International Journal of
Computer Science Languages

ISSN: 3048-944X
Volume 4, Issue 1, 2026
January–June
DOI (Journal): 10.37591/IJCSL

https://journals.stmjournals.com/ijcsl

Review IJCSL

Optimized Sentiment Analysis Through TextBlob and Hybrid RNN Models

Pulkit Tiwari[1,*], Bhavna Sharma[2]

Abstract

In today's world, analyzing people's feelings from what they write online has become very important. This is because there is a large amount of content created by users. To make this analysis accurate and fast, we present a method. This method uses a mix of two approaches: one that looks up words in a dictionary and another that uses computer learning. TextBlob is an affordable tool for getting an initial idea of how people feel. However, it sometimes misses complex emotions because it uses simple rules. To fix this, our system combines TextBlob with a special kind of computer learning model. This model incorporates two kinds of recurrent neural networks: long short-term memory (LSTM) and gated recurrent unit (GRU) architectures. These networks are good at understanding text because they can look at the sentence, not just individual words. When we prepare our data, we use TextBlob to get an idea of how positive or negative the text's. We then use this information to help our model understand the context better. By doing this, our model becomes more accurate and faster to train. We tested our approach. Found that it works well, especially when we use certain types of computer learning embeddings like LSTM, BiLSTM, and GRU. Our method provides a balance between being accurate and not taking too much computer power. This means it can handle a lot of data and still give results. Overall, our approach is a solution for analyzing people's feelings from what they write online. It helps us understand and interpret volumes of user-generated content more effectively. Sentiment analysis has become essential for understanding and interpreting volumes of user-generated content. The proposed system combines TextBlob with a deep learning architecture built on LSTM and GRU networks for sentiment analysis. A sentiment analysis framework that integrates lexicon-based methods with deep learning techniques offers a balanced solution that improves generalization while maintaining computational efficiency.

Keywords: Sentiment analysis, TextBlob, NLP, opinion mining, recurrent neural network (RNN), deep learning models, feature embedding

INTRODUCTION

The generation and proliferation of user-generated content through various online sites, such as social networks, blogs, forums, review sites, and news portals, has experienced an unprecedented explosion in recent years. Understanding and analyzing this unstructured textual data has become increasingly important for academic research as well as commercial applications since people's opinions, thoughts, and feelings are freely available on the Web. Sentiment analysis (SA), also referred to as opinion mining, is an important area that has emerged for this requirement. It is a computerized treatment of subjective data to determine polarity, such as positive, negative, or neutral, on the sentiment that a piece of text expresses.

***Author for Correspondence**
Pulkit Tiwari
E-mail: vhans2011@gmail.com

[1]Student, Department of Computer Science and Engineering, JECRC University, Jaipur, Rajasthan, India
[2]Associate Professor, Department of Computer Science and Engineering, JECRC University, Jaipur, Rajasthan, India

Received Date: March 18, 2025
Accepted Date: March 31, 2026
Published Date: April 27, 2026

Citation: Pulkit Tiwari, Bhavna Sharma. Optimized Sentiment Analysis through TextBlob and Hybrid RNN Models. International Journal of Computer Science Languages. 2026; 4(1): 29–38p.

SA is a crossover between computational linguistics, machine learning (ML), and natural language processing (NLP). Its uses span consumer feedback systems and recommendation engines for product and service evaluations, brand monitoring, and election outcome projection. Conventional SA techniques mostly depend on established lexicons or rule-based models, such as TextBlob, which examine sentiment using dictionaries of words with given polarity scores. Although these lexicon-based instruments provide simplicity and explainability, they may find it difficult to grasp complicated verbal events such as sarcasm, irony, slang, or context-dependent emotion changes.

Researchers have been looking more and more at deep learning methods, especially recurrent neural networks (RNNs) and their derivatives, long short-term memory (LSTM), and gated recurrent unit (GRU), to solve the limits of lexicon-based SA (Figure 1). These models represent long-term relationships in text and fit for processing sequential data. Though they outperform conventional ML techniques, RNN-based models are not without flaws; however, among them are high computing cost, overfitting risk, and susceptibility to data noise.

