ISSN: 2229-6956 (ONLINE)
DOI: 10.21917/ijsc.2018.0224

A FRAMEWORK FOR ASPECT BASED SENTIMENT ANALYSIS USING FUZZY LOGIC

A. Jenifer Jothi Mary[1] and L. Arockiam[2]
Department of Computer Science, St. Joseph's College (Autonomous), India

Abstract
Sentiment Analysis (SA) is the study of people's opinions, emotions, and appraisals toward products and events. In the past years, it fascinated a great deal of attentions from both industry and academia for a variety of applications. Opinions are significant, because people need to make decisions. It is helpful not only for the individuals but also for the business organizations. Fuzzy logic can provide a quick way to solve the haziness present in most of the natural languages. The techniques are less explored in sentiment analysis. In this paper, Aspect based Sentiment Summarization (ASFuL) is proposed with fuzzy logic by classifying opinions polarity as strong positive, positive, negative and strong negative. It also integrates the non-opinionated sentences using Imputation of Missing Sentiment (IMS) mechanism which plays a vital role in generating precise results. The researchers used Fuzzy Logic to find sentiment classes in the review. The results show that the mechanism is viable to extract opinions in an efficient manner.

Keywords:
ASFuL, Sentiment Analysis, Aspect, Sentiment Summarization, Fuzzy Logic

1. INTRODUCTION

The development of web and social networking sites have radically changed the way that people articulate themselves and interact with others. People can now contribute their opinions in terms of reviews on the products and services as they use on the merchant sites. People share their opinions, ideas and interests based on the reviews acquired from the websites. Business people depend more on these sites to collect opinions on their product's development for their business improvement.

Sentiment analysis is an innovative computing technology to improve the decision making process. Opinions are always expressed in free text as reviews, assessment or comments. It uses sentences related with relevant products and their attributes or aspects hauled out from the reviews as the basis for summarization. Sentiment analysis is used to mine these opinions that come out in the web to check the attitude of an individual who posted about a product or event.

Fuzzy logic is a tool for the management of vague and heterogeneous information [14]. It is a form of multi-valued logic and it deals with reasoning that is close to the actual rather than fixed and exact. Generally traditional logic have different values between true or false, but in the fuzzy logic variables can have a truth value that ranges between 0 and 1. Fuzzy logic is used to classify the sentiment by defining membership functions.

Though product reviews have great impact on business development and decision making, gathering of these product reviews from different websites are time consuming and checking the polarity of them needs more efforts. Moreover, these reviews should be classified into positive, negative, and neutral to support the business development. So, classifying the opinionated data and identification of polarities is also very vital in Sentiment Analysis. Thus, a novel technique needs to be developed to classify the data and identify the polarities of these classified data in a faster and accurate manner. Thus, the researchers proposed a technique called ASFuL, Aspect based Sentiment Summarization using Fuzzy Logic to classify the reviews into polarity classes, for e.g. strong positive, strong negative etc. using fuzzy system.

The primary aim of this article is to propose a new mechanism for summarizing the sentiments which are given in the reviews. Aspect extraction method and fuzzy logic have been taken up to build this mechanism.

This paper is going along with the following sections. Section 2 presents a brief overview on literature review. Section 3 and 4 give the motivation and objective of the research work respectively. Section 5 elaborates the framework of the proposed mechanism and section 6 concludes the paper.

2. RELATED WORKS

A hybrid approach to SA was presented by Appel et al. [1]. The authors used semantic rules, unsupervised ML (Machine Learning) techniques, a sentiment lexicon and fuzzy sets which were enhanced by Senti-WordNet. A Hybrid Standard Classification was performed and improved into a Hybrid Advanced approach. It incorporated linguistic classification of semantic polarity provided by means of fuzzy sets. The polarity of a given sentence was computed using the new SA methodology for the Movie Review Dataset.

Suresh et al. [2] presented a new fuzzy clustering model for analyzing tweets relating to the sentiments of a particular brand name. The authors used the original dataset composed of 12 months. A comparative analysis was prepared with the previous K Means partitioning clustering techniques and also Expectation Maximization algorithms. The metrics precision, execution time, recall and accuracy were used. The approach proposed was experimented and it was shown prominent twitter SA results.

