ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 8 , August 2017

Multiclass Sentiment Classification On Product Reviews

[1] Milan Gaonkar, [2] Prof. Amit Patil
[1][2] Computer Engineering Department, Goa College of Engineering

***Abstract*— Sentiment Analysis is an evolving field of research in text mining field. Nowadays, the opinions expressed through reviews are rising day by day on the web. It becomes practically unfeasible to analyze and extract opinions from such huge number of reviews manually. To overcome this problem an automated sentiment classification approach is needed. One aspect of study which is considered in this paper is to classify a given review tweet or paragraph whether it is of Positive[True-positive ,False-positive] or Negative[True-negative, False-negative] sentiment . In this paper a new approach is been proposed that uses lexicon database to assign each word in a text a value called score . The score is nothing but how a single word is affecting the whole sentence in which it is used. Every word in a sentence has its own strength and it tries to influence the overall semantic of the sentence. Higher the value of score of a word in the sentence, the more influential it is. The method proposed in this paper makes use of lexicon based approach as well as machine based learning. It uses AFINN lexicon database to assign score to words which is useful for sentiment classification and for testing and training the model the Support Vector Machine (SVM) and Naïve Bayes classifier (NB) machine learning algorithms are used.**

***Keywords*— AFINN, Machine learning,Product reviews,Sentiment analysis.**

I. INTRODUCTION

Sentiment analysis has been carried out on a variety of topics. For example, there are sentiment analysis studies for movie reviews product reviews , and news and blogs.[15]Research reveals that sentiment analysis is more difficult than traditional topic- based text classification, despite the fact that the number of classes in sentiment analysis are less than the number of classes in topic-based classification . In sentiment analysis, the classes to which a piece of text is assigned are usually negative or positive. They can also be other binary classes or multi-valued classes like classification into positive negative and neutral but still they are less than the number of classes in topic-based classification. Sentiment analysis is tougher compared to topic-based classification as the latter relies on keywords for classification. Whereas in the case of sentiment analysis keywords a variety of features have to be taken into account. The main reason that sentiment analysis is more difficult than topic-based text classification is that topic-based classification can be done with the use of keywords while this does not work well in sentiment analysis .Other reasons for difficulty are: sentiment can be expressed in subtle ways without any perceived use of negative words; it is difficult to determine whether a given text is objective or subjective; it is difficult to determine the opinion holder ;there are other factors such as dependency on domain and on order of words . Other challenges of sentiment analysis are to deal with sarcasm, irony ,negation and so on.

This paper is organized as follows. Section II describes the literature survey related to sentiment analysis. Section III presents the proposed work. Section IV shows the experimental results and section V concludes the work .

A. Problem statement

Given a review tweet/paragraph of a particular product the paper aims to classify it as Positive sentiment [True-positive, False-positive] or Negative sentiment[True-negative ,False- negative].

II. LITERATURE SURVEY

Sentiment analysis is the field of study that analyzes people's opinions, sentiments, evaluations, attitudes, and emotions towards entities such as products, services, organizations, individuals, issues, events, topics which are discussed on web. It represents a large problem space. There are also many names and slightly different tasks, e.g., sentiment analysis, opinion mining, opinion extraction, sentiment mining, subjectivity analysis, affect analysis, emotion analysis, review mining, etc. While in industry, the term sentiment analysis is more commonly used, but in academia both sentiment analysis and opinion mining are frequently employed. They basically represent the same field of study. Sentiment analysis and opinion mining mainly focuses on opinions which expresses positive or negative sentiments.

In [1] author used data mining techniques for the purpose of classification to perform sentiment analysis on the views people have shared in Twitter. The data was collected from

twitter that are in natural language and apply text mining

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)

Vol 4, Issue 8, August 2017

techniques –tokenization, POS tagging etc. to convert them into useful form and then use it for building sentiment classifier that is able to predict happy, sad and neutral sentiments for a particular tweet. Rapid Miner tool is being used, that helps in building the classifier as well as able to apply it to the testing dataset.

In [2] paper focused on aspect level opinion mining and proposed a new syntactic based approach for it, which uses syntactic dependency, collective score of opinion words, SentiWordNet and aspect table together for opinion mining. The experimental work was done on restaurant reviews. The dataset of restaurant reviews was collected from web and tagged manually.

In [3] a method which performs 3-class classification of tweet sentiment in Twitter has been proposed. An end to end system which can determine the sentiment of a tweet at two levels- phrase level and message level.

