**IAES International Journal of Artificial Intelligence (IJ-AI)**
Vol. 14, No. 3, June 2025, pp. 2001~2011
ISSN: 2252-8938, DOI: 10.11591/ijai.v14.i3.pp2001-2011 

# Hybrid semantic model based on machine learning for sentiment classification of consumer reviews

**Palaniraj Rajidurai Parvathy[1], Nagarajan Mohankumar[2], Rajendran Shobiga[3], Gour Sundar Mitra Thakur[4], Mamatha Bandaru[5], Velusamy Sujatha[6], Shanmugam Sujatha[7]**
[1]Project Manager, Mphasis Corporation, Chandler, United States of America
[2]Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India
[3]Department of Electronics and Communication Engineering, J. J. College of Engineering and Technology, Tiruchirappalli, India
[4]Department of Computer Science and Engineering (AIML), Dr. B. C. Roy Engineering College, Durgapur, India
[5]Department of Electronics and Communication Engineering, Koneru Lakshmaiah Education Foundation, Guntur, India
[6]Department of Electronics and Communication Engineering, Shree Sathyam College of Engineering and Technology, Sangagiri, India
[7]Department of Biomedical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, India

**Article Info**

***Article history:***

Received Feb 27, 2024
Revised Dec 20, 2024
Accepted Jan 27, 2025

***Keywords:***

Consumer reviews
Ensemble categorization
Lexicon model
Sentiment analysis
Sentiment score

**ABSTRACT**

Digital information is regularly produced from a variety of sources, including social media and customer service reviews. For the purpose of increasing customer happiness, this written data must be processed to extract user comments. Consumers typically share comments and thoughts about consumable items, technological goods, and services supplied for payment in the modern period of consumerism with simple access to social networking globe. Each object has a plethora of remarks or thoughts that demand special attention due to their sentimental worth, especially in the written portions. The goal of the current project is to do sentiment prediction on the Amazon Electronics, Kindle, and Gift Card datasets. In order to predict sentiment and evaluate utilizing many executions evaluates admitting accuracy, recall, and F1-score, a hybrid soft voting ensemble method that combines lexical and ensemble methodologies is proposed in this study. In addition to calculating a subjectivity score and sentiment score, this study also suggests a non-interpretive sentiment class label that may be used to assess the sign of the evaluations applying suggested method for sentiment categorization. The effectiveness of our suggested ensemble model is examined using datasets from Amazon customer product reviews, and we found an improvement of 2-5% in accuracy compared to the current state-of-the-art ensemble method.

*This is an open access article under the CC BY-SA license.*

***Corresponding Author:***

Nagarajan Mohankumar
Symbiosis Institute of Technology, Symbiosis International (Deemed University)
Nagpur Campus, Pune, India
Email: nmkprofessor@gmail.com

## 1. INTRODUCTION

Numerous sectors for data analysis have emerged as a result of the enormous amount of details created by popular social media platforms. The goal of sentiment analysis (SA), in particular, is to glean user thoughts about a good or service from the relevant text. The raw text must be analyzed using machine learning (ML) algorithms along with natural language processing (NLP) techniques due to the enormous amount of information generated [1]. Sentiment examination made possible through the detail that social media is a necessary device for people and that they regularly distribute their notions. Expressions of sentiment in texts are predicted by SA studies. People's texts are evaluated for their positivity, negativity, or neutrality. Businesses can do a preliminary analysis of new goods, and films thanks to SA [2]. On the other hand, in order to achieve a significant increase in accuracy, this methodology is costly, time-utilizing, and may need advanced technology [3].

SA and NLP are becoming increasingly popular as a result. SA uses tools and methods from NLP, ML, and statistics to determine the contextual polarity of text content [4]. The internet is one of the information sources for SA that contains an ever-growing amount of data in a variety of multimedia formats that can be useful to governmental bodies, commercial organizations, and individual decision-makers [5]. Initially, a quick study was conducted to identify a generalized consumer review pattern on textual customer comments shared for the products they bought online [6]. These online reviews are posted on social media platforms or on the websites of reputable companies. As indicated in Figure 1, consumer reviews pattern can be divided into two distinct categories based on the inclinations of the reviewers' points of view: polar reviews and non-polar reviews.

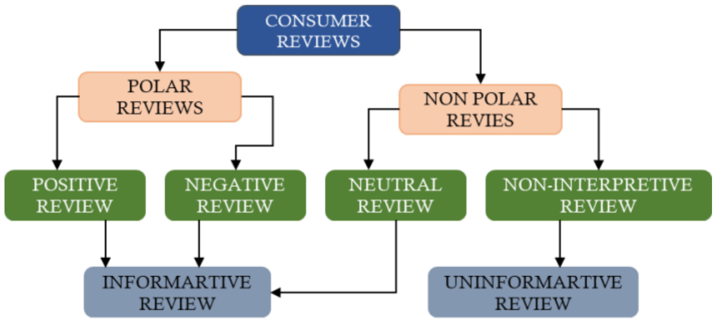


Figure 1. Illustration of consumer reviews pattern

Due to their attitudinal viewpoint slant, polar evaluations include both positive and negative reviews, even if both of them are interpretative and useful from the perspective of the text material. Although they are in the non-polar review class, neutral evaluations contain educational language. Therefore, a category of reviews known as informative reviews includes all three types of evaluations: positive, negative, and neutral. After discovering a lack of interpretive components in the review text, non-polar reviews were finally given a separate entity for the first time known as non-interpretive reviews. Additionally, they are labelled as "uninformative reviews". Sentence-level analysis seeks to decide if a specific sentence represents positive, unfavorable (negative) or neutral opinions. However, further in-depth examination at an entity or aspect level might completely expose attitudes or opinions about many characteristics or aspects of a specific entity. Some of the newly determined non-interpretive reviews are also given separately with customer comments on Amazon Kindle and Electronics. Table 1 provides a brief summary of text content trends for widely accepted sentiment groups, namely positive, negative, and neutral. Examples of recently discovered non-interpretive reviews are also provided in isolation with consumer comments from the Amazon Kindle dataset and the electronic product dataset [7] respectively. Product reviews should always rationally relate to a product's quality or utility value from the perspective of the consumer. It might also represent his ranking of merits and demerits. In addition, the extreme polar review remarks are likewise deserving of a high subjectivity rating.

