Scientific Journal of Impact Factor (SJIF): 4.72
e-ISSN (O): 2348-4470
p-ISSN (P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 4, Issue 4, April -2017

A Comprehensive Analysis of various techniques used for Sentiment Classification

Foram K. Kalyani[1], Prof. Tosal M. Bhalodia[2]

[1]*M.E. Student, Atmiya Institute of Technology and Science, Rajkot, Gujarat, India*
[2]*Department of Computer Engineering, Atmiya Institute of Technology and Science, Rajkot, Gujarat, India*

Abstract — *With the beginning of Web, various social media sites such as review websites, blogs, and community websites have been developed to gather the opinions of the user. This websites, blogs determine the sentiment or opinion of the user towards product or service. Recent survey conveys that online reviews of user have greater impact than offline reviews. People lean more on the opinions of the user. This paper presents a detailed analysis of approaches or techniques used in classification of sentiments. The paper describes various approaches of Machine learning and Neural network used for sentiment classification.*

Keywords – *Classification, Emotion AI, Machine Learning, Neural Network, Opinion Mining, Sentiment Analysis*

I. INTRODUCTION

With the growing use of Internet and online activities such as social media, chatting, ticket booking, shopping, etc. leads to mining, transforming, loading and analyzing of data using various mining techniques. Various blogs and websites act as platform to express opinion which can be later used for understanding the sentiments of the user.
To achieve this, researchers are working on sentiment analysis from last many years. Sentiment Analysis (sometimes referred as Opinion Mining or Emotion AI) refers to the study of opinions, sentiments, emotions expressed in content towards an entity. Sentiment analysis aims to determine the mind-set of the orator or writer with respect to some subject or document or overall polarity.
Nowadays, customers on e – commerce websites rely more on opinions from existing users, producers and service providers which are used to analyze the customer's opinion about the product or service to improve the quality. Various e – commerce website such as Amazon, IMDb, Yelp, etc. influences the users' choice in buying the product.
This paper provides detailed study of approaches used in sentiment classification. The paper is further divided into 8 main sections. Section II describes Literature Survey. Section III represents Sentiment Analysis. Section IV describes Feature selection Section V describes Classification Approaches. Section VI describes Evaluation of Sentiment Classification. Section VII describes Issues and Future directions. Section VIII provides Conclusions.

II. LITERATURE SURVEY

Pang and Lee [1] executed a survey of many papers by covering applications, challenges for sentiment analysis, and tasks of opinion mining such as opinion extraction, sentiment classification, polarity determination, and summarization.
Then, Tang et al. [2] discussed four problems related to opinion mining, i.e., subjectivity classification, word sentiment classification, document sentiment classification and opinion extraction.
For subjectivity classification, they highlighted some approaches like similarity dependent, NB classifier, Multiple NB classifier, and cut-based classifier [3].
Montoyo et al. [4] gave some issues along with achievements gained in the area of subjectivity analysis and sentiment analysis. Tsytsarau and Palpanas [5] presented a survey on Sentiment Analysis which described opinion mining, opinion aggregation including spam detection and contradiction analysis. They even represented comparison of various opinion mining methods, which were employed on some common dataset.
Liu [6] represented different tasks and works published in Sentiment Analysis and opinion mining. Major tasks described includes subjectivity and sentiment classification, feature level Sentiment Analysis, sentiment lexicon generation, opinion summarization, analysis of comparative opinions, opinion search and retrieval, opinion spam detection and quality of reviews.
Feldman [7] paid more attention on five specific problems in the field of Sentiment Analysis: Document-level, Sentence-level, Aspect-based, Comparative and, sentiment lexicon acquisition. They also described issues like Sentiment Analysis of composition statement, Automatic entity recognition, Multi-entity in same review, sarcasm detection and subjectivity classification at finer level.
Medhat et al. [8] represented a survey on feature selection and sentiment classification methods. A very brief description about methods is presented.

M. Ghiassia, J. Skinnerb, D. Zimbraa [9] combined Artificial Neural Network (ANN) with Natural Language Processing (NLP) and proved that ANN gives better accuracy and Speed compared to Machine Learning.
Anuj Sharma, Shubhamoy Dey [10] used Self Organized Maps (SOM) for analyzing the sentiments and found that supervised learning gives better results than unsupervised.
Soumith Chintala, [11] classified sentiment involving recognition of emotions from text and described that negative comments can be detected on blog by using Neural Network.