LITERATURE REVIEW

SA, often referred to 575 as opinion mining, has been an active research area in the field of NLP and has gained significant interest in investigating public opinions, consumer feedback, and social media texts. In the last decade, several approaches, from rule-based lexicon-based methods to deep learning algorithms, have been developed for accurate sentiment classification. This literature review aims to provide a summarized guide of previous works that address either one of the following: (1) Lexicon-based SA with TextBlob, (2) Deep learning RNN models, and (3) Hybrid or ensemble methods of sentiment classification.

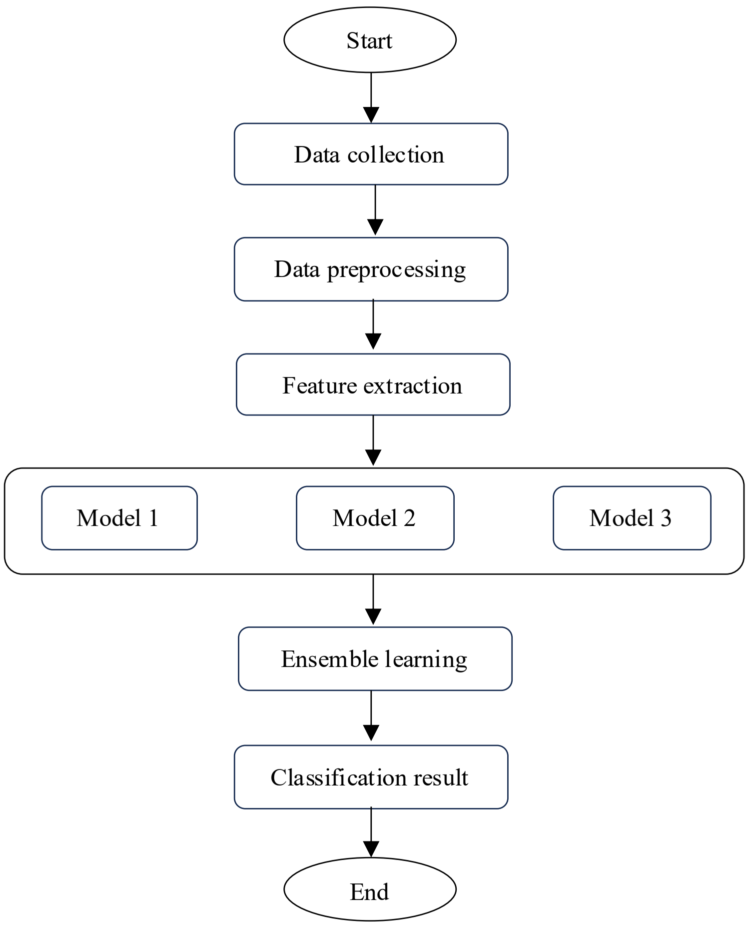

Figure 1. Workflow of the hybrid sentiment classification system.

Lexicon-Based Sentiment Analysis with TextBlob

Lexicon-based SA involves calculating sentiment polarity based on predefined dictionaries of opinionated terms. TextBlob, a Python library for text processing, is widely used because of its simplicity and ease of integration. It leverages a combination of Naïve Bayes and pattern-based analysis, providing sentiment polarity (ranging from -1 to 1) and subjectivity scores.

- They showed an example with TextBlob for movie review SA that obtained reasonable accuracy with the least amount of computation.
- Tested TextBlob with Twitter datasets and concluded that it was appropriate for short and informal texts. However, the drawback is that it works poorly on sentences that are context-dependent and sarcastic.
- Combining TextBlob with word embeddings or neural networks is suggested for better performance in classifying domain texts.

Deep Learning Approaches in Sentiment Classification

SA has been highly transformed by deep learning, especially RNNs, which can learn a long-term dependency between the words in a sentence. Models such as LSTM and GRU are used to overcome the vanishing gradient challenge in RNN, and understandably model complex language aspects.

- LSTM networks were applied to IMDB movie reviews, outperforming traditional models such as SVM and Naïve Bayes.
- Highlighted that GRUs provide comparable performance to LSTMs while being computationally lighter.
- They emphasized that bidirectional LSTMs capture context from both past and future tokens, improving the classification of nuanced sentiments.

Despite their good performance, RNN-based models require large training data and are time-consuming to train. Hence, adding them along with lightweight models, such as TextBlob, can offer an efficient performance tradeoff.