Aniello et al. [3] presented a new method to maintain urban planners and decision makers for creating awareness on city issues and its assets such as environment, mobility and security. The researchers used Fuzzy sets and it's Cognitive Maps for accomplishing the goal. To find out the relation between two fuzzy sets, signatures were obtained and different signatures were grouped to characterize an area. Moreover, the authors used subjective concept of a Fuzzy Cognitive Map and performed if-then analysis. It was applied for three POIs of Salerno city by gathering data and involving some real citizens. The result was encouraging and created good awareness on urban areas quality which created effect on issues related to city.

Indhuja et al. [4] proposed an approach for extraction of features from the product reviews. The researcher classified the features into positive (+ve), negative (-ve) and neutral (0) using feature based sentiment classification method. Preprocessing was done for removing noise, feature extractions and corresponding descriptors. It was extended to encompass the effect of linguistic hedges and fuzzy functions to imitate the outcome of concentrators, modifiers and also dilators. The method was assessed on SFU corpus and the outcome showed that fuzzy logic performed perfectly in SA.

Luneva et al. [5] projected a new method for the evaluation of the social network users' sentiment. The proposed method considered the user's influence and authorship of a person for several messages. The results were used to analyze the reviews on products and user communication interfaces designed for robotics. It was deployed in developing analytic tools and made use of Web Observatories datasets.

Dragoni et al. [6] modeled the fuzzy logic for concept polarities. The researchers explained the uncertainty associated with the fuzzy logic applied to different domains. The proposed approach combined two linguistic resources, namely SenticNet and WordNet. Later, the derived knowledge graph was manipulated by a graph-propagation algorithm. It was propagated sentiment of categorized (labeled and unlabelled) datasets. The proposed work was implemented and performed on the Blitzer dataset. The results demonstrated its feasibility in original problems.

Reshma et al. [7] proposed a mechanism to understand the overall opinion about aspects of products from customers and manufacturers. The authors presented a more precise and realistic value of opinion that was retrieved through Naïve Bayesian classifier and fuzzy method. A hybrid approach was proposed for sentiment classification and used Sentiwordnet, Naïve Bayesian classifier and fuzzy function to determine value of opinion. Not only the opinions identification, but also linguistic hedges were also performed in the proposed mechanism. Fuzzy rules were applied to magnify the effect of opinion to provide varying degrees of values to describe vague and imprecise information.

Jenifer et al. [8] proposed a mechanism for considering the aspects of the objective sentences while calculating the polarity of a review. The authors suggested adding negative polarity of the words in the sentiment analysis of a product review. The researchers argued that the polarity of the objective sentences presented in the aspects will be helpful in improving the accuracy of the aspect polarity to support the customer. And also the authors proved their argument with the product reviews of the Samsung Galaxy note 7.

Supriya et al. [12] presented a three step algorithm for analyzing the public sentiment in Twitter tweets. The algorithm consists of Cleaning, Entity identification, and Classification for sentiment analysis. The results shown in the form of graphs and the performance calculated using precision, recall and accuracy. It is easy and the experiment was yet to be performed.

3. PROPOSED FRAMEWORK OF ASFUL MECHANISM

The objective of this research work is to propose a novel mechanism to reduce the neutral score using the Imputation of Missing Sentiment technique (IMS) in sentiment classification using fuzzy logic. The proposed system is explained as follows: System architecture represents the framework and Methodology to present the procedure for the ASFuL mechanism. The Work flow gives the flow and describes the steps in the proposed mechanism.

