American Journal of Engineering Research (AJER)
e-ISSN: 2320-0847 p-ISSN : 2320-0936
Volume-7, Issue-4, pp-45-53
www.ajer.org
Research Paper Open Access

SVM and Fuzzy SVM Based Opinion Mining In Tamil Using R

[1]Sharmista A, Ramaswami M [2]
[1]*Madurai Kamaraj University Tamil Nadu, India*
[2]*Madurai Kamaraj University Tamil Nadu, India*
Corresponding author:sharmista A

ABSTRACT: *Opinion Mining is growing popular nowadays in different domains. Huge number of reviews and feedbacks given by the customers are available online for each and every product they purchase. Products purchased through the shopping website like flipkart, amazon and snap deal reviews were examined in this paper. The main aim is to identify Tamil opinions for mobile products and examine which product is best among all the products. Sentence level analysis is done and each feature is extracted by applying POS tagging. From this we can able to classify the reviews as* நேர்மறை *(Positive),* மிகவும்சாதகமான) *Most Positive),* எதிர்மறை (*Negative),* மிகவும்எதிர்மறை *(Most Negative) and* யாரும் *(None). For this to be implemented we have applied machine learning approach like SVM and Fuzzy SVM classification. We compare the performance of both approaches and conclude which is giving more accuracy. The experimental result shows that fuzzy svm classification gains more accuracy than svm classification.*

KEYWORDS *-Data mining, Feature Extraction, Fuzzy SVM, Opinion mining, POS tagging, Sentiment analysis, SVM.*

Date of Submission: 23-03-2018 Date of acceptance: 07-04-2018

I. INTRODUCTION

Nowadays Opinion Mining is an arena of study of Information Retrieval (IR) and Natural Language Processing (NLP) task and shares some characteristics with other disciplines such as text mining and Information Extraction. Opinion mining is a technique to detect and extract subjective information in text documents [1]. In this paper we examined especially Tamil reviews at the sentence level extraction. The sentiment may be his or her judgment, attitude or appraisal. In recent years, "opinion mining" has been drawing more attention. Sentiment analysis has different names like opinion mining, sentiment extraction, or affective rating. All subjective content was extracted for future analysis. The subjective content consists of all sentiment sentences. A sentiment sentence is the one that contains, at least, one positive or negative word. All of the sentences were firstly tokenized into separated Tamil words [12]. Every word of a sentence has its syntactic role that defines how the word is used. The syntactic roles are also known as the parts of speech. There are 8 parts of speech in English: the verb, the noun, the pronoun, the adjective, the adverb, the preposition, the conjunction, and the interjection. In natural language processing, part-of-speech (POS) taggers have been developed to classify words based on their parts of speech. For sentiment analysis, a POS tagger is very useful because of the following two reasons: 1) Words like nouns and pronouns usually do not contain any sentiment, filter out such words with the help of a POS tagger; 2) A POS tagger can also be used to distinguish words that can be used in different parts of speech.

Sentiment analysis (also known as opinion mining) is the process by which text is analyzed to extract opinion and assign a relevant sentiment, usually positive, most positive, negative, most negative or neutral. The NLP process is a combination of pre-processing steps and applying POS tagging techniques, combined with appropriate machine learning classification techniques like SVM, Naïve bayes, Decision tree and Roughset [2]. In contrast to these, our proposed fuzzy SVM and SVM approach can extract the feature to which a sentence contains a specific sentiment. Based on the predicted classes (e.g. நேர்மறை and எதிர்மறை) and the corresponding fuzzy membership values, the proposed approach allows the inference of more refined categories (e.g. யாரும்) or intensities (e.g. மிகவும்சாதகமான, மிகவும்எதிர்மறை) of sentiment without the need to define more classes. We compare the performance of our proposed SVM and Fuzzy SVM approach with the common techniques used in sentiment analysis mentioned above.

The remainder of the paper is structured as follows: Section II covers related studies that outline previous work on sentiment analysis. Section III outlines the proposed architecture of our system using Tamil sentiment analysis. Section IV covers the algorithmic approach of the fuzzy classifier with a SVM classifier. Section V presents the experimental result analysis of overall dataset that has collected for this paper work. Section VI gives the comparative result analysis of SVM and Fuzzy SVM. Section VII concludes the paper and outlines further research in this area.