III. SENTIMENT ANALYSIS

Sentiments or Opinions depicts the behavior of human. People lean to know others' opinions whenever they want to make decision. Many multinational companies and organizations always get the peoples' opinion about their product or services, even while political election, the opinion of people about the leader is important before making decision.
The Opinion is individuals' view, idea or judgment about any topic or item. The opinion or sentiment can be positive or negative. The sentiment contains three main features: source of the opinion, object or feature, and polarity of opinion i.e. positive, negative or Neutral. The opinion mining can be performed at three levels:

3.1 Document Level
In Document Level sentiment analysis, the opinion is considered as a document and then it is classified as positive, negative or neutral. For Example, consider a product review; based on the reviews of the user, we classify the product review as positive, negative or Neutral. The main problem with this level is that entire review is considered as a single subject. Thus, the review containing multiple subjects cannot be classified.

3.2 Sentence Level
In Sentence Level sentiment analysis, the opinion is considered as a sentence. The sentence is then classified as positive, negative or neutral polarity. Neutral polarity contains sentences that have unrelated words. The sentence level opinions contain subjectivity and objectivity which describes the factual information of sentence and subjective opinions. But, there are certain complex sentences in which the subjectivity/objectivity is not properly defined; hence, sentiment classification fails in such cases.

3.3 Feature or Aspect Level
Feature/ Aspect Level classification, also referred as phrase level classification, performs fine – grained analysis. Feature level analysis is directly performed on the opinion rather than on sentence or paragraph. The feature level analysis is done on specific feature of entity rather than analysis done on entire entity. For Example, the statement "The processor of phone is very fast but the camera quality is not good" has both positive and negative opinions. The above statement is positive for the phone but due to camera quality it is negative. The main aim of feature level analysis is to identify the polarity of entities and provides structured overview of opinion about entity. Hence, feature level analysis is used for converting unstructured data into structured data.

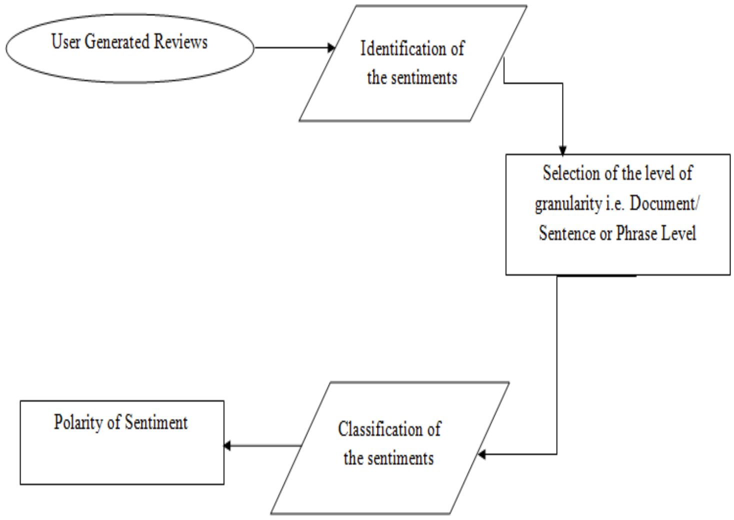

Figure 1. The Process of Sentiment Classification

3.3.1 Sentiment Analysis Tasks
The feature level analysis is performed by three main tasks:

3.3.1.1 Aspect Identification

The aspect identification identifies and extracts relevant subject from the entire text. In [13], Hu and Liu presented a technique based on NLP. In their proposed system, syntax tree parsing and part-of-speech (POS) tagging are used to detect nouns and noun phrases (NP). Then, the most frequent nouns and NPs are identified by using frequent item set mining. Then using distinct linguistic rules, the discovered sets of nouns and NPs are filtered [14].

3.3.1.2 Sentiment Classification

After aspect identification, now the sentiment classification is done which describes the semantic orientation of aspect. Ding et al. [15] proposed system uses a lexicon and rule-based approach. Their method depends on opinion words, a list of positive and negative words contained in sentiment word dictionary which was used to determine the semantic orientation of words in the opinion. Linguistic rules have been provided to change the orientation which handles negative word such as “no” or “not”.

