International Journal of Information & Computation Technology.
ISSN 0974-2239 Volume 8, Number 1 (2018), pp. 1-10

http://www. irphouse.com

Sentimental Analysis Using Fuzzy Logic for Cloud Service Feedback Evaluation

Jamilah R. Alharbi [1], Wadee S. Alhalabi [1,2]

[1] Department of Computer Science, Faculty of Computing and Information Technology, King Abdul-Aziz University KAU, Jeddah, Kingdom of Saudi Arabia.
[2]Department of computer Science, Effat University, Jeddah, Kingdom of Saudi Arabia.

Abstract

Sentimental Analysis (SA) is used to analyze users' posts, called tweets, in social networks and understand users’ opinion and their feeling about specific topic. This process produces a description label that can be positive or negative and can also produces emotional labels, such as: angry, happy and sad. Accordingly, SA depends on deep analysis of the posts and the polarity of their contents. Given that, this process faces massive challenges embodied in word sense disambiguation, analysis of comparison, negation, intensive and sarcasms sentences and deal with grammatical mistakes, which is complicated task. To overcome these challenges, in this paper, SA system is developed by combining multiple inputs of different forms, which can be extracted from the text and its associated attachments (e.g.: emoji). Fuzzy logic is used to combine the inputs together for an accurate identification of user emotions. The outputs of the proposed system can be one of the following emotional labels: Very Positive, Positive, Neutral, Negative, Very Negative. The results showed that, over user satisfaction tweets about Google, Amazon and Microsoft cloud services, the proposed system gives significant results of 83% precision, 89% recall and 83% F-score.

Keywords: Sentimental Analysis, Fuzzy Logic, Cloud Service Provides, Amazon, Google, Microsoft.

INTRODUCTION

Social networks and social media are used extensively by users to express their thoughts, opinions and emotions, thus, analyzing the content of these sources provides a useful information for understanding users' mindsets. In fact, nothing can give better and up-to-date information about user mindsets, likes and dislikes such as social networks. This analysis can be used for the following goals: 1) Identify, analyze and understand user opinion and satisfaction about services and products and analyze user experience. 2) Identify market characteristics, demographic and strength in market research. 3) Identify topics and trends that interest users and triggers their emotions. 4) Evaluate the effective of users' interactions and influences on each other [1]. Given that, shallow and deep analysis of social media gained great interest for its benefits in decision making process for companies, organizations and even government plans. Sentimental Analysis (SA) is one of the social networks analysis tasks that process posts in order to develop understanding for the opinions and the feelings about specific topic. Accordingly, SA produces a description label that can be positive or negative and can also produces emotional labels, such as: angry, happy and sad (Figure 1) [2].

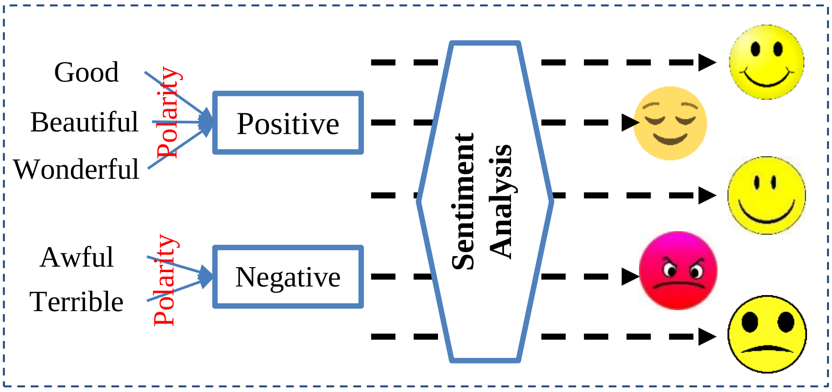

Figure 1: Sentimental Analysis: Input-Output and Components

SA is implemented using document processing, natural language processing (NLP) and machine learning with reference to dictionaries. Accordingly, existing SA approaches can be classified into dictionary-based approach, machine learning approach and hybrid approach. The dictionary-based approach depends merely on using dictionary(s) to obtain the polarity of the words, a special word-polarity form of dictionary is available for this purpose. SentiWordNet and SentiStrength are the best among other available dictionaries according to Islam and Zibran [1] and Jhaveri, Chaudhari [3]. Machine learning approach depends on using classification algorithms with a set of words in order to find the correct label of the analyzed text. Hybrid approach, which proved to give the best results, depends both on polarity dictionary and classification algorithms.

