Opinion Mining Techniques for Non-English Languages: An Overview

Mhaske N.T. *neelimamhaske@gmail.com*
School of Computer Sciences,
North Maharashtra University, Jalgaon,
Maharashtra, 425001, India

Patil A.S. *ajaypatil.nmu@gmail.com*
School of Computer Sciences,
North Maharashtra University, Jalgaon,
Maharashtra, 425001, India

Abstract

The amount of user-generated data on web is increasing day by day giving rise to necessity of automatic tools to analyze huge data and extract useful information from it. Opinion Mining is an emerging area of research concerning with extracting and analyzing opinions expressed in texts. It is a language and domain dependent task having number of applications like recommender systems, review analysis, marketing systems, etc. Early research in the field of opinion mining has concentrated on English language. Many opinion mining tools and linguistic resources have been built for English language. Availability of information in regional languages has motivated researchers to develop tools and resources for non-English languages. In this paper we present a survey on the opinion mining research for non-English languages.

Keywords: Opinion Mining, Sentiment Analysis, Natural Language Processing.

1. INTRODUCTION

The rapid growth of information over web has given rise to the need of tools and techniques to automatically analyze this huge information and represent it in the form that will be convenient for average users to work with. The information over web is available in different forms like numeric data, textual data, images, multimedia contents, etc. How to process and use this information depends on the application for which it is being used.

In this paper we discuss one of the popular areas of research - Opinion Mining / Sentiment Analysis. Being social by nature, people value opinions of other people. What other people think has always been a matter of concern, particularly when making decisions. Before the web, people used to consult with friends, colleagues, family, etc. However nowadays people prefer to use Internet for getting as well as sharing information. Many e-commerce sites provide platform for users to express their opinions about the services and products. Newspaper sites give platform for users to comment on news articles. The information collected from these various user generated data provides insight into what the users think on that matter. This information is useful for commercial industries to understand the user expectations about their services and products, for political parties to analyze common man's opinions on their party or certain candidate, for individuals to make decisions, etc. While using the information on web, the major problem is size of the available data. For a single issue there might be hundreds to thousands of comments / reviews available. Reading all this data manually and analyzing opinions is not feasible and time efficient. Another issue is that of language. The language used in user generated contents is not necessarily grammatically correct. Along with opinions there might also be other information like facts, observations, etc. that needs to be separated from the opinions. The problems in manually analyzing opinion information has lead researchers to develop automatic tools and techniques

that can efficiently process this data and produce opinion summaries. In this paper we focus on the opinion mining work for non-English languages. The papers are chosen so as to explore research carried out to develop resources for languages other than English.

Rest of the paper is organized as follows: Section 2 provide a brief information of the problem definition of opinion mining. Section 3 discuss some previous work. Section 4 focus on construction of linguistic resources used for opinion mining. Section 5 and 6 discuss opinion mining related works in non-English and Indian languages.

2. THE PROBLEM

Opinion mining, also known as sentiment analysis is the area of research that deals with automatic extraction and analysis of subjective information in the text. The subjective information may be in the form of opinions, sentiments, beliefs, emotions, attitudes, etc. Opinion mining is a language and domain dependent task. The sentiment analysis results are influenced by the differences in grammar and usage of language [1]. This problem gives rise to the need of developing new systems and resources for new languages. English is the most studied language in the field of opinion mining. Many linguistic resources are available for analyzing opinions in English language. However the increasing need of automatic opinion mining systems have motivated many researchers to study different languages.

The document on which opinion mining is to be applied, needs to be subjective in nature. The types of documents used for automatic opinion mining includes:

Domain	**Used in**
Product Reviews	[2][3][4][5][6][7]
Movie Reviews	[8][9][10][11]
News articles	[12][13][14][15][16][17]
Social Media Contents	[18][19]

TABLE 1: Domains for Opinion Mining.

The workflow of opinion mining follows the steps:

- Feature Extraction - This step is applied for aspect-based opinion mining where the opinions are extracted around a particular feature / aspect of entity under consideration.
- Subjectivity Classification - This task is related to identifying opinions and facts. The opinion mining systems require input documents to be opinionated. However one cannot be sure if a document contains only opinions. So input documents are first processed to filter opinion information.
- Sentiment Analysis - Sentiment analysis is performed on subjective data. It deals with identifying the overall polarity of subjective text.

For each of the above tasks, different methodologies have been adopted by researchers.

