

# Feature level Sentiment Analysis on Movie Reviews

**Pallavi Sharma**

M.Tech Scholar,  
Poornima University, Jaipur, India  
Pallavisharma.ps34@gmail.com

**Nidhi Mishra**

Associate Professor,  
Poornima University, Jaipur, India  
Nidhi.mishra@poornima.edu.in

**Abstract** - Sentiment Analysis is a very challenging task in these days, which is a great requirement in every field as in Political field, Marketing field and in social field mainly. Many researchers have been taken a lot of interest in this area and proposed their work. The Sentiment analysis is done on the basis of Document, Sentence and feature levels. But the first two levels didn't consider object features that have been commented in a sentence. So the feature level sentiment analysis is more appropriate compare to both. Different types of tools and approaches have been used by the researchers for pre-processing, tagging, semantic orientation, and finally for calculating scores for deriving sentiments of the reviews. But the problems, found in analysis generally are, where reviews contain the negative, intensifier, conjunctive and synonyms words. And other problems are coreference resolution; anaphora resolution, named-entity recognition, and word-sense disambiguation. These problem are unresolved sometimes so that the performance of sentiment analysis decrease. A lot of work is done mainly in Product, News, Sports domains. We proposed our work in Movie Reviews because it's a more attractive area, for now-a-days generation where multiple sites allow users to submit reviews describing what they either liked or disliked about a particular movie. In this work we have proposed a system which classifies the polarity of the movie reviews on the basis of features by handling negation, intensifier, conjunction and synonyms with appropriate pre-processing steps. We have used SentiWordNet tool for calculating the scores of reviews.

**Keywords**— *Sentiment Analysis, Sentiment classification, SentiWordNet, Feature Selection*

## I. INTRODUCTION

Decision Making is a very essential part of human life, so that a good and right decision can be taken for our routine life. By knowing other's opinion on a particular thing, makes easy to take decisions. Manual efforts have been decreased for thinking and wasting time to take any kind of decision. Many movie sites take the reviews from users for movies, and on the basis of reviews they give rating to the movies. As the same, many social sites like Amazon, Flip cart etc have their product reviews, on those basis customers decide what should be more appropriate for purchasing. Same happens in all the fields like in politics, it can be used to predict election results. So Sentiment Analysis became a very demanding phase, so that things can be easier for everyone.

Generally three types of feedbacks can be considered from users as positive, negative and neutral for the reviews. These feedbacks can be derived using sentiment analysis on different levels as described below. The main purpose of sentiment analysis is to extract the main opinions, on which basis the decision can be made very right. Many techniques can be found for sentiment analysis. It can be done by classifying the text on document level, sentence level and feature level. Document level only gives whole document polarity, it considers the opinion of only one opinion holder. Sentence level deals with multiple sentences, but can distinguish only subjectivity or objectivity of sentences which is not equivalent to derive sentiment of particular feature. It can only determine that sentence's polarity that it is positive, negative and neutral. If there are more than one objects and features with complex sentences then above both levels failed to get accurate sentiment analysis. For it, Feature level sentiment analysis is adopted which not only deals with the polarity classification of particular features but also handle many problems as coordinating conjunctive sentences and comparative sentences which are quite tough to extract the opinion.

In our work, we have adopted feature level sentiment analysis on movie reviews where multiple problems are handled, as negative words, intensifiers, coordinating conjunctions and synonyms. We have proposed a system where sentiment scores of reviews is calculated using SentiWordNet. SentiWordNet is publically available library which contains the words as synset terms, their scores and is used for sentiment analysis tasks.

The rest of the paper is organized as Section-II presents the related research of the proposed work. Section- III describes the process flow of proposed work. Section-IV elaborates the methodologies of proposed system. Section-V highlights the experimental set up and performance evaluation with results and finally Section-VI concludes the proposed work.