[4] Deals with fundamental problem of sentiment analysis, sentiment polarity categorization. Online product reviews from Amazon.com are selected as data used for study. A sentiment polarity categorization process has been proposed. Experiments for both sentence-level categorization and review-level categorization have been performed.

In [5] author have examined how well ANEW and other word lists performs for the detection of sentiment strength in microblog posts in comparison with a new word list specifically constructed for microblogs. Author have used manually labeled postings from Twitter scored for sentiment. Using a simple word matching and showed that the new word list may perform better than ANEW, though not as good as the more elaborate approach found in SentiStrength.

In [16] authors have proposed the model that uses lexicon database to assign each word in a text a value called as ‘Impact Factor’. Every word in a sentence has its own Impact Factor and it tries to influence the overall semantic of the sentence. Higher the value of Impact Factor of a word in the sentence, the more influential it is. The approach proposed in this paper makes use of lexicon based approach as well as machine based learning. It uses AFINN lexicon database to assign Impact Factor to words and Support Vector Machine(SVM), k-Nearest Neighbors (KNN) and Naïve Bayesian (NB) machine learning algorithms for testing the model.

III PROPOSED APPROACH

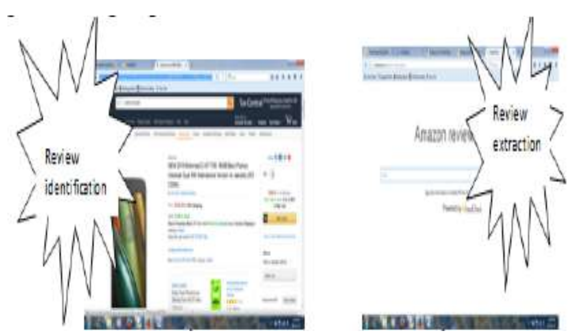
This section contains the detailed process flow of sentiment analysis. Fig. 1 describes a brief outline of the algorithm that was followed in sentiment analysis.
The reviews sources are mainly review sites. SA is not only applied on product reviews but can also be applied on stock markets and , news articles, or political debates . They are also used as data sources in the SA process.The tweets/paragraph that are in natural language are taken as input data .Next text mining techniques are applied on data–like Tokenization, Stopword elimination ,POS tagging to convert data into useful form for feature extraction that is able to classify Positive[True-positive,False-positive] or Negative[True-negative,False-negative] sentiments for a particular tweet or paragraph.

A. Data collection

Data used is a set of product reviews collected from amazon.com. on mobile phones of Samsung Brand. FeedCheck amazon review exporter is an API is used to extract reviews. The reviews extracted are in free format i.e. no restriction of writing reviews in PROS and CONS separately.

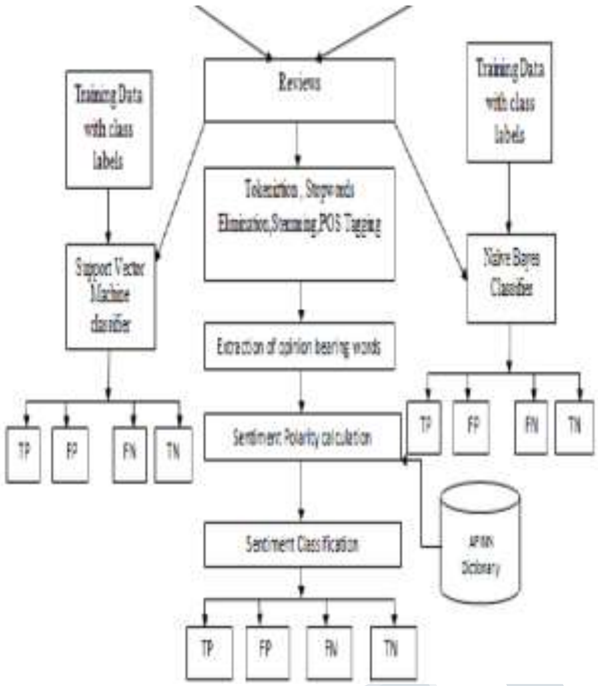
B. Data Preprocessing:

As the dataset is from Amazon.com, the data is in the form of text. The text data is highly prone to inconsistencies. This step is very important as it extract out unwanted words from tweets. To make the data more relevant for analysis, text preprocessing is performed.



ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 8, August 2017



TP –> True-Positive TN–> True-Negative
FP –>False-Positive FN–>False-Negative

Fig. 1: System architecture

1) Tokenization:
The process of breaking a stream of text up into phrases, words, symbols, or other meaningful elements called tokens. The goal of the tokenization is the exploration of the words in a sentence.