Table 1. Example of informative consumer review taken from Amazon kindle dataset

| Consumer reviews and comments | Sentiment classification |
|---|---|
| Great results. I loved it. | Positive |
| Nice stuff all around fantastic tablet for the money. The navigation is simple. | Positive |
| Poor visuals, unstable wireless connection, and overall dissatisfaction made me regret purchasing this product. | Negative |
| I won't ever suggest this to anyone. | Negative |
| Sometimes it simply remains in sleep mode, requiring a restart. | Non-Interpretive |
| Nothing to say about the product | Non-Interpretive |
| Could not understand it | Non-Interpretive |
| Very lagging and suitable for kids only | Neutral |
| Consider an Android tablet if you require more than basic model. | Neutral |

As a result, it is suggested in this research that the reviews are added to a new category of "non-interpretive" reviews. Reviews that fall into the "non-interpretive" category are "non-polar" and devoid of information. As a result, they are also categorized as "uninformative" reviews. The following is a list of this study's main contributions: i) to determine how subjectively the unstructured consumer review words are expressed; ii) to develop a new sentiment category called "non-interpretive sentiment" based on the subjective expression score obtained from reviews; and iii) by merging Lexicon-based as well as ensemble learning with soft voting for sentiment categorization, propose a unique hybrid model for sentiment prediction.

## 2. RELATED WORK

Nowadays, sentiment evaluation is increasingly popular with businesses, governments, and organizations [8]. ML methods reach superior accuracy, sentiment categorization [9]. Balaji *et al.* [10] used powerful ML techniques to thoroughly examine the various social media analysis and algorithms. Ohana and Tierney [11] evaluated how well the senti-wordnet opinion lexicon classified the sentiment in movie reviews. Jain *et al.* [12] describes the generation of several classifiers through training with various feature sets, followed by the selection and combination of component classifiers using a number of preset combination criteria. A useful classification method for many domains is the ensemble methodology, which integrates the results of various base classification models to provide an integrated output [13]. A combination of various classification algorithms (K-nearest neighbor (KNN) and Bayesian classifiers) yields superior results than any single type of classifier in the early work [14]. Turkish SA techniques for customer reviews of hotels were developed by Erşahin *et al.* [15] using a combination of sentiment dictionaries and ML algorithms. Coban *et al.* [16] recently did a SA study using the innovative extreme learning machine (ELM) to extract opinions from Turkish tweets [16].

To increase the precision of the opinion mining process, Alfrjani *et al.* [17] suggests a hybrid semantic knowledge-based ML technique. Bhoir and Kolte [18] presents a rule-based approach employing two techniques: a lexicon applied to internet movie database (IMDb) movie reviews and a naïve Bayes classifier on ML. The naïve Bayes classifier is defended in the study [19]. A reinforcement learning technique basis of document level aspect-based sentiment categorization (DASC), Hasan *et al.* [20] suggests a hierarchical reinforcement learning strategy. The accuracy of the ensemble learning models [21] can be significantly improved by using the right ensemble learning algorithm [22], [23]. For the purpose of classifying tweet sentiment, Rehioui and Idrissi [24] suggests a brand-new clustering technique that makes use of K-means and Denclue. Al-Saqqa *et al.* [25] offers cutting-edge methodology using a group of classifiers to ascertain the polarity of the Arabic text's opinions.

Behera *et al.* [26] talks on numerous ensemble learning techniques used to enhance a model's performance in terms of classification, function approximation, prediction, and performance. In order to improve the accuracy, performance, and speed at which the algorithm is executed, Sadhasivam and Kalivaradhan [27] suggests a majority voting-based ensemble technique that combines support vector machine (SVM) [28] with naïve Bayes. In order to address the binary classification problem based on tweets, movie reviews, and product reviews, Rajeswari *et al.* [29] proposed a hybrid strategy employing lexical and ML methods. To ascertain the sentiment of TripAdvisor user reviews, Maheswari *et al.* [30] use homogeneous ensemble classifiers including bagging decision tree, bagged multilayer perceptrons (MLP), random forests (RF), and logistic model trees. The research claims that the hybrid technique is more efficient and gets over the drawbacks of each of the original lexicon method and ensemble method-based methods. A hybrid bidirectional long short-term memory (Bi-LSTM)-artificial neural network (ANN) model based on both temporal and hypernym characteristics is implemented by Sridhar and Sanagavarapu [31]. The previous research employs a hybrid approach to SA, combining the sentiment lexicon with ML algorithms like SVM [32] and KNN, and finds that SVM executes better on news comments than KNN [33], [34]. In this work, a hybrid model based on ensemble and lexicon models using soft voting method is developed for sentiment classification trained on Amazon Kindle and Electronics dataset. Cloud computing allows the proposed seizure prediction mechanism is more accessible and scalable [35]. SVM's classification algorithm to categorizes patients' risks and forecast therapy. Recurrent neural networks algorithm to recogonizes trends in patient symptoms and drug [36]. Cloud-based decision-support system [37] applying KNN algorithm with IoT structure system integrates IoT devices to collect and transmit the data [38].

## 3. PROPOSED METHOD

This paper proposes a hybrid model based soft voting model by combining lexicon approach and an ensembling technique as shown in Figure 2. The lexicon and ML to create predictions by weighted average.

The open-source library named NumPy, SCIKIT learn [39] senti-wordnet, and natural language toolkit (NLTK) are also installed to perform this simulation.