II. RELATED STUDIES

In this section we give an overview of related work to sentiment analysis, including types of classification tasks, types of data, pre-processing techniques and machine learning algorithms. We also give a comprehensive overview of fuzzy approaches for text processing.

Most approaches in the sentiment analysis area focus on polarity or opinion classification into positive or negative classes; some researchers also include a neutral class [3]. Other classification tasks focus on subjectivity vs. objectivity, on predicting categories of emotions (e.g. anger, fear joy), or the strength of sentiment. The research in this area has been applied to a variety of data sources, such as movie reviews, product reviews, Facebook data and micro-blog data. In this research reported in this paper we use product reviews data on mobile.

To prepare the data for classification tasks, several pre- processing techniques are typically used, such as spelling corrections, tokenization (splitting the text in tokens such as words), and removal of numbers, punctuation and repeated letters. Machine learning approaches have been successfully used for sentiment analysis [10] and a number of algorithms have been shown to perform well for sentiment analysis tasks: Naïve Bayes, Support Vector Machines, Maximum Entropy and Decision Trees.

In recent years, fuzzy approaches have started to emerge for text processing [4]. In 2012, a review of fuzzy approaches for natural language processing highlighted that the percentage of papers relating to fuzzy approaches is very low over all the papers in the literature of natural language processing despite the suitability of fuzzy approaches for text processing and classification. Since then, a number of fuzzy approaches have been proposed for a variety of applications, as outlined below. A fuzzy rule based approach was proposed in [5], which was shown to lead to a reduction in computational complexity while maintaining a similar performance to other well-known machine learning approaches which works on the concept of transportation and city traffic controlling for safe travelling.

In this paper we build on the work to discuss how the membership degree values can be used for more refined outputs, including different intensities or strengths of sentiment. We compare the performance of fuzzy approach with machine learning algorithms known to perform well on sentiment analysis tasks such as SVM. Unlike previous approaches, we do not only look at the classification performance, but at ways in which fuzzy approaches can be used to provide more refined, interpretable outputs.

The aim of paper is to find best effective features which provide better result and also provide better feature selection method. They have also express that how unigram feature set can be reduced to get better result. Pre-processing steps such as stop word removal, stemming, postagging is performed. After pre-processing, this pre-processed data is converted into numerical vector using scaling techniques [11]. Support vector machine and fuzzy SVM classifiers are used to classify numerical vector.

III. ALGORITHMIC APPROACH

Support Vector Machines

Support Vector Machine (SVM) is one of the majorly experimented data mining techniques in sentiment analysis. It is a kernel-based supervised learning technique. The use of SVM in sentiment analysis can be said to be popular because of its non-linear nature which makes it simple to evaluate both theoretically and computationally. SVM has the greatest efficiency at traditional text categorization when compared with other classification techniques like Naïve Bayes and Maximum Entropy.

Given a set of training examples, each marked as belonging to one of two categories, an SVM training algorithm builds a model that assigns new examples into one category or the other, making it a non-probabilistic binary linear classifier. Every data represented as a vector is classified in a particular class. Now the task is to find a margin between two classes that is far from any document. The distance defines the margin of the classifier, maximizing the margin reduces indecisive decisions. SVM also supports classification and regression which are useful for statistical learning theory and it helps recognizing the factors precisely, that needs to be taken into account, to understand it successfully.

In this work, SVM is used for sentiment classification. First module is sentiment analysis and Support vector machines perform sentiment classification task on mobile product review data. The goal of a Support Vector Machine (SVM) classifier is to find a linear hyperplane (decision boundary) that separates the data in such a way that the margin is maximized. Each boundary has an associated margin. If we choose the one that maximizes the margin we are less likely to misclassify unknown items in the future.