Hybrid and Ensemble-Based Sentiment Classification Models

Hybrid solutions also mix lexicon-based systems with ML or deep learning systems to enhance the accuracy and robustness of the system. In addition, the ensemble approach of combining several models has been particularly successful in mitigating overfitting and improving generalization.

- A hybrid approach of the SVM classifier and TextBlob, achieving better precision in classifying small texts.
- Developed a CNN-LSTM ensemble for Twitter SA and demonstrated superior F1 scores compared to single models.
- Presented a hybrid ensemble that combined BERT embeddings with LSTM and TextBlob output through majority voting, which achieved competitive performance on customer feedback datasets.

These results validate the ensemble approach, particularly for deep models and light NLP models such as TextBlob, which have proven effective for efficient and accurate sentiment classification.

Summary of Related Works

The hybrid lexicon and transformer-based SA method uses a domain-specific educational sentiment lexicon together with TF-IDF and BERT-based features to capture sentiment in student feedback, including automatically transcribed spoken feedback. This is the best-precedence method for automatic intrusion detection compared with prior ML in educational evaluation situations; however, it is slowed down by domain-specific needs and numerous computing requirements to perform real-time analysis [1].

Harris Hawks optimization combined with ML was also applied to the Kaggle Twitter US Airline Sentiment dataset, using hybrid TF-IDF and FastText for feature representation with optimal feature selection by HHO. Classification via ETC reached an accuracy of 96.35%, but the high similarity with feature selection and sensitivity to dataset specifics decreased operational manageability [2].

A BERT-based fake review detection model uses the BERT base-uncased model in combination with TF-IDF and word frequency features on IMDB movie reviews to obtain a classification accuracy of 93%. The advantage is that large multi-label datasets are relatively large-scale subsets of existing annotated images with a large number of tags [3].

The hybrid machine learning-based SA model works for Roman-Urdu and English tweets, in which subjectivity and polarity features are extracted, and NB, RF, DT, and SVM classifiers are utilized to achieve 94% accuracy. Its performance is very sensitive to extensive preprocessing, which aims to compensate for multilingual noise and does not have immediate cross-lingual or cross-platform transferability [4].

ABSA is a task of SA in which the target is to detect and interpret sentiments from reviews, extract feature-based aspects, and propose context-aware languages that use BERT and LSTM architectures to capture the feelings present in Amazon and TripAdvisor reviews with accuracies between 91% and 93%. This is based on sophisticated aspect extraction and domain-based lexicons, which create more computational overhead and annotation efforts [5].

Ensemble classifiers, applied to Amazon product review analysis, combined with the outputs of the different ML classifiers using TF-IDF features, yielded better accuracy than individual models. The ensemble approach requires higher computational resources and has greater system complexity, making it difficult to apply in large-scale real-world applications. SA techniques applied to services such as Amazon, TripAdvisor, Yelp, and similar services using reviews are surveyed, considering a range of deep learning models from BERT to LSTM.

SVM with reported accuracy of up to 99.4%. The accuracy and generalizability of these survey results depend heavily on the underlying data and design decisions, making them difficult to apply directly for new purposes [6].

Surveys from the NLP perspective for review analysis include emerging techniques such as BERT, LSTM, and ensemble classifiers applied to Amazon or Yelp data; they often claim up to 99.3% accuracy. However, these results are subject to context-specific biases, and differences in methodology from each study also potentially reduce comparability and overall validity [7].

Hybrid ML models for Roman-Urdu and English Twitter SA also yielded a 94% accuracy employing NB, RF, DT, and SVM classifiers, which indicates that thorough preprocessing is crucial towards successful cross-lingual sentiment extraction [8].

A similar model for IMDb movie review classification using recurrent BERT is also near state-of-the-art, with 93% accuracy; however, it has issues in domain adaptation and requires large amounts of labeled data to work across multiple domains [9].

Research Gap and Motivation

Although deep learning (DL)-based models, such as LSTM and GRU, have achieved high accuracy in sentiment classification, they tend to be computationally expensive and rely on large amounts of labeled data. It is like I want comprehension over speed, so TextBlob applies but does not work... The fusion of two approaches via ensemble learning has rarely been investigated for leveraging their cooperative mechanisms.

Thus, this study proposes a hybrid sentiment classification method that integrates TextBlob-derived sentiment features with deep learning models (LSTM and GRU) using an ensemble strategy, aiming for an optimal balance between efficiency and accuracy [10].