3.3.1.3 Summary generation

After classification is performed, the results are summarized and presented in such a way that it is easily understandable by user. Bing Liu defines a aspect-based opinion summary [16], [17], which shows various bar charts containing number of positive and negative reviews about each and every aspect of a single entity.

3.3.2 Feature Selection

Feature selection includes many methods like TF – IDF (Term Frequency – Inverse Document Frequency), Information Gain, Mutual Information, Chi – Square, Count vector, Feature Vector, Unigram, Bigram, and N – gram methods.

3.3.2.1 Count Vector

Count Vector specifies the number of occurrences of feature in review or comment. In [18], the author used two feature selection methods: TF – IDF and Chi – Square. TF – IDF is used for balancing the weights while Chi – Square provides good result for both positive and negative category.

3.3.2.2 TF – IDF

TF – IDF is product of frequency of word in review (TF) and frequency of word in corpus (IDF). The formula for TF – IDF is:

$$TF - IDF_i = t_{ij} * \log(N/df_i) \tag{1}$$

Where, $TF - IDF_i$ is weight of i terms. t_{ij} is frequency of i terms in j samples. N is total number of samples in corpus and df_i is number of samples containing i term [19].

3.3.2.3 Information Gain

Information Gain is used to identify relevant feature for sentiment analysis. The formula for Information Gain is:

$$IG(f,c) = - \Sigma_{c,c} P(c) \log P(c) + \Sigma_{f,f} P(f) \Sigma_{c,c} P(c|f) \log P(c|f) \tag{2}$$

Where, P(c|f) is joint probability, c is class and f is feature, and P(c) is marginal probability [19].

3.3.2.4 Mutual Information

Mutual Information is used to select feature that are not equally distributed across different classes. The formula for Mutual Information is:

$$MI(f,c) = \Sigma_{c \epsilon C} \Sigma_f P(f,c) \log P(f,c) / P(f) P(c) \tag{3}$$

Where, P(f,c) is joint probability distribution function. C is positive and negative classes. P(f) and P(c) is marginal probability distribution of f and c [19].

3.3.2.5 Chi – Square

Chi – Square method is measure of observed count and expected count and analyzes the deviation among them. The formula for Chi- square is:

$$\lambda^2(f,c) = N(WZ - YX)^2 / (W+Y) * (X+Z) * (W+X) \quad * \ (Y+Z) \tag{4}$$

Where, W, X, Y, Z is frequencies. f is feature and c is class. N represents the presence or absence of feature in sample [19].

IV. CLASSIFICATION APPROACHES

The sentiment classification can be performed by three main approaches:

4.1 Machine Learning approach

Machine Learning consists of four stages. Firstly, the training data is obtained. Second, the reviews of the user are then presented in the form of feature vector. Third, using sentiment features, the training of sentiment is performed. Finally, the trained classifier is used to classify the sentiments of user.

The Machine Learning Approach is further classified into two main approaches: Supervised Learning approach and unsupervised Learning approach

4.1.1 Supervised Learning approach

Supervised Learning approach infers a function from labeled training data. The major algorithms contained in Supervised Learning includes: Probabilistic classifiers, linear classifiers, decision tree classifiers and rule – based classifiers.

While, Unsupervised Learning includes: K – Means, Hebbian Learning, Principal Component Analysis and Independent Component Analysis.

4.1.2 Probabilistic Classifier

The probabilistic classifier uses probability to predict the hidden data. It includes classifier such as: Naïve Bayes (NB), Bayesian Networks (BN) and Maximum Entropy (ME).

4.1.2.1 Naïve Bayes

The Naïve Bayes approach uses probabilistic model to classify the text. It contains two steps: training step, where training data is used to estimate probability, and prediction step, where training data is used to determine probability of hidden data. In [20], the authors classified Arabic Facebook posts normally and in various other languages using NB Classifier. Results showed better results when it is used for binary classification.

In [21], the authors used NB classifier on a dataset which contained 2591 tweets/comments. The authors compared the results with other techniques such as SVM and KNN classifiers.