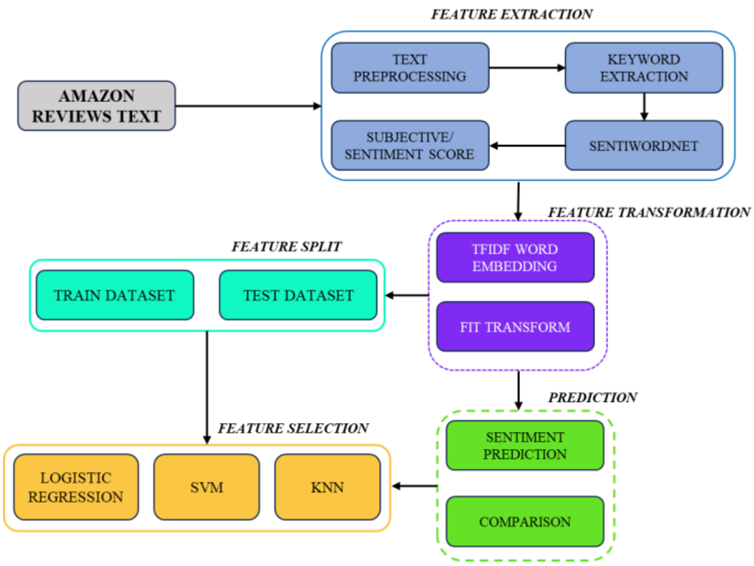


Figure 2. Hybrid ensemble method architecture

### 3.1. Datasets

This work is done by consumer electronic product review (CEPR) from Kaggle website. The CEPR dataset consist of the list of 34,661 Amazon Electronic consumer products reviews specifically Kindle and Fire TV in .csv format. It admits the product details, comments text review, and rate. Likewise, the electronic products and Gift Cards datasets include 2,972 as well as 2,375 reviews correspondingly.

### 3.2. Proposed model

Senti-wordnet and ensemble-based ML techniques are used to create a suggested hybrid ensemble model. The introduced method reaches improved functioning and increased precision are generally explicated by the following two stages:

#### 3.2.1. Lexicon-based technique

This stage involves pre-processing the provided datasets to remove the keywords utilizing tokenization as well as lemmatization for keyword recognition and removal of extraneous stop words, determinants, and prepositions. Using the NLTK package, which is based on NLP, the part-of-speech (POS) tagging is calculated. Finally, using the semantic information gleaned from the senti-wordnet dictionary, the sentiment (SS) and subjectivity (ST) score is being calculated. The computational methods required to calculate the entire SS score as well as ST score, correspondingly, by the content of Amazon review comments are described in (1). The initial split of the dataset is 90:10 to the train and 80:20 for the test dataset. The senti-wordnet dictionary's internal keywords' semantic properties are used to forecast each person's sentiment score. In order to determine the polarity of opinions, such train dataset is utilized to calculate sentiment as well as ST scores as well as to identify subjective statements that have a significant impact on the SS score as a whole. To calculate the SS score and establish the SS polarity of these review condemnations, the training data is employed. In (1) is used to calculate each review sentence's overall sentiment score (SS).

$$SS = \sqrt[3]{\frac{\sum_{w=1}^{n}(p)_w^3 - \sum_{w=1}^{n}(n)_w^3}{n}} \tag{1}$$

Where $p_w$ is the positive SS score as computed from the senti-wordnet dictionary; $n_w$ is the negative SS score as computed from the senti-wordnet dictionary. The overall ST per review sentence can be computed using in (2).

$$ST = \sum_{words=1}^{n} \frac{1-(objective\_score)_{words}}{n} \quad (2)$$

Similar to the objective score, senti-wordnet dictionary's semantic properties are used to calculate the subjective score based on the objective score. In order to determine if the SS connected to each review sentence are good, neutral, or negative, the individual scores are combined to provide an overall sentence SS score.

### 3.2.2. Ensemble learning technique

In this module, it applies an ensemble learning to build a SS predictions. This method decides the SS forecasting on comments employing the Amazon reviews dataset of customer electronic products via calculating the weighted average voting of individual classifier. In the introduced weights are allotted to relevant five classifiers and the weighted average of possibility voting is being calculated as in (3).

$$W = \frac{\mathrm{w1pc1+w2pc2+w3pc3}}{\mathrm{w1+w2+w3}} \quad (3)$$

Where, pc1 represents the predicted possibility decided by SVM classifier (supervised learning), pc2 represents the predicted possibility decided by logistic regression (LR), classifier (supervised learning), pc3 is predicted possibility decided by KNN classifier (unsupervised learning), w1 indicates the weight allotted for LR classifier, w2 denotes the weight allotted for SVM classifier, w3 indicates the weight allotted for KNN classifier.

## 4. RESULTS AND DISCUSSION

The proposed mechanism is employed to calculate SS study for Amazon's customer products approximating Gift Cards, Magazines, Fire TV, and Kindle. The ST and SS scores are calculated in the proposed method and the result received equated with TextBlob library that also offers native API to compute the SS and ST score. TextBlob is accepted python library intended on lexicon-based method and is habitually applied for text mining and treating of the textual data. The experiment is carried out in Google Colab Pro+ which is completely cloud-based platform with NVIDIA Tesla T4 GPUs and 8 GB of high-RAM. The comparison of SS and ST scores are shown in Table 2. The proposed mechanism executes the ensemble learning through compounding three classifiers explicitly LR, SVM, and KNN correspondingly.