METHODOLOGY

For each word in the input text, TextBlob looks up the corresponding polarity and subjectivity scores in its lexicon. The overall sentiment of the text is then calculated by aggregating these individual word scores, which are calculated in two ways:

- *Polarity calculation:* The polarity of a sentence or text is generally computed as the **average** of the polarity scores of all sentiment-bearing words within it. Let sentence T consist of N words, $w_1, w_2, \ldots, w_N$. If only a subset of these words, that is, $w_{s1}, w_{s2}, \ldots, w_{sk}$, have sentiment scores in the lexicon, then the overall polarity P_T is calculated as:

$$P_T = \frac{1}{k}\sum_{i=1}^{k} pol(w_{si}) \tag{1}$$

- *Subjectivity calculation:* Similarly, the subjectivity of a sentence or text is computed as the average of the subjectivity scores of all sentiment-bearing words.

$$S_T = \frac{1}{k}\sum_{i=1}^{k} subj(w_{si}) \tag{2}$$

This study presents a basic innovation in merging conventional SA using TextBlob with a neural network model, including GRU and bidirectional long short-term memory (bi-LSTM). Later, the hybrid of both NN model techniques increased the sentiment categorization accuracy. The performance of the hybrid model was assessed holistically and compared with that of conventional models. This creative assessment clearly shows the useful advantages of the integrated strategy, as the hybrid model greatly improves the classification accuracy [11–15].

Algorithm Composition

1. Download the reviews data set and import it into your program environment.
2. Retrieve the reviews for the two products we want to compare.
3. Transform the dataset as follows:
   a. Create a new target variable, target_rating, with reviews. rating indicating ‘positive’ or ‘negative.’
   b. Fill null values on a column basis if required.
4. Inspect the data and compare the number of reviews for each product, review counts, rating count percentage, and positives/negatives.
5. Convert the text of reviews into a format for NLP:
   a. Change the text to lower-case.
   b. Eliminate digits and special characters.
   c. Remove stop words based on the NLTK dictionary.
   d. Tokenize and lemmatize the words to convert the reviews into a standard lexicon.
6. Use the TextBlob algorithm for SA.
7. Divide the dataset and scores into train-test sets using an appropriate technique, e.g., k-fold cross-validation or holdout validation.
8. Train the RNN (LSTM, BiLSTM, GRU) model.
9. Implement a new embedded model that sees the dataset from two sources (word vectorized and sentiment scores) and gives one classification output.
10. Analysis of the model performance with accuracy, precision, recall, F1-score, and confusion matrix performance of the model calculated through the following formula:
    i. Accuracy = (TP + TN)/(TP + TN + FP + FN)
    ii. Precision= TP/(TP + FP)
    iii. Recall = TP/(TP + FN)
    iv. F1-score = 2*Precision*Recall/(Precision+Recall)

11. The confusion matrix is a summary of the model's performance, which displays the count of true positive, false negative, false positive, and true negative predictions [16–20].
12. Compare the model's performance.

Flow Layout

Starting with loading a dataset of product reviews, the suggested layout in Figure 2 is to be specified, in which fundamental and NLP transformations follow in extracting and preparing the reviews. The first SA was performed using TextBlob; TF-IDF vectorization transforms text into numerical values. The training and testing sets were then separated. Using the training data and sentiment ratings, deep learning models such as LSTM, BiLSTM, and GRU were either independently or in a hybrid manner. The performance is assessed to complete the process, and attitudes on the test data are projected using the trained model [21–26].

RESULT

The many scaling techniques used in this work are assessed in the outcome part, along with the sentiment ratings obtained using TextBlob on the RNN algorithm. Therefore, the efficiency of any approach for precisely translating SA is under discussion.

Figure 3 indicates the sentiment categories of positive, negative, and neutral types. In the positive sentiment, the overwhelmingly dominant sentiment, represented by the largest segment of the pie chart in teal color, accounted for 93.33% of the total, whereas in neutral sentiment, a smaller portion of approximately 4.33% is shown in light blue/cyan, with negative sentiment at 2.34%.