4.1.2.2 Maximum Entropy

Maximum Entropy optimizes features of training data by using search based technique. Kamal Nigam et.al. Proposed [22, 23] that Maximum Entropy suits best for text classification and compared with Naïve Bayes.

4.1.2.3 Neural Network

Neural Network is a collection of artificial neurons that is used for analysis of computational model. Artificial Neural Network is categorized in to two main categories: FeedForward Neural Network and Recurrent Neural Network. Other types of ANN are: Radial Basis Function (RBF), Kohonen self – organizing, Learning vector quantization, Neuro – Fuzzy, etc. Various algorithms included in ANN are: Multi – Layer Perceptron (MLP), Back Propagation Network (BPN), Probabilistic Neural Network (PNN), etc.

In [24], sentiment classification is performed using Back – Propagation Neural Network (BPN). This approach uses information gain, BPN and subjectivity knowledge present in sentiment lexicon. The result showed reduced dimensionality and accuracy in classification performed on movie and hotel review.

4.1.3 Linear Classifier

The Support Vector Machine (SVM) is a linear classifier. The SVM is used to analyze data and pattern which are used for classification and regression. An SVM-based classifier is used to classify news comments on Facebook for Arabic slang language in [25]. It includes three main steps: data preparation, data preprocessing and data classification which increased the performance to 86.66% from 75.35%

In [26], the authors used low level light stemming and compared it with stemming and light stemming. They classified phrases into 5 main classes (excellent, very good, middling, weak and horrible). These were then used to label each comment. Linear programming (SVM) is then used to get the optimal value.

4.1.4 Decision Tree

Decision Tree classification depends on hierarchical decomposition of training data where tree nodes are labeled with words, branch nodes are labeled with weight and leaf nodes are labeled with category names.

Harrag et al [27] classifies Arabic text documents using decision tree. The authors basically used preprocessing phase, a vector containing all documents of corpus, a subset vector and lastly determines the semantics of document.

The authors in [28] classified Arabic YouTube pages using various Machine Learning algorithms. The authors focused on the Sentiword Lexicon based on the Corpus that they gathered to evaluate and compare each algorithm used.

4.1.5 Rule – Based Classifier

Rule based classification uses basic rules to rate document and reviews. Oraby et al [29] used rule based approach and listed problems of sentiment analysis and opinion strength measurement. Set of grammar rules are used to parse corpus document and mapped with polar words. These words will be used to create lexicons in opinion count calculation phase. Based on the count, the document will be classified as positive or negative.

4.2 Lexicon Based approach

Lexicon based is an unsupervised learning approach for sentiment classification and it mainly contains two categories: Dictionary based and Corpus based.

4.2.1 Dictionary based

Dictionary based approach consists of list of words, each having its own semantic or polarity value. This dictionary of words is used for determining the polarity of sentence or document.
The authors of [30] used three lexicon techniques, one manual ant two automatic. They also designed an Arabic sentiment analysis tool. The accuracy of implementing the Arabic Sentiment Analysis tool reached to 74.6%.
The authors in [31] presented colloquial Non – Standard Arabic – Modern Standard Analysis tool to determine the polarity of reviews and comments. They used K – Nearest Neighbor (KNN) classifier for classification and experimental results showed 90% of accuracy.
Another lexicon-based approach is used in [32], where a lexicon of 120,000 Arabic terms is used along with a Sentiment Analysis tool. Experimental results showed an accuracy of 86.89%.
The authors of [33] used Decision Tree, Naive Bayes (NB), KNN, and SVM for classification on 18 lexicons. The authors also used subjectivity algorithm that yielded 93.9% accuracy, polarity evaluation algorithm which yielded 90% accuracy and intensity algorithm which yielded 96.6% accuracy.

4.2.2 Corpus based

Corpus based approach is used to train classifiers. The authors of [34] proposed two approaches: syntactic approach for sentiment classification and semantic approach for document polarity. The SVM classifier was used on 29 sentences and other method were used on 44 sentences. The syntactic approach yielded 89.3% accuracy, the semantic approach yielded 62% accuracy while 87% accuracy reached for document level classification.
The authors in [35] used two main components, for providing information of local businesses to users, namely review classifer and sentiment analyzer. The authors of [36] represented a multi domain data using corpus which uses feature based sentiment level classification. The result yielded 95%accuracy.
In [37], the subjectivity and sentence level tagging system has been developed. The performance of system improved when Stemming level and language specific features were added in the system.

