# SENTIMENT ANALYSIS ON TAMIL REVIEWS USING ENSEMBLE CLASSIFICATION TECHNIQUES WITH MULTI-CLASS ANALYSIS

**A. SHARMISTA**
Research Scholar,
Department of Computer Applications,
*Madurai Kamaraj University*
*Tamil Nadu, India.*
ssharmistasaravananpkn@gmail.com

**M. RAMASWAMI**
Associate Professor,
Department of Computer Applications,
*Madurai Kamaraj University*
*Tamil Nadu, India.*
mrswami123@gmail.com

***Abstract***— **Opinion Mining or Sentiment Analysis is a Natural Language Processing and Information Extraction task that identifies the user's views or opinions explained in the form of positive, very positive, negative, very negative or neutral comments. Nowadays social network plays an important role for the customers to take decision about the product they purchase. In this paper we have analysed Tamil mobile product reviews for the task of sentiment analysis. Reviews are collected from social media platform like Twitter, Facebook and blogs. Sentiments of each posted reviews will be analysed and conclude by telling which product is the perfect match to buy. Then we propose ensemble learning algorithms to construct prediction model with multi class instances in order to achieve more accuracy. Sentiment lexicon is developed that contains lakhs and lakhs of Tamil words in it related with its synonyms and antonyms. To show the applicability of the technique on other languages, we experiment our technique on Tamil language. The whole work is implemented using 'R' language.**

**Keywords** — *Sentiment Analysis, Opinion Mining, Feature Selection, Opinion Lexicon, Boosting, Bagging, Stacking, Classification.*

## I. INTRODUCTION

Text mining can solve the above-mentioned problems. Text mining is gaining more importance due to the increasing number of readily available electronic information (digital libraries, electronic mail, and blogs). Text mining is a knowledge discovery technique that provides computational intelligence [7]. This technique comprises of multidisciplinary fields, such as information retrieval, text analysis, natural language processing, and information classification based on logical and non-trivial patterns from large data sets. Text mining techniques become more complex as compared to data mining due to unstructured and fuzzy nature of natural language text [5]. Social networking platform, such as Twitter and Facebook are rich in texts that enable user to create various text contents in the form of comments, wall posts, social media, and blogs. These user comments are used for finding the sentiments and also add financial, commercial and social values. Due to global use of social networks in recent years, an enormous amount of data is available via the Web. People interact with each other, share their ideas, opinions, interests and personal information. Using this data, we can extract information from social network platform that gives different views of a person globally. Moreover, text mining techniques in conjunction with social networks can be used for finding general opinion about any specific subject, for example, purchase of any mobile product online gives different opinion from different people and it is useful to examine those reviews to conclude the result of a particular product [19]. The decision boundary that separates data from different classes may be too complex, or lie outside the space of functions that can be implemented by the chosen classifier model. A linear classifier, one that is capable of learning linear boundaries, cannot learn this complex non-linear boundary. However, appropriate combination of an ensemble of such linear classifiers can learn this (or any other, for that matter) non-linear boundary [3]. In a sense, the classification system follows a divide-and-conquer approach by dividing the data space into smaller and easier to learn partitions [4], where each classifier learns only one of the simpler partitions. The underlying complex decision boundary can then be approximated by an appropriate combination of different classifiers. If we collect information from different sources, which contains different features it is difficult to learn all the data with a single classifier hence data fusion is used in ensemble concepts [18]. Especially for combining data from different sources data fusion application can be used and this concept is applied in ensemble approaches successfully. Several machine learning techniques do this by learning an ensemble of models and in this paper, we combine popular components called bagging, boosting, and stacking with multi- class instances like positive, very positive, negative, very negative or neutral. They can all, more often than not, increase predictive performance over a single model [2]. Then we need to combine the output of individual classifiers that make up the ensemble in such a way that the correct decisions are amplified and incorrect ones are cancelled out.

## II. RELATED WORK

The importance of opinion and sentiments mining is described by [17]. Opinion mining is distributed into four sub- categories: sentence level, document level, feature level and compound level respectively (and has discussed the various tools and mining methods like precision, recall and F-measure on movie reviews or product reviews). Aspect level classification can provide greater detail for analysis.

The distribution categories of classification are subjective expressions for better understanding of sentiments.

