Journal of
Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.776948

# A COMBINED MACHINE LEARNING APPROACH FOR SENTIMENT ANALYSIS ON PRODUCT REVIEWS

**Uttej Kumar Nannapaneni[1], A V R Mayuri[2], V Sujatha Lakshmi[3], D Naga Malleswari[4], Sajja Radharani[1], Deevi Radha Rani[1]**
[1]Department of CSE, VFSTR Deemed to be University, Guntur, India
[2]School of Computing Science and Engineering, VIT Bhopal University, Bhopal, India
[3]Department of Computer Applications, RVR & JC College of Engineering, Guntur, India
[4]Department of CSE, KoneruLakshmaiah Education Foundation, Vaddeswaram, India

**Abstract:**
Sentimental Analysis has been popular in recent years with all online product companies. The number of people who utilize a specific product has grown, prompting the industry to improve the product's features. Many people who utilize websites, blogs, and online shopping these days leave reviews on the things they have used. Other people took these reviews into account while looking for items. As a result, most of the companies decided that delivering the good product to the consumer depends on user evaluations employing the emotional analysis approach. In Sentimental Analysis a group of users providing reviews are collected and processed to suggest the users. The reviews that are offered are lengthy and contain many paragraphs of substance. The data for this paper was gathered from the Amazon website. The dataset if pre-processed and categorized using the Naive Bayes and SVM algorithms. These previous methods delivered precision that was insufficient. As a result, an ensemble technique was used to improve the accuracy of the reviews. Theproposed classification method combines SVM and Naïve Bayes algorithms and calculates the mode value for each algorithm using the vote reference. We developed an Ensemble technique that improves on the current algorithm's accuracy. After the accuracy is computed, the user is suggested a certain product based on the reviews.
**Keywords:** Machine Learning, SVM, Naive Bayes, Ensemble, Product Reviews

## I. Introduction:

Sentiment analysis is a popular academic topic that is being pursued by a number of scholars. Currently, many approaches, methods, tools, and strategies have been developed in this subject to improve the effectiveness and accuracy of product or data recommendation. In comparison to traditional text papers, online evaluations have less characteristics Text evaluations are often made up of a variety of idioms and traditional terms gathered from a variety of sources. Scoring algorithms may be used to determine the overall sentiment of a text by identifying weighted phrases from the preceding side and adding them up in a proprietary fashion. Social networks and blogsare examples of places where people may share their own opinions on certain issues. Using computational algorithms to extract public opinion from text reviews on multiple websites and analyze it is a popular research topic these days. These studies make use of a group of root terms derived from conventional and slag languages. These researchers provide an efficient method for acquiring relevant data and assisting with data processing. Sentiment expressions are detected as a result of these methods. The review datasets are divided into two categories: supervised and unsupervised datasets.

Supervised learning have a collection of labeled data, which means they know the values of the inputs and outputs. The goal of machine learning is to discover the exact connection between them, which is known as a model in mathematics. In machine learning, there are a variety of methods that may be used to create a model of the data. The researchers' goal, and how machine learning might help, is to predict the result given a new input after the model is known. We don't have the data labeled in unsupervised learning. We have the inputs, but not the outputs, as it were. As a result, the goal is to find a pattern in the data. It is required to obtain a model in this case.