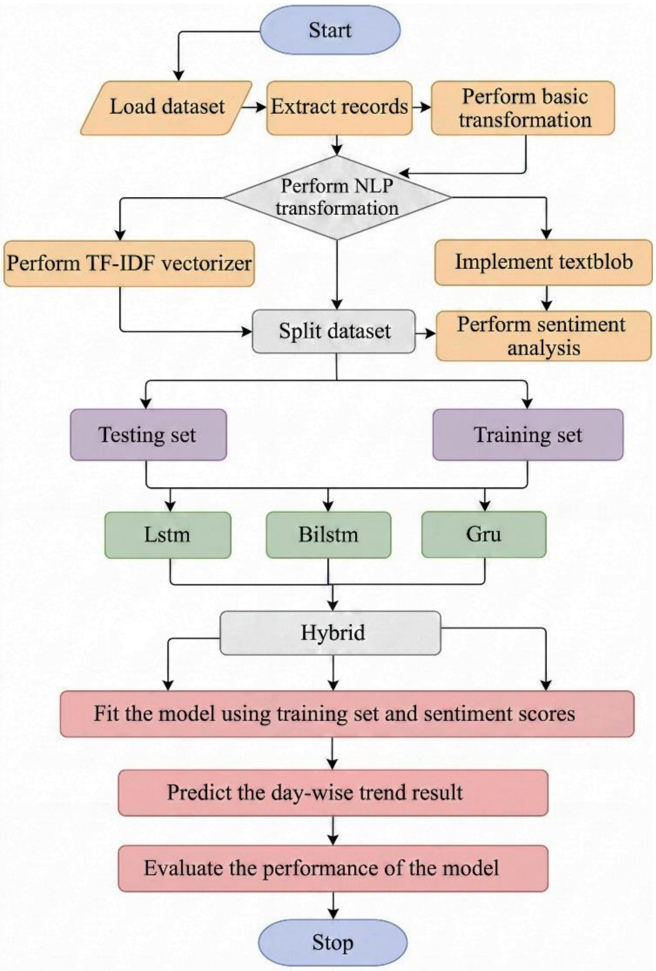

Figure 2. Proposed flow layout. Sentiment-based prediction workflow.

In Figure 4, the bar charts indicate the frequency of ASINs (Amazon Standard Identification Numbers). One chat showed the raw count of each ASIN, and others indicated a highly skewed distribution, where some product frequencies were transformed by a base-10 logarithm.

Figure 5 shows the comparison of products with a median sentiment of approximately 0.3, with the majority of sentiment scores falling between approximately 0.25 and 0.5. Another product shows a slightly higher median sentiment of 0.4, and its interquartile range is also slightly higher, between roughly 0.3 and 0.5.

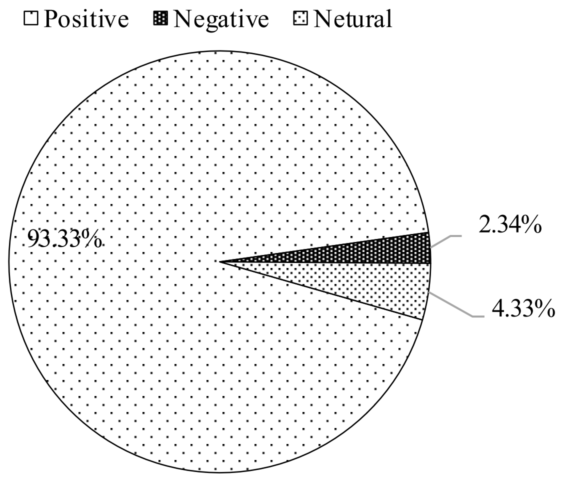

Figure 3. Distribution of sentiments.