3. RELATED WORK

Opinion mining is a very popular and challenging research area. Many researchers are working on opinion mining systems for different languages with different issues and challenges. The research work in the field of opinion mining have been presented in many surveys. Pang and Lee (2008) presented a comprehensive survey on opinion mining and sentiment analysis research covering the application areas, issues and challenges, techniques used for performing various intermediate steps, etc. [20]. Montoyo et al. (2012) presented an overview of current state of the research in the field of sentiment analysis [21]. The review focus on four major categories of approaches namely, resource construction, text classification, opinion extraction and sentiment analysis. The detailed survey of Tsysarau and Palpanas (2012) presented an overview of web mining and subjectivity analysis algorithms. The study reviewed the most prominent approaches

for the problems of opinion mining and opinion aggregation, contradiction analysis, etc. [22]. In survey on opinion mining framework, Selvam and Abirami (2013) discussed the various steps involved in the opinion mining and provided a brief listing of different techniques used in each step [23]. The survey of Medhat et. al. (2014) provided sophisticated categorization of a large number of techniques used for sentiment analysis [24]. They also discussed available benchmark data sets and categorized them according to their use in certain applications. The latest review of Rana and Cheah (2016) presented a comprehensive overview of different aspect extraction techniques and approaches [25]. The review focus on various techniques for implicit and explicit aspect extraction along with techniques for aspect categorization. The research in the field of opinion mining is continuously increasing with new approaches and new languages.

Indian languages are also gaining interest among researchers to develop natural language processing resources and systems like named entity recognition [26][27], stemmers[28], etc. The purpose of present study is to explore the research in opinion mining that has been carried out for the non-English languages.

4. SENTIMENT LEXICON

Development of opinion mining systems require the knowledge of how opinions are expressed in the target language. Opinions may be expressed at word, sentence, paragraph or document level.

The most common way to express opinions in any language is by using sentiment bearing words like 'excellent', 'good', 'bad', etc. These words can easily be classified as positive or negative. The collection of such sentiment bearing words and phrases provide a valuable information for building opinion mining systems. Sentiment Lexicon is a basic lexical resource used in many systems. The sentiment lexicon is a collection of sentiment words along with their polarity information. The lexicons include:

1. Sentiment Word / Phrase
2. Polarity (Positive/Negative/Neutral)
3. Strength of Polarity (numeric values / degree of intensity in form of strong - weak range)

Sentiment lexicons play an important role in building classifiers that rely on the presence of lexicon entries in the text. The sentiment lexicon may be domain dependent or domain independent. As word senses may change according to contexts, domain oriented lexicons produce better results for the selected domain. Manually constructing the sentiment lexicon for new language is very time consuming and laborious task. So semi-automatic and fully automatic techniques are used for the task. Table 2 summarizes lexicon constructions techniques used for different languages:

Authors (Year)	**Language**	**Resources Used**	**Technique**	**Polarity**
Hiroshi, Tetsuya and Hideo (2004) [29]	Japanese	Product reviews + Bilingual lexicon	Corpus Based	Extraction of Sentiment Units
Mihalcea, Banea and Wiebe (2007) [30]	Romanian	SemCor corpus + Romanian documents	Corpus Based	Projection
Mihalcea, Banea and Wiebe (2007) [30]	Romanian	OpinionFinder	Machine Translation	Inherited from the English Resource
Banea, Mihal-cea and Wiebe	Romanian	Online Romanian dictionary	Wordnet Based	Bootstraping

(2008) [31]				
Bestgen (2008) [32]	French	French lexicon + Newspaper Articles	Corpus Based	SO-LSA DI-LSA
Kim, et al. (2009) [33]	Korean	OpinionFinder + SentiWordNet	Machine Translation	Inherited from the English Resource
Jijkoun and Hofmann (2009) [34]	Dutch	Dutch Wordnet	Wordnet Based	PageRank like algorithm
Waltinger U. (2010) [35]	German	Subjectivity Clues + SentiSpin	Machine Translation	Inherited from the English Resource
Remus, Quasthoff and Heyer (2010) [36]	German	General Inquirer	Machine Translation	PMI technique
Abdul-Mageed and Diab (2012) [37]	Arabic	SentiWordNet + Youtube lexicon + General Inquirer	Machine Translation	From scores of words in English lexicon
Paulo-Santos, Ramos and Marques (2011) [38]	Portuguese	Portuguese Wordnet + Common online Dictionary	Wordnet Based	Graph algorithm

TABLE 2: Sentiment Lexicon Construction Methods.