Fuzzy based semantic knowledge

An ontology is shared knowledge of a specific domain among people and systems [6]. It is written in a specific language called a web language (WL). To achieve efficiency for the proposed system, a classic ontology was designed using fuzzy svm. The fuzzy svm plug-in is used to assert fuzzy terms in the ontology. The classes, instances, concepts, and axioms of a classic and a fuzzy ontology are the same. However, all the concept values of a classic ontology are blurry terms. A classic ontology cannot handle uncertainty. A fuzzy svm is generally defined to express vague knowledge using fuzzy concepts [4]. Therefore, this system needs fuzzy ontology based semantic knowledge to handle any kind of situation related to sentiment analysis. Useful transportation knowledge is accumulated regarding mobile product reviews like positive, negative, most positive, most negative and neutral.

A fuzzy set, F, over the universe of discourse A can be shown by its membership function μF, which presents an element 'A' in the interval [0,1]. $\mu F(A):A'[0,1]$ shows that, A′ belongs to A and μF presents the membership degree by which $A' \in A$. A′ is considered a full member of set A if $\mu F(A) = 1$. A is considered a partial member of set A if $\mu F(A)$ is between zero and one (e.g., 0.63).

A fuzzy ontology exchanges the knowledge among feature extractions, reviews classifications, and performs feature polarity identification. Therefore, the description of polarity for feature classification using a fuzzy svm expedites the proposed transportation sentiment analysis system. This system categorizes the features and extracts the correct feature polarity terms. The fuzzy svm defines the concepts of product feature extraction in the collection of reviews.

The fuzzy svm efficiently employed for review and tweet classification [2]. The main task is precise data collection, which can accelerate the fuzzy ontology development process. Ontology is a set of classes or concepts, properties (data types), instances, and relationships (object properties). There are two types of relation; fuzzy object relations and fuzzy data type relations. Fuzzy object relations are used to link instances with a certain degree. A fuzzy data type is used to assign an instance that includes the fuzzy predicate of product price. We gathered all information, like ontology classes, object properties, data properties, and fuzzy data types. Data and object properties define the relationships of the classes connected to the basic data types. The fuzzy data type shows the interval of membership variables. Sentiment analysis employs this fuzzy ontology-based semantic knowledge and evaluates pairs of product features and transportation activity polarity. In the last tasks, these feature polarity reviews are gathered from all the tweets and reviews. The final sentiment analysis result and polarity values are obtained for customer attitudes and product feature polarity [8].

IV. PROPOSED ARCHITECTURE

Sentiment analysis is the process of determining whether social media publications are positive or negative. Due to ambiguities inherent in language it can be very challenging to program software analysis tools to accurately resolve whether a word is positive or negative. Most of the sentiment analysis materials available are in English. So, to interpret sentiment in Tamil, for example, which is spoken by approximately 20 per cent of the population, involves a time-consuming and often unreliable process of machine translation before analysis can take place". Hence the need for sentiment analysis in Tamil is essential for a particular person to take decision on whether to buy a particular product or not [15]. The overall architecture of the system is shown in figure 1.

The figure below shows how the data is retrieved and some pre-processing steps to be done to clean up unwanted data. Features are extracted from the sentence and matched with opinion lexicon. Then by applying svm and fuzzy svm algorithm, classification will be done and opinions are summarized such as நேர்மறை (Positive), யாரும் (None) and எதிர்மறை (Negative) for three class problem and மிகவும்சாதகமான (More Positive), நேர்மறை (Positive),யாரும் (None), மிகவும்எதிர்மறை (More Negative), and எதிர்மறை (Negative) for five class problem.

