposium: Commonsense Knowledge*, Vol. 10, 2010.

[43] J.C. de Albornoz, L. Plaza, P. Gervas, "Sentisense: An easily scalable concept-based affective lexicon for sentiment analysis", in*Proc. of the 8th Int. Conf. on Language Resources and Evaluation*, pp. 3562–3567, 2012.

[44] A. Ghose and P. G. Ipeirotis, "Estimating the Helpfulness and Economic Impact of Product Reviews: Mining Text and Reviewer Characteristics", *IEEE Transactions on Knowledge and Data Engineering*, Vol. 23, No. 10, 2011.

[45] P. Racherla, W. Friske, "Perceived 'usefulness' of online consumer reviews: An exploratory investigation across three services categories", *Electronic Commerce Research and Applications*,Vol. 11, pp. 548–559, 2012.

[46] S.M. Mudambi, & D. Schuff, "What makes a helpful online review? A study of customer reviews on amazon.com",*MIS Quarterly*, Vol. 34, pp. 185–200, 2010.

[47] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination", in *Proc. of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*, Vol. 1, pp. 309-319, 2011.

[48] D. H. Fusilier, M. Montez-y-Gomez, P. Rosso, R. G. Cabrera,"Detecting positive and negative deceptive opinions using PU-learning", *Information Processing and Management*, Vol. 5, pp. 433- 443, 2014.

[49] G.A. Miller, "WordNet: a lexical database for English", *Communications of the ACM*,Vol. 38, No. 11, pp. 39-41, 1995.

[50] J.C. de Albornoz, L. Plaza, P. Gervas, "Sentisense: An easily scalable concept-based affective lexicon for sentiment analysis", in *Proc. of the 8th International Conference on Language Resources and Evaluation*, pp. 3562–3567, 2012.

[51] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, A. Kappas, "Sentiment Strength Detection in Short Informal Text", *Journal of the American Society for Information Science and Technology*, Vol. 61, no. 12, pp. 2544–2558, 2010.

[52] M. Thelwall, K. Buckley, G. Paltoglou, "Sentiment Strength Detection for the SocialWeb", *Journal of the American Society for Information Science and Technology*, Vol. 63, pp. 163–173, 2012.

[53] A. C. R. Tsai, C. E. Wu, R. T. H. Tsai, J. Y. Hsu, "Building a Concept-Level Sentiment Dictionary Based on common sense knowledge, *IEEE Intelligent Systems*, Vol. 28, pp. 22-30, 2013.

[54] E. Cambria, C. Havasi, A. Hussain, "SenticNet 2: A Semantic and Affective Resource for Opinion Mining and Sentiment Analysis", in *Proc. 25th Int'l Florida Artificial Intelligence Research Society Conf., AAAI*, pp. 202–207, 2012.

[55] S. Poria, E. Cambria, G. Winterstein, G.-B. Huang, "Sentic patterns: Dependency-based rules for concept-level sentiment analysis", *Knowledge-Based Systems*,Vol. 69, pp. 45-63, 2014.

[56] A. Balahur, Jesu´s M. Hermida, Andre´s Montoyo, "Building and Exploiting EmotiNet, a Knowledge Base for Emotion Detection Based on the Appraisal Theory Model", *IEEE Transactions on Affective Computing*, Vol. 3, No. 1, 2012.

[57] L. Qiu, H. Rui, A. Whinston, "Social network-embedded prediction markets: The effects of information acquisition and communication on predictions", *Decision Support Systems,*Vol. 55, pp. 978–987, 2013.

[58] Y. Yu, W. Duan, Q. Cao, "The impact of social and conventional media on firm equity value: A sentiment analysis approach", *Decision Support Systems*,Vol. 55, pp. 919–926, 2013.

[59] K. Coussement, D. V. Poel, "Improving customer attrition prediction by integrating emotions from client/company interaction emails and evaluating multiple classifiers", *Expert Systems with Applications,*Vol. 36, pp. 6127–6134, 2009.

[60] H. Rui, Y. Liu, A. Whinston, "Whose and what chatter matters? The effect of tweets on movie sales", *Decision Support Systems*,Vol. 55, pp. 863–870, 2013.