4.2.3 Hybrid approach

The hybrid approach is a combination of Lexicon based and Machine Learning approach. The researcher in [38] presented a hybrid approach which included three methods mainly lexicon based maximum entropy and K – Nearest method for classification of documents and achieved 80% of accuracy.
In [39], the researchers merged lexicons with semantic orientation algorithm. The feature selection was performed by using Information Gain. 84% accuracy was obtained from polarity classification and 81% accuracy was obtained from subjectivity classification.
The authors of [40], presented a hybrid approach which combines lexicon based approach with NB classifier and achieved 91% of accuracy.

Table 1. Comparison of various techniques

Techniques	Pros.	Cons.
Naïve Bayes	Execution is faster, improved performance, better computational Complexity.	When database size increases, its performance decreases
Neural Network	Can be implemented for any application. Even if one element fails, then its execution does not stop.	Requires large training time and high processing time when there are large Neural Networks
Support Vector Machine	When data size is small, it works very well as its kernel based framework is very powerful.	When data size increases, its performance decreases and two classes SVM must be combined as it does not support multi class SVM.
Decision Tree	Execution is faster as data is segmented.	When the number of segmentation increases, the interpretability decreases.

V. EVALUATION OF SENTIMENT CLASSIFICATION

The performance of the sentiment classification can be assessed using four parameters: Accuracy, Precision, Recall and F1.

$$\text{Accuracy} = (TP+TN)/(TP+TN+FP+FN) \quad (5)$$
$$\text{Precision} = TP / (TP + FP) \quad (6)$$
$$\text{Recall} = TP / (TP + FN) \quad (7)$$
$$F1=(2*\text{Precision}*\text{Recall})/(\text{Precision}+\text{Recall}) \quad (8)$$

Where, TN, TP, FN, FP represents True Negative, True Positive, False Negative, and False Positive. The confusion matrix from these parameters will be:

Table 2. Confusion Matrix

	Positives	Negatives
Positives	True Positives	False Negative
Negatives	False Positives	True Negatives

VI. ISSUES AND FUTURE DIRECTIONS

There are certain issues regarding sentiment analysis such as:

1. The data source can contain noisy, unstructured data. No automatic spelling correction and noise removal system exists. Hence, Noisy and unstructured data must be first handled.
2. Data preprocessing step uses maximum time for converting raw data into structured format which in turn reduces the accuracy of system.
3. During Political discussions, the words often contain derisive sentences. Little or no work has been done for such sentences. So we require more approach to solve such problem.
4. In the sentiment classification, document level and sentence level work exists but very less work for feature / aspect/ phrase level work is done. Hence, for comparing the entity with respect to features of that entity must be done.
5. Machine learning has helped to get close to data. While, the use of ontology has also been successful in the field of sentiment analysis. By applying machine learning along with ontology can help to resolve the problem of scalability.
6. There exists many machine learning methods for feature selection but no hybridization of machine learning along with optimization technique is present.
7. Sentiment Analysis has been widely applied to English language and of Arabic language but there is a lack of work in non – English languages.
8. There is similar product available for which sentiment classification need to be performed at feature level so that comparison between similar products can be done.
9. The major problem with sentiment analysis lies in the validation of the methods used for classification.
10. Finally, some issues of subjectivity, streaming of text, and feature based classification for Facebook posts also need to be solved.

VII. CONCLUSION

This paper presents a comprehensive analysis on the various sentiment classification approached used. The paper is analyzed on various dimensions such as subjectivity, objectivity, sentiment classification approaches, and lexicon. Forty papers were surveyed.

This paper also showed the various techniques of sentiment classification such as Machine Learning, Lexicon based and Hybrid approach. The most common approaches used in Machine Learning were SVM, NB and Neural Network which were applied on different types of dataset. Based on the survey of various techniques, it has been found that Neural Network provides better accurate results than other approaches but it also requires more training time.