Most of the sentiment analysis or opinion mining was performed at the document level, but only some researchers have examined opinion mining at the more fine-grained sentence in recent years [13]. This analysis is often performed at the feature level for product reviews to provide in-depth analytics for the target product. Public influence across different issues / topics to determine the influence author used in degree re-tweets and mentions approaches on a large dataset of twitter was done by [16]. Spearman's rank correlation coefficient has been used for evaluating the influence of people against three different topics. Author concludes the result in terms of influence against each topic. The sentiment mining analysis on Facebook textual dataset for news channel by using Recall and Precision approach was implemented by [9]. In this approach, it is very difficult to identify what is relevant to one person might not be relevant to another person and vice versa. Precision and Recall method is used. Naïve Bayes can also be used for better results. Analysing sentiments in Chinese product reviews by introducing a novel approach which based on the fuzzy domain sentiment ontology was done by [10]. Sentiment polarity along with its weight is identified for each word extracted from the corpus. Final decision is taken from precision and recall values. An attribute-based sentiment analysis system on customer reviews to make a decision on products purchased through E- commerce was proposed by [11]. Polarity analysis based on sentence or document level was analysed and classified. Final visualization shows the consumers to arrive at the understanding of all reviews. New Avenues in Opinion Mining and Sentiment Analysis focus on collecting reviews from valuable resources in order to buy the best products. Bayesian inference and SVM on text classification is implemented by [8] to find out the best result.

Opinion mining on YouTube comments that rely on the tree kernel technology to automatically extract and learn features with better generalization power than bag-of-words was proposed by [1]. It highlights the benefits of structural models in a cross-domain setting. Ensemble based opinion mining system is analysed that classify the documents as positive, negative and neutral was proposed by [19]. Experimental results using reviews of movies show the effectiveness of the system. Extracting a lot of unstructured text data proposed by [12]. Customer reviews are not only helpful for potential customers, but also are helpful for the manufacturers of the products to raise their products and services. Here they propose a supervised lazy learning model utilizing syntactic rules for the product features and opinion words extraction in subjective review sentences. In lazy learning algorithm, i.e. K-NN with k=3 is used for the review sentences classification into two classes (subjective, objective). This method can improve the performance of existing work in terms of average precision, recall and f-score for the extraction of opinion sentences and product features. Focus on subjective texts was proposed by [23]. Main focus is on framework and lexicon construction, feature extraction, and polarity determination. In-depth analysis on research and application in business and Blog sphere has been concentrated in this sentiment survey.

## III. ENSEMBLE BASED CLASSIFICATION

### *A. Boosting*

Boosting is a general method for improving the performance of any learning algorithm [22]. Boosting works by repeatedly running a given learning algorithm on various distributions over the training data, and then combining the classifiers produced by the weak learner into a single composite classifier. Boosting, the algorithm is now considered as one of the most important developments in the recent history of machine learning. Boosting also creates an ensemble of classifiers by resampling the data. This method extends an efficient frequent pattern mining method, FP-growth, and mines large database efficiently [14]. At last Boosting is applied on multiple classifiers and collect relevant reviews. Cart and gbm algorithms are applied in case of boosting. It shows that accuracy achieved is less and hence we move onto bagging.

### *B. Bagging*

Bagging is simplest to implement resulting in good performance. In this concept training data are randomly drawn by replacing entire training data. Each training data subset is used to train a different classifier of the same type. Individual classifiers are then combined by taking majority positive reviews of their decisions. For any given instance, the class chosen by most classifiers is the ensemble decision. Bagging is particularly appealing when available data is of limited size [7]. Decision tree algorithms like treebag and random forest is applied in this paper. It is proved that it gives best prediction by achieving more accuracy. We move on to stacking in order to achieve more accuracy than bagging.

### *C. Stacking*

Stacking is used to increase the predictive force of the classifier. Stacking is the combining process of multiple classifiers [6] generated by different learning algorithms L1…Ln on a single dataset. The stacking ensemble approach (x, y) is left out of the training set of hi, after training is completed for hi, the output y can still be used to assess the model's error. In fact, since (x, y) was not in the training set of hi, hi(x) may differ from the desired output y. A new classifier then can be trained to estimate this discrepancy, given by y – hi(x). In essence, a second classifier is trained to learn the error the first classifier has made. Adding the estimated errors to the outputs of the first classifier can provide an improved final classification decision. Stacking mainly produce strong models less biased than their components even if the variance can also be used. Stacking is combined with (c50, cart, random forest) and also stacking with (lda, knn, svm). Result of this combination of algorithms with stacking achieves more accuracy.