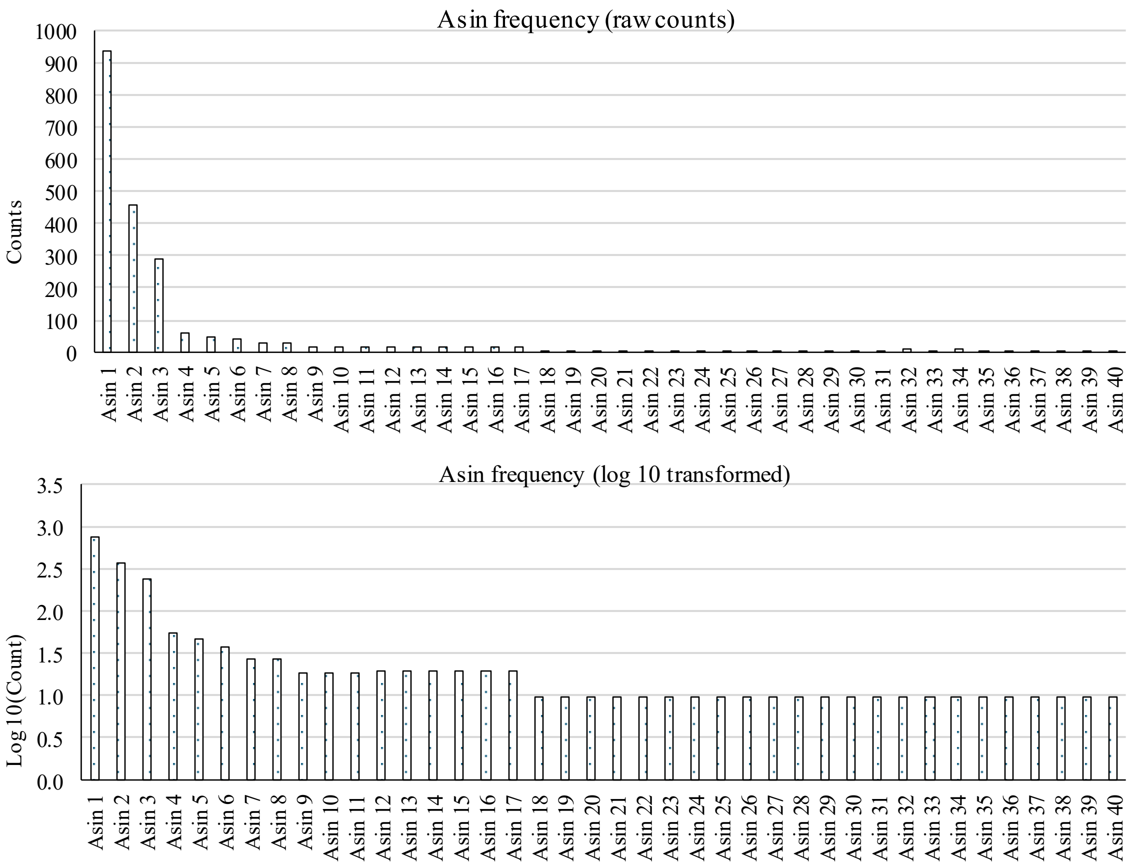

Figure 4. ASIN frequency.

In Figure 6, a histogram with a kernel density estimate (KDE) shows the spread of sentiment scores.

A comparison of the results of all ML models is presented in Table 1. The output is based on their training and validation accuracies. The embedded model optimization demonstrated the highest performance in both training and validation, suggesting that it generalizes best to unseen data.

Table 1. Model comparison.

Models	Train accuracy	Validation accuracy
LSTM	0.818	0.816
BiLSTM	0.833	0.806
GRU	0.803	0.76.6
Embedded	0.874	0.881

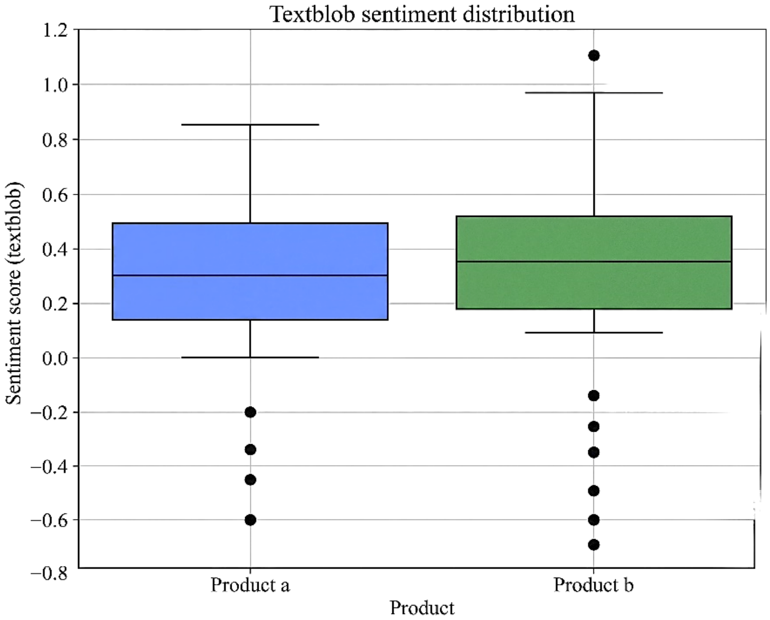

Figure 5. TextBlob sentiment distribution.