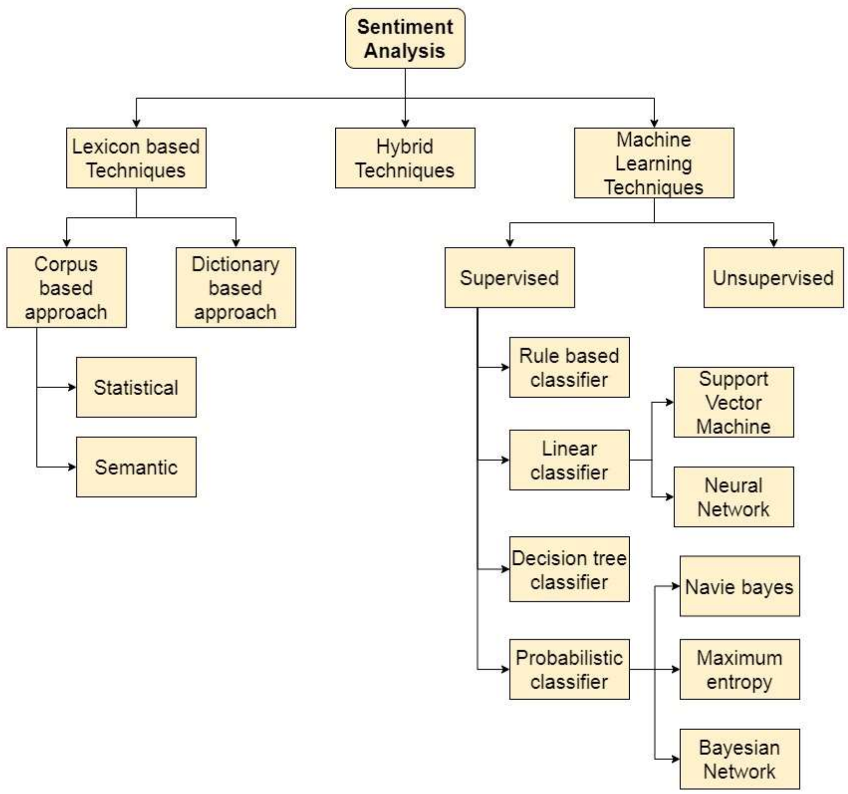


Figure 1. Classification of Sentiment analysis Techniques

Sentiment Analysis allows users to communicate each other via reviews. Other consumers use these reviews and responses to decide whether or not to purchase the product. One of the major issues in sentiment analysis is increasing accuracy. The majority of algorithms [1] [2] [3] [4] were successful in classifying the dataset, but not in achieving accuracy. As a result, the system is unable to provide any additional recommendations. Combining the several algorithms was a considerable problem when using the ensemble technique. Each algorithm has its own set of features and operates in a unique way. As a result, merging these algorithms and categorizing the data took a lot of time and effort, and this voting method (Ensemble) must be more accurate than the current system. The current system has a lesser accuracy, and the suggested technique must have a higher accuracy than the current system algorithms.

Because there are several current ways to sentiment analysis (Figure 1), such as Naive Bayes and Logistic regression for categorizing product reviews, these approaches are effective but not accurate. Many research investigations are being conducted in this subject in attempt to improve accuracy. To improve the classification accuracy, a novel approach is used in our work.

The rest of the paper consist is organized as follows: literature survey section discusses the existing methods of sentiment analysis; proposed approach sections gives the phases of sentiment analysis and each phase in detail; Implementation section describes the algorithm used in the proposed approach; Experimental results section presents the results obtained from our proposed approach; finally our work is concluded in conclusion section.

**II. Literature Survey:**

In [5] the classifier is used to categorise the tweets as positive, negative, or neutral, and the twitter data was pre-processed using POS (Parts of Speech) tagging. In NLP programmes for information retrieval and extraction, the POS tagger is utilised as a pre-processing approach. Sequential classification, a pre-classification process, is used in the POS tagging method.An enhanced technique of HMM model to examine the outcome of emotional subjectivity, i.e., POS tagging, was used in the syntactic level job. Incorporate the procedure into an interval-type HMM. The proposed technique improved the performance of POS tagging, according to the experimental analysis and findings on datasets of public tagging. This approach has the advantage of improving POS tagging performance over other methods.It outperforms the human classification training set. Pre-classification aids in the presentation of models for training set. The disadvantage of this strategy is that it can only be used for certain characteristics and not for all other features. Processing time is longer than for other types of algorithms. When choosing on algorithms, it takes longer.

In [6] authors created a machine learning strategy after considerable thought which depends on quality of various reviews. The goal is to build a machine learning technique to automatically analyse and rate the usefulness of reviews. Automatic categorization aids in the analysis of reviews that are poorly written and lack the necessary information for the consumer. The classifier finds and analyses trends in the data, and makes review publication easier.