Table 2. Comparison of SS/ST score calculated utilizing proposed ensemble model and TextBlob

| Review comments | SS score | ST score | SS score TextBlob | ST score TextBlob |
|---|---|---|---|---|
| Great product for such price | 0.3012 | 0.6531 | 0.6 | 0.4475 |
| I strongly suggest to others | 0.4472 | 0.7343 | 0.7 | 0.81 |
| The tab quitted functioning after functioning for two weeks. | 0.0521 | 0.452 | 0.1853 | 0.3317 |
| Absolute waste of money. It did not join to Wi-Fi at my home. Junk! | -0.0026 | 0.5124 | 0.1853 | 0.3802 |

The weights are currently allotted to the classifiers derived from how well each individual classifier performed, with the best performing classifier receiving more weights. The training dataset's review comments serve as the basis for calculating opinion polarity, which is then used to train the suggested model. Several current common execution measures to evaluate the information recovery method include precision, recall, and F1-score. The computation formulations for these evaluation indices are listed in (3) to (7).

$$Accuracy\ (Acc) = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

$$Precision\ (PR) = \frac{TP}{TP+FP} \quad (5)$$

$$Recall\ (RE) = \frac{TP}{TP+FN} \quad (6)$$

$$F1 - measure = \frac{2 \times PR \times RE}{PR + RE} \quad (7)$$

Where TP is true positives, TN is true negatives, FP is false positives, and FN is false negatives.

Table 3 shows that the F1-score of proposed models, which is based on soft voting, outperforms the current state-of-the-art ensemble model. The "non-interpretative sentiment" is used in the proposed approach to categorize review comments that lack any substantive opinions. For our suggested model to perform better and attain greater accuracy in sentiment classification, a large number of datasets must be used for training. In order to give the sluggish and underachieving learners less weight, we must use ensemble learning with weighted voting. The class label with the highest average possibilities from the group of classifiers is returned by the weighted average.

Table 3. Relative analysis use precision, recall, and F1-score for proposed model

| Dataset | Algorithm | Precision | Recal | F1-score |
|---|---|---|---|---|
| Kindle | Bagging ensemble | 0.63 | 0.51 | 0.5636 |
| | Boosting ensemble | 0.65 | 0.52 | 0.5777 |
| | Proposed ensemble model | 0.81 | 0.52 | 0.6333 |
| Gift-Cards | Bagging ensemble | 0.87 | 0.67 | 0.7570 |
| | Boosting ensemble | 0.74 | 0.59 | 0.6565 |
| | Proposed ensemble model | 0.98 | 0.61 | 0.7519 |
| Electronic products | Bagging ensemble | 0.76 | 0.62 | 0.6828 |
| | Boosting ensemble | 0.79 | 0.61 | 0.6884 |
| | Proposed ensemble model | 0.81 | 0.61 | 0.6959 |

## 4.1. Performance evaluation

The accuracy score shows a slight increase throughout the evaluation of the suggested model's performance in comparison to the current ensemble-based mode. The suggested ensemble model has the best accuracy score of more than 80% across diverse datasets, as shown in the Figure 3. The currently employed ensemble classifier is established on majority voting, and it typically forecasts target class labels based on the mode of anticipated individual class labels.

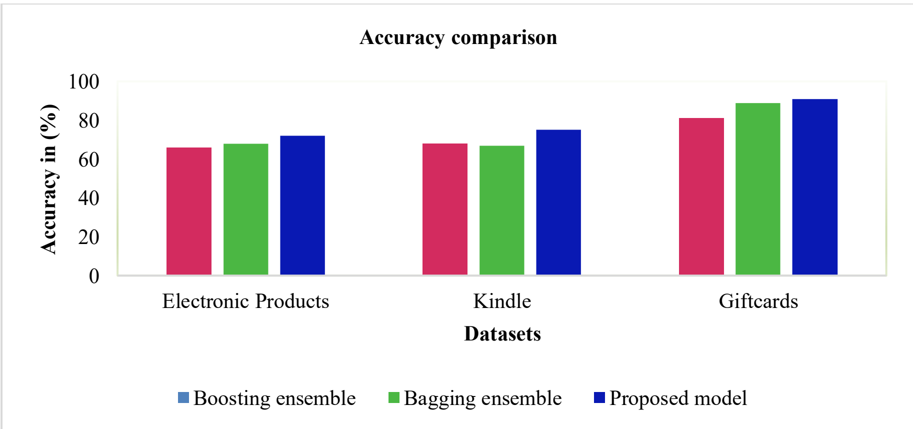


Figure 3. Accuracy of proposed ensemble method for SA trained on Amazon Kindle dataset

As illustrated in Figure 4, the proposed hybrid ensemble method is assessed utilizing the receiving operating characteristic (ROC) curve on many class labels, including positive, neutral, and negative SS as well as non-interpretive sentiment. A possibility-based curve among true positive rate (TPR) and false positive rate (FPR) at different thresholds is known as the ROC curve. The classification system's efficacy is explained by the sharper slope between TPR and FPR.

According to Figure 4, the area under the curve (AUC) values for positive, neutral, negative, and non-interpretive sentiment for the Kindle dataset respectively. The TPR and FPR values between 0.0 to 1.0. Better performance outcomes can be explained by AUC's higher value. Compare to all curve micro average ROC curve has 0.92.

According to Figure 5, the AUC values for positive, neutral, negative, and non-interpretive sentiment for the Gift Cards dataset are 0.92, 0.89, 0.89, and 0.91, respectively. Better performance outcomes can be explained by AUC's higher value. The micro average ROC curve has 0.97 value than all other curves.