REFERENCES

[1] B. Pang, L. Lee, “Opinion mining and sentiment analysis”, Foundation and Trends in Information Retrieval (2008) 1–135.
[2] H. Tang, S. Tan, X. Cheng, “A survey on sentiment detection of reviews”, Expert System Application 36 (2009) 10760–10773.
[3] Kumar Ravi, Vadlamani Ravi, “A Survey on opinion mining and sentiment analysis: Tasks, approaches and applications”, Journal of Knowledge – Based System 89 (2015) 14–46, Elsevier,
[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 System 53 (2012) 675–679.
[5] M. Tsytsarau, T. Palpanas, “Survey on mining subjective data on the web”, Data Mining and Knowledge Discovery 24 (2012) 478–514, http://dx.doi.org/10.1007/s10618-011- 0238-6.
[6] B. Liu, “Sentiment Analysis and Opinion Mining”, Morgan and Claypool publishers, 2012 (May).
[7] R. Feldman, “Techniques and applications for sentiment analysis”, Communication ACM 56 (4) (2013) 82–89 (Review Articles).
[8] W. Medhat et al., “Sentiment analysis algorithms and applications: a survey”, Ain Shams Eng. J. (2014), http://dx.doi.org/10.1016/j.asej.2014.04.011.
[9] M. Ghiassia, J. Skinnerb, D. Zimbraa (2013), “Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network”, Journal of Expert Systems with Applications, ELSEVIER Volume 40, Issue 16, 6266–6282.
[10] Anuj Sharma, Shubhamoy Dey, “Using Self- Organizing Maps for Sentiment Analysis”, CoRR (2013).
[11] Soumith Chintala, “Sentiment Analysis using neural architectures”, New York, NY 10012-2012.
[12] Foram Kalyani, “Neural Network based Classification of Sentiments: An Analysis”, 2016 International Conference on Engineering and Technology (ICET), 265 – 269 , Volume 3, 978 – 1 – 5090 – 3212 – 9.
[13] Hu, Minqing, and Liu Bing. “Mining opinion features in customer reviews.” In AAAI, vol. 4, pp. 755-760, 2004.
[14] Solanki Yogesh, Bhumika Shah, “Feature based opinion mining: A Survey”, 978-1-4799-8047-5/15, IEEE 2015.
[15] Ding, Xiaowen, Bing Liu, and Philip S. Yu, “A holistic lexicon-based approach to opinion mining”, In Proceedings of the 2008 International Conference on Web Search and Data Mining, ACM, pp. 231-240, 2008.
[16] Hu, Minqing, and Liu Bing, “Mining and summarizing customer reviews”, In Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, pp. 168 – 177, 2004.
[17] Hu, Minqing, and Bing Liu., “Opinion extraction and summarization on the web”, In AAAI, vol. 7, pp. 1621-1624, 2006.
[18] Tripathy, A., Agrawal, A. and Rath, S.K., 2015, “Classification of Sentimental Reviews Using Machine Learning Techniques.”, Procedia Computer Science, 57, pp.821-829.
[19] Vimalkumar Vaghela, Bhumika Jadav, “Analysis of various Sentiment Classification Techniques”, International Journal of Computer Applications (0975 – 8887) Volume 140 – No.3, April 2016.
[20] Maher M. Itani, Lama Hamandi, Rached N. Zantout and Islam Elkabani, “Classifying Sentiment in Arabic Social Networks: Naïve Search versus Naïve Bayes”, 2012 2nd International Conference on Advances in Computational Tools for Engineering Applications (ACTEA).
[21] Rehab M. Duwairi and Islam Qarqaz, “Arabic Sentiment Analysis using Supervised Classification”, The 1st International Workshop on Social Networks Analysis, Management and Security (SNAMS - 2014), August 2014, Barcelona, Spain.
[22] McCallum and Kamal Nigam. 1998. “A Comparison of Event Models for Naive Bayes Text Classification”. In Procedings of the AAAI-98 Workshop on Learning for Text Categorization, pages 41–48.
[23] Stanley Chen and Ronald Rosenfeld. 2000. “A Survey of Smoothing Techniques for ME Models”. IEEE Transactions on Speech and Audio Processing, 8(1):37–50.
[24] Wei Jin, Hung Hay Ho, Rohini K. Srihari, “OpinionMiner: A Novel Machine Learning System for WebOpinion Mining and Extraction”, In Proceedings of KDD’09, June 28–July 1, 2009, Paris, France.
[25] 25. Taysir Hassan A. Soliman1, M. Ali M., Abdel Rahman Hedar3, M. M. Doss4, “MINING SOCIAL NETWORKS’ ARABIC SLANG COMMENTS”, In Proceedings of IADIS European Conference on Data Mining 2013 (ECDM'13), 22 – 24 July, Prague, Czech Republic.
[26] Walid Cherif, Abdellah Madani and Mohamed Kissi, “A new modeling approach for Arabic opinion mining recognition”, Intelligent Systems and Computer Vision (ISCV), 2015.
[27] Fouzi Harrag, Eyas El-Qawasmeh and Pit Pichappan, “Improving Arabic Text Categorization using Decision Trees”, Networked Digital Technologies, 2009. NDT '09.
[28] Rasheed M. Elawady, Sherif Barakat and Nora M. Elrashidy, “Sentiment Analyzer for Arabic Comments”, International Journal of Information Science and Intelligent System, 3(4): 73-86, 2014.
[29] Shereen Oraby, Yasser El-Sonbaty, and Mohamad Abou El-Nasr, “Finding Opinion Strength Using Rule-Based Parsing for Arabic Sentiment Analysis”, 12th Mexican International Conference on Artificial Intelligence, MICAI 2013, Mexico City, Mexico, November 24-30, 2013, Proceedings, Part II.

