IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15454

A COMPARATIVE STUDY OF GENRE-BASED SENTIMENT ANALYSIS

Dr. Angelpreethi A[1], Hepshiba Sherly P A[2], P Anitha[3]

Assistant Professor, Department of Computer Science, St.Joseph's College (Autonomous), Affliated to Bharathidasan University, Tiruchirappalli[1]

PG student, St.Joseph's College (Autonomous), Affliated to Bharathidasan University, Tiruchirappalli[2]

Assistant Professor, Department of Computer Science,

Dr.Umayal Ramanathan College for Women, Affiliated to Alagappa University, Karaikudi[3]

Abstract: The emergence of multiple online platforms has generated a huge volume of user-generated movie reviews from various sources like IMDb, Kaggle repositories, and Twitter. Due to the unstructured nature and diversity in using language, manual analysis is hard to conduct. Sentiment analysis hence serves as an effective means of automatic opinion polarity detection from textual data. However, sentiment expression in movie reviews is pretty much influenced by genre and dataset characteristics. This paper hence represents a comparison study of genre-oriented sentiment analysis based on both the lexicon-based approach and the machine learning-based approach. In this context, sentiment classification tasks are performed on a number of movie genres such as Naïve Bayes, Support Vector Machine, and Random Forest. The performances of these approaches are compared across a number of datasets, and the result reports that machine learning-based methods usually tend to gain higher accuracy, particularly in informal social media data.

Keywords: Sentiment Analysis, Movie Reviews, Genre-Wise Analysis, Machine Learning, LexiconBased Approach, IMDb, Twitter.

I. INTRODUCTION

The necessity of sentiment analysis in movie reviews lies in its ability to automatically extract and summarize public opinion from large volumes of user-generated content, enabling audiences to make informed viewing decisions and helping filmmakers, producers, and platforms understand viewer preferences, emotional responses, and overall reception of films. By classifying reviews into positive, negative, or neutral sentiments, sentiment analysis reduces the time and effort required to manually interpret thousands of reviews, identifies trends and patterns in audience feedback, and supports datadriven decision-making in marketing, recommendation systems, and content improvement, thereby playing a crucial role in enhancing both user experience and the film industry's strategic planning.

The rapid growth of the internet or various online media has created a huge number in usergenerated data in the form of text information of movies, comments, and opinions from various sites like IMDb, repositories from Kaggle, tweets from Twitter, etc., which are useful to understand audience perception, though it becomes challenging to analyze huge volumes of unstructured information manually. In this context, automated algorithms to analyze sentiment are very important to seek significant insights from online reviews of movies [1].

In the movie domain, sentiment expressions may vary regarding genre. Words that are appearing negative in general contexts might convey positive sentiments when genres like horror or thriller are concerned, and romance and comedy genres often tend to exhibit clear sentiment polarity. Moreover, the performance of sentiment classification also varies with datasets due to variations in review length, structure of language, and formality of speech [2]. This paper presents a comparative study in genre-wise sentiment analysis over IMDb, Kaggle, and Twitter movie review datasets by applying lexicon-based and machine learning-based approaches

II. LITERATURE REVIEW

A study by Kouloumpis, E [3] explore and improve sentiment analysis specifically for Twitter by identifying the unique challenges posed by informal language, abbreviations, emoticons, and hashtags, and to evaluate the effectiveness of different linguistic features and machine learning techniques in accurately classifying tweet sentiments, thereby determining which features contribute most to better performance in social media sentiment analysis. Another study by

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.15454

A. Angelpreethi and S. B. Ramesh Kumar [4] focuses on improving the accuracy of opinion mining on big data using a dictionary-based approach. The research emphasizes the importance of structured sentiment lexicons in effectively processing large-scale textual data. The study highlights that lexicon-based techniques can serve as a strong foundation for sentiment classification tasks. Although the work primarily deals with big data, its relevance to movie review sentiment analysis is significant, especially when working with structured datasets such as IMDb.

The authors Kolchyna, et. al., [5] investigate and compare different approaches for sentiment analysis on Twitter data, specifically lexicon-based methods, machine learning techniques, and a hybrid combination of both, in order to evaluate their effectiveness and accuracy in classifying tweet sentiments and to determine whether combining these approaches can improve overall performance compared to using each method individually. In a survey conducted by Kharde and Sonawane [6] regarding various techniques for Twitter sentiment analysis, it has been found that machine learning classifier techniques are often superior to lexicon-based techniques for the same purpose.