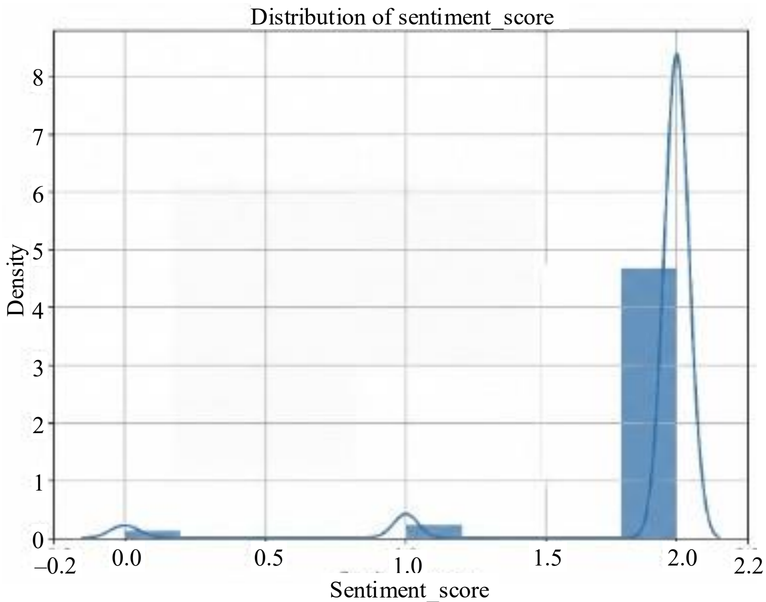

Figure 6. Distribution of sentiment score.

CONCLUSION

The results of this study emphasize the importance of proper scaling techniques for enhancing the reliability and validity of sentiment-based rating schemes. By selecting the appropriate scaling approach, businesses and platforms can gain a much-improved understanding or utilization of sentiment data to make informed decisions, tailor marketing strategies, and enhance consumer experiences. The planted model had an accuracy of approximately 87%, which is significantly superior to that of previous models.

In addition, further work could focus on extending the dataset to multilingual reviews and cross-cultural contexts to examine the scalability and generalization of the proposed scaling methods globally.

REFERENCES

1. Maria H, Subhashni R. Hybrid lexicon and transformer-based sentiment analysis of student feedback for faculty evaluation: A speech-to-text approach. Commun Appl Nonlinear Anal. 2025;32(2):610–622.
2. Zahraa Z, Attia H, Ali YH. Optimized and comprehensive fake review detection based on Harris Hawks optimization integrated with machine learning techniques. J Cybersecurity Inf Manag. 2025;15(1):11–21. doi:10.54216/JCIM.150102.
3. Tayal N, Singh M, Raj M. A BERT-based technique on IMDb for false movie review detection. J Electr Syst. 2024;20(10s):5322–5333.
4. Malik F, Fuzail M, Aslam N, Sarwar R, Abid K, Maqbool MS, et al. A hybrid machine learning model to predict sentiment analysis on X. J Comput Biomed Inform. 2024;6(2):64–79.
5. Sablok B, Tyagi A, Dharmander D, Malik MHS, Chaturvedi DS, Sureshwati D. NLP models for sentiment analysis of Twitter data. Int J Res Appl Sci Eng Technol. 2025;13(4):3544–3552. doi:10.22214/ijraset.2025.68682.
6. İsenkul ME. Optimizing the sentiment recognition in Spotify playlists through ensemble-based approaches. Acta Infol. 2025;9:74–89. doi:10.26650/acin.1604272.
7. Mamillapalli A, Ogunleye B, Timoteo Inacio ST, Shobayo O. GRUvader: Sentiment-informed stock market prediction. Mathematics. 2024;12(23):3801. doi:10.3390/math12233801.
8. Malik N, Bilal M. Natural language processing for analyzing online customer reviews: A survey, taxonomy, and open research challenges. PeerJ Comput Sci. 2024;10:e2203. doi:10.7717/peerj-cs.2203. PubMed PMID: 39145232.
9. Ashbaugh L, Zhang Y. A comparative study of sentiment analysis on customer reviews using machine learning and deep learning. Computers. 2024;13:340. doi:10.3390/computers13120340.
10. Alsubaie MS, Atawneh SH, Abual-Rub MS. Building machine learning model with hybrid feature selection technique for keylogger detection. Int J Adv Soft Comput Appl. 2023 Jul;15(2). doi:10.15849/IJASCA.230720.03.
11. Elsadig MA, Gafar A. An ensemble model to detect packet length covert channels. Int J Electr Comput Eng. 2023;13(5):5296–5304. doi:10.11591/ijece.v13i5.pp5296-5304.
12. Sharma P, Sharma A. Online K-means clustering with adaptive dual cost functions. 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT); Kannur, India. 2017. p. 793–799. doi:10.1109/ICICICT1.2017.8342665.
13. Sharma A, Sharma A. KNN-DBSCAN: Using k-nearest neighbor information for parameter-free density based clustering. 2017 International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT); Kannur, India. 2017. p. 787–792. doi:10.1109/ICICICT1.2017.8342664.
14. Magade K, Sharma A. Significant role of IoT in cyber-physical systems, context awareness, and ambient intelligence. In: Hussain A, Obaid AJ, Tyagi G, Sharma A, editors. The Next Generation Innovation in IoT and Cloud Computing with Applications. Boca Raton: CRC Press; 2024. p. 16–34. doi:10.1201/9781003406723-2.
15. Kumar V, Gupta SK, Hussain A, Sharma A. A systematic approach to prevent threats using IDS in IoT-based devices. GMSARN Int J. 2025;19(1):107–112.