Regardless of the type, SA faces massive challenges, some of these challenges are: identifying word sense, or what so called word sense disambiguation, because SA depends on analyzing words as the media of sentimental expression. Even worst, analyzing of sentences with comparison, negation, intensive and sarcasms sentences and deal with grammatical mistakes. Besides, these challenges, real-time analysis and disagreement and variations in opinions are still challenging for SA[4]. To overcome these challenges, every aspect of the input media should be considered in the analysis process, side by side with the main source, which is the text. However, to extract such variety aspects and combine them is anything but trivial.

Accordingly, two major problems are addressed in this paper, these are: how identify and utilize the entire content of the posts to enhance the accuracy of SA and how to combine these contents in its varied forms (words, symbols, emoji and etc.) using machine learning techniques. In this paper, a system for SA is developed by combining multiple inputs of different forms, which can be extracted from the text and associated attachments (e.g.: emoji) in order to overcome the limitations of the current sentiment analysis. Fuzzy logic is used to fuse the different form of input together for an accurate identification of user emotions. Fuzzy logic is an artificial intelligent approach that mimics the human way of making decisions, by converting the crisp inputs into fuzzy terms (short, long, medium, etc.) and using a set of rules to link these terms with appropriate actions or decisions. Generally, it incorporates the term of uncertainty, by which the human brain is activated in difficult decision-making process. Accordingly, it has been successfully utilized in various fields such as medical, science, engineering and finance. Overall, fuzzy logic allows the proposed system to work with the expected uncertainty in SA, combine multiple and varied inputs and mimics human decision-making process using a set of established rules that are equal the rules used by human for sentimental analysis [5].

In order to validate, verify and evaluate the proposed system, a selected sample dataset should be prepared and utilized. Tweeter is selected to be the source of the data in this paper. Twitter stand as one of the most utilized source as it has over 600 million users worldwide [6], which makes financial, industrial and services providers use its tweets as a basic source for market and user information. As dataset formation depends on the target field and the target topic, in this paper user satisfaction with the cloud services provided is determined as the target field. The target topic that is addressed in this paper, is satisfaction with clouds provided by Google, Microsoft and Amazon. The significant of the created dataset stands with the fact that one of the evolved area in understanding user satisfaction about cloud services is using their feedbacks via their posts. The rest of this paper is organized as follows: Section 2 discusses the existing literature on SA, the classification of the literature, advantages and disadvantages of the existing approaches. Section 3 presents the proposed work and discuss the utilized resources. Results are presented in Section 4 and the conclusion is given in Section 5.

LITERATURE REVIEW

Dictionary-based approach rely on using word's polarity combine multiple polarities of the input to come out with final polarity that can be easily mapped into a sentimental category. Advanced dictionary-based approach uses other lexical resources that provide antonyms and synonyms in order to increase the chance of matching words with the entities of the dictionaries [7].

Liu, Yang [8] proposed a SA system to analyze the reviews of micro videos and differentiate between positive and negative words based on their polarities in the dictionary together with calculating the effects of adverbs on the polarities. The dictionary was evolved manually and the polarities were assigned as positive and negative. Based on the developed system, the final output can be positive, negative or neutral. The positive category occurs if the sum of the positive polarities in the post is greater than the sum of negative polarities and vice versa. Neutral is seldom to be presented as an output and it occurs if the negative and positive polarities are equal. Similarly, Bidulya and Brunova [9] analyzed posts and feedbacks about bank services' and differentiated between positive and negative words based on their polarities in the dictionary together with calculating the effects of negative words and adverbs on the polarities. Based on the polarity of the words, which can be positive or negative, the outputs are assigned either positive or negative labels, however, polarities of the output are increased or decreased with the presence of adverbs, such as very, absolutely and never and are reversed with the present of negative words. To combine these forms, a set of rules is developed to evaluate input and produce proper output. Several other systems that belongs to this category have been developed as presented by Jhaveri, Chaudhari [3] and Sheeba and Vivekanandan [10]. Dictionary-based approach is straight-forward that can be implemented in wide-range of domains to produce satisfactory results. However, it is clearly that dictionary-based approach is not effective with complicated text that includes compound sentences unless all the possible sentence formation is identified, which might be difficult. Moreover, according to the presented literature, dictionary-based approach has been used to produce limited annotation with two categories positive and negative and sometimes neutral, while sentimental analysis required a variation of these categories as human observations of the sentimental content.