### *D. Feature Selection Methods*

Embedded methods perform feature selection during the modelling algorithm's execution [21]. These methods are thus embedded in the algorithm either as its normal or extended functionality. Some embedded methods perform feature weighting based on regularization models with objective functions that minimize fitting errors and, in the meantime, force the feature coefficients to be small or to be exact zero. Hybrid methods were proposed to combine the best properties of filters and wrappers. First, a filter method is used in order to reduce the feature space dimension space, possibly obtaining several candidate subsets. Then, a wrapper is employed to find the best candidate subset. Hybrid methods usually achieve high accuracy that is characteristic to wrappers and high efficiency characteristic to filters. It is found that stacking method is particularly better for combining multiple different types of models.

## IV. PROPOSED METHOD

Sentiment analysis is the process of determining whether social media publications are 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 [20]. The overall architecture of the system is shown in fig.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 boosting, bagging and stacking algorithm, classification will be done and opinions are summarized such as நேர்மறை (Positive), யாரும் (None) and எதிர்மறை (Negative).

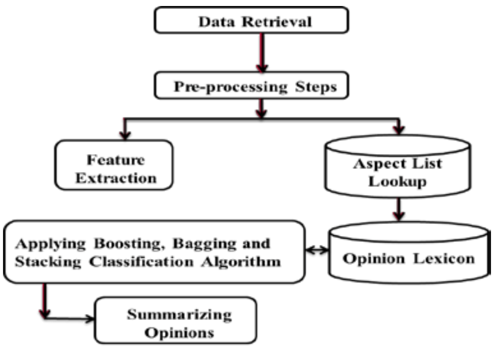


Fig.1: Flow of Summarizing Opinions from Collection of Reviews

Final result will be easier to choose the product. For example, நல்ல, சிறந்த are positive terms while கெட்ட, தவறு and மோசமான are negative terms. செங்குத்து, மஞ்சள் are objective terms. சிறந்த and மோசமான are more intense than நல்ல and மோசமான.

### A. Document Pre-Processing

**By applying part of NLP steps for obtaining the token of words are 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. 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). 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.

### B. 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 [15]. This section proposes an unsupervised learning method. Sequential pattern mining is carried out in all 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. EXPERIMENTS AND DISCUSSIONS

Ensemble classification algorithms are applied under five classes namely மிகவும்சாதகமான (Very Positive), நேர்மறை (Positive), எதிர்மறை (Negative), மிகவும்எதிர்மறை (Very Negative) and நடுநிலை (Neutral). The ensemble techniques like boosting (cart and gbm), bagging (treebag and random forest), stacking (c50, cart, random forest) and stacking (lda, knn, svm) is executed to predict the best product by identifying the total number of positive opinions and also, we show that which algorithm achieves more accuracy. The total number of instances used for analysis of product data for five classes is 5000. Training data is considered as 70% from the total dataset size (i.e., dataset considered is 3500) and testing data is considered as 30% from the total dataset size (i.e. dataset considered is 1500). The below table shown the confusion matrix for both training and testing data.

Table I. Shows result of training dataset on boosting (cart and gbm) classification matched with opinion lexicon for product reviews dataset run on R platform. It is predicted that there are 245 மிகவும்சாதகமான (Very Positive) reviews, 397 நேர்மறை (Positive) classified reviews, 54 மிகவும்எதிர்மறை (Very Negative) reviews and 198 எதிர்மறை (Negative) reviews in case of boosting analysis when matched with opinion lexicon.