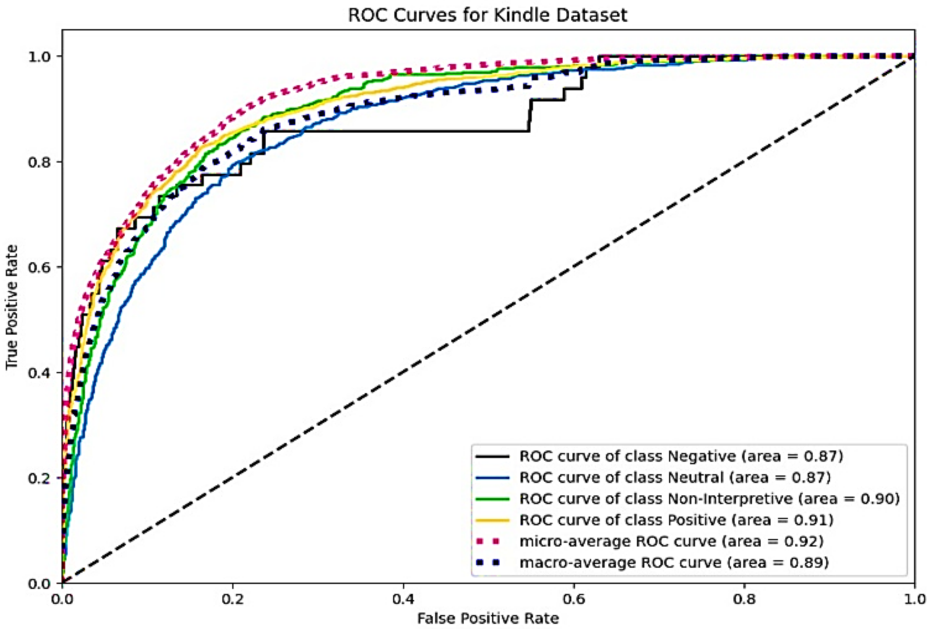


Figure 4. ROC curve of proposed ensembling approach using Kindle dataset and Gift Card dataset

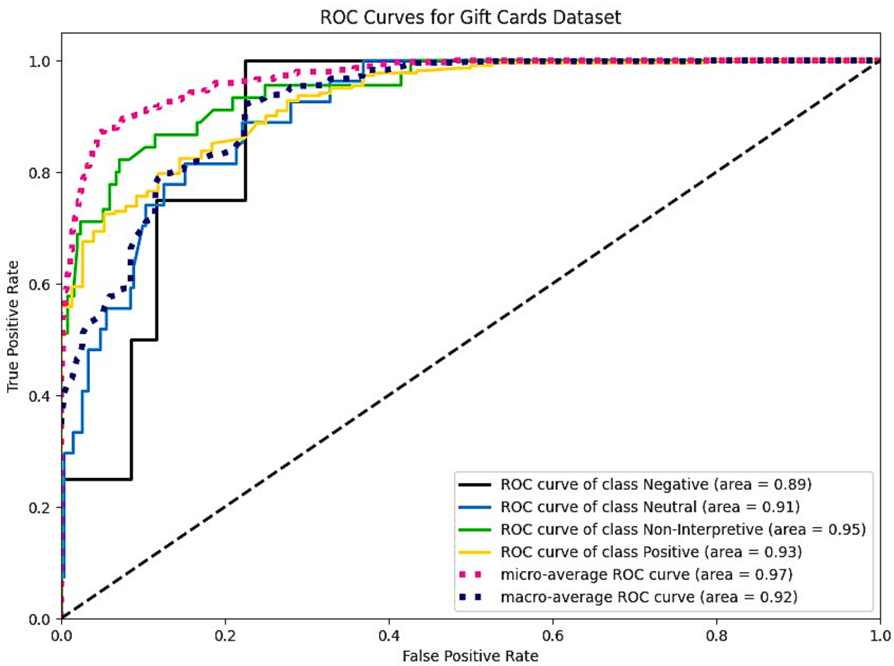


Figure 5. ROC curve of proposed ensemble approach using Gift Card dataset

## 5. CONCLUSION

The purpose of this study is called a hybrid ensemble that combines lexicon and ensemble learning with the use of soft voting and an appropriate distribution of weights to the classifiers. The features are then removed from the review text utilizing term frequency-inverse document frequency vectorizer word embeddings utilizing a mixture of unigram as well as bigram text features, following the initial pre-processing. The subjectivity score, which has been utilized as an additional measure to predict the subjective expressions to identify opinionated material, has been computed using a novel approach that has been proposed in this research. A separate class label called "non-interpretive sentiment" is applied to review comments that lack substantive opinions based on the subjectivity score. The key difficulty in the proposed hybrid ensemble model is figuring out an effective method for giving the best-performing classifier the right weights during the ensemble learning process. We also point out that syntactic relations are important characteristics for sentiment categorization, but they also present a high computational complexity issue. Future research will broaden its focus as we seek to advance enhance the intend of the existing method and develop deep neural networks for handling spam, duplicate content, and fraudulent reviews.

## FUNDING INFORMATION

Funding information is not available.

## AUTHOR CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

| Name of Author | C | M | So | Va | Fo | I | R | D | O | E | Vi | Su | P | Fu |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Palaniraj Rajidurai Parvathy | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ | ✓ | ✓ | | | ✓ | |
| Nagarajan Mohankumar | | ✓ | | | | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | | ✓ |
| Rajendran Shobiga | ✓ | | ✓ | ✓ | | ✓ | | | ✓ | | ✓ | | ✓ | |
| Gour Sundar Mitra Thakur | | ✓ | | ✓ | | ✓ | | ✓ | ✓ | ✓ | | ✓ | | |
| Mamatha Bandaru | | | ✓ | | ✓ | | ✓ | | | ✓ | | ✓ | ✓ | ✓ |
| Velusamy Sujatha | | ✓ | ✓ | | | ✓ | | ✓ | | ✓ | ✓ | ✓ | | |
| Shanmugam Sujatha | | | | | ✓ | | | | | ✓ | ✓ | ✓ | ✓ | |

| | | |
|---|---|---|
| C : **C**onceptualization | I : **I**nvestigation | Vi : **Vi**sualization |
| M : **M**ethodology | R : **R**esources | Su : **Su**pervision |
| So : **So**ftware | D : **D**ata Curation | P : **P**roject administration |
| Va : **Va**lidation | O : Writing - **O**riginal Draft | Fu : **Fu**nding acquisition |
| Fo : **Fo**rmal analysis | E : Writing - Review & **E**diting | |

## CONFLICT OF INTEREST STATEMENT

The authors have no conflict of interest relevant to this paper.