4.1. Machine Translation Approach

With this approach, existing resources from source language are translated into target language. The translation approach can easily be adopted for languages for which bilingual dictionaries / translation tools are available. As many useful resources are available for English, it is the most preferred source language. Although many sentiment lexicons are constructed for English language, following are the lexicons that are used by many researchers:

- Sentiwordnet[1]: it is a publicly available lexical resource for supporting sentiment classification and opinion mining applications. It is based on Word Net synsets. The strength of polarity is represented in the form of numeric values.
- OpinionFinder (MPQA Subjectivity Lexicon)[2] : this system aims to identify subjective sentences and to mark various aspects of the subjectivity in these sentences. In the subjectivity lexicon each word is assigned its part of speech category with polarity as positive/ negative/neutral and subjectivity strength as weak or strong.
- General Inquirer[3] : is a lexicon attaching syntactic, semantic, and pragmatic information to part-of-speech tagged words.

The polarity information for the target language lexicon can be inherited from the source resource.

The translation based approach is simple and easy to follow. A large size lexicon can be constructed in less time with minimum human efforts. However, the major problem with this

1 http://sentiwordnet.isti.cnr.it/
2 http://mpqa.cs.pitt.edu/lexicons/
3 http://www.wjh.harvard.edu/~inquirer/

approach is that, sense information might be lost in translation. Also there may arise another problem like translation of multiword expressions, translation ambiguity, handling of inflections, etc. that needs to be addressed differently for different languages [30].

4.2. Wordnet/Thesaurus Based Approach

This is another popular approach applicable to the languages for which wordnet / thesaurus is available. Starting with an initial set of seed words with known polarity values, wordnet / thesaurus is used to expand the initial seed set by using lexical relations and propagate the polarity values to unknown words. With wordnet / thesaurus method a large sized lexicon can be constructed. The lexicon may be used for general purpose as well as domain specific applications. The method is computationally difficult and requires the availability of wordnet / thesaurus.

4.3. Corpus Based Method

In the corpus based methods, the lexicon is constructed using corpus in target language and language specific features. The corpus based method relies on association of words. e.g. Turney and Littman (2003) introduced a method for inferring the semantic orientation of a word from its statistical association with a set of positive and negative paradigm words [39]. The statistical measures used are Pointwise Mutual Information (PMI) and Latent Semantic Analysis (LSA). This method is more suited for domain specific applications.

5. APPROACHES TO OPINION MINING

A large number of approaches have been implemented for various tasks in the opinion mining systems. The approaches can broadly be classified into **Lexicon/Rule based** and **Machine Learning based.**

5.1. Lexicon/Rule Based and Statistical Approaches

The lexicon based approaches use sentiment lexicon and set of rules to extract and classify subjective text. The language specific issues, handling of negation, identification of features etc. are modeled in the form of rules. Most of the rule based techniques use term presence and frequency as the basis for classification. The problem with this technique is that, quality of the underlying lexicon affects the performance of the classifier. Sentiment lexicons suffer from inability to consider the specific context in which the words are used [14]. To handle this problem Ding et al. (2008) proposed a holistic approach that exploit external evidences and linguistic conventions of natural language expressions [5].

Use of statistical measures is another popular approach to classify opinions. This approach employ corpus and statistical measures to find co-occurrence patterns to identify and extract opinions. Measures like PMI (Pointwise Mutual Information), LSA (Latent Semantic Analysis), Semantic Orientation, etc. are used for this task. For efficient systems, the statistical techniques require availability of large corpus in the target language which may not be easily available. This approach is used for very few non-English languages. Some of the lexicon based and statistical approaches are summarized in table 3:

Authors (Year)	Language	Task	Technique	Result
Fujii and Ishikawa (2006) [40]	Japanese	Opinion Summarization	Rule Based	Graphical Results
Wang and Araki (2007) [2]	Japanese	Opinion Mining	Modified SO-PMI	Accuracy (Pos:78%, Neg:72%)
Zubaryeva and Savoy (2009) [17]	Japanese Chinese English	Opinion Detection	Statistical Approach using Z Score and Logistic Model	F1 Measure (English:0.54, Chinese:0.82, Japanese:0.62)

Wang and Fu (2010) [16]	Chinese	Sentiment Classification	Sentiment Morphemes based	F Measure(0.40)
Bal et al. (2011) [1]	Dutch and English	Multilingual Sentiment Analysis	Lexicon Based	Accuracy (Dutch:79%, English:71%)

TABLE 3: Lexicon Based / Statistical Approaches to Opinion Mining Tasks.