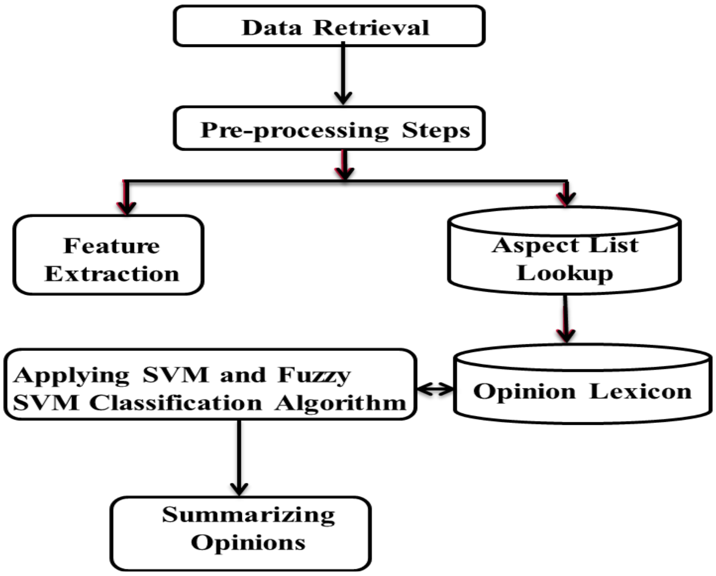

Figure 1: Flow for generating opinion from feedback

Sentiment related properties like subjectivity and orientation are well defined in this approach. Most researches were focused mainly on these domains of sentiment analysis to find out the best opinion from the list of reviews. Final result will be easier to choose the product by them. For example, நல்ல,சிறந்த are positive terms while கெட்ட,தவறு and மோசமான are negative terms. செங்குத்து, மஞ்சள் and திரவ are objective terms. சிறந்த and மோசமான are more intense than நல்ல and மோசமான.

A. Document Pre-processing

Currently, input is of plain text format. By applying part of NLP steps for obtaining the token of words is as follows:

Statement extraction: firstly, from each post, sentences are individuated, that are parts of text ending with a full stop, comma, question mark, exclamation mark or semicolon. Subsequently, conjunctions are analysed for dividing sentences into statements, which are parts of text expressing only one meaning.

Tokenization: In this step, each statement is divided into tokens, which are parts of text bounded by a separator (space, tab or end of line).

Stemming and Lemmatization: In order to reduce the number of different terms, each token is transformed reducing its inflectional forms to a common base form. The main difference among stemming and lemmatization‘s that the former extract ―brutally the root of a word (e.g. bio is the stem of biology, biocatalyst and biochemical); while the latter uses a vocabulary for returning the dictionary form of a word, that is the lemma.

Tagging and Stop words elimination: Some word categories are too common to be useful to distinguish among statements. Hence, in this step articles, prepositions and conjunctions are first recognized and then removed. At the same time, we also remove proper nouns, which usually don‘t have an affective content.

B. Word Sense Disambiguation

In automatic text summarization, word sense disambiguation is important and many different approaches have been taken [7]. The nouns and the verbs are first extracted from each sentence together with their senses given in Lexicon as the input to the following process for sense disambiguation:

1) For a word to be disambiguated, the process first scores the semantic relatedness between any two senses, one for this word and the other for any other word in the same sentence.

2) Note each sense in Lexicon is semantically related with a set of similar senses. The score computed in the previous step reflects the direct relatedness between any two senses. It is like a local link.

3) The final score of each sense for the word is the sum of the score given by Step 1.) And the half value of the score given by Step 2.).

4) Among all the senses of the word, the sense with the highest score is selected as the candidate sense of the word. After all words in a sentence are disambiguated, this phase builds and reports the sense representation for the sentence in terms of Lexicon senses to indicate what concept the sentence may cover.

C. Natural Language Parsing Techniques

The main idea for entity discovery is to discover linguistic patterns and then use the patterns to extract entity names. However, basic methods need a large number of training samples, and it is very time consuming. This section proposes an unsupervised learning method [9]. The most common idea of the algorithm is that the user starts with a few attributes. The system bootstraps from them to find more attributes in a set of documents (or posts). Sequential pattern mining is carried out at each duplication to find more attributes based on already found attributes. The iterative process ends when no new attribute names are found.

Pruning methods are also proposed to remove those unlikely entities. Given a set of seed entities E = {e1, e2 …en}, the algorithm consists of the following iterative steps:

Step 1 – data preparation for sequential pattern Mining
Step 2 – Sequential pattern mining
Step 3 – Pattern matching to extract candidate Entities
Step 4 –Candidate pruning
Step 5 –Pruning using relations among whole review dataset.

V. EXPERIMENTAL RESULT ANALYSIS

R is a sophisticated statistical software package, which provides new approaches to data mining. We have analyzed an effect of product review dataset obtained from Opinion Lexicon [14]. The SVM and Fuzzy SVM classification algorithm is executed to predict the best product by identifying the total number of positive opinions. The number of instances used for analysis of product data is 5000.