TABLE I :
CONFUSION MATRIX FOR TRAINING DATASET ON BOOSTING (CART AND GBM) CLASSIFICATION (FIVE CLASS) MATCHED WITH OPINION LEXICON

| | Predicted Class | | | | |
|---|---|---|---|---|---|
| Actual Class | மிகவும்சாதகமான **(Very Positive)** | நேர்மறை (Positive) | நடுநிலை (Neutral) | மிகவும்எதிர் மறை (Very Negative) | எதிர்மறை (Negative) |
| | **245** | 79 | 10 | 2 | 15 |
| | 47 | **1090** | 71 | 20 | 102 |
| | 12 | 152 | **293** | 16 | 168 |
| | 6 | 41 | 15 | **237** | 20 |
| | 15 | 205 | 132 | 42 | **465** |

TABLE II
CONFUSION MATRIX FOR TESTING DATASET ON BOOSTING (CART AND GBM) CLASSIFICATION (FIVE CLASS) MATCHED WITH OPINION LEXICON

| | Predicted Class | | | | |
|---|---|---|---|---|---|
| Actual Class | மிகவும்சாதகமான **(Very Positive)** | நேர்மறை (Positive) | **நடுநிலை (Neutral)** | மிகவும்எதி ர்மறை (Very Negative) | எதிர்மறை (Negative) |
| | **51** | 15 | **2** | 0 | 2 |
| | 29 | **397** | **37** | 17 | 68 |
| | 9 | 103 | **150** | 10 | 108 |
| | 1 | 12 | 4 | **54** | 5 |
| | 5 | 105 | 87 | 31 | **198** |

Table II. Shows result of testing dataset on boosting (cart and gbm) classification matched with opinion lexicon for product reviews dataset run on R platform. It is predicted that there are 51 மிகவும்சாதகமான (Very Positive) reviews, 397 நேர்மறை (Positive) classified reviews, 150 நடுநிலை (Neutral) reviews, 54 மிகவும்எதிர்மறை (Very Negative) reviews and 198 எதிர்மறை (Negative) reviews in case of boosting analysis when matched with opinion lexicon.

Table III. Shows result of training dataset on bagging (treebag and random forest) classification matched with opinion lexicon for product reviews dataset run on R platform. It is predicted that there are 343 மிகவும்சாதகமான (Very Positive) reviews, 1258 நேர்மறை (Positive) classified reviews, 492 நடுநிலை (Neutral) reviews, 294 மிகவும்எதிர்மறை (Very Negative) reviews and 673 எதிர்மறை (Negative) reviews in case of bagging analysis
when matched with opinion lexicon.

TABL III
CONFUSION MATRIX FOR TRAINING DATASET ON BAGGING (TREEBAG AND RANDOM FOREST) CLASSIFICATION (FIVE CLASS) MATCHED WITH OPINION

| | Predicted Class | | | | |
|---|---|---|---|---|---|
| Actual Class | மிகவும்சாதகமான **(Very Positive)** | நேர்மறை (Positive) | நடுநிலை (Neutral) | மிகவும்எதிர் மறை (Very Negative) | எதிர்மறை (Negative) |
| | **343** | 4 | 2 | 0 | 2 |
| | 24 | **1258** | 19 | 4 | 25 |
| | 4 | 73 | **492** | 12 | 60 |
| | 0 | 15 | 4 | **294** | 6 |
| | 0 | 79 | 88 | 19 | **673** |

TABLE IV
CONFUSION MATRIX FOR TESTING DATASET ON BAGGING (TREEBAG AND RANDOM FOREST) CLASSIFICATION (FIVE CLASS) MATCHED WITH OPINION LEXICON

| | Predicted Class | | | | |
|---|---|---|---|---|---|
| Actual Class | மிகவும்சாதகமான **(Very Positive)** | நேர்மறை (Positive) | நடுநிலை (Neutral) | மிகவும்எதிர் மறை (Very Negative) | எதிர்மறை (Negative) |
| | **70** | 0 | 0 | 0 | 0 |
| | 12 | **495** | 14 | 4 | 23 |
| | 3 | 51 | **276** | 7 | 43 |
| | 0 | 3 | 0 | **71** | 2 |
| | 0 | 51 | 53 | 12 | **310** |

Table IV. Shows result of testing dataset on bagging (treebag and random forest) classification matched with opinion lexicon for product reviews dataset run on R platform. It is predicted that there are 70 மிகவும்சாதகமான (Very Positive) reviews, 495 நேர்மறை (Positive) classified reviews, 276 நடுநிலை (Neutral) reviews, 71 மிகவும்எதிர்மறை (Very Negative) reviews and 310 எதிர்மறை (Negative) reviews in case of bagging analysis when matched with opinion lexicon.