## II. BACKGROUND AND RELATED WORK

Sentiment analysis has been a good research area from a long time. Many have presented their proposed work in a large volume, they have been solved many problems like anaphora resolution, thwarting, sarcasm, word sense disambiguation, intensifiers, negation etc. In our study, we have focused on sentence and feature level sentiment analysis, problems occur

in extraction of features and how to deal with complex sentences. Many researchers have used various types of tools like WordNet, Opinion Lexicon, MPQA, SentiWordNet etc for calculating the scores and used machine learning and rule based techniques for resolving their problems. Some have been proposed the systems for multilingual text and cross domain analysis also. We have summarized their work below.

[Peifeng Li, et-al, 2011] proposed a Dependency Tree. They used SVM Classifier to identify the polarity. The proposed sentiment word-based pruning strategy for dependency tree reduced the noise and showed good performance. [13]

[Zhongchao Fei, et-al, 2004] introduced the basic task of sentiment classification that how to get sentiment information in text using phrase patterns algorithms. They decided to use machine learning technology to evaluate the polarity and strength of phrase patterns. Their experiment achieved an 86% precision rate. [16]

[Farah Benamara, et-al, 2007] proposed an AAC-based (Adverb-Adjective Combinations) sentiment analysis technique that used a linguistic analysis of adverbs of degree. Their results showed that using adverbs and AACs produced significantly higher precision and recall. [7]

[Ann Devitt, et-al, 2007] presented a lexical cohesion based metric method of sentiment intensity and polarity in text. The results on polarity tag assignment, the best performers were the relation type and node specificity metrics using only modifiers, significant to the 0.05 level. [2]

[Bruno Ohana, et-al, 2009] proposed a method for applying SentiWordNet to derive a data set of document metrics and other relevant features, and performed an experiment on sentiment classification of film reviews using the polarity data set. The proposed method yielded an overall accuracy of 65.85% in SentiWordNet features. [13]

[Si Li Hao Zhang, et-al, 2011] presented a novel combined model based on phrase and sentence level's analysis. The improvements over sentence-level from phrase-level are found to be statistically significant for precision, recall and F-measure with large margin. [14]

[V.K. Singh, et-al, 2013] had implemented two Machine Learning based classifiers (Naïve Bayes and SVM), the Unsupervised Semantic Orientation approach (SO-PMI-IR algorithm) and the SentiWordNet approaches for sentiment classification of movie reviews. The accuracy of classification by NB was marginally better than the SVM and was close to the SO-PMI-IR algorithm. [15].

### III. PROCESS FLOW OF PROPOSED WORK

Figure 1 presents the process flow diagram of our proposed system of feature level sentiment analysis. In this work we have performed data pre-processing as the first step of sentiment analysis after collecting the reviews.

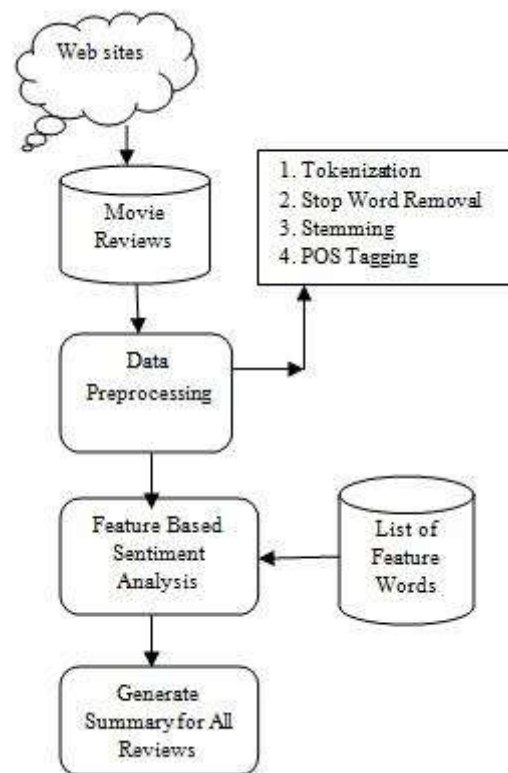


Fig.1. Process flow diagram

The overall description of the design specification as shown in above figure 1 is explained below.