5.2. Machine Learning Based Approaches

In machine learning approaches the opinion mining task is viewed as text classification problem and various machine learning algorithms are employed using different features. This approach involves construction of training data for the classification task. The supervised approaches use data annotated with sentiment labels while the unsupervised approaches use unannotated data. The next step is to represent the training data in the form of feature vectors which are used to train the classifier. Finally, the trained classifier is used to predict opinions in unseen documents / texts. Table 4 summarizes machine learning techniques adopted for different languages.

Authors (Year)	**Language**	**Task**	**Technique**[4]	**Result**
Kobayashi et al. (2005) [3]	Japanese	Opinion Extraction	SVM	F-measure (Attr.-Value Pairs:58)
Wang and Araki (2007) [7]	Japanese	Opinion Mining	NB + modified SO-PMI	Graphical Results
Kanamaru et al. (2007) [13]	Japanese	Opinion Extraction	SVM	Accuracy (Opinion Extraction:42.88%, Op. holder extraction:14.31%, Polarity:19.90%, Relevance judgment:63.15%)
Atteveldt et al. (2008) [14]	Dutch	Sentiment Analysis	ME	F1 Score(0.63)
Zhang et al. (2008) [41]	Chinese	Sentiment Classification	Lexicon Based + SVM + NB + DT	Accuracy (SVM:81.65%, Lexicon Based:76.26%)
Mellebeek et al. (2010) [6]	Spanish	Opinion Mining	Classifiers implemented in Mallet and Weka	Use of non-expert annotations through crowd sourcing is a viable and cost effective alternative to the use of expert annotations
Martinez-Camara et al. (2011) [10]	Spanish	Opinion Classification	SVM+NB+BBR+ KNN+C4.5	Highest Precision with BBR (87%)

[4] SVM-Support Vector Machine, NB-Naive Bayes, ME-Maximum Entropy, DT-Decision Tree, BBR-Bayesian Binary Regression, KNN-K Nearest Neighbor

Lee and Renganathan (2011) [4]	Chinese	Sentiment Analysis	ME	Accuracy (0.87%)
Rushdi- Saleh et al. (2011) [9]	Arabic + English	Opinion Mining	SVM + NB	F1 Measure (SVM without stemmer:0.90%)
Hamouda and El-taher (2013) [15]	Arabic	Sentiment Analyzer	DT + SVM + NB	Accuracy (SVM:73.4%)
Yang and Chao (2014) [11]	Chinese	Sentiment Analysis	PMI + SVM	Average balanced accuracy rate (0.7)
Duwairi and Qarqaz (2014) [19]	Arabic	Sentiment Analysis	SVM + NB + KNN	Highest Precision (SVM:75.25%), Highest Recall (KNN: K=10, 69.04%)
Ibrahim et al. (2015) [18]	Arabic	Sentiment Analysis	SVM	Accuracy (95%)

TABLE 4: Machine Learning Techniques for Opinion Mining.

The quality and size of training data highly affect the performance of the classifier. While working with machine learning algorithms it is necessary to define useful features for the training and testing instances. The selection of features affects the classification results. The most commonly used features for opinion mining tasks are bag of words, appraisal phrases, n-grams, sentiment words, term frequency, term presence, sentiment word position, valence shifters, etc. Although many types of machine learning algorithms have been experimented for the opinion mining task, the two popular algorithms used in English as well as non-English languages are:

Support Vector Machines	[3][9][10][11][13][15][18][19][41] It is a supervised binary classifier that takes as input set of examples with associated class labels. The machine construct a hyperplane that separate the classes. The new examples are categorized based on this hyperplane.
Naive Bayes	[7][9][10][15][19][41] It is a simple probabilistic classifier that estimates the class-conditional probability by assuming that the attributes are conditionally independent. The algorithm make use of Bayes Theorem to predict the probability of unknown instances.

TABLE 5: Popular Machine Learning Algorithms for Opinion Mining.