16. Hussain A, Obaid AJ, Tyagi G, Sharma A. The Next Generation Innovation in IoT and Cloud Computing with Applications. Boca Raton (FL): CRC Press; 2024. doi:10.1201/9781003406723.
17. Garg P, Sharma A. A distributed algorithm for local decision of cluster heads in wireless sensor networks. In: 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI); Chennai, India. 2017. p. 2411–2415. doi:10.1109/ICPCSI.2017.8392150.
18. Hammood BAK, Sadiq AT. Ensemble machine learning approach for IoT intrusion detection systems. Iraqi J Comput Inform. 2023;49(2):93–99. doi:10.25195/ijci.v49i2.458.
19. Ding W, Popoola S. An ensemble learning voting technique for network intrusion detection systems [Preprint]. 2023. Research Square. doi:10.21203/rs.3.rs-2989954/v1.
20. Afrifa S, Varadarajan V, Appiahene P, Zhang T, Domfeh EA. Ensemble machine learning techniques for accurate and efficient detection of botnet attacks in connected computers. Eng. 2023;4(1):650–664. doi:10.3390/eng4010039.
21. Titl V, Mazrekaj D, Schiltz F. Identifying politically connected firms: A machine learning approach. Oxf Bull Econ Stat. 2024;86(1):137–155. doi:10.1111/obes.12586.
22. Afriyie JK, Tawiah K, Pels WA, Addai-Henne S, Dwamena HA, Owiredu EO, et al. A supervised machine learning algorithm for detecting and predicting fraud in credit card transactions. Decis Anal J. 2023;6:100163. doi:10.1016/j.dajour.2023.100163.
23. Talman A, Apidianaki M, Chatzikyriakidis S, Tiedemann J. How does data corruption affect natural language understanding models? A study on GLUE datasets. Proceedings of the 11th Joint Conference on Lexical and Computational Semantics; Seattle, WA, USA. Association for Computational Linguistics; 2022. p. 226–233. doi:10.18653/v1/2022.starsem-1.20.
24. El Alami H, Najid A. Energy-efficient fuzzy logic cluster head selection in wireless sensor networks. In: 2016 International Conference on Information Technology for Organizations Development (IT4OD); 2016. IEEE; 2016. p. 1–7. doi:10.1109/IT4OD.2016.7479300.
25. Kaissar A, Nassif AB, Injadat M. A survey on network intrusion detection using convolutional neural network. ITM Web Conf. 2022;43:01003. doi:10.1051/itmconf/20224301003.
26. Yang D, Mai Ngoc KM, Shin I, Lee KH, Hwang M. Ensemble-based out-of-distribution detection. Electronics. 2021;10(5):567. doi:10.3390/electronics10050567.