Machine learning approach depends on features extraction from the input text and using classification algorithms for identifying the category of the input media. Rohini, Thomas [11] proposed a SA system to analyze movies' reviews using decision tree classifier. First, features in the forms of terms frequency-inverse document frequency are extracted from the input text. Then, these features are used as inputs for the decision tree, which produce positive or negative label as an output. Although, it is known that terms frequency-inverse document frequency features have bad performance in short text [12, 13], the proposed system gave satisfied results. Similarly, Hegde and Padma [14] analyzed product reviews using random forest, an extension to decision tree, classifier. The output of the developed system can be one of four categories, negative (bad), neutral (ok), positive (good) and very positive (best). Shakeel and Karwal [15] analyzed financial budget feedback to the tax association in India using random forest. The output of the developed system can be one of three categories, negative, neutral and positive. Nurwidyantoro [16] analyzed news feedback using random forest with term frequency features and produced two categories, negative or positive. The advantage of the machine learning approach is the ability to deal with as many categories as required, however, using frequency features in short sentences leads to inaccurate results and cannot be used in wide domain.

Hybrid approaches were proposed in order to gain the advantages of both dictionary and machine learning approaches and enable combination of multiple content-related features and polarities. Mishra, Rajnish [17] proposed a system to analyze Smart-City reviews by differentiating between positive and negative words as presented in the SentiWordNet dictionary with calculating term frequency, term co-occurrence and POS features. Using these features in a classification technique, the produced outputs are positive, neutral or negative. Appel, Chiclana [18] proposed a robust system to analyze movies' reviews by differentiating between positive and negative words as presented in SentiWordNet dictionary and using rules to add the effect of negative words and compound sentences. Rules are applied prior to the final classification in order to adjust the polarity of the involved words based on the form of the sentence. Naïve Bayesian and maximum entropy are used for the classification purpose. Although, the literature on hybrid approach is limited, this approach showed capabilities in solving the limitation of the other approaches. A summary of the reviewed literature is given in Table 1.

Table 1. Summary of the Related Work for Sentimental Analysis

Reference	Machine Learning	Dictionary	Polarity Features	Textual Features	Output Cat.	Domain
Liu, Yang [8]	No	Costume Dictionary	Words	Adverbs	3	Micro-video Review
Bidulya and Brunova [9]	No	Costume Dictionary	Words	Adverbs and Negative	2	Bank Service Quality
Jhaveri, Chaudhari [3]	No	Afinn Dictionary	Words	Adverbs, Negative and Hashtag	3	E-commerce
Sheeba and Vivekanandan [10]	No	Costume Dictionary	Words	Topic Extraction	2	Product Review
Rohini, Thomas [11]	Decision Tree	No	No	Term Frequency	2	Movies Review
Hegde and Padma [14]	Random Forest	No	No	Terms	4	Product Review
Shakeel and Karwal [15]	Random Forest	No	No	Term Frequency	3	Financial Budget
Nurwidyantoro [16]	Random Forest	No	No	Term Frequency	2	Economic News
Mishra, Rajnish [17]	Undetermined	SentiWordNet	Words	Term Frequency, Co-occurrence and POS	3	Smart-City
Appel, Chiclana [18]	Naïve Bayes and Maximum Entropy	SentiWordNet	Words	Negative and Compound	2	Movies Review

METHODOLOGY

A hybrid approach is proposed in this paper in order gain the advantages of both dictionary-based and machine learning based and to overcome the limitations of the existing hybrid systems, as can be concluded from Table 1, which embodied in using limited content information, classify input into two or three categories and using classical classification algorithms, such as naïve Bayesian and decision tree. Accordingly, in the proposed approach, features of various forms are extracted and aggregated together using fuzzy-logic system in-order to classify tweets into five categories, these are: “very negative”, “negative”, “neutral”, “positive” and “very positive”. The proposed method, as illustrated in Figure 2, is built up with four processing phases, these are: data collection, pre-processing, feature extraction and model construction.