Table V. Shows result of training dataset on stacking (c50, cart and random forest) classification matched with opinion lexicon for product reviews dataset run on R platform. It is predicted that there are 339 மிகவும்சாதகமான (Very Positive) reviews, 1260 நேர்மறை (Positive) classified reviews, 496 நடுநிலை (Neutral) reviews, 300 மிகவும்எதிர்மறை (Very Negative) reviews and 665 எதிர்மறை (Negative) reviews in case of stacking analysis when matched with opinion lexicon.

TABLE V
CONFUSION MATRIX FOR TRAINING DATASET ON STACKING (C50, CART AND RANDOM FOREST) CLASSIFICATION (FIVE CLASS) MATCHED WITH OPINION LEXICON

| | Predicted Class | | | | |
|---|---|---|---|---|---|
| Actual Class | மிகவும்சாதகமான **(Very Positive)** | நேர்மறை (Positive) | நடுநிலை (Neutral) | மிகவும்எதிர்மறை (Very Negative) | எதிர்மறை (Negative) |
| | **339** | 8 | 2 | 0 | 2 |
| | 20 | **1260** | 19 | 4 | 27 |
| | 4 | 73 | **496** | 14 | 54 |
| | 0 | 15 | 4 | **300** | 0 |
| | 0 | 77 | 92 | 25 | **665** |

Table VI. Shows result of testing dataset on stacking (c50, cart and random forest) classification matched with opinion lexicon for product reviews dataset run on R platform. It is predicted that there are 69 மிகவும்சாதகமான (Very Positive) reviews, 496 நேர்மறை (Positive) classified reviews, 279 நடுநிலை (Neutral) reviews, 73 மிகவும்எதிர்மறை (Very Negative) reviews and 301 எதிர்மறை (Negative) reviews in case of stacking analysis when matched with opinion lexicon.

Table VII. Shows result of training dataset on stacking (lda, knn and svm) classification matched with opinion lexicon for product reviews dataset run on R platform. It is predicted that there are 343 மிகவும்சாதகமான (Very Positive) reviews, 1258 நேர்மறை (Positive) classified reviews, 496 நடுநிலை (Neutral) reviews, 300 மிகவும்எதிர்மறை (Very Negative) reviews and 663 எதிர்மறை (Negative) reviews in case of stacking analysis when matched with opinion lexicon.

TABLE VI

CONFUSION MATRIX FOR TESTING DATASET ON STACKING (C50, CART AND RANDOM FOREST) CLASSIFICATION (FIVE CLASS) MATCHED WITH OPINION LEXICON

| | Predicted Class | | | | |
|---|---|---|---|---|---|
| Actual Class | மிகவும்சாதகமா ன **(Very Positive)** | நேர்மறை (Positive) | நடுநிலை (Neutral) | மிகவும் எதிர்ம றை (Very | எதிர்மறை (Negative) |
| | **69** | 1 | 0 | 0 | 0 |
| | 10 | **496** | 14 | 4 | 24 |
| | 3 | 51 | **279** | 9 | 38 |
| | 0 | 3 | 0 | **73** | 0 |
| | 0 | 50 | 58 | 17 | **301** |

TABL
E VII
CONFUSION MATRIX FOR TRAINING DATASET ON STACKING (LDA, KNN AND SVM) CLASSIFICATION (FIVE CLASS) MATCHED WITH OPINION LEXICON

| | Predicted Class | | | | |
|---|---|---|---|---|---|
| Actual Class | மிகவும்சாதகமா ன **(Very Positive)** | நேர்மறை (Positive) | நடுநிலை (Neutral) | மிகவும் எதிர்ம றை (Very | எதிர்மறை (Negative) |
| | **343** | 4 | 2 | 0 | 2 |
| | 24 | **1258** | 19 | 4 | 25 |
| | 4 | 73 | **496** | 14 | 54 |
| | 0 | 15 | 4 | **300** | 0 |
| | 0 | 79 | 92 | 25 | **663** |

Table VIII. Shows result of testing dataset on stacking (lda, knn and svm) classification matched with opinion lexicon for product reviews dataset run on R platform. It is predicted that there are 70 மிகவும்சாதகமான (Very Positive) reviews with 0 error rate, 495 நேர்மறை (Positive) classified reviews with 53 error rate, 278 நடுநிலை (Neutral) reviews with 102 error rate, 73 மிகவும்எதிர்மறை (Very Negative) reviews and 301 எதிர்மறை (Negative) reviews in case of stacking analysis when matched with opinion lexicon.