#### A. Web Sites

A very large volume of user-generated reviews are available on internet in the form of various websites like imdb.com, twitter.com, amazone.com, flipcart.com, ndtv.com etc., these are used to collect the reviews.

#### B. Movie Reviews

To determine the polarity of the sentences based on features, large numbers of reviews are collected from the www.imdb.com site for movies.

#### C. Data Pre-processing

After collecting the reviews, we did pre-processing before sentiment analysis. Unnecessary noise can be removed from the reviews which are not the part of analysis in this step. We did following pre-processing steps for preparing our data.

1) *Tokenization*: After collecting the data, all sentences or clauses are tokenized in the reviews according to some defined delimiter like full stop, exclamatory mark, question marks etc. The term tokenization is the process of removing the comma (,), semi-column (;), and period (.) from the document lines.

2) *Stop Word Removal*: After tokenization step, the system removed the stop words from the sentence if there is any. The stop words are propositions, irrelevant words like is,

were, the etc which are not the part of sentiment analysis, are removed on tokenized data. These stop words create unnecessary noise on analysis

3) *Stemming*: Stemming is the process to find up the root word (like the word „performing“ will be stemmed as „perform,,).It is helpful for finding main key words for feature extraction. It refers to get the base form of the word and for this we have make use of the WordNet. First we have installed WordNet 2.0 and WordNet 2.1 and then we have given the reference of the WordNet dictionary in the source code file.

4) *POS (Part Of Speech) Tagging*: Next step include the tagging of data, which is very important step for information retrieval. For POS tagging we used the Stanford POS tagger, it has good performance and pre-trained POS tagging models. It parses a string of words (e.g., a sentence) and tags each term with its part of speech.

#### D. List of Feature Words

We have used the following list of feature words of movie reviews for our proposed work.

TABLE 1. LIST OF FEATURES

Feature Words	story, casting, music, character, songs, film, role, act, lyrics, direction, production, dress-up, make-up
---------------	--

#### E. Feature Based Sentiment Analysis

In this step we dealt with the negation handling, intensifier handling, conjunction handling and synonyms handling. We will cover these problems in next section. After that, the polarity of the review is determined for each feature. Polarity is determined on the basis of final score of review, if the score of review is greater than zero, then the polarity of the feature is commented as positive otherwise the polarity is negative and if score is zero then the feature commented as neutral polarity. The scores of Words (except the feature of movies) are computed by using SentiWordNet 3.0.

#### F. Generate Summary of Reviews

The summary generated after all scoring of the input reviews is represented in the following figure 2.

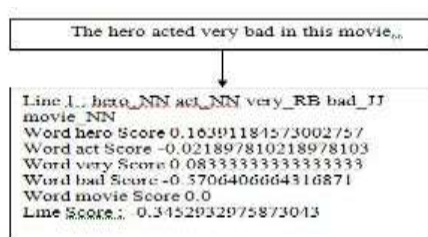


Fig. 2 Output summaries generated after sentiment analysis

SentiWordNet: - SentiWordNet is a lexical resource [16] which is publically available for research works.

SentiWordNet is an extension of WordNet [1] in which all synsets can be associated with a value concerning the negative, positive or objective connotation. Each synset is labeled with a value between 0.0 and 1.0. The sum of the three values is always 1.0, so each synset can have a nonzero value for each sentiment, after preprocessing the text is prepared in a normalized form. For each of the words, SentiWordNet retrieves the synsets that contain that word. If SentiWordNet does not find any suiting synset, the sentiment scores for this word becomes zero.

#### IV. METHODOLOGY

In this section we discussed about the methodologies used in the proposed work in detail which include Negation, Intensifier, Coordinating conjunction and synonyms handling. We have created a Graphical User Interface where we selected a file containing reviews, and then compute it. We considered two types of cases, as if we give single review as input then, only negation, intensifier and conjunction handling is done, but if we give multiple reviews as input then along with them, synonyms handling also done.