Table I. shows result of svm classification (3 class) with opinion lexicon for product reviews dataset run on R platform. It depicts the நேர்மறை (Positive),யாரும் (None), எதிர்மறை (Negative) reviews which are classified based on the number of instances 5000. It is predicted that there are 1853 நேர்மறை (Positive) reviews, 160 யாரும் (None) classified reviews with 1035 error rate and 1005 எதிர்மறை (Negative) reviews in case when matched with opinion lexicon.

Table I: Confusion Matrix of SVM Classification (3 Class) With Opinion Lexicon

Predicted Class

Actual Class	நேர்மறை (Positive)	யாரும் (None)	எதிர்மறை (Negative)
	1853	36	245
	532	**160**	503
	570	96	**1005**

Table II. Shows results of SVM classification (5 Class) with opinion lexicon for product reviews dataset run on R platform. It depicts the மிகவும்சாதகமான (More Positive)**,**நேர்மறை(Positive),யாரும்(None), மிகவும்எதிர்மறை(More Negative), எதிர்மறை (Negative) reviews which are classified based on the number of instances 5000. It is predicted that there are 79 மிகவும்சாதகமான(More Positive) reviews, 1560 நேர்மறை(Positive) classified reviews , 239 யாரும்(None) reviews, 124 மிகவும்எதிர்மறை(More Negative) reviews and 685 எதிர்மறை (Negative) reviews in case when matched with opinion lexicon.

Table II: Confusion Matrix of SVM Classification (5 Class) With Opinion Lexicon

Predicted Class

Actual Class	மிகவும் சாதகமான (More Positive)	நேர்மறை (Positive)	யாரும் (None)	மிகவும் எதிர்மறை (More Negative)	எதிர்மறை (Negative)
	79	313	0	3	26
	28	**1560**	54	4	232
	17	346	**239**	0	419
	15	117	8	**124**	131
	11	356	146	87	**685**

Table III. shows result of fuzzy svm classification (3 class) with opinion lexicon for product reviews dataset run on R platform. It depicts the நேர்மறை (Positive),யாரும் (None), எதிர்மறை (Negative) reviews which are classified based on the number of instances 5000. It is predicted that there are 1720 நேர்மறை (Positive) reviews, 60 யாரும் (None) classified reviews with 1135 error rate and 852 எதிர்மறை (Negative) reviews in case when matched with opinion lexicon.

Table III: Confusion Matrix of Fuzzy SVM Classification (3 Class) With Opinion Lexicon

Predicted Class

Actual Class	நேர்மறை (Positive)	யாரும் (None)	எதிர்மறை (Negative)
	1720	32	382
	585	**60**	550
	758	61	**852**

Table IV. Shows results of fuzzy svm classification (5 Class) with opinion lexicon for product reviews dataset run on R platform. It depicts the மிகவும்சாதகமான (More Positive),நேர்மறை(Positive),யாரும்(None), மிகவும்எதிர்மறை(More Negative), எதிர்மறை (Negative) reviews which are classified based on the number of instances 5000. It is predicted that there are 93 மிகவும்சாதகமான(More Positive) reviews, 1552 நேர்மறை(Positive) classified reviews , 209 யாரும்(None) reviews, 67 மிகவும்எதிர்மறை(More Negative) reviews and 644 எதிர்மறை (Negative) reviews in case when matched with opinion lexicon.

Table IV: Confusion Matrix of Fuzzy SVM Classification (5 Class) With Opinion Lexicon

Predicted Class

Actual Class	மிகவும் சாதகமான (More Positive)	நேர்மறை (Positive)	யாரும் (None)	மிகவும் எதிர்மறை (More Negative)	எதிர்மறை (Negative)
	93	287	0	26	15
	18	**1552**	70	3	235
	6	349	**209**	8	449
	0	198	35	**67**	95
	3	479	130	29	**644**

VI. COMPARATIVE RESULTS

In general, the performance of sentiment classification is evaluated by using eight indexes. They are Specificity, Sensitivity, Precision, Recall, F1-measure, G-measure, Detection rate and Accuracy [13]. The common way for computing these indexes is based on the result shown below:

Table V: Overall Performance of SVM and Fuzzy SVM Classification (3 Class) with Opinion Lexicon