## DATA AVAILABILITY

The data that support the findings of this study are available on request from the corresponding author, [NM]. The data, which contain information that could compromise the privacy of research participants, are not publicly available due to certain restrictions.

## REFERENCES

[1] R. A. Jacobs, M. I. Jordan, S. J. Nowlan, and G. E. Hinton, "Adaptive mixture of local experts," *Neural Computation*, vol. 3, no. 1, pp. 79–87, 1991, doi: 10.1162/neco.1991.3.1.79.

[2] L. Augustyniak, P. Szymański, T. Kajdanowicz, and W. Tuliglowicz, "Comprehensive study on lexicon-based ensemble classification sentiment analysis," *Entropy*, vol. 18, no. 1, 2016, doi: 10.3390/e18010004.

[3] W. Medhat, A. Hassan, and H. Korashy, "Sentiment analysis algorithms and applications: a survey," *Ain Shams Engineering Journal*, vol. 5, no. 4, pp. 1093–1113, 2014, doi: 10.1016/j.asej.2014.04.011.

[4] D. Tiwari, B. Nagpal, B. S. Bhati, A. Mishra, and M. Kumar, "A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques," *Artificial Intelligence Review*, vol. 56, no. 11, pp. 13407–13461, 2023, doi: 10.1007/s10462-023-10472-w.

[5] A. Salih, M. Salih, and A. Abraham, "Novel ensemble decision support and health care monitoring system," *Journal of Network and Innovative Computing*, vol. 2, no. 2014, pp. 41–51, 2014.

[6] R. Xia, C. Zong, and S. Li, "Ensemble of feature sets and classification algorithms for sentiment classification," *Information Sciences*, vol. 181, no. 6, pp. 1138–1152, 2011, doi: 10.1016/j.ins.2010.11.023.

[7] B. Srigiriraju, "Amazon reviews: Kindle store category," *Kaggle*. 2018. Accessed: Jun. 14, 2023. [Online]. Available: https://www.kaggle.com/datasets/bharadwaj6/kindle-reviews

[8] M. Wankhade, A. C. S. Rao, and C. Kulkarni, "A survey on sentiment analysis methods, applications, and challenges," *Artificial Intelligence Review*, vol. 55, no. 7, pp. 5731–5780, 2022, doi: 10.1007/s10462-022-10144-1.

[9] M. B. Imani, M. R. Keyvanpour, and R. Azmi, "A novel embedded feature selection method: A comparative study in the application of text categorization," *Applied Artificial Intelligence*, vol. 27, no. 5, pp. 408–427, 2013, doi: 10.1080/08839514.2013.774211.

[10] T. K. Balaji, C. S. R. Annavarapu, and A. Bablani, "Machine learning algorithms for social media analysis: A survey," *Computer Science Review*, vol. 40, 2021, doi: 10.1016/j.cosrev.2021.100395.