2) Stopword Elimination
The most common words that unlikely to help text mining such as prepositions, articles, and pro-nouns can be considered as stopwords. Since every text document deals with these words which are not necessary for application of text mining. All these words are eliminated. A new list of stop words was created that eliminated only those words that did not contribute to opinion mining.

3) POS Tagging
The Part-Of-Speech of a word is a linguistic category that is defined by its syntactic or morphological behaviour. Common POS categories in English grammar are: noun, verb, adjective, adverb, pronoun, preposition, conjunction, and interjection.POS tagging is the task of labeling each word in a sentence with its appropriate part of speech. POS tagging is an important phase of opinion mining, it is essential to determine the features and opinion words from the reviews.

C. Extraction of opinion bearing words in an around feature

First, a Set of opinion words (adjectives, as they are normally used to express opinions) is identified. If an adjective appears near a product feature in a sentence, then it is regarded as an opinion word.

D. Sentiment polarity calculation

Sentiment oriented words will have prior polarity which is obtained from AFINN dictionary. This dictionary is of about 2490 English language words assigns every word a score between -5 (Negative) and +5 (Positive).In this processing the extracted opinion bearing words will be compared with the words in AFINN dictionary. For classifying each review the corresponding weights of each opinion bearing words are counted. Sentiment polarities are divided into four categories like True-positive & False-positive And True-negative & False-negative.

E. Classifiers

1) Naive Bayes Classifier(NB)

Naive Bayes classifiers are studying the classification task from a Statistical point of view. The starting point is that the probability of a class C is given by the posterior probability P(C|D) given a training document D. Here D refers to all of the text in the entire training set. It is given by D =(d1,d2,….dn) , where di is the ith attribute (word) of document D.
Using Bayes' rule, this posterior probability can be rewritten as:

$$P(C = c_i|D) = \frac{P(D|C=c_i).P(C=c_i)}{P(D)} \quad (1)$$

Since the marginal probability P(D) is equal for all classes, it can be disregarded and the equation becomes:

$$P(C = c_i|D) = P(D|C = c_i).P(C = c_i) \quad (2)$$

The document D belongs to the class C which maximizes this probability, so:

$$C_{NB} = argmaxP(D|C).P(C) \quad (3)$$

$$C_{NB} = argmaxP\left(d_1,d_2,\ldots,d_n\middle|C\right).P(C) \quad (4)$$

Assuming conditional independence of the words di , this equation simplifies to:

IFERP
ISSN (Online) 2394-2320
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 8, August 2017

$$C_{NB} = argmaxP(d_1|C).P(d_2|C)\ldots.(d_n|C).P(C) \quad (5)$$

$$C_{NB} = argmaxP(C).\textstyle\prod_i P(d_i|C) \quad (6)$$

Here is the conditional probability that word i belongs to class C. For the purpose of text classification, this probability can simply be calculated by calculating the frequency of word i in class C relative to the total number of words in class C.

$$P(d_i|C) = \frac{count(d_i,C)}{\sum_i count(d_i,C)} \quad (7)$$

We have seen that we need to multiply the class probability with all of the prior-probabilities of the individual words belonging to that class. The question then is, how do we know what the prior-probabilities of the words are? Here we need to remember that this is a supervised machine learning algorithm: we can estimate the prior-probabilities with a training set with documents that are already labeled with their classes. With this training set we can train the model and obtain values for the prior probabilities. This trained model can then be used for classifying unlabeled documents.

2) Support Vector Machine

A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyper plane. In other words, given labeled training data, the algorithm outputs an optimal hyper plane which categorizes new examples.

To deal with datasets with more than two classes usually the dataset is reduced to a binary class dataset with which the SVM can work. There are two approaches for decomposing a multiclass classification problem to a binary classification

problem: the one-vs-all and one-vs-one approach. In the one-vs-all approach one SVM Classifier is build per class. This Classifier takes that one class as the positive class and the rest of the classes as the negative class. A datapoint is then only classified within a specific class if it is accepted by that Class' Classifier and rejected by all other classifiers. Although this can lead to accurate results(if the dataset is clustered), a lot of datapoints can also be left unclassified In the one-vs-one approach, you build one SVM Classifier per chosen pair of classes. Since there are 0.5N(N-1) possible pair combinations for a set of N classes, this means you have to construct more Classifiers. Datapoints are then categorized in the class for which they have received the most points.