In [7] demonstrates how we rated aspects according to their value. Our system accepts free text reviews as input. Free reviews are analysed using natural language processing (NLP), which detects the product's features. We utilised the supervised classifier SVM to categorise attitudes, and then we used the probabilistic aspect ranking method to rank natural aspects. It has the advantage of swiftly classifying. The accuracy of logistic regression is superior than accuracy of accuracy of accuracy of accuracy of accuracy of accuracy of accuracy of accuracy of accuracy It only works on specific aspects, which is a disadvantage. If the ranking isn't done correctly, it might lead to misunderstanding.

Table 1.1: Literature Review

| S No | Studies | Algorithm / Method Used | Description |
|---|---|---|---|
| 1 | Twitter part-of-speech tagging using pre-classification hidden Markov model [5] | Hidden Markov Model | It generally classifies the tweet as positive, negative or neutral |
| 2 | Predicting Helpfulness Ratings of Amazon Product Reviews [6] | SVM and Naive Bayes | It covers automata features of data and also classified based on token and syntactic analysis |
| 3 | Survey on Product Review Sentiment Analysis with Aspect Ranking [7] | SVM | Classification is done based on aspects of the product. Every aspect are provided with aspect ranking |
| 4 | Automatically detecting and rating product aspects from textual customer reviews [8] | Aspectator | Special algorithm for aspect-based classification |
| 5 | Unsupervised Opinion Mining From Text Reviews Using SentiWordNet [9] | Sentiword Net | The classification is done based on aspect level which finds out aggregate scores for a particular aspect (Fixed Syntactic patterns) |
| 6 | Ontology-based sentiment analysis of twitter posts. Expert Systems with Applications [10] | Ensemble Approach | Classification is done based on polarity |
| 7 | Sentiment analysis using product review data. Journal of Big Data [11] | Naive Bayes | Extracting subjective content and tackling polarity categorization problem |
| 8 | Sentiment Analysis for Movie Reviews [12] | Random Forest | Classified by counting the number of words repeated |
| 9 | User Bias Removal in Fine Grained Sentiment Analysis [13] | SVM | Normalizing each user review score with respect to mean and standard deviation of all products rated by the user |
| 10 | A Twitter Sentiment Analysis for Cloud Providers: A Case Study of Azure [14] | Naive Bayes | AWS and Azure. The Opinion of customers around each one of them |

In [8] assess the aspect-based sentiments, an innovative algorithm called aspectator was utilised, which gave accuracy depending on domains. To uncover opinions revealed towards candidate characteristics, handcrafted dependency routes were matched in each unique sentence. Then, using a WordNet-based similarity metric, it groups together the same characteristic stated in several locations. Finally, it analyses each element and generates a sentiment score for each, which reveals the general evolving opinion of a group of customers about a certain feature of a product. Because this strategy solely involves the self-sentiment lexicon, domain-specific information and seed words are not required. The specific method used in this article delivers improved performance and classification. Sentiment ratings are produced in a way that is more predictive than typical algorithms. The accuracy of this document is low for certain characteristics, and it is not applicable for all features. In aspect ranking, matching takes longer than normal.

In [9] employed aspect-level sentiment analysis, which considers the fact that a text may include opinions on several aspects of one or more things. The categorization is carried out at the aspect level, which determines aggregate scores for a certain aspect. SentiWordNet extracts domain-specific characteristics from a message using Senti-features, which may construct a vocabulary using abbreviations, slang words, and emoji before expanding the message with other features. The benefit of this paper has resulted in favourable outcomes in more ways than one. SentiWordNet values are classified via feature extraction, which plays an important role. The downside of this work is that other algorithms are unable to categorise the characteristics. Every other input has a different level of accuracy.

In [11] has carried out emotional analysis using Naive Bayes classifier to extract subjective content and solve the polarity categorization problem. It focuses on both the sentence and document levels. To categorise sentiment polarity, a generic approach is presented with suitable process descriptions. The dataset for this study was created by analysing online product reviews on Amazon. The downside of this approach is that if Naive Bayes does not yield enough accuracy, algorithms such as "aspectator", "sentiwordnet", and others can be used instead. Because of the low accuracy, it is not recommended for decision-making.