[11] B. Ohana and B. Tierney, "Sentiment classication of reviews using SentiWordNet," in *The 9th. IT & T Conference*, 2009, doi: 10.21427/D77S56.
[12] P. K. Jain, R. Pamula, and G. Srivastava, "A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews," *Computer Science Review*, vol. 41, pp. 1-17, Aug. 2021, doi: 10.1016/j.cosrev.2021.100413.
[13] S. Li, C. Zong, and X. Wang, "Sentiment classification through combining classifiers with multiple feature sets," in *2007 International Conference on Natural Language Processing and Knowledge Engineering*, 2007, pp. 135–140, doi: 10.1109/NLPKE.2007.4368024.
[14] B. H. Juang and S. Katagiri, "Discriminative learning for minimum error classification," *IEEE Transactions on Signal Processing*, vol. 40, no. 12, pp. 3043–3054, 1992, doi: 10.1109/78.175747.
[15] B. Erşahin, Ö. Aktaş, D. Kilinç, and M. Erşahin, "A hybrid sentiment analysis method for Turkish," *Turkish Journal of Electrical Engineering and Computer Sciences*, vol. 27, no. 3, pp. 1780–1793, 2019, doi: 10.3906/elk-1808-189.
[16] O. Coban, B. M. Ozyildirim, and S. A. Ozel, "An empirical study of the extreme learning machine for Twitter sentiment analysis," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 3, no. 6, pp. 178–184, 2018, doi: 10.18201/ijisae.2018644774.
[17] R. Alfrjani, T. Osman, and G. Cosma, "A hybrid semantic knowledgebase-machine learning approach for opinion mining," *Data and Knowledge Engineering*, vol. 121, pp. 88–108, 2019, doi: 10.1016/j.datak.2019.05.002.
[18] P. Bhoir and S. Kolte, "Sentiment analysis of movie reviews using lexicon approach," in *2015 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, Madurai, India, Dec. 2015, pp. 1-6, doi: 10.1109/ICCIC.2015.7435796.
[19] E. Cambria, "An introduction to concept-level sentiment analysis," in *Advances in Soft Computing and Its Applications*, Berlin, Heidelberg: Springer, 2013, pp. 478–483, doi: 10.1007/978-3-642-45111-9_41.
[20] M. D. A. Hasan, K. Balasubadra, G. Vadivel, N. Arunfred, M. V. Ishwarya, and S. Murugan, "IoT-driven image recognition for microplastic analysis in water systems using convolutional neural networks," in *2024 2nd International Conference on Computer, Communication and Control (IC4)*, 2024, pp. 1–6, doi: 10.1109/IC457434.2024.10486490.
[21] K. Ravi and V. Ravi, "A survey on opinion mining and sentiment analysis: Tasks, approaches and applications," *Knowledge-Based Systems*, vol. 89, pp. 14–46, 2015, doi: 10.1016/j.knosys.2015.06.015.
[22] M. Whitehead and L. Yaeger, "Sentiment mining using ensemble classification models," *Innovations and Advances in Computer Sciences and Engineering*, pp. 509–514, 2010, doi: 10.1007/978-90-481-3658-2_89.
[23] R. I. Permatasari, M. A. Fauzi, P. P. Adikara, and E. D. L. Sari, "Twitter sentiment analysis of movie reviews using ensemble features based naïve Bayes," in *2018 International Conference on Sustainable Information Engineering and Technology (SIET)*, 2018, pp. 92–95, doi: 10.1109/SIET.2018.8693195.
[24] H. Rehioui and A. Idrissi, "New clustering algorithms for twitter sentiment analysis," *IEEE Systems Journal*, vol. 14, no. 1, pp. 530–537, 2020, doi: 10.1109/JSYST.2019.2912759.
[25] S. Al-Saqqa, N. Obeid, and A. Awajan, "Sentiment analysis for Arabic text using ensemble learning," in *2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA)*, Aqaba, Jordan, 2018, pp. 1-7, doi: 10.1109/AICCSA.2018.8612804.
[26] R. N. Behera, M. Roy, and S. Dash, "Ensemble based hybrid machine learning approach for sentiment classification – a review," *International Journal of Computer Applications*, vol. 146, no. 6, pp. 31–36, 2016, doi: 10.5120/ijca2016910813.
[27] J. Sadhasivam and R. B. Kalivaradhan, "Sentiment analysis of Amazon products using ensemble machine learning algorithm," *International Journal of Mathematical, Engineering and Management Sciences*, vol. 4, no. 2, pp. 508–520, 2019, doi: 10.33889/ijmems.2019.4.2-041.
[28] K. S. Rajesh, S. Kalaiselvi, P. Sathyanathan, K. Sudha, G. Kavitha, and S. Murugan, "Smart vending machines: Improving service quality with integrated controllers and SVM algorithm," in *2024 6th International Conference on Energy, Power and Environment (ICEPE)*, 2024, pp. 1–6, doi: 10.1109/ICEPE63236.2024.10668878.
[29] A. M. Rajeswari, M. Mahalakshmi, R. Nithyashree, and G. Nalini, "Sentiment analysis for predicting customer reviews using a hybrid approach," in *2020 Advanced Computing and Communication Technologies for High Performance Applications (ACCTHPA)*, 2020, pp. 200–205, doi: 10.1109/ACCTHPA49271.2020.9213236.
[30] P. Maheswari, S. Gowriswari, S. Balasubramani, A. R. Babu, N. K. Jijith, and S. Murugan, "Intelligent headlights for adapting beam patterns with raspberry pi and convolutional neural networks," in *2024 2nd International Conference on Device Intelligence, Computing and Communication Technologies (DICCT)*, 2024, pp. 182–187, doi: 10.1109/DICCT61038.2024.10533159.
[31] S. Sridhar and S. Sanagavarapu, "Sentiment analysis using ensemble-hybrid model with hypernym based feature engineering," in *2020 FORTEI-International Conference on Electrical Engineering (FORTEI-ICEE)*, 2020, pp. 13–18, doi: 10.1109/FORTEI-ICEE50915.2020.9249945.
[32] B. J. Ganesh, P. Vijayan, V. Vaidehi, S. Murugan, R. Meenakshi, and M. Rajmohan, "SVM-based predictive modeling of drowsiness in hospital staff for occupational safety solution via IoT infrastructure," in *2024 2nd International Conference on Computer, Communication and Control (IC4)*, 2024, pp. 1–5, doi: 10.1109/IC457434.2024.10486429.
[33] A. Mukwazvure and K. . Supreethi, "A hybrid approach to sentiment analysis of news comments," in *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO) (Trends and Future Directions)*, 2015, pp. 1–6, doi: 10.1109/ICRITO.2015.7359282.
[34] V. V. Baskar, S. Sekar, K. S. Rajesh, N. C. Sendhilkumar, T. R, and S. Murugan, "Cloud-based decision support systems for securing farm-to-table traceability using IoT and KNN algorithm," in *2024 Second International Conference on Intelligent Cyber Physical Systems and Internet of Things (ICoICI)*, Coimbatore, India, 2024, pp. 443-448, doi: 10.1109/ICoICI62503.2024.10696659.
[35] G. Thahniyath *et al.*, "Cloud based prediction of epileptic seizures using real-time electroencephalograms analysis," *International Journal of Electrical and Computer Engineering (IJECE)*, vol. 14, no. 5, pp. 6047–6056, Oct. 2024, doi: 10.11591/ijece.v14i5.pp6047-6056.
[36] N. Mohankumar *et al.*, "Advancing chronic pain relief cloud-based remote management with machine learning in healthcare," *Indonesian Journal of Electrical Engineering and Computer Science*, vol. 37, no. 2, pp. 1042–1052, 2025, doi: 10.11591/ijeecs.v37.i2.pp1042-1052.
[37] J. Ramasamy, E. Srividhya, V. Vaidehi, S. Vimaladevi, N. Mohankumar, and S. Murugan, "Cloud-enabled isolation forest for anomaly detection in UAV-based power line inspection," in *2024 2nd International Conference on Networking and Communications (ICNWC)*, 2024, pp. 1–6, doi: 10.1109/ICNWC60771.2024.10537407.

[38] J. Jegan, M. R. Suguna, M. Shobana, H. Azath, S. Murugan, and M. Rajmohan, “IoT-enabled black box for driver behavior analysis using cloud computing,” *2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS)*, Chennai, India, 2024, pp. 1-6, doi: 10.1109/ADICS58448.2024.10533471.
[39] A. A. K. Tahir and S. Anghelus, “Improving iris recognition accuracy using gabor kernels with near-horizontal orientations,” *International Journal of Advances in Signal and Image Sciences*, vol. 8, no. 1, pp. 25–39, 2022, doi: 10.29284/ijasis.8.1.2022.25-39.