3.1 SYSTEM ARCHITECTURE OF ASFUL MECHANISM

The high level design model of the proposed ASFuL mechanism is presented in Fig.1. The building blocks of the ASFuL mechanism is listed below:

- Data Collection
- Sentence Separation
- Sentence Label Propagation
- Fuzzification
- Defuzzification
- Aspect-wise Summarization

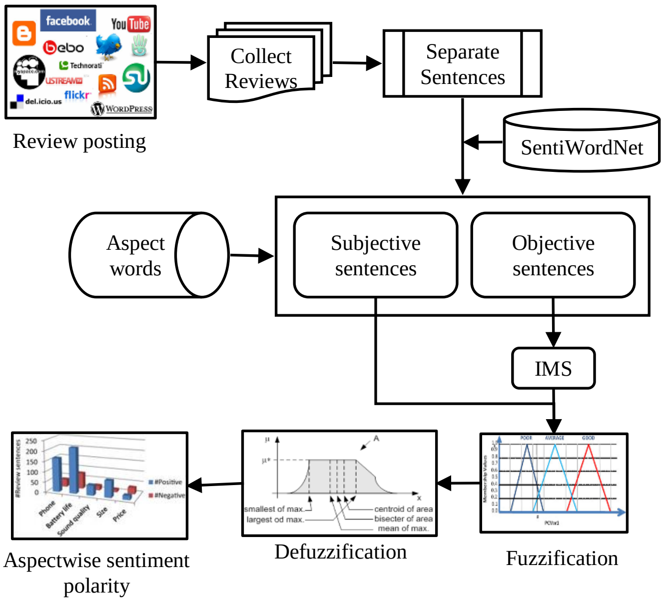

Fig.1. Architecture of the proposed ASFuL mechanism

These operations will be explained in the forthcoming sections.

3.2 METHODOLOGY OF PROPOSED ASFUL MECHANISM

The methodology for ASFuL mechanism is given as an algorithm named as ASFuL Algorithm. This is helpful to understand the system architecture of the proposed system in a better way. The ASFuL algorithm is presented below.

Step 1: Start

Step 2: Collect the Reviews

Step 3: Call the Sentence_Count()

Step 4: Call Label_Probagation()

Step 5: Check Label

Step 6: If Label='Opin' then

Seed Aspect words

Get the Sentiment value using Senti_WordNet

else

Seed Aspect Word

Call IMS()

End if

Step 7: Apply Fuzzification rule

Step 8: Compute the result by Defuzzification

Step 9: Generate Aspect-wise Summary

Step 10: Stop the Process

3.3 WORKFLOW OF PROPOSED ASFUL MECHANISM

The workflow of the ASFuL is given in Fig.2. It is explained in the following section with detailed description.

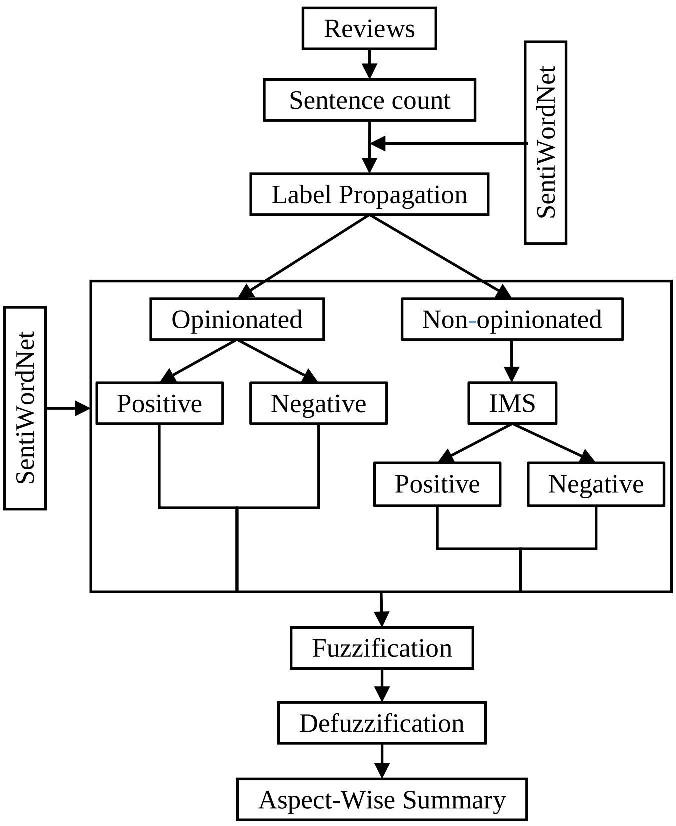

Fig.2. Workflow of Proposed mechanism

3.3.1 Data Collection:

It is nothing but collection of online reviews. The advent of internet made the users to post their reviews on products and events through social networking sites such as, blogs and forums. This can be large in volume and helps in decision making. Usually data can be collected using surveys like censuses, sample surveys and administrative data.