In [12] By measuring the number of times the word appeared in the text, a random forest classifier was used to create a movie review. " It may be used to identify the reviewer's attitude toward certain issues as well as the overall polarity of the review. Using sentiment analysis, determine the reviewer's emotional state when writing the review and determine if the individual was "glad," "sad," "angry," and so on. Sentiment Analysis is a method of analysing a series of movie reviews to determine what the reviewers' general reaction to the film was, i.e., whether they liked or disliked it. The links between the phrases in the review to anticipate the review's overall polarity." The benefit of this strategy is that it is quite efficient and simple to apply. The great precision and scalability of memory-based approaches are standard features. When there is limited data regarding user ratings, this method's downside is that it is inaccurate. Systems are not content aware, which means that when they make suggestions, they do not take into account information about the things.

In [13] authors used the natural language processing field has recently become interested in fine-grained sentiment analysis of text reviews. The majority of previous work has focused on developing effective feature representations of text reviews for categorization. During fine-grained sentiment categorization, the algorithms often neglect other common features such as user identification, product identity, and helpfulness rating. The existence of user bias in review ratings causes noise, which is a serious problem with existing classification techniques. The benefit of this study is that it forecasts a review's score as the mode of practically all reviews. Performance is improved by using hyper-parameters. Produces consistent results regardless of the testing input. The paper's drawback is that it focuses solely on feature representation strategies. Bigram, when used to classify Unigram, makes processing more challenging.

In [14] authors claimed that social networking sites are web-based activities centred on hobbies, connections, family, and so on. This allows individuals to speak on a variety of topics, including exchanging ideas, information, feelings, and events. Text mining is the practise of collecting information from social media sites, public blogs on a certain topic. In the fields of information retrieval and natural language processing, this strategy is used. Dataset extraction from Amazon's cloud products must be studied in order to determine client thoughts about the company. The benefit of this study is that it may be used with any input and for any sort of dataset. For larger datasets, reliable cloud storage of classified assessments is required. The drawback of this strategy is that it does not perform well with sentiment net texts. Not is a good choice for words that aren't used very often. Other data is less accurate than emoticons.

In the current system, several algorithms are employed in the categorization of reviews. These algorithms have a lower level of accuracy. Accuracy must be assessed using a variety of indicators. In general, reviews are categorised as good, negative, or neutral. As a result, when the accuracy is poor, the algorithm is unable to determine whether a review is favourable, negative, or neutral. These reviews can't be used for any subsequent recommendations. Each method takes a short amount of time to categorise the data.

**III. Proposed Approach for Sentiment Analysis:**

In our proposed approach named Ensemble we combined algorithm to classify the dataset by using votes. The construction of the approach is more complex as we are trying to improve the accuracy of the algorithm. The algorithms that are combined are allotted with a mode value as the dataset is labeled dataset. Ensemble is efficient classification algorithm compared to existing algorithms.

To categorize the review, the suggested method use an Ensemble technique. Majority voting or voting classifier are other names for the ensemble technique. The ensemble approach is a set of algorithms that work together. The outcome is based on the predictors' vote, and one or more algorithms are categorized (mode values of all the algorithms). The Ensemble technique is utilized in this case, which combines Naive Bayes with SVM. As a result, when compared to current algorithms that are less accurate than the proposed one, the suggested technique would generate better accuracy. The ensemble technique, in general, delivers higher accuracy since it integrates algorithms that aid in efficient operation.This approach outperforms any of the algorithms by a significant margin. Every model makes a forecast for each situation, with the final output prediction receiving more than ½ of the votes.

Advantages

- The Ensemble approach is more accurate than other algorithms that work on their own.
- It is commonly used to create many hypotheses from a single base learner.
- It avoids over fitting when used in conjunction with more than two algorithms.
- The ensemble technique decreases variation and averages out biases, making it possible to run all of the algorithms concurrently rather than one at a time.

Our proposed approach for sentiment analysis consists of 3 phases (Figure 2): Data collection phase, Pre-processing phase and Sentiment classification phase.