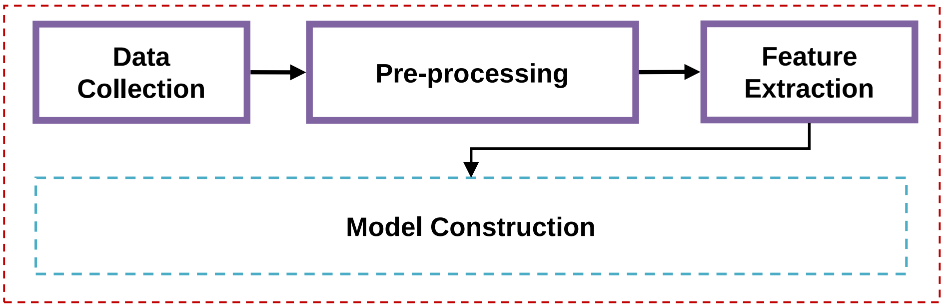

Figure 2. Flowchart of the Proposed Sentimental Analysis Approach

In the data collection, the dataset that is utilized in the proposed work contains 9263 tweets of users’ satisfaction about the cloud services that are provided by Amazon, Microsoft and Google that are collected from their official page of these companies on Tweeter.

In the pre-processing stage, punctuations, stop-words, URLs, numbers, non-English words and tweeter terms are removed. Then, the text is tokenized. Followed by disjunction words, hashtag, emoji, negative, adverbs, exclamation and intensive words are identified.

In feature extraction, the features that are extracted from the input text are as follows: words polarity as given in SentiWordNet dictionary, words polarity as given in SentiStrength dictionary, emoji polarity, negation, adverb, intensive and exclamation polarity, hashtag and retweets polarity. The term polarity from the SentiWordNet is numerical number in the range [-1-1] and it has linear effects on the output results. Accordingly, with a value of -1, the polarity tends to contribute to an output of “very negative” category while a value of 1, the polarity tends to contribute to an output of “very positive” category and so one. Similarly, the term polarity from the SentiStrength is numerical number in the range [-5-5] and it also has linear effects on the output results. The emoji are classified into six groups, these are: joy, smile, savoring, smirking, confound and winking, each of which have positive, negative or neutral influenced on the output, as will be clarified in the model construction. The intensive words, including negation, adverb, intensive and exclamation, each of which have positive, negative or neutral influenced on the output results. The hashtag influence has either positive, negative or neutral influence and retweet influence has either positive, negative or neutral. The influence of the hashtag is calculated as the ratio of the positive tweets in the hashtag to the total number of tweet. Positive hashtag is the one with greater than or equal 60% positive tweets, while negative hashtag is the one with greater than or equal 60% negative tweets and the rest in neutral. For retweets or likes, it can be positive, negative or neutral. Positive occurs with the number of retweets or likes is greater than or equal 200 with a tweet that has polarity of positive words, while negative occurs with the number of retweets or likes is greater than or equal 200 with a tweet that has polarity of negative words otherwise, it is neutral.

Model construction is implemented as fuzzy logic system for two reasons, the first is the fuzziness influence of the input variable on the output and the second is the ability of fuzzy systems to incorporate inputs of various forms. Accordingly, Takagi-Sugeno-Kang (TSK) *fuzzy logic* control system is used in this paper as the model of sentimental classification. TSK has more flexibility that required in the developed sentimental analysis model compared to other fuzzy logic systems. The developed model is developed in four phases, these are parameter initialization, fuzzification, rule evaluation and aggregation. In the parameter initialization, six parameters are identified, which are average polarity based on SentiWordNet, word polarity based on SentiStrength, emoji, intensive, hashtag and retweet. The fuzzy set for each of these parameters is identified based on the terms mentioned before and the membership functions are formulated around the values mentioned before as well using trial and error approach in which the final sets and functions are finalized as the experiments are conducted. The fuzzification process is implemented based on TSK system by convert the crisp value for inputs with crisp values, such as polarities and hashtag into fuzzy terms. The set of rules that are used in the rule evaluation step is created using trial and error process. A set of initial rules is developed, as given in Table 2, however, the set is finalized as the experiments are conducted. The aggregation process is based on using the output of the rule with highest confidence.