##### A. Negation Handling

As we know, the reviews can give any type of sense as negative, positive and neutral. But the problem occurs when sentence has negative words as no, never, neither, hardly etc. These terms when appear in the review, the system is not able to identify the review's actual sense. We handled two cases in it, which are following.

1) *Negative word used in positive sentence makes negative sense*: In this case, if the sentence is positive and before the opinion word (good, attractive etc) there is any negative word , then the sentence will be converted into negative sense.

**“Story of the movie is good”**

It means that the story of the movie, talking about, is good. But now we consider the review.

**“Story of the movie is not good.”**

Good always results in the positive score, so we have to analyze the negative word before it, which change the overall concept of the review.

2) *Negative word used in negative sentence makes positive sense*: In this case, if the sentence is negative and before the opinion word (bad, worst etc) there is any negative word, then the sentence will be converted into positive sense.

**“Story of the movie is bad.”**

It means that the story of the movie, talking about, is bad. But now we consider the sentence

**“Story of the movie is not bad.”**

Bad always results in the negative score, so we have to analyze the negative word before it, which change the overall concept of the review.

For handling these cases we have used the following algorithm as

Step 1: Start

Step 2: Check the Previous word of every word of the sentence.

Step 3: IF (Previous word is negative word) THEN

Score of word is multiplied by (-

1) Step 4: End

This algorithm will change the overall sentiment of the review.

### B. Intensifier Handling

The intensifiers are the word which pushes the overall score of the sentence and its meaning. We have stored all the intensifiers like “Too, Very, Sorely, Most, More” in a file and then followed the algorithm defined below.

Step1: Start

Step2: Check the Previous word of the

review. Step 3: IF (Previous word is intensifier word) THEN

Score of intensifier is added with the score of its opinion word Step 4: End

This algorithm increase the weight of review, in comparison to non intensifier sentence score.

### C. Synonyms handling

As we have discussed already that synonym are handled only in multiple reviews, single reviews didn’t handle synonyms. For it we use the features list, and divide the categories for their synonyms as following.

TABLE II. LIST OF SYNONYMS FOR SPECIFIC FEATURES

Feature	Synonyms
Movie	Film , Story
Acting	Role, Character
Songs	Music, Lyrics
Dress-up	Make-up
Direction	Production

From the above table II, we can understand that the particular features have their respective synonyms. We have created a combo box, in our GUI, where all features listed in table II, are given as options for selecting. If any feature is selected from them, then only those reviews are extracted from the whole reviews, which have that selected feature and it’s synonyms also. Then the overall scores of extracted reviews will be calculated for that particular feature only.

We take an example in the following figure, where multiple reviews are given as input which is following,

**“Role of actor is not bad. Songs are good to listen. Music is very bad. The direction of movie was taken very attractively.**

**Acting of heroes was sufficient to see the movie for one time.”**

If we selected „songs” feature from combo box as shown in figure 3

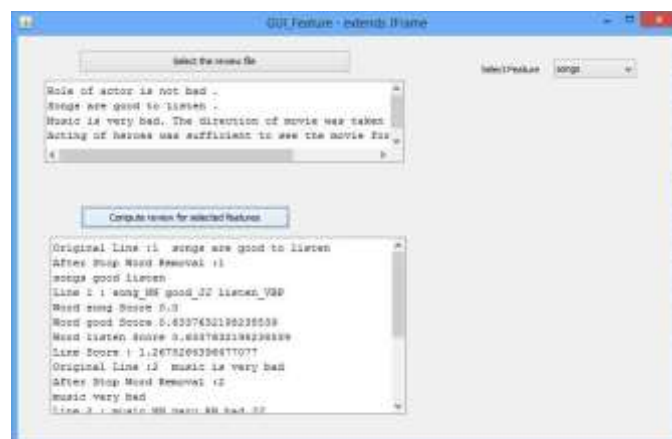


Fig. 3. GUI for sentiment analysis

And then, we compute the reviews for selected feature then it only extract or select the reviews of „songs” feature and its respective synonyms as listed in table 1 (Music, Lyrics) if there is any in the file. If any synonym is found here, then same score of selected feature is provided to it.