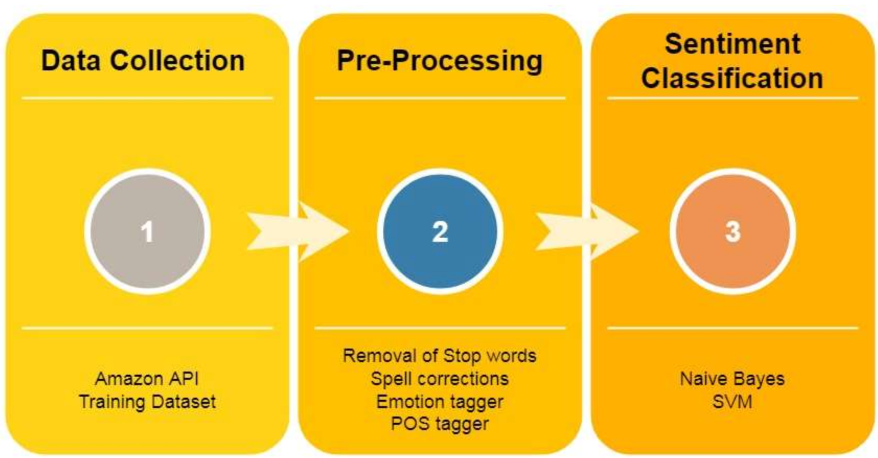


Figure 2. Proposed approach for sentiment analysis

Data Collection Phase:

The Training dataset will be built in this phase by gathering data using the Amazon API (Figure 3). A scraping tool may also be used to build a dataset, which is then saved in Excel. There will be no ambiguity in the data because it will be provided in a systematic manner. The information retrieved can be saved in a file, printed, or displayed on a computer screen. Datasets are gathered based on our system requirements, as some of the datasets may be large and require analysis before use. It's a supervised classification, and there will be labels for categorization in the dataset. Data will be represented by label values.

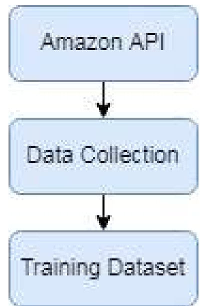


Figure 3. Data Collection Phase

Pre-Processing Phase:

After the data has been collected, it should be pre-processed to eliminate any undesired phrases, words, or symbols. Before executing an analysis, the representation and quality of data must first be addressed. Symbols such as '@' and '#' will be eliminated first, and then spell check and emotion tagging is done. The tagging of POS was the next step in the process (Parts Of Speech). All of the required POS words and phrases are retrieved here, including verbs, adverbs, and adjectives. These words will be taken into account when calculating the weighted ratings (Figure 4).

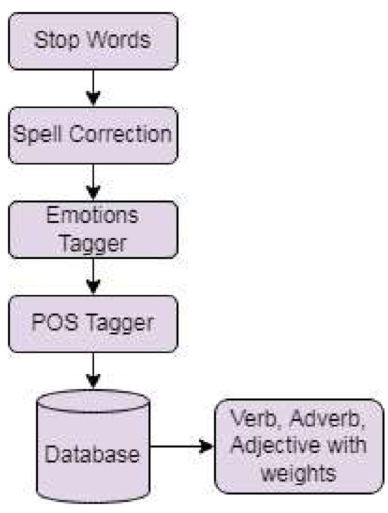


Figure 4. Pre-Processing Phase

Sentiment Classification Phase:

Once training data is gathered, Naive Bayes and SVM classifiers are used (Figure 5). Supervised and Unsupervised classifications are the two most common forms of categorization. Labels will be given for supervised categorization. Unsupervised classification will not be given any labels because the training data is made up of a collection of training instances. Some supervised classifiers, such as Naive Bayes and SVM, have been proposed. Naive Bayes is a basic approach for building classifiers that give class labels to problem cases represented as vectors of feature values, with the class labels selected from a finite set. The state is not immediately visible with Support Vector Machine (SVM), but the output depends on the state is. Each state is assigned a probability distribution across the potential token outputs. As a result, the token sequence created by an SVM contains some information about the state sequence. Finally, all of the foregoing approaches are combined to create an Ensemble approach. Once the classification is complete, the majority vote (Accuracy) is taken, and the classified system delivers greater accuracy than the current system.