6. OPINION MINING FOR INDIAN LANGUAGES

Although many non-English languages are being studied for building automated opinion mining systems, very little work in this field has been reported for Indian languages. Das and Bandyopadhyay (2009) initiated a study in sentiment analysis of Bengali language on the news and blog corpus [12]. Using the translation approach sentiment lexicon was constructed and rule based technique was employed for subjectivity annotation. The final classifier was built using conditional random fields that resulted in precision values of 72.16% and 74.6% for the news and blog domains respectively. Continuing the work, they implemented single-document opinion

summarization system for Bengali with topic based document-level theme relational graphical representation [42].

For Hindi language, Bakliwal et al. (2012) proposed a graph based method to generate the Hindi subjectivity lexicon [43]. The Hindi WordNet, synonym and antonym relations and simple graph traversal were exploited to construct the subjectivity lexicon. The proposed algorithm achieved 79% accuracy on classification of reviews and 70.47% agreement with human annotators. Another lexical resource for Hindi, Hindi-SentiWordNet (H-SWN), presented by Joshi et al. (2010) was constructed by linking Hindi wordnet with English SentiWordNet [8]. To evaluate the lexicon, three techniques viz. in language, machine translation based and resource-based sentiment analysis were applied on movie reviews. Out of the three techniques, In-language approach out performed others with accuracy of 78.14%.

Kaur and Gupta (2014) proposed a lexicon based algorithm of sentiment analysis for Punjabi text [44]. Translation based approach was adopted for lexicon construction with Hindi Subjectivity lexicon as the source lexicon. The resulting system achieved F1 score of 0.67.

Mhaske and Patil (2016) reported various issues and challenges in analyzing opinions in Marathi text [45]. Other languages are also being explored for developing opinion mining resources and systems.

7. CONCLUSION

Research in development of opinion mining systems is continuously increasing with the growing availability of subjective data in different languages. With new approaches, new issues and challenges get emerged. Although the English language has dominance in the field of opinion mining, developing opinion mining resources for non-English languages is also gaining interest among researchers.

This paper presented an overview of the opinion mining tasks and techniques implemented to construct opinion mining resources and systems for non-English languages. The purpose of the present study is to review different techniques employed for different languages so as to provide direction to develop resources and systems for new languages. The area of opinion mining has attracted many researchers due to its practical applications and need to automate the analysis process. The major challenge in opinion mining of non-English languages is the unavailability of linguistic resources. Present review discuss various methodologies adopted by researchers to develop opinion mining related resources for different languages. The study will be helpful to decide on the adoption of technique to be applied for new languages.

8. REFERENCES

[1] D. Bal, M. Bal, A. Bunningen, A. Hogenboom, F. Hogenboom, and F. Frasincar. “Sentiment Analysis with a Multilingual pipeline," LNCS, no. 6997, pp. 129-142, 2011.

[2] G. Wang and K. Araki. “Modifying SO-PMI for Japanese Weblog Opinion Mining by using a Balancing Factor and Detecting Neutral Expressions," in Proceedings of NAACL HLT 2007, Companion Volume, Association for Computational Linguistics, 2007, pp. 189-192.

[3] N. Kobayashi, R. Iida, K. Inui, and Y. Matsumoto. “Opinion Extraction Using a Learning-based Anaphora Resolution Technique," in The Second International Joint Conference on Natural Language Processing (IJCNLP), Companion Volume to the Proceeding of Conference including Posters/Demos and Tutorial Abstracts, 2005, pp. 175-180.

[4] H. Y. Lee and H. Renganathan. “Chinese Sentiment Analysis using Maximum Entropy," in Proceedings of the Workshop on Sentiment Analysis where AI meets Psychology (SAAIP), IJCNLP, 2011, pp. 89-93.

[5] X. Ding, B. Liu, and P. S. Yu. “A Holistic Lexicon-Based Approach to Opinion Mining," in Proceedings of the 2008 International Conference on Web Search and Data Mining, WSDM '08, ACM, 2008, pp. 231-240.

[6] B. Mellebeek, F. Benavent, J. Grivolla, J. Codina, M. R. Costa-jussa, and R. Banchs. “Opinion Mining of Spanish Customer Comments with Non-Expert Annotations on Mechanical Turk," in Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk, Association for Computational Linguistics, 2010, pp. 114-121.

[7] G. Wang and K. Araki. “OMS-J: An Opinion Mining System for Japanese Weblog Reviews using a Combination of Supervised and Unsupervised Approaches," in NAACL HLT Demonstration Program, Association for Computational Linguistics, 2007, pp. 19-20.