TABLE VIII
CONFUSION MATRIX FOR TESTING DATASET ON STACKING (LDA, KNN AND SVM) CLASSIFICATION (FIVE CLASS) MATCHED WITH OPINION LEXICON

| | Predicted Class | | | | |
|---|---|---|---|---|---|
| Actual Class | மிகவும்சாதகமா ன **(Very Positive)** | நேர்மறை (Positive) | நடுநிலை (Neutral) | மிகவும் எதிர்ம றை (Very | எதிர்மறை (Negative) |
| | **70** | 0 | 0 | 0 | 0 |
| | 12 | **495** | 14 | 4 | 23 |
| | 3 | 51 | **278** | 9 | 39 |
| | 0 | 3 | 0 | **73** | 0 |
| | 0 | 51 | 57 | 17 | **301** |

**Performance Measure for Five-Class**

The performance of each ensemble algorithms will be measured on the basis of accuracy, specificity, sensitivity, precision, recall and AUC (Area under the Curve). From these measures we can conclude that which algorithm performs efficiently by checking its overall accuracy.

Table IX. Shown training performance measure on ensemble algorithms. It shows that bagging (treebag and random forest), stacking (c50, cart, random forest) and stacking (lda, knn, svm) achieves more accuracy with 0.8743. In case of AUC measure it shows that stacking (lda, knn, svm) algorithm achieves more accuracy with 0.9243 from training dataset.

TABLE IX
OVERALL TRAINING DATASET PERFORMANCE MEASURE OF ENSEMBLE A LGORITHMS (BOOSTING, BAGGING AND STACKING - FIVE CLASS) MATCHED WITH OPINION LEXICON

| Performance Measure | Ensemble Classification Method | | | |
|---|---|---|---|---|
| | **Boosting (Cart and gbm)** | **Bagging (Treebag and Random** | **Stacking (C50, Cart and Random Forest)** | **Stacking (LDA, KNN and SV** |
| **Accuracy** | 0.6657 | **0.8743** | **0.8743** | **0.8743** |
| **Specificity** | 0.9111 | 0.9671 | 0.9671 | 0.9671 |
| **Sensitivity** | 0.6727 | 0.8781 | 0.8772 | 0.8781 |
| **Precision** | 0.6518 | 0.8791 | 0.8803 | 0.8791 |
| **Recall** | 0.6727 | 0.8781 | 0.8772 | 0.8781 |
| **AUC** | 0.7845 | 0.9236 | 0.9239 | 0.9243 |

Table X. Shows testing performance measure on ensemble algorithms. It shows that bagging (treebag and random forest) achieves more accuracy with 0.8147. In case of AUC measure it shows that bagging (treebag and random forest) algorithm achieves more accuracy with 0.9078 from testing result. Hence bagging (treebag and random forest) is chosen to be the best predicted algorithm for buying mobile products in case of five class analysis.

TABLE X
OVERALL TESTING DATASET PERFORMANCE MEASURE OF ENSEMBLEALGORITHMS (BOOSTING, BAGGING AND STACKING - FIVE CLASS) MATCHED WITH OPINION LEXICON

T

| Performance Measure | Ensemble Classification Method | | | |
|---|---|---|---|---|
| | **Boosting (Cart and gbm)** | **Bagging (Treebag and Random Forest)** | **Stacking (C50, Cart and Random Forest)** | **Stacking (LDA, KNN and** |
| **Accuracy** | 0.5667 | **0.8147** | **0.812** | **0.8113** |
| **Specificity** | 0.8809 | 0.9489 | 0.9483 | 0.9481 |
| **Sensitivity** | 0.5405 | 0.8057 | 0.7999 | 0.7966 |
| **Precision** | 0.6046 | 0.8583 | 0.8584 | 0.8604 |
| **Recall** | 0.5405 | 0.8057 | 0.7999 | 0.7966 |
| **AUC** | 0.7660 | **0.9078** | **0.9070** | **0.9077** |

**Graphical Representation:**

Fig.2. Shown graphical bar chart representation of testing result on overall accuracy of algorithms like boosting, bagging and stacking. From this it clearly shows that bagging achieves more accuracy with 0.8147 and it is predicted to be the best algorithm.