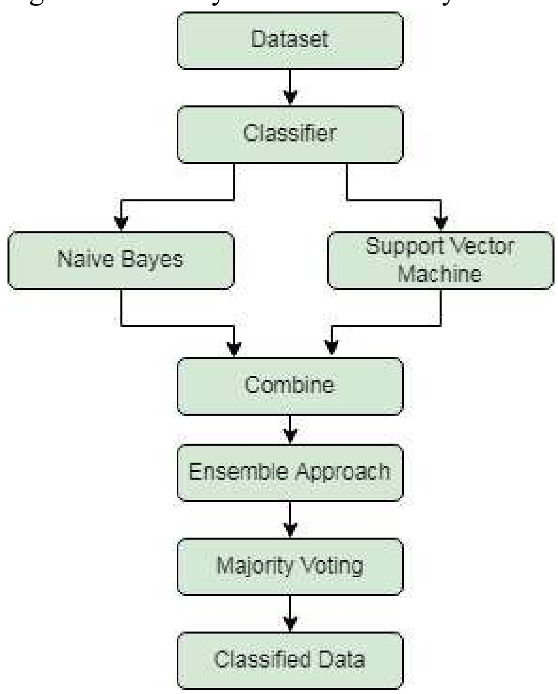


Figure 5. Sentiment Classification Phase

**IV. Implementation:**

1. Naïve Bayes

A basic probability-based technique known as a Naive Bayes classifier. It employs the Bayes theorem but makes the unreasonable assumption that the cases are independent of one another in the real world. In difficult real-world circumstances, the Naive Bayes classifier performs well.

Algorithm 1. Naïve Bayes

Step 1: Split data into k subsets of equal size.
Step 2: Use each subset in turn for testing, the remainder for training.
Step 3: Find the total positive and total negative probability of the data.
Step 4: For every word in a dataset, finds its positive probability and negative probability.
Step 5: Multiply all the positive and all the negative probability values.
Step 6: Multiply those values with the total probability values.
Step 7: Select the sentiment for which probability has more significant value.

2. Support Vector Machine

For a collection of training sets with varied forms and categories, SVM is utilised in regression analysis and classification. Calculate the similarity using just a portion of the training points obtained during the training phase. As a result, we fix a selection of training points from which we compute the similarity of any test point throughout the training phase. These selected training points are referred to as support vectors since only these points will help us decide which test point class to use. We expect that our training step uncovers as few support vectors as possible, reducing the amount of similarities to calculate.

Algorithm 2. SVM

Step 1: Find weight for each training point, and those points whose weight becomes zero are not support vectors i.e. their importance is zero during test time, and rest are support vectors.
Step 2: Transform each of the vectors into some other higher dimensional vector, and then take dot product between those two complex higher-dimensional support vectors.
Step 3: Find the maximum of vectors as a sentiment value.

3. Ensemble

To improve the outcomes, the suggested technique combines Naive Bayes and Support Vector Machine classification methods. Figure 3 depicts an ensemble method that demonstrates the idea of data processing from the dataset to the final output. Data for training and testing had been cleaned and preprocessed before being fed into machine learning algorithms. The review is classified using an ensemble technique. Majority voting or voting classifier are other names for the ensemble technique. The ensemble approach is a set of algorithms that work together. The outcome is based on the predictors' vote, and one or more algorithms are categorized. The Ensemble technique is utilized in this case, which combines Naive Bayes with SVM. As a result, when compared to existing algorithms, the proposed algorithms generate better accuracy. The ensemble technique, in general, delivers higher accuracy since it integrates algorithms that aid in efficient operation. This approach outperforms any of the algorithms by a significant margin. At each instance, each model performs a prediction analysis which can predict 50% of votes. The ensemble approach could not provide a reliable forecast in this case if none of the predictions received more than 1/2 of the votes.