### D. Fragmentation of movie reviews into different clauses on the basis of coordinating conjunction (Conjunction Handling)

The conjunction is used where the multiple features are commented in a review. We have taken some conjunctions in consideration which are, “For, and, nor, But, Or, Yet, So”. We have stored these conjunctions in a list and when any review contains any of above words we split the sentence into two clauses on the basis of the conjunction found and computed the sentiment polarity of review.

TABLE III. CONJUNCTION HANDLING

Input Review	Output
The story of movie was interesting but dialogues were very cheap	Line 1: the story of movie was interesting
	Line 2: dialogues were very cheap

From the above table III, it is clear that the whole input review is splitted on the basis of conjunction found in it as “but”, and then scores are calculated for both the clauses separately.

### E. Sentiment Polarity and Score Calculation using SentiWordNet 3.0

After applying above methodologies for handling the problems, we calculated the scores from SentiWordNet 3.0 library, SentiWordNet contains a large text file where it



maintains the score of each and every word in the dictionary. It gives three terms associated with the score as the word itself, its POS tag and word score. The all words' scores are added and final score is computed.

If the total score is found as greater than zero then the polarity of feature is said to be positive and if its less than zero, the polarity of feature is said to be negative otherwise it will be neutral.

## V. EXPERIMENT AND RESULTS

In our experiment set-up we have taken total 100 reviews in which 55 reviews are positive and 45 are negative. These reviews are collected from www.imdb.com website which is a famous web site for posting and considering the reviews about all movies. For evaluation of proposed work, we have taken four classifier labels as True Positive (TP), False Positive (FP), True Negative (TN) and False Negative(FN). These labels are used for classifying the reviews' polarity given by our proposed system. The performance measurement parameters are taken as Accuracy, Precision, Recall and F-Measure which are standard parameters for sentiment analysis. The formulas used for calculating these measurements are defined below

**Accuracy (A):** Accuracy is the ratio of correctly classified classifiers to the total number of classifiers.

$$\text{Accuracy (A)} = (\text{TP} + \text{TN}) / (\text{TP} + \text{TN} + \text{FP} + \text{FN})$$

**Precision (P):** Precision is the ratio of true positives among all retrieved instances. Precision measures the exactness of the system.

$$\text{Precision (p)} = \text{TP} / (\text{TP} + \text{FP})$$

**Recall (R):** The ratio of true positives among all positive instances. Recall measures the completeness of the system.

$$\text{Recall} = \text{TP} / (\text{TP} + \text{FN})$$

**F-measure (F):** The harmonic mean of recall and precision. F-measure is more useful in sentiment analysis compare to accuracy. If false positive and false negative have same cost then it works best or if they have different cost it's better to look precision and recall.

$$F = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

We have compared our proposed system to three sentiment analysis tools which are freely available on internet named as WordNet [8], SentiWordNet [8] and Opinion Lexicon [8]. The scores and polarity is calculated for each review by all systems. The classifiers give the following comparison table.

TABLE IV. COMPARISON IN CLASSIFIER PARAMETERS

Classifier Parameters	Word Net[22]	SentiWord-Net [22]	Opinion Lexicon [22]	Proposed System [22]
TP	28	45	41	50
FP	27	10	14	05
TN	29	18	29	23
FN	16	27	16	22

**True positive** is the classifiers, which classified the reviews correctly positive. **False Positive** is the classifiers which classified the reviews incorrectly positive, means they didn't classify the reviews as true positive which should be true. **True Negative** is the classifiers which classified the reviews correctly negative. **False Negative** is the classifiers which didn't classify the reviews as correctly negative which should be true.