Algorithm1:

Input: Document text X
Output: Sentiment being conveyed by each review in X
Algorithm Sentidetect (R , A)

```
{
For each review Ri extract opinion bearing feature
word and compare it with AFINN dictionary .
Fetch the corresponding ratings beside each opinion
bearing word from AFINN dictionary and store it w.r.t.
each review.
pol[Ri]<-0
OBWSc<-pol[Ri]+pol[w]
If ( OBWSc > = 3)Then
        Print " R conveys True-Positive sentiment"
Else if OBWSc = 2 or OBWSc = 1) Then
        Print"R conveys False-Positive sentiment"
Else if ( OBWSc= -2 or OBWSc = -1) Then
        Print"R conveys False-Negative sentiment"
Else if (OBWSc >= -3 ) Then
        Print"R conveys True-Negative sentiment"
}
* OBWSc –Opinion Bearing word Score
```

IV EXPERIMENTAL RESULTS

Experiments were performed on a dataset obtained by extracting product reviews from Amazon.com. We focused on the mobile phone domain. Considering reviews of one product at a time sentiment of the reviews were classified into four categories namely True-positive sentiment, False-positive sentiment, True-negative sentiment, False-negative sentiment.

Table 1: Shows the results of the algorithms when applied on the given data set.

Class Labels	Methods Applied		
	Lexicon Dictionary (1000 reviews)	SVM Classifier (1000 reviews)	NB Classifier (1000 reviews)
True-positive	42%	98%	86.6%
False-positive	43%	0.5%	3.9%
True-negative	2%	0.7%	4.7%
False-negative	11%	0.4%	4.8%

ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 4, Issue 8, August 2017

Fig 2,Fig 3 and Fig 4 Shows graphical representation of Sentiment Classification using each method

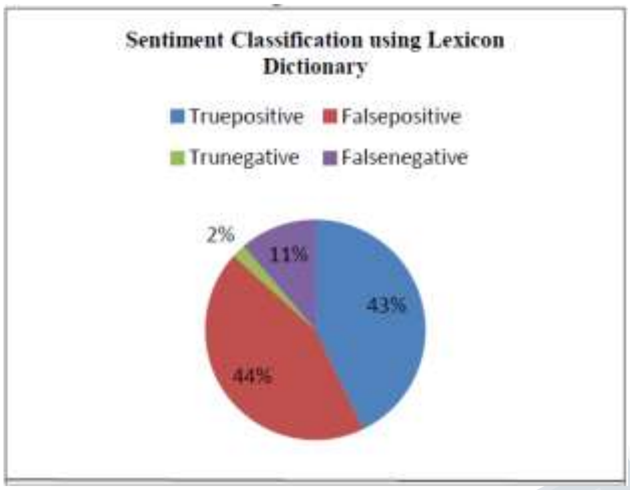


Fig 2: Pie-chart representation of Sentiment Classification using Lexicon dictionary