Machine Learning Approach						
Algorithm	SVM (3 Class)			Fuzzy SVM (3 Class)		
Class	நேர்மறை (Positive)	யாரும் (None)	எதிர்மறை (Negative)	நேர்மறை (Positive)	யாரும் (None)	எதிர்மறை (Negative)
Specificity	0.8626	0.7802	0.7949	0.7863	0.7658	0.7453
Sensitivity	0.6271	0.5479	0.5733	0.5615	0.3922	0.4776
Precision	0.591	0.0584	0.3506	0.6126	0.0306	0.3568
Recall	0.3706	0.032	0.201	0.344	0.012	0.1704
F1-measure	0.7282	0.2152	0.5870	0.6619	0.0890	0.4932
G-measure	0.7379	0.2709	0.5872	0.6728	0.1403	0.4935
Detection Rate	0.3706	0.032	0.201	0.344	0.012	0.1704
Accuracy	0.7448	0.6641	0.6841	0.7639	0.7790	0.7115

Table VI: Overall Performance of SVM Classification (5 Class) with Opinion Lexicon

Machine Learning Approach					
Algorithm	SVM(5 Class)				
Class	மிகவும் சாதகமான (More Positive)	நேர்மறை (Positive)	யாரும் (None)	மிகவும் எதிர்மறை (More Negative)	எதிர்மறை (Negative)
Specificity	0.9295	0.8622	0.8282	0.9433	0.8289
Sensitivity	0.5267	0.5795	0.5347	0.5688	0.4588
Precision	0.03	0.5384	0.0894	0.0436	0.2986
Recall	0.0158	0.312	0.0478	0.0248	0.137
F1-measure	0.2767	0.6827	0.3256	0.4046	0.4932
G-measure	0.3144	0.6938	0.3538	0.4226	0.4945
Detection Rate	0.0158	0.312	0.0478	0.0248	0.137
Accuracy	0.7281	0.7209	0.6815	0.7561	0.6439

Table VII: Overall Performance of Fuzzy SVM Classification (5 Class) with Opinion Lexicon

Machine Learning Approach					
Algorithm	Fuzzy SVM (5 Class)				
Class	மிகவும் சாதகமான (More Positive)	நேர்மறை (Positive)	யாரும் (None)	மிகவும் எதிர்மறை (More Negative)	எதிர்மறை (Negative)
Specificity	0.9328	0.8473	0.8218	0.9326	0.8200
Sensitivity	0.775	0.5417	0.4707	0.5038	0.4478
Precision	0.024	0.573	0.0888	0.0266	0.2876
Recall	0.0186	0.3104	0.0418	0.0134	0.1288
F1-measure	0.3438	0.6544	0.2853	0.2538	0.4730
G-measure	0.4138	0.6691	0.3104	0.2923	0.4738
Detection Rate	0.0186	0.3104	0.0418	0.0134	0.1288
Accuracy	0.8539	0.6945	0.6462	0.7182	0.6339

We calculate accuracy of classifier for both SVM (3 Class and 5 Class) and Fuzzy SVM (3 Class and 5 Class) classification, by training the algorithm on 5000 sentences for each class of pre-classification dataset and applying it on the rest of the remaining datasets.

By comparing the algorithm SVM and Fuzzy SVM for 3 class we achieved 0.7448 value of correct classification of opinions for SVM (3 Class) and 0.7639 for Fuzzy SVM (3 Class). From the result we conclude that Fuzzy SVM classification has more accuracy than SVM (3 Class) classification. By analyzing for 5 class we achieved 0.7281 for SVM (5 Class) and 0.8539 for Fuzzy SVM classification. From these 5 Class comparisons we conclude Fuzzy SVM achieves more accuracy.

VII.CONCLUSION

In this paper, a SVM and Fuzzy SVM approach was proposed for sentiment analysis, with a focus on polarity classification. We compared the accuracy of SVM and Fuzzy SVM approach which are known to be among the best performing techniques for sentiment analysis. In addition to performing well in terms of accuracy, the proposed approach has the advantage of more refined outputs based on the fuzzy membership degree values. From the result it shows that Fuzzy SVM achieves greater accuracy than SVM classification.