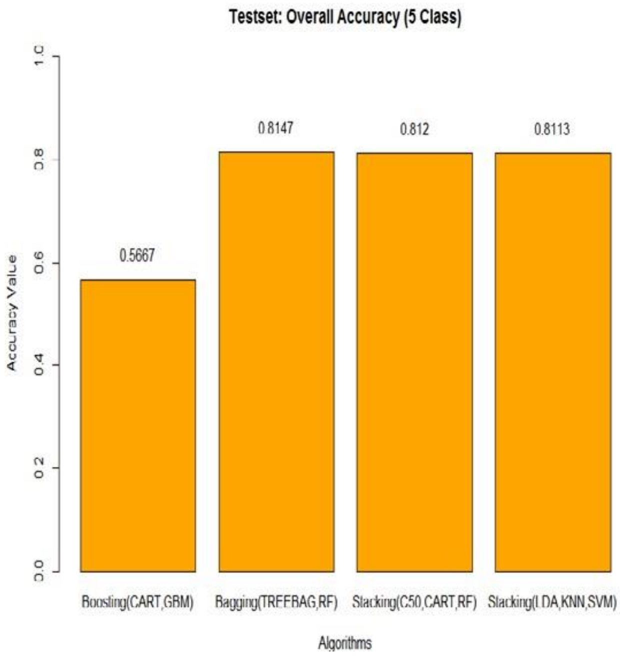


Fig.2. Ensemble Classifications - Overall Accuracy on Testing Result for – (5 Class)

Fig.3. Shows graphical bar chart representation of testing result on overall AUC prediction of algorithms like boosting, bagging and stacking. From this it clearly shows that bagging algorithm achieves more accuracy with 0.9078 and it is predicted to be the best algorithm.

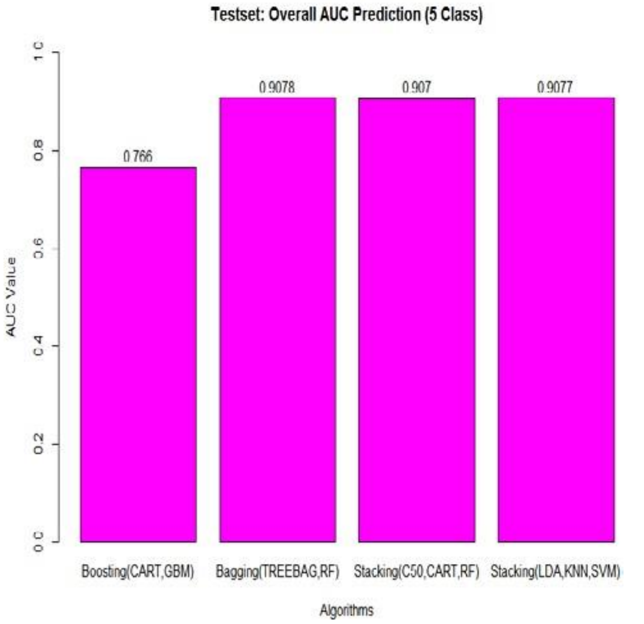


Fig 3. Ensemble Classifications - Overall AUC Prediction on Testing Result for – (5 Class)

## VI. CONCLUSION

This paper suggests the need for making an ensemble of classifiers and the various methods for making it. One of the most recent questions in front of data mining researchers today is "Whether a combined classifier model gives better performance than the best among the base level classifiers". In this work, the base level classifiers are stacked with many Meta learning algorithms like bagging, boosting and stacking algorithms, and it was observed that when multi response model algorithm is used at the Meta level the model is giving much better performance. Hence through this work, it is strongly suggested that a combined ensemble approach gives a better performance than selecting the best base level classifier. Ensemble concept effectively predicts new class labels with high classification accuracy.

## References


[1] Aliaksei Severyn, Alessandro Moschitti, Olga Uryupina, Barbara Plank and Katja Filippova. (2014) 'Opinion Mining on YouTube' Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics, pages 1252–1261, Baltimore, Maryland, USA.