Algorithm 3. Ensemble

Step 1: Implement existing algorithms such as Naive Bayes and SVM.
Step 2: Create a class called Voting classifier and inherit the list of classifiers used such as Naive Bayes and SVM.
Step 3: Classification of the training set based on this voting classifier and this would calculate the votes for every algorithm that classifying the dataset.
Step 4: In voting classifier class, create our own classify method. Now iterating through our list of classifiers objects. Then, for each one, we ask it to classify based on the features.
Step 5: After we are done iterating, we then return the mode (votes), which is just returning the

## V. Experimental Results and Discussion:

In this section, we presented experimental results of 3 machine learning techniques to perform sentiment analysis of Amazon product review dataset. The evaluation parameters considered are accuracy, precision, recall and F1 Score.

$$Accuracy = \frac{TP + TN}{TotalNo.ofObservations}$$

$$Precision = \frac{TP}{TP + FP}$$

$$Recall = \frac{TP}{TP + FN}$$

$$F1\ Score = \frac{2 * TP}{2 * TP + FP + FN}$$

The Table 2 shows the classification results of all evaluation parameters for all the techniques.

Table 2. Summary of Classification Results

| Technique | Accuracy (%) | Precision (%) | Recall (%) | F1 Score (%) |
|---|---|---|---|---|
| Naïve Bayes | 87 | 88 | 98 | 92 |
| SVM | 89 | 90 | 98 | 94 |
| Ensemble | 90 | 91 | 97 | 94 |

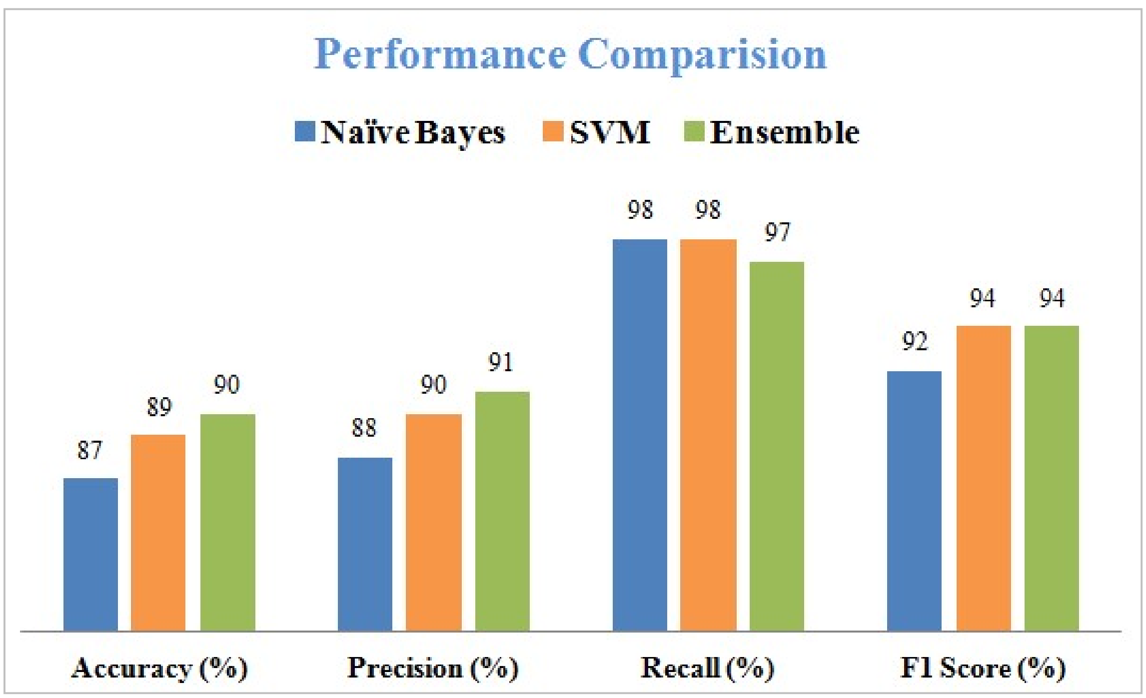


Figure 6. Comparison of accuracy, precision, recall and F1 score of 3 classifiers

**Conclusion:**

The suggested Ensemble approach improves upon the present algorithm's accuracy. We also successfully addressed the algorithm's performance, accuracy, and speed of execution. The suggested system outperforms the current method in terms of accuracy. The accuracy level, on the other hand, changes with the number of classifiers we combine to generate a projected review result. We have developed a strategy to improve the accuracy of review categorization in this study. If the accuracy is higher, the system can be used to make recommendations to users.