3.3.2 Sentence Separation:

A separation is a break, or the place where a split happens. A review document has been made up of many sentences. These sentences are separated by considering dots, commas, and semicolons. And also the number of sentences in a review is to be counted in this operation. Each sentence should be separated as words and stored in a word array for further processing. The procedure for Sentence Separation is given in Fig.3.

```
Function Sentence_Count()
{
Ch1='y', i=0,sen_count=0, word=0, sen[], word_array[]
While ch1=='y' do
  Read a character ch
    if ch=='\t' || ch==' ' then
     word=word+1
       if ch=='.' or ch=='!' or ch=='?' or ch=='\n' then
          sen_count=sen_count+1
       else  sen[i]= ch
        end if
    end if
   word_array[i]=ch
   i=i+1
   print 'Want to Continue?'
   read ch1
End while
Print sen_count
}
```

Fig.3. Function Sentence_Count()

3.3.3 Sentence Label Propagation:

It is the process of labeling all the sentences in a review. In this process, the separated sentences are labeled using SENTIWORDNET. The sentences are labeled as "Opinionated" if it is with a sentiment word, otherwise labeled as "Non-Opinionated". The Label propagation function is depicted in Fig.4.

Function Label_Propagation()
Step 1: Seed SentiWordNet
Step 2: If Sentiment Word present
Label 'Opin'
Else
Label 'Non-Opin'
End if

Fig.4. Function Label Propagation

The opinionated sentences are called as 'Subjective sentences' and Non-opinionated sentences are called as 'Objective sentences'. Aspect word list is used to find the aspect words that are presented in the review. If aspects are presented then, the corresponding strength of the sentiment retrieved from the SENTWORDNET dictionary. Normally, the objective sentences are factual sentences which do not imply any impact because of their nature, but IMS (Imputation of missing sentiment) function helps to fill the missing sentiment for the existing aspects and get the strength of the sentiment from the dictionary. The function IMS is given in Fig.5.

```
Function IMS()
{
If aspect present
        Impute the sentiment
        Goto step 6
Else
        Exit;
}
```

Fig.5. Function IMS()

3.3.4 Fuzzification:

In a fuzzy system, the first step includes the identification of input and output variables. Once the input variables and Membership Function are defined, the rule-base design (or decision matrix of the fuzzy knowledge-base) has to be defined which is composed of expert IF-THEN rules. These rules transform the input variables into an output. This will tell the risk of operational problems usually low, normal and high. The fuzzification mechanism is explained in Fig.6.

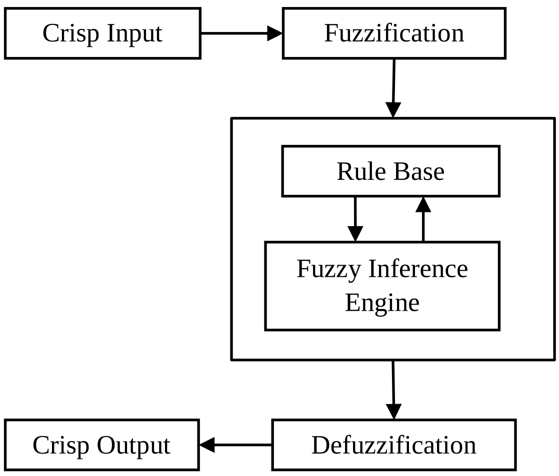

Fig.6. Working mechanism of Fuzzification and Defuzzification

3.3.5 Membership Function Design:

The crisp values of the linguistic variables are transformed to the given fuzzy sets by the function known as Membership Function (MF) [9]. The inputs are mapped into the degree of membership by suitable membership functions for the partitions of linguistic variables. There are many forms of membership functions. The selection of appropriate membership functions for fuzzy sets is significant in a fuzzy system which perfectly represents the fuzzy modeling. In the proposed research, the classification labels for the intensity of semantic orientation and positive or negative polarity of a given sentence are identified as:

- Strong Negative (SN)
- Negative (NN)
- Positive (PP)
- Strong Positive (SP)

Either the sentence is classified as objective or subjective. Then, if it is considered to be subjective, it could either be negative or positive, and then again either strong positive or positive, or strong negative or negative. In this work, triangular and trapezoidal membership functions are used for respective fuzzy terms by carefully analyzing the scores of the lexicon provided in SentiWordNet. In case, the MF are trapezoids (in this case ‘Weak’ or ‘Strong’), the MF can be defined in Eq.(1) and Eq.(2).

$$\mu_{weak}(trap)\left(x; p_0, p_1, p_2, p_3\right) = \max\left(\min\left(\frac{x-p_0}{p_1-p_0}, 1, \frac{p_3-x}{p_3-p_2}\right), 0\right) \tag{1}$$

$$\mu_{strong}(trap)\left(x; p_3, p_4, p_5, p_6\right) = \max\left(\min\left(\frac{x-p_3}{p_4-p_3}, 1, \frac{p_6-x}{p_6-p_5}\right), 0\right) \tag{2}$$

In case, the MF are triangles (in this case ‘Positive’ or ‘Negative’), the MF can be defined in Eq.(3).

$$\mu_{Pos/Neg}(tri)\left(x; p_2, p_3, p_4\right) = \max\left(\min\left(\frac{x-p_2}{p_3-p_2}, \frac{p_4-x}{p_4-p_3}\right), 0\right) \tag{3}$$

The rules for applying fuzzy sets in the ASFuL are explained in Table.1.

Table.1. Fuzzy Controller Rules

P N	SP	PP	NN	SN
SP	SP	PP	NN	NN
PP	PP	SP	NN	NN
NN	NN	NN	SN	SN
SN	NN	SN	SN	PP

Example: For a given variable x involved in the development of a problem (the risk output has its own MF, low, normal and high risk), for example, rule can “say” that [11]:

- IF *x* is low THEN risk of problem is low.
- IF *x* is normal THEN risk of problem is normal.
- IF *x* is high THEN risk of problem is high.

According to these rules, suppose the degree of membership for *x* is 0.6 to the MF low, then the risk of the problem will be 0.6 low, too.

3.3.6 Defuzzification:

Defuzzification is the conversion of a fuzzy quantity to a specific quantity. Three steps involves in calculating the area of the resulting figures for each MF. Taking into account that these areas will be triangles and trapezoids most of the times. If the degree of membership is not equal to one, the area will have a trapezoidal shape instead of a triangle. An example of three MF for a given input is shown in Fig.7.

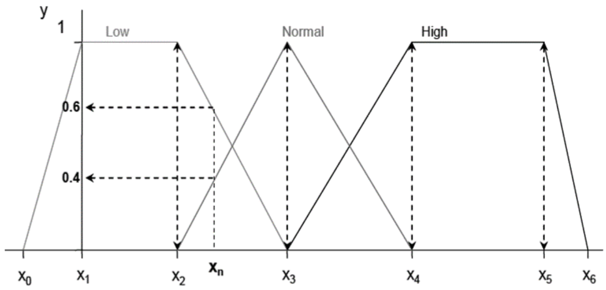

Fig.7. Example of the three MF for a given input

Rules for defuzzification are given below:

if ($x \leq 0.25$) then

y = 'Strong Negative'

if ($x \geq 0.25$ and $x \leq 0.5$) then

y = 'Negative'

if ($x \geq 0.5$ and $x \leq 0.75$) then

y = 'Positive'

if ($x \geq 0.75$ and $x \leq 1$) then

y = 'Strong Positive'

3.3.7 Aspect-wise Summarization:

The fuzzy MF shows the aspect wise result for better decision. This helps the user and a company to improve the weaker section.