From the above table 3 we can see that our system gave highest true positive (TP), compared to other tools. But the Highest True Negative (TN) were given by Opinion Lexicon [8] and WordNet [8], compare to our system.

TABLE V. COMPARISONS IN PERFORMANCE PARAMETERS CALCULATED FOR DIFFERENT SYSTEMS

Performance Parameters	Word-Net[8]	SentiWord Net [8]	Opinion Lexicon [8]	Proposed System
Accuracy	57%	63%	70%	81%
Precision	50%	81 %	74%	90%
Recall	63%	62%	58%	69%
F-Measure	55%	70%	64%	78%

From the above table 4, we can see that accuracy, precision, recall and F-measure have given highest percentage in compared to all systems". The accuracy decreased of our system because of not handling the True Negative reviews.

The graph for proposed system evaluation is given below , where performance parameters are compared for all the tools.

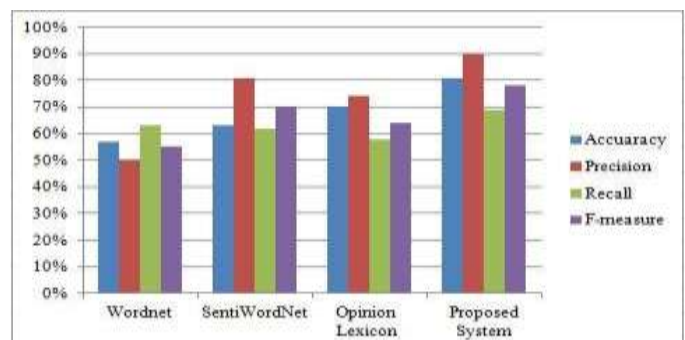


Fig.5 Performance Parameters Graph with Comparison

From the above figure 5, it's cleared that our proposed system gave highest accuracy, precision, recall and F-measure. The graph shows that SentiWordNet and our proposed system gave close precision values as 81% and 90%, which are higher to Precision values of WordNet and Opinion Lexicon, it means SentiWordNet is very helpful dictionary for classifying the positive reviews more correctly. Proposed system's performance decrease in terms of recall (69%), because it was not able to classify the negative reviews more correctly compare to other existing systems, here if we see in table IV, WordNet and Opinion Lexicon classified the True Negatives (TN) reviews more than SentiWordNet and proposed system. Proposed system classified the TN and FN almost equally, so F-measure also used for performance parameter, which also gave highest value as compare to others'. Accuracy of the system depends on the correctly classified reviews, which are more then to other systems, so that our system got highest accuracy as 81%.

## VI. CONCLUSION

In this paper, we have proposed a system, which deals with feature level sentiment analysis on movie reviews by handling various problems occur in reviews classification as having negation, intensifier, coordinating conjunctive and synonyms words in reviews. We have applied four methodologies for handling above problems, along with them we have also done pre-processing which prepare our data and make easy it for further analysis. For Stemming, we have used WordNet which give the sense of words. For calculating the scores and classifying the feature polarity, we have used SentiWordNet 3.0 tool.

Our proposed work have given highest accuracy as 81% , the performance is decreased because of not handling the True Negative reviews correctly, the reason behind it, is that our proposed system was not able to understand the negative words, like if the review is "*Gangaaajal is a highly underrated movie .*" Then our system gave final score of this review as positive. That is incorrect. So our system failed here. When we did stemming in data pre-processing ,then it converted the words in its base form like "performing" in "perform" , but it didn't handle some improper words which are mostly written by users in reviews as" supperbb" which should be changed into "super". These words can be used for giving better opinion about any feature. So in our next work, we will apply machine learning technique to solve above problem. We have dealt with only single features on sentence, except it we can also handle multiple features .We can do grouping here using Semantic Role Labeller. We can also handle the sentiment analysis on comparative sentence which is also a challenging task in it. We also found in our work that when we use double intensifiers in the sentence like (very very, so much etc), then the proposed system don't give high scores of the sentence compared to other systems. Handling double intensifiers in our work can increase our performance of the system