## BIOGRAPHIES OF AUTHORS

**Mr. Palaniraj Rajidurai Parvathy** is a Project Manager at Mphasis Corporation in Chandler, Arizona, USA. He has 16+ years of IT experience in the BI and analytics domain with a focus on data modeling, integration and visualization (Snowflake, Azure, AWS, GCP, Azure Data Factory, Databricks, Tableau, Power BI, Python, R, SAP BO, Altreyx, Xceptor (RPA)). Rewarded from customer for providing “customer value addition” for performance tuning on Schedule. He has been rewarded the “Star performer” award of the quarter for a support project by Hexaware leadership. Also, he received the “Stat performer” award of the quarter for migration project from Hexaware leadership. Moreover, rewarded “Most valuable player” support project from wipro- best buy account leadership. Furthermore, rewarded a “Feather in my cap” for outstanding contribution to the project business group hierarchy iteration. He was rewarded with a “Feather in my cap” award for his outstanding contribution to Project Business Group Hierarchy Iteration 1. He can be contacted at email: palanirajrps@gmail.com.

**Dr. Nagarajan Mohankumar** was born in India in 1978. He received his B.E. degree from Bharathiyar University, Tamilnadu, India in 2000 and M.E. and Ph.D. degree from Jadavpur University, Kolkata in 2004 & 2010. He joined the Nano Device Simulation Laboratory in 2007 and worked as a senior research fellow under CSIR direct Scheme till September 2009. Later he joined SKP Engineering College as a Professor to develop research activities in the field of VLSI and NANO technology. He is currently working as Research Professor at Symbiosis Institute of Technology, Nagpur Campus, (Symbiosis International Deemed University), Pune, India. He is a senior member of IEEE. He has about 85 international journal publications in reputed journals and about 50 international conference proceedings. He received the carrier award for young teachers (CAYT) from AICTE, New Delhi in the year of 2012-2014. His research interest includes modeling and simulation study of HEMTs, optimization of devices for RF applications and characterization of advanced HEMT architecture, Terahertz electronics, high frequency imaging, sensors and communication. He can be contacted at email: nmkprofessor@gmail.com.

**Ms. Rajendran Shobiga** received her B.E. in electronics and communication engineering from Greentech College of Engineering for Womens (Anna University) in 2014 and her M.E. from Karpagam College of Engineering (Anna University) in 2016. At present; she is working as an Assistant Professor at J. J. College of Engineering and Technology, Trichy. The author has international approved master engineer in communication system by world education services Canada. She has working experience as administrator cum planner in how united services LLC, Doha, Qatar and Planner cum estimator in spectrum cube technical services LLC, Dubai, UAE. She can be contacted at email: shobigaraj123@gmail.com.

**Dr. Gour Sundar Mitra Thakur** is educator, AI enthusiast, researcher, and mentor. He is AI enthusiast with 16 years of teaching and research experience. After pursuing B.Tech. and M.Tech. in computer science and engineering, he completed Ph.D., with the thesis titled “Intelligent systems in manifold decision making” from National Institute of Technology Durgapur, India. Currently, he associated with Dr. B. C. Roy Engineering College Durgapur as Associate Professor in the Department of Computer Science and Engineering (AIML). He was able to publish research findings in 22 international journals, national and international conferences and book chapters till date. He working as a reviewer of many reputed international journals like, Knowledge Based Systems, Applied Soft Computing, International Journal of Fuzzy Systems, and Frontiers in Oncology. He passionate about research and curious about applying different AI/ML/DL algorithms to solve various real-life problems in biology, agriculture, environment and manufacturing industries. He has the ability to collaborate with other experts from different fields to achieve a common goal. He believes in and is committed to develop AI systems that are fair, transparent, and unbiased. He has developed various modern ML/DL models using Tensorflow and Pytorch and developed some expertise in ideating ML pipelines for complex real-life problems. He can be contacted at email: gour.mitrathakur@bcrec.ac.in.

**Ms. Mamatha Bandaru** working as an Assistant Professor in the Department of Electronics and Communication Engineering at the Koneru Lakshmaiah Education Foundation located in Vaddeswaram, Guntur District, Andhra Pradesh. She has 12 years of teaching experience and pursuing her Ph.D. in electronics and communication engineering from VIT-AP. Her research interests include VLSI architecture design, low-power VLSI, digital system design, post-quantum cryptography, machine learning, and artificial intelligence. She is an active life member of the Indian Society for Technical Education (ISTE) and lifetime member of the Institution of Electronics and Telecommunication Engineers (IETE). Additionally, she has participated in numerous training sessions and conferences that focus on advancements in post quantum cryptography, quantum cryptography, machine learning, and artificial intelligence. She can be contacted at email: mamatha.bandaru@gmail.com.

**Dr. Velusamy Sujatha** received her doctoral degree and post-graduation from Anna University, Chennai, and her undergraduate degree from Madras University. Her research focuses on designing charge pump phase-locked loops, with a keen interest in analog VLSI. She began her career as a lecturer in 1993 and has served as an academician in various reputed institutions, holding several positions and guiding numerous UG and PG projects. Currently, she is a Professor in the Department of Electronics and Communication Engineering (ECE) at Shree Sathyam College of Engineering and Technology, Sankari, Tamil Nadu. She can be contacted at email: info@shreesathyam.edu.in.

**Ms. Shanmugam Sujatha** is an Adjunct Professor, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamilnadu, India. She published her research articles in many international and national conferences and journals. Her research areas include network security and machine learning. She can be contacted at email: sujathasmvr@gmail.com.