[30] Abdulla, N.; Majdalawi, R.; Mohammed, S.; Al-Ayyoub, M.; and Al- Kabi, M., "Automatic Lexicon Construction for Arabic Sentiment Analysis", 2014 International Conference on Future Internet of Things and Cloud (FiCloud-2014), 27-29 August 2014, Barcelona, Spain, pp. 547-552.

[31] Mohammed Al-Kabi, Amal Gigieh, Izzat Alsmadi, Heider Wahsheh, Mohamad Haidar, "An Opinion Analysis Tool for Colloquial and Standard Arabic". ICICS'13, April 23–25, 2013, Irbid, Jordan.

[32] Al-Ayyoub, M., Essa, S. B., & Alsmadi, I. (2015), "Lexicon-based sentiment analysis of arabic tweets", International Journal of Social Network Mining,2(2), 101-114.

[33] Mohammed Al-Kabi, Amal Gigieh, Izzat Alsmadi, Heider Wahsheh, Mohamad Haidar, "Opinion Mining and Analysis for Arabic Language", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 5, No. 5, 2014.

[34] Noura Farra, Elie Challita, Rawad Abou Assi, Hazem Hajj, "Sentence level and Document-level Sentiment Mining for Arabic Texts", 2010 IEEE International Conference on Data Mining Workshops, 978-0-7695-4257-7/10.

[35] Mohamed Elhawary and Mohamed Elfeky, "Mining Arabic Business Reviews". 2010 IEEE International Conference on Data Mining Workshops.

[36] Hossam S. Ibrahim, Sherif M. Abdou and Mervat Gheith, "SENTIMENT ANALYSIS FOR MODERN STANDARD ARABIC AND COLLOQUIAL", International Journal on Natural Language Computing (IJNLC) Vol. 4, No.2, April 2015.

[37] Muhammad Abdul-Mageed, Mona T Diab, and Mohammed Korayem. 2011, "Subjectivity and Sentiment Analysis of Modern Standard Arabic", In Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics, number 3, pages 587–591, Portland, Oregon, US.

[38] El-Halees, A., "Arabic Opinion Mining Using Combined Classification Approach", In Proceedings of the International Arab Conference on Information Technology, ACIT (2011)

[39] El-Makky, N., Nagi, K., El-Ebshihy, A., Apady, E., Hafez, O., Mostafa, S., Ibrahim, S., "Sentiment analysis of colloquial Arabic tweets", In ASE BigData/SocialInformatics/PASSAT/BioMedCom 2014 Conference, Harvard University, Dec 2014.

[40] Khalifa, K., & Omar, N. (2014), "A hybrid method using lexicon-based approach and naive Bayes classifier for Arabic opinion question answering", Journal of Computer Science, 10(10), 1961.