4. CONCLUSION

In this paper, a novel, 3-Phase mechanism called ASFuL is proposed for aspect based sentiment summarization. It uses Fuzzy Logic to classify sentiments from the product reviews. The first step shows the architecture, the second step identifies methodology and the last step expresses the workflow of classifying the aspects as related MF. It is well suited for product reviews or events which will produce a high level of accuracy. It also incorporates the non-opinionated sentences using IMS mechanism which plays a significant role in producing accurate result. This mechanism is also useful to reduce the neutral score in sentiment polarity. The proposed ASFuL mechanism is simple and expects better performance when compared to the already existing methods however the experiment is yet to be performed.

REFERENCES

[1] O. Appel, F. Chiclana, J. Carter and H. Fujita, "A Hybrid Approach to Sentiment Analysis", *Proceedings of IEEE Congress on Evolutionary Computation*, pp. 4950-4957, 2016.

[2] Suresh Hima and S. Gladston Raj, "An Unsupervised Fuzzy Clustering Method for Twitter Sentiment Analysis", *Proceedings of IEEE International Conference on Computational Systems and Information Systems for Sustainable Solutions*, pp. 80-85, 2016.

[3] Giuseppe D. Aniello, Angelo Gaeta, Matteo Gaeta, Vincenzo Loia and Marek Z. Reformat, "Collective Awareness in Smart City with Fuzzy Cognitive Maps and Fuzzy sets", *Proceedings of IEEE Sponsored International Conference on Fuzzy Systems*, pp. 1554-1561, 2016.

[4] K. Indhuja and P.C. Reghu Raj, "Fuzzy Logic Based Sentiment Analysis of Product Review Documents", *Proceedings of 1st International Conference on Computational Systems and Communications*, pp. 18-22, 2014.

[5] E.E. Luneva, P.I. Banokin, A.A. Yefremov and T. Tiropanis, "Method of Evaluation of Social Network User Sentiments Based on Fuzzy Logic", *International Journal of Key Engineering Materials,* Vol. 685, pp. 847-851, 2016.

[6] Mauro Dragoni, Andrea G. B. Tettamanzi and Celia Da Costa Pereira, "Propagating and Aggregating Fuzzy Polarities for Concept-Level Sentiment Analysis", *Cognitive Computation*, Vol. 7, No. 2, pp. 186-197, 2015.

[7] V. Reshma and Ansamma John, "Aspect Based Summarization of Reviews Using Naive Bayesian Classifier and Fuzzy Logic", *Proceedings of IEEE International Conference on Control, Communication and Computing*, pp. 617-621, 2015.

[8] A. Jenifer Jothi Mary and L Arockiam, "IMS: Imputation of Missing Sentiment in the Aspect based Sentiment Analysis", *Proceedings of International Conference on IoT, Data Science and Security*, pp. 263-273, 2017.

[9] S. Zhang, C. Zhu, J.K.O. Sin and P.K.T. Mok, "A Novel Ultrathin Elevated Channel Low-Temperature Poly-Si TFT", *IEEE Electron Device Letters*, Vol. 20, No. 11, pp. 569-571, 1999.

[10] Jorge Ropero and Ariel Gomez, "A Fuzzy Logic Intelligent Agent for Information Extraction Introducing a New Fuzzy Logic-based Term Weighting Scheme", *Expert Systems with Applications*, Vol. 39, No. 4, pp. 4567-458, 2012.

[11] A Fuzzy Implementation, http://apps.ensic.inpl-nancy.fr/benchmarkWWTP/RiskAnalysis/RiskWeb/RiskModule_070423_fichiers/Fuzzy_implementation_070423.pdf.

[12] B.N. Supriya, V. Kallimani, S. Prakash and C.B. Akki, "Twitter Sentiment Analysis Using Binary Classification Technique", *Proceedings of International Conference on Nature of Computation and Communication*, pp. 391-396, 2016.