[8] A. Joshi, A. R. Balamurali, and P. Bhattacharyya. “A Fall-Back Strategy for Sentiment Analysis in Hindi: A Case Study," in Proceedings of ICON 2010: 8th International Conference on Natural Language Processing, Macmillan Publishers, 2010.

[9] M. Rushdi-Saleh, M. T. Martin-Valdivia, L. A. Urena-Lopez, and J. M. Perea-Ortega. “Bilingual Experiments with an Arabic-English Corpus for Opinion Mining.," in Proceedings of Recent Advances in Natural Language Processing, 2011, pp. 740-745.

[10] E. Martinez-Camara, M. T. Martin-Valdivia, and L. A. Urena-Lopez. “Opinion Classification Techniques applied to a Spanish Corpus," LNCS, vol. 6716, pp. 169-176, 2011.

[11] H. Yang and A. Chao. “Sentiment Analysis for Chinese Reviews of Movies in Multi-Genre based on Morpheme-based Features and Collocations," Information Systems Frontiers, vol. 17, no. 6, pp. 1335-1352, 2014.

[12] A. Das and S. Bandyopadhyay. “Subjectivity Detection in English and Bengali: A CRF-based Approach," in Proceedings of ICON-2009: 7th International Conference on Natural Language Processing, Macmillan Publishers, 2009.

[13] T. Kanamaru, M. Murata, and H. Isahara. “Japanese Opinion Extraction System for Japanese Newspapers using Machine-Learning Method," in Proceedings of NTCIR-6 Workshop Meeting, 2007, pp. 301-307.

[14] W. Atteveldt, J. Kleinnijenhuis, N. Ruigrok, and S. Schlobach. “Good News or Bad News? Conducting Sentiment Analysis on Dutch Text to Distinguish between Positive and Negative Relations," Journal of Information Technology & Politics, vol. 5, no. 1, pp. 73-94, 2008.

[15] A. Hamouda and F. El-taher. “Sentiment Analyzer for Arabic Comments System," International Journal of Advanced Computer Science and Applications, vol. 4, no. 3, pp. 99-103, 2013.

[16] X. Wang and G. Fu. “Chinese Sentence-Level Sentiment Classification based on Sentiment Morphemes," in International Conference on Asian Language Processing, IEEE Computer Society, 2010, pp. 203-206.

[17] O. Zubaryeva and J. Savoy. “Investigation in Statistical Language-Independent Approaches for Opinion Detection in English, Chinese and Japanese," in Proceedings of CLIAWS3, Third International Cross Lingual Information Access Workshop, Association for Computational Linguistics, 2009, pp. 38-45.

[18] H. S. Ibrahim, S. M. Abdou, and M. Gheith. “Sentiment Analysis for Modern Standard Arabic and Colloquial," International Journal on Natural Language Computing, vol. 4, no. 2, pp. 95-109, 2015.

[19] R. M. Duwairi and I. Qarqaz. “Arabic Sentiment Analysis using Supervised Classification," in The 1st International Workshop on Social Networks Analysis, Management and Security, 2014, pp. 1-10.

[20] B. Pang and L. Lee. “Opinion Mining and Sentiment Analysis," Foundations and Trends in Information Retrieval, vol. 2, no. 1-2, pp. 1-135, 2008.

[21] A. Montoyo, P. Martinez-Barco, and A. Balahur. “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, 2012.

[22] M. Tsytsarau and T. Palpanas. “Survey on Mining Subjective Data on the Web," Data Mining and Knowledge Discovery, vol. 24, no. 3, pp. 478-514, 2012.

[23] B. Selvam and S. Abirami. “A Survey on Opinion Mining Framework," International Journal of Advanced Research in Computer and Communication Engineering, vol. 2, no. 9, pp. 3544-3549, 2013.

[24] W. Medhat, A. Hassan, and H. Korashy. “Sentiment Analysis Algorithms and Applications: A Survey," Ain Shams Engineering Journal, vol. 5, pp. 1093-1113, 2014.

[25] T. A. Rana and Y. Cheah. “Aspect Extraction in Sentiment Analysis: Comparative Analysis and Survey," Artificial Intelligence Review, pp. 1-25, 2016.

[26] N. Patil, A. Patil, and B. Pawar, “Survey of Named Entity Recognition Systems with respect to Indian and Foreign Languages”, International Journal of Computer Applications, vol. 134, no. 16, pp.21-26, 2016.