## REFERENCES

- [1]. Alena Neviarouskaya, Helmut Prendinger, and Mitsuru Ishizuka,"SentiFul: A Lexicon for Sentiment Analysis" IEEE Transaction on Affective Computing, Volume. 2, Issue No. 1, January-March 2011, 1949-3045/11/\$26.00 \_ 2011 IEEE
- [2]. Ann Devitt, Khurshid Ahmad " Sentiment Polarity Identification in Financial News: A Cohesion-based Approach" Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics, pages 984-991, Prague, Czech Republic, June 2007. © 2007 Association for

- Computational Linguistics
- [3]. Aurangzeb khan Baharum Baharudin, "Sentiment Classification Using Sentence-level Semantic Orientation of Opinion Terms from Blogs",2011 978-1- 4577-1884-7/11/\$26.00 ©2011 IEEE.
- [4]. Chinsha T C, Shibily Joseph," A Syntactic Approach for Aspect Based Opinion Mining " Proceedings of the IEEE 9th International Conference on Semantic Computing , 2015
- [5]. Chinsha T C, Shibily Joseph," A Syntactic Approach for Aspect Based Opinion Mining " Proceedings of the IEEE 9th International Conference on Semantic Computing , 2015
- [6]. Muhammad Shahbaz1, Dr. Aziz Guergachi2, Rana Tanzeel Rehman3," Sentiment Miner: A Prototype for Sentiment Analysis of Unstructured Data and Text" 978-1-4799-3010-9/14/\$31.00 ©2014 IEEE
- [7]. Farah Benamara, Carmine Cesarano, and Diego Reforgiato ,"Sentiment Analysis: Adjectives and Adverbs are better than Adjectives " ICWSM Boulder, CO USA, 2017
- [8]. [http://sentiment.christopherpotts.net/lexicon/textscores\\_results/](http://sentiment.christopherpotts.net/lexicon/textscores_results/)
- [9]. Kerstin Denecke "Using SentiWordNet for Multilingual Sentiment Analysis" 2008.
- [10]. Lizhen Liu, Xinhui Nie, Hanshi Wang," Toward a Fuzzy Domain Sentiment Ontology Tree for Sentiment Analysis" 2012 5th International Congress on Image and Signal Processing ,CISP 2012
- [11]. Monalisa Ghosh, Animesh Kar ," Unsupervised Linguistic Approach for Sentiment Classification from Online Reviews Using SentiWordNet 3.0" International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 9, September – 2013
- [12]. Ohana, B. & Tierney, " Sentiment classification of reviews using SentiWordNet" 9th. IT&T Conference, Dublin Institute of Technology, Dublin, Ireland, 22-23 October,2009
- [13]. Peifeng Li, Qiaoming Zhu, Wei Zhang, "A Dependency Tree based Approach for Sentence-level Sentiment Classification 12th ACIS International Conference on Software Engineering, Artificial Intelligence, Networking and Parallel/Distributed Computing ,2013
- [14]. Si Li, Hao Zhang, Weiran Xu, Guang Chen and Jun Guo," Exploiting Combined Multi-level Model for Document Sentiment Analysis" 2010 International Conference on Pattern Recognition, 2010 1051- 4651/10 \$26.00 © 2010 IEEE
- [15]. V.K. Singh, R. Piriyani, A. Uddin," Sentiment Analysis of Textual Reviews" 5th International Conference on Knowledge and Smart Technology (KST), 978-1-4673-4853-9/13/\$31.00 ©2013 IEEE, 2013
- [16]. Zhongchao Fei, Jian Liu, Gengfeng Wu , "Sentiment Classification Using Phrase Patterns" Proceedings of the Fourth International Conference on Computer and Information Technology, 2004