Table 2. Examples of the Set of Fuzzy Logic Rule

Rules
IF ***TP1*** is missing AND ***TP2*** is missing THEN ***Out*** is Neutral
IF ***TP1*** is missing AND ***TP2*** is V. Negative THEN ***Out*** is V. Negative
IF ***TP1*** is missing AND ***TP2*** is Negative THEN ***Out*** is Negative

EXPERIMENT AND RESULTS

As mentioned, the data were collected from three companies official page on Tweeters, these tweets are annotated manually. A total of 500 tweets are used in the experiments, 100 tweets in each category. Precision, Recall and F-Score are used to evaluate the results of the proposed system. The confusion matrix of the results of the proposed approach are given in Table 3 and the Precision, Recall and F-Score are given in Table 4.

Table 3. Confusion Matrix Results of the Proposed Approach

	VP	P	N	NG	VN
VP	85	0	0	0	0
P	4	94	2	2	4
N	11	6	97	26	27
NG	0	0	1	72	0
VN	0	0	0	0	69
Total	100	100	100	100	100

Table 4. Performance Results of the Proposed Approach

Categories	**Precision**	**Recall**	**F-score**
Very Positive	85%	100%	91%
Positive	94%	88%	90%
Natural	97%	58%	72%
Negative	72%	98%	83%
Very Negative	69%	100%	81%
average	**83%**	**89%**	**83%**

As given in Table 3 and Table 4, the results of the proposed approach can be summarized as follows: High precision value is obtained using the developed methodology with percentage of 83%. Accurate recall value is obtained with percentage of 89% and F-score of 83%. Besides, the results of the proposed approach on the collected data showed that the proposed approach achieved precision of 83%, recall of 89% and F-score of 83%. The results also revealed that the proposed approach has a high accuracy in recalling “very positive” and “very negative” categories with 100% percentage. Slightly less recall for “negative” category with 98% and marginally less for “positive” category with 88%. Although the recall of the “neutral” category is low with 58%, the proposed approach is considered to be robust, as analyzing sentimental of such domain required identifying bad feedback, and sometime good feedback, in order to improve their services. Accordingly, having low “neutral” recall does not influence the usability of the proposed approach in SA of cloud service providers as it is the less beneficial category in the proposed approach. As for the precision, good results were obtained for all the categories, except of the “very negative” as the proposed approach labeled some of the “neutral” category as negative, which is considered a limitation of the proposed approach and will be considered in the future work.

CONCLUSION

In this paper, a new sentimental analysis approach based on diverse extracted features and using fuzzy logic system is proposed. The proposed approach, which is classified as hybrid of machine learning and dictionary-based, uses two dictionaries to extract the polarities of the words in the input text, the polarity of the emoji, hashtag and the intensive words are together combined in the fuzzy logic. Accordingly, a fuzzy system, which is implemented in three stages, fuzzification, rule evaluation and aggregation is implemented and the output of the system, for each input can be one of the five categories, very positive, positive, neutral, negative and very negative. A dataset of users' satisfaction about cloud service provides, of three well-known providers, these are: Amazon, Google and Microsoft were collected. The results of the proposed approach on the collected data showed that the proposed approach achieved precision of 83%, recall of 89% and F-score of 83%. The results also revealed that the proposed approach has a high accuracy in recalling "very positive" and "very negative" categories with 100% percentage. Slightly less recall for "negative" category with 98% and marginally less for "positive" category with 88%. Although the recall of the "neutral" category is low with 58%, the proposed approach is considered to be robust, as analyzing sentimental of such domain required identifying bad feedback, and sometime good feedback, in order to improve their services. Accordingly, having low "neutral" recall does not influence the usability of the proposed approach in SA of cloud service providers as it is the less benefecial category in the proposed approach. As for the precision, good results were obtained for all the categories, except of the "very negative" as the proposed approach labeled some of the "neutral" category as negative, which is considered a limitation of the proposed approach and will be considered in the future work.