[2] Bing Liu. (2012) 'Sentiment Analysis and Opinion Mining' Morgan & Claypool Publishers, May 2012.

[3] Bo Pang and Lillian Lee. (2004) 'A Sentiment Education : Sentient Analysis Using Subjectivity' Proceedings of ACL, PP. 271- 278, 2004.

[4] Bravo-Marquez F., Mendoza M., Poblete B. (2014) 'Meta-level Sentiment Models for Big Social Data Analysis' Knowledge-Based Systems, 69, 86–99 (2014).

[5] da Silva N.F.F., Hruschka E.R., Hruschka E.R. (2014) 'Tweet Sentiment Analysis with Classifier Ensembles' Decision Support Systems, 66, 170–179 (2014).

[6] Dietterich, T.G. (2000) 'Ensemble Methods in Machine Learning. In: Multiple Classifier Systems' pp. 1-15. Springer, Berlin Heidelberg (2000).

[7] Divya Ramani, Harshita Kanani, Chirag Pandya. (2013) 'Ensemble ofClassifiers Based on Association Rule Mining' ISSN: 2278 – 1323 International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 11, November 2013.

[8] Erik Cambria, Bjorn Schuller, Yunqing Xia and Catherine Havasi. (2013) 'New Avenues in Opinion Mining and Sentiment Analysis' IEEE 15 Published by the IEEE Computer Society, 1541-1672/13.

[9] Federico Neri, Carlo Aliprandi, Federico Capeci, Montserrat Cuadros and Tomas. (2012) 'Sentiment Analysis on Social Media' IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining.

[10] Hanshi Wang, Xinhui Nie, Lizhen Liu. (2013) 'A Fuzzy Domain Sentiment Ontology based Opinion Mining Approach for Chinese Online Product Reviews' JOURNAL OF COMPUTERS, VOL. 8, NO. 9, SEPTEMBER 2013.

[11] Hanxiao Shi, Guodong Zhou and Peide Qian. (2010) 'An Attribute-based Sentiment Analysis System'. Information Technology Journal, 9: 1607-1614.

[12] Jawad Khan and Byeong Soo Jeong. (2016) 'Summarizing customer review based on product feature and opinion' International Conference on Machine Learning and Cybernetics (ICMLC), IEEE, ISSN: 2160- 1348.

[13] 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.

[14] Khan F.H., Bashir S., Qamar U. (2014) 'TOM: Twitter Opinion Mining Framework Using Hybrid Classification Scheme' Decision Support Systems, 57, 245–257 (2014).

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

[16] Meeyoung Cha, Hamed Haddadi, Fabricio Benevenuto and Krishna Gummad P. (2010) 'Weblogs and Social Media Measuring User Influence in Twitter' Proceedings of the Fourth International AAAI Conference on Weblogs, CS Dept., Federal University of Minas Gerais (UFMG), Brazil.

[17] Nidhi Mishra and Jha C.K. (2013) 'Opinion Mining from Text in Movie domain' International Journal of Computer Science Engineering and Information Technology Research (IJCSEITR) ISSN 2249-6831 Vol. 3, Issue 4, Oct 2013, 121-128 © TJPRC Pvt. Ltd.

[18] Rana Alnashwan, Adrian O'Riordan, Humphrey Sorensen, and Cathal Hoare. (2015) 'Improving Sentiment Analysis Through Ensemble Learning of Meta-level Features' literature survey, 2015.

[19] Robi Polikar. (2006) 'Ensemble based system in decision making' IEEE circuits and systems magazine-2006.

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

[21] Velu C.M., Kishana Kashwan R. (2012) 'Performance Analysis for Visual Data Mining Classification Techniques of Decision Tree, Ensemble and SOM' International Journal of Computer Applications (0975 – 8887) Volume 57– No.22, November 2012.

[22] Yoav Freund, Robert Schapire E. (1996) 'Experiments with a New Boosting Algorithm' AT&T Laboratories 600 Mountain Avenue Murray Hill, NJ 07974-0636 Machine Learning: Proceedings of the Thirteenth International Conference, 1996.

Zhu Nanli, Zou Ping and Li Weiguo. (2012) 'Sentiment Analysis: A literature review' International Symposium on Management of Technology (ISMOT), ISBN: 978-1-4673-4593-4.