In future work, the fuzzy approach for polarity classification will be used for identifying the relationships between classes, through looking at how the membership degrees of different classes are correlated. We will also extend our fuzzy approach for the classification of categories of emotions, an emerging subarea of sentiment analysis.

REFERENCES

[1] Akash, AshleshMithur, Meghana, NehaKamath, Sathisha M.S., “Opinion Mining of Text Reviews Using Machine Learning Algorithm”International Journal of Emerging Research in Management &Technology ISSN: 2278-9359 (Volume-6, Issue-5) May 2017.

[2] BholaneSavita D and Deipali Gore., “Sentiment Analysis on Twitter Data Using Support Vector Machine” International Journal of Computer Science Trends and Technology (IJCST) – Volume 4 Issue 3, Page 365, May - Jun 2016 ISSN: 2347-8578.

[3] Bhumika M. Jadav and Vimalkumar B. Vaghela., “Sentiment Analysis using Support Vector Machine based on Feature Selection and Semantic Analysis”,International Journal of Computer Applications (0975 – 8887) Volume 146 – No.13, July 2016.

[4] Chris Jefferson, Han Liu, IEEE and MihaelaCocea., “Fuzzy Approach for Sentiment Analysis” Expert Syst. Appl., vol. 40, pp. 4241–4251, 2013.

[5] Farman Ali, DaehanKwak, Pervez Khan, S. M. Riazul Islam, Kye Hyun Kim, K. S. Kwak., “Fuzzy Ontology-based Sentiment Analysis of Transportation and City Feature Reviews for Safe Traveling” survey paper, 2016.

[6] Farman Ali, Kyung-sup Kwak, Yong-Gi Kim., “Opinion Mining based on fuzzy domain ontology and Support Vector Machine”. ELSEVIER publication issue on June, 2016.

[7] JayashriKhairnar and MayuraKinikar., “Machine Learning Algorithms for Opinion Mining and Sentiment Classification” International Journal of Scientific and Research Publications, Volume 3, Issue 6, June 2013,ISSN 2250-3153.

[8] ManveeChauhan and DivakarYadav.," Sentimental Analysis of Product Based Reviews Using Machine Learning Approaches" Journal of Network Communications and Emerging Technologies (JNCET), Volume 5, Special Issue 2, December (2015).

[9] Mariam Adedoyin-Olowe, Mohamed MedhatGaber and Frederic Stah., "A Survey of Data Mining Techniques for Social Media Analysis" survey paper, 2017.

[10] Muhammad Afzaal, Muhammad Usman, A. C. M. Fong, Simon Fong and Yan Zhuang., "Fuzzy Aspect Based Opinion Classification System for Mining Tourist Reviews", Research Article, Hindawi Publishing Corporation Advances in Fuzzy Systems Volume 2016, Article ID 6965725, 14 pages http://dx.doi.org/10.1155/2016/6965725.

[11] Poobana S, SashiRekha k., "Opinion Mining From Text Reviews Using Machine Learning Algorithm" International Journal of Innovative Research in Computer and Communication Engineering, Vol. 3, Issue 3, March 2015, ISSN(Online): 2320-9801, ISSN (Print): 2320-9798.

[12] Sharmista A and Ramaswami M – Naïve Bayes and NBTree Algorithm for opinion mining in Tamil, "IJCA", 2017.

[13] Sharmista A and Ramaswami M – Tree based opinion mining in Tamil for product recommendations using R, "IJCII",Vol 6, No.2, 2016.

[14] Sharmista A and Ramaswami M., "Rough Set based Opinion Mining in Tamil", International Journal Of Engineering Research And Development, e-ISSN: 2278-067X, p-ISSN: 2278-800X, Volume 13, Issue 9 (Sep., 2017), PP.43-51.

[15] Sharmista A, Ramaswami M and SharmilaViaz J., "Roughset based sentiment analysis in tamil", 8th International Conference preceedings in saveetha engineering college, Chennai.

Sharmista A "SVM and Fuzzy SVM Based Opinion Mining In Tamil Using R." American Journal of Engineering Research (AJER), vol. 7, no. 4, 2018, pp.45-53.