REFERENCES

1. Islam, M.R. and M.F. Zibran. *A Comparison of Dictionary Building Methods for Sentiment Analysis in Software Engineering Text*. in *2017 ACM/IEEE International Symposium on Empirical Software Engineering and Measurement (ESEM)*. 2017. IEEE.
2. Bhonde, R., et al., *Sentiment Analysis Based on Dictionary Approach.* International Journal of Emerging Engineering Research and Technology, 2015. **3**(1): p. 51-55.
3. Jhaveri, D., A. Chaudhari, and L. Kurup, *Twitter sentiment analysis on e-commerce websites in India.* Int J Comput Appl, 2011: p. 0975-8887.
4. Balaji, P., O. Nagaraju, and D. Haritha. *Levels of sentiment analysis and its challenges: A literature review*. in *Big Data Analytics and Computational Intelligence (ICBDAC), 2017 International Conference on*. 2017. IEEE.
5. Liu, H. and M. Cocea. *Fuzzy rule based systems for interpretable sentiment analysis*. in *Advanced Computational Intelligence (ICACI), 2017 Ninth International Conference on*. 2017. IEEE.
6. millions, N. *Twitter: number of active users 2010-2017 | Statista*. 2017 [cited

2018 May, 5]; Available from: http://www.statista.com/statistics /282087/number-of-monthlyactive-twitter-users/.
7. Mestry, P., et al., *A Survey on Twitter Sentiment Analysis with Various Algorithms.* 2016.
8. Liu, Z., N. Yang, and S. Cao. *Sentiment-analysis of review text for micro-video*. in *Computer and Communications (ICCC), 2016 2nd IEEE International Conference on*. 2016. IEEE.
9. Bidulya, Y. and E. Brunova. *Sentiment analysis for bank service quality: A rule-based classifier*. in *Application of Information and Communication Technologies (AICT), 2016 IEEE 10th International Conference on*. 2016. IEEE.
10. Sheeba, J. and K. Vivekanandan, *A fuzzy logic based on sentiment classification.* International Journal of Data Mining & Knowledge Management Process, 2014. **4**(4): p. 27.
11. Rohini, V., M. Thomas, and C. Latha. *Domain based sentiment analysis in regional Language-Kannada using machine learning algorithm*. in *Recent Trends in Electronics, Information & Communication Technology (RTEICT), IEEE International Conference on*. 2016. IEEE.
12. Sahami, M. and T.D. Heilman. *A web-based kernel function for measuring the similarity of short text snippets*. in *Proceedings of the 15th international conference on World Wide Web*. 2006. AcM.
13. Banerjee, S., K. Ramanathan, and A. Gupta. *Clustering short texts using wikipedia*. in *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*. 2007. ACM.
14. Hegde, Y. and S. Padma. *Sentiment Analysis Using Random Forest Ensemble for Mobile Product Reviews in Kannada*. in *Advance Computing Conference (IACC), 2017 IEEE 7th International*. 2017. IEEE.
15. Shakeel, M. and V. Karwal. *Lexicon-based sentiment analysis of Indian Union Budget 2016–17*. in *Signal Processing and Communication (ICSC), 2016 International Conference on*. 2016. IEEE.
16. Nurwidyantoro, A. *Sentiment analysis of economic news in Bahasa Indonesia using majority vote classifier*. in *Data and Software Engineering (ICoDSE), 2016 International Conference on*. 2016. IEEE.
17. Mishra, P., R. Rajnish, and P. Kumar. *Sentiment analysis of Twitter data: Case study on digital India*. in *Information Technology (InCITe)-The Next Generation IT Summit on the Theme-Internet of Things: Connect your Worlds, International Conference on*. 2016. IEEE.
18. Appel, O., et al., *Successes and challenges in developing a hybrid approach to sentiment analysis.* Applied Intelligence, 2018. **48**(5): p. 1176-1188.