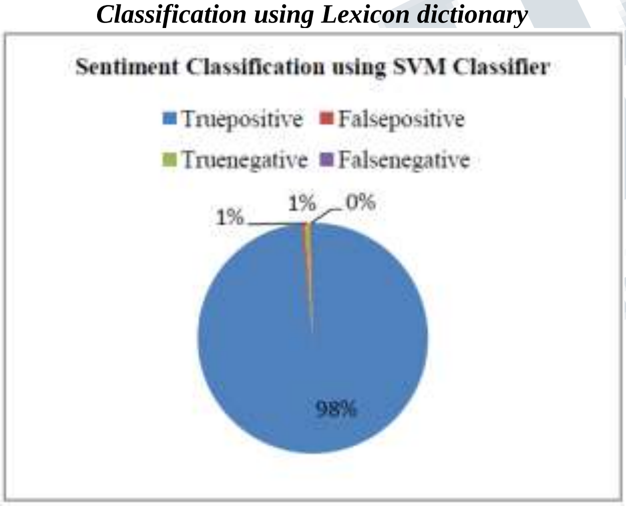


Fig 3: Pie-chart representation of Sentiment Classification using SVM Classifier

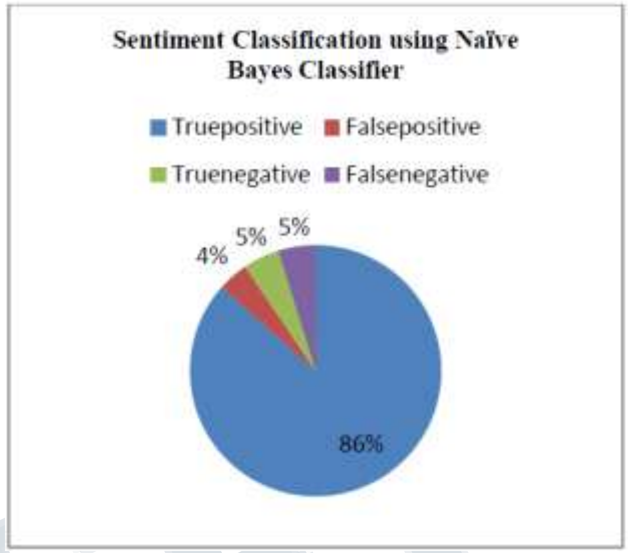


Fig 4: Pie-chart representation of Sentiment Classification using Naïve Bayes Classifier

CONCLUSION

A proposal for developing an sentiment classification system for product reviews was suggested. Various technical papers were briefly described that suggested relevant methods to perform sentiment classification in English. The detailed design along with the algorithms to be performed at each step of the implementation was explicitly stated. The implementation of the sentiment classification system in Python was done. Lastly, the classification results were described using pie-chart which concludes that using SVM classifier we get better result for True-positive class as compared to Lexicon dictionary and Naïve Bayes Classifier. Whereas for other classes like False-positive and False-negative Lexicon dictionary found to be better than SVM and Naïve Bayes classifier. And Naïve bayes classifier gave better result for True-negative class than both the other methods.

REFERENCES

[1] P. Tripathi, S. Kr. Vishwakarma, A. Lala ,"Sentiment Analysis of English Tweets Using RapidMiner," 2015IEEE International Conference on Computational Intelligence and Communication Networks.

[2] T.C. Chinsha , S. Joseph " A Syntactic Approach

for Aspect Based Opinion Mining," Proceedings of the 2015 IEEE 9th International Conference on Semantic Computing.

[3] M. Gupta, A. Dalmia, A. Jaiswal and C. T. Reddy ,"Sentiment Analysis in Twitter".

[4] X. Fang and J. Zhan," Sentiment analysis using product review data" Springer. 2015.

[5] F. A. Nielsen," A new ANEW: Evaluation of a word list for sentiment analysis in microblogs," DTU Informatics, Technical University of Denmark, Lyngby, Denmark.

[6] S. H. Ghorashi, R. Ibrahim, S. Noekhah and N. S. Dastjerdi ,"A Frequent Pattern Mining Algorithm for Feature Extraction of Customer Reviews".

[7] S. Mukherjee, P. Bhattacharyya,"Feature Specific Sentiment Analysis for Product Reviews".

[8] C. Fiarni1, H. Maharani1 ,R. Pratama1,"Sentiment Analysis System for Indonesia Online Retail Shop Review Using Hierarchy Naive Bayes Technique," 2016 Fourth International Conference on Information and Communication Technologies (ICoICT)

[9] J. Jotheeswaran, Loganathan R. and Madhu S. B.,"Feature Reduction using Principal Component Analysis for Opinion Mining,"International Journal of Computer Science and Telecommunications Volume 3, Issue 5, May 2012.

[10] D. V. N. Devi, C. K. Kumar ,S. Prasad ,"A Feature Based Approach for Sentiment Analysis by Using Support Vector Machine," 2016 IEEE 6th International Conference on Advanced Computing.

[11] V. Singh, R. Piryani, A. Uddin, and P. Waila, "Sentiment analysis of movie reviews: A new feature-based heuristic for aspect-level sentiment classification,"Automation, Computing, Communication, Control and Compressed Sensing (iMac4s), 2013 International Multi-Conference on.IEEE, 2013, pp. 712-717.

[12] Govin Gaikwad, Deepali J. Joshi. "Multiclass mood classification on Twitter using lexicon dictionary and machine learning algorithms", 2016 International Conference on Inventive Computation Technologies (ICICT), 2016

[13] S. Gojali, M. L. Khodra," Aspect Based Sentiment Analysis for Review Rating Prediction," 2016 IEEE.

[14] Govin Gaikwad, Deepali J. Joshi. "Multiclass mood classification on Twitter using lexicon dictionary and machine learning algorithms", 2016 International Conference on Inventive Computation Technologies (ICICT), 2016

[15] B. Liu. Sentiment Analysis and Opinion Mining, Morgan & Claypool Publishers, May 2012.