[27] N. Patil, A. Patil, and B. Pawar, “Issues and Challenges in Marathi Named Entity Recognition”, International Journal on Natural Language Computing (IJNLC), vol. 5, no. 1, pp.15-30, 2016.

[28] H. Patil, B. Pawar, and A. Patil, “A Comprehensive Analysis of Stemmers Available for Indic Languages”, International Journal on Natural Language Computing (IJNLC), vol. 5, no. 1, pp.45-55, 2016.

[29] K. Hiroshi, N. Tetsuya, and W. Hideo. “Deeper Sentiment Analysis using Machine Translation Technology," in Proceedings of the 20th International Conference on Computational Linguistics, COLING '04, Association for Computational Linguistics, 2004.

[30] R. Mihalcea, C. Banea, and J. Wiebe. “Learning Multilingual Subjective Language via Cross-Lingual Projections," in Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, Association for Computational Linguistics, 2007, pp. 976-983.

[31] C. Banea, R. Mihalcea, and J. Wiebe. “A Bootstrapping Method for Building Subjectivity Lexicons for Languages with Scarce Resources," in Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08), European Language Resources Association (ELRA), 2008.

[32] Y. Bestgen. “Building Affective Lexicons from Specific Corpora for Automatic Sentiment Analysis," in Proceedings of the Sixth International Conference on Language Resources and Evaluation, 2008, pp. 496-500.

[33] J. Kim, H. Jung, S. Nam, Y. Lee, and J. Lee. “Found in Translation: Conveying Subjectivity of a Lexicon of one Language into another using a Bilingual Dictionary and A Link Analysis Algorithm," ICCPOL 2009, LNAI 5459, 2009, pp. 112-121.

[34] V. Jijkoun and K. Hofmann. “Generating a non-English Subjectivity Lexicon: Relations that Matter," in Proceedings of the 12th Conference of the European Chapter of the ACL, Association for Computational Linguistics, 2009, pp. 398-405.

[35] U. Waltinger. “GermanPolarityClues: A Lexical Resource for German Sentiment Analysis," in Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC), 2010, pp. 1638-1642.

[36] R. Remus, U. Quasthoff, and G. Heyer. “SentiWS - a Publicly available German-Language Resource for Sentiment Analysis," in Proceedings of the Seventh International Conference on Language Resources and Evaluation (LREC'10), European Language Resources Association (ELRA), 2010.

[37] M. Abdul-Mageed and M. T. Diab. “Toward Building a Large-Scale Arabic Sentiment Lexicon," in Proceedings of the Sixth International Global Wordnet Conference, 2012, pp. 18-22.

[38] A. Paulo-Santos, C. Ramos, and N. C. Marques. “Determining the Polarity of Words through a Common Online Dictionary," EPIA 2011, LNAI 7026, 2011, pp. 649-663.

[39] P. D. Turney and M. L. Littman. “Measuring Praise and Criticism: Inference of Semantic Orientation from Association," ACM Trans. Inf. Syst., vol. 21, no. 4, pp. 315-346, 2003.

[40] A. Fujii and T. Ishikawa. “A System for Summarizing and Visualizing Arguments in Subjective Documents: Toward Supporting Decision Making," in Proceedings of the Workshop on Sentiment and Subjectivity in Text, Association for Computational Linguistics, 2006, pp. 15-22.

[41] C. Zhang, D. Zeng, Q. Xu, X. Xin, W. Mao, and Fei-Yue. “Polarity Classification of Public Health Opinions in Chinese," ISI 2008 Workshops, LNCS 5075, 2008, pp. 449-454.

[42] A. Das and S. Bandyopadhyay. “Opinion Summarization in Bengali: A Theme Network Model," in IEEE International Conference on Social Computing/PASSAT, IEEE Computer Society, 2010, pp. 675-682.

[43] A. Bakliwal, P. Arora, and V. Varma. “Hindi Subjective Lexicon: A Lexical Resource for Hindi Polarity Classification," in 24th International Conference on Computational Linguistics (CICLing), 2012, pp. 1189-1196.

[44] A. Kaur and V. Gupta. “Proposed Algorithm of Sentiment Analysis for Punjabi Text," Journal of Emerging Technologies in Web Intelligence, vol. 6, no. 2, pp. 180-183, 2014.

[45] N. Mhaske and A. Patil. “Issues and Challenges in Analyzing Opinions in Marathi Text," IJCSI International Journal of Computer Science Issues, vol. 13, no. 2, pp. 19-25, 2016.