The aspect level classification will be used to guide future development. Aspect level classification is different from standard categorization in that it specifies the review. Features are defined and processed for the algorithms in aspect level classification. As a result, review-based categorization is well-developed these days, and it aids in the product suggestion to consumers based on the user's feedback. We may also practice on conveying emotions in the evaluation using emotions.

**References**


1. Mahendran, N., and T. Mekala. "A survey: sentiment analysis using machine learning techniques for social media analytics." International Journal of Pure and Applied Mathematics 118.8 (2018): 419-423.
2. Nguyen, Heidi, et al. "Comparative study of sentiment analysis with product reviews using machine learning and lexicon-based approaches." SMU Data Science Review 1.4 (2018): 7.
3. Rahman, Atiqur, and Md Sharif Hossen. "Sentiment analysis on movie review data using machine learning approach." 2019 International Conference on Bangla Speech and Language Processing (ICBSLP). IEEE, 2019.
4. Mehta, Pooja, and Sharnil Pandya. "A review on sentiment analysis methodologies, practices and applications." International Journal of Scientific and Technology Research 9.2 (2020): 601-609.
5. Sun, S., Liu, H., Lin, H., & Abraham, A. (2012, October). Twitter part-of-speech tagging using pre-classification Hidden Markov model. In *Systems, Man, and Cybernetics (SMC), 2012 IEEE International Conference on* (pp. 1118-1123). IEEE.
6. Rodak, J., Xiao, M., & Longoria, S. (2014). Predicting helpfulness ratings of amazon product reviews. Available at http://cs229.stanford.edu/proj2014/Jordan%20Rodak,%20Minna%20Xiao,%20Steven%20Longoria,%20Predicting%20Helpfulness%20Ratings%20of%20Amazon%20Product%20Reviews.pdf
7. Patil, H., & Mane, P. M. (2016). Survey on product review sentiment analysis with aspect ranking. *International Journal of Science and Research*, *5*(9), 749–752.
8. Bancken, W., Alfarone, D., & Davis, J. (2014, August). Automatically detecting and rating product aspects from textual customer reviews. In *Proceedings of the 1st international workshop on interactions between data mining and natural language processing at ECML/PKDD,* (pp. 1-16).
9. Soni, V., & Patel, M. R. (2014). Unsupervised opinion mining from text reviews using sentiwordnet. *International Journal of Computer Trends and Technology, 11*(5), 234–238.
10. Kontopoulos, E., Berberidis, C., Dergiades, T., &Bassiliades, N. (2013). Ontology-based sentiment analysis of twitter posts. *Expert Systems with Applications*, *40*(10), 4065–4074.

11. Fang, X., & Zhan, J. (2015). Sentiment analysis using product review data. *Journal of Big Data*, *2*(1), 1-14.
12. Goyal, A., &Parulekar, A. (2015). Sentiment analysis for movie reviews. Available at https://cseweb.ucsd.edu/classes/wi15/cse255-a/reports/fa15/003.pdf.
13. Wadbude, R., Gupta, V., Mekala, D., Jindal, J., &Karnick, H. (2016). User bias removal in fine grained sentiment analysis. In *European Chapter of the Association for Computational Linguistics*, arXiv:1612.06821v1 [cs.CL] 20 Dec 2016.
14. Qaisi, L. M., &Aljarah, I. (2016, July). A twitter sentiment analysis for cloud providers: a case study of Azure vs. AWS. In *Computer Science and Information Technology (CSIT), 2016 7th International Conference on* (pp. 1-6). IEEE.