A study by A. Angelpreethi and S. Britto Ramesh Kumar. in [7] has suggested an advanced approach for opinion mining based on feature extraction, giving more importance to the extraction of domain-specific features and relating them to expressions containing sentiment. The significance of the above-mentioned study lies in the fact that it gives importance to the extraction of features based on context, which helps in improving the accuracy of classification of sentiments. Even though the abovementioned study is mainly focused on product review datasets, the relevance of the study to genre-based sentiment analysis in the movie industry is quite high.

The authors of [8] have examined the sentiment from IMDb Movie Reviews dataset to find the polarity of the movie reviews on a scale of 0 (highly disliked) to 4 (highly liked). Then they have performed feature extraction, followed by training a multilabel classifier to classify the reviews into its correct label. They have obtained an accuracy of around 88.95%. The objective of the paper by A. Angelpreethi and S. B. Ramesh Kumar [9] is to analyze the issues and challenges involved in visualizing large-scale data in big data mining. It also aims to highlight opportunities for improving data understanding and decision-making through effective visualization techniques.

A study by R. Merlin Packiam and Sinthu Janita Prakash [10] is to propose a multilevel sparse dimension selection approach using taxonomy to improve the efficiency of big data processing by reducing data dimensionality and enhancing the relevance of selected features for better performance and accuracy. The paper by A. Angelpreethi and S. Britto Ramesh Kumar [11] is to develop a hybrid approach for opinion mining in big data by integrating different techniques to enhance the accuracy and efficiency of sentiment analysis, while also addressing challenges such as data volume, variety, and complexity in large-scale datasets. A study by R. Merlin Packiam and V. Sinthu Janita Prakash [12] is to propose a relevance-based feature selection algorithm using Hidden Markov Models (HMM) to improve the preprocessing of textual data by selecting the most significant features, thereby enhancing the efficiency and accuracy of text mining and classification tasks.

The paper by A. Angelpreethi and S. B. Ramesh Kumar [13] introduces NIC_LBA, a lexiconbased approach designed to handle negations and intensifiers in microblog data. The study aims to improve sentiment classification accuracy by effectively capturing contextual modifiers that influence sentiment polarity. It highlights that ignoring negation and intensity can lead to incorrect sentiment interpretation. This approach is particularly useful for analyzing Twitter data, where informal language and expressive patterns are commonly used. The authors of [14] have applied sentiment analysis on IMDb

Movie Reviews Dataset in which, they have applied various steps of text processing and feature selection, followed by classifying them into positive or negative reviews. Furthermore, they have evaluated the model with eight different classifiers using five different evaluation metrics. A study by A. Angelpreethi and S. B. Ramesh Kumar [15] is to introduce Dom_Classi, an enhanced weighting mechanism for domain-specific words using frequency-based probability. It aims to improve the accuracy of text classification by assigning more effective weights to important domain-related terms. The paper by Noura O. Aljehane [16] is to apply Long Short-Term Memory (LSTM) neural networks for Twitter sentiment analysis, aiming to improve the accuracy of classifying tweet sentiments by effectively capturing sequential patterns and contextual information in textual data.

Another study by M. Gayathri, S. Shajun Nisha, and M. Mohamad Sathik [17] is to provide a comprehensive survey of Twitter sentiment analysis by reviewing various existing techniques, tools, and approaches, and to highlight their strengths, limitations, and challenges in order to guide future research in this field. D. O. Ratmana, [18] is to analyze and evaluate various feature selection techniques used in movie review sentiment analysis. The study focuses on how different methods impact the performance of classification models. It aims to identify the most relevant features that contribute to

Impact Factor 8.471 ⪤ Peer-reviewed & Refereed journal ⪤ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15454

accurate sentiment prediction. By comparing multiple techniques, the paper seeks to improve the efficiency and effectiveness of sentiment classification systems. The authors of [19] have applied four different classification algorithms on IMDb movie reviews dataset and evaluated their performance using Accuracy Score, FI Score and *AVC* Score.

A study by A. Angelpreethi [20] presents a comparative analysis of lexicon-based and machine learning techniques for sentiment classification. The research aims to evaluate the effectiveness of both approaches in identifying user opinions. The findings indicate that machine learning techniques outperform lexicon-based methods in capturing contextual relationships and improving classification accuracy. This supports the use of machine learning models in this study for better sentiment prediction.

The paper by F. Sayeedunnisa [21] is to enhance sentiment analysis of social media data by incorporating slang words and emoticons as important features. It aims to improve the accuracy of sentiment classification by capturing informal language patterns commonly used in online communication. The study emphasizes the role of non-standard expressions in better understanding user opinions.

Angelpreethi et al. [22] proposed a fuzzy-based sentiment classification approach using fuzzy linguistic hedges to improve decision-making. The study aims to handle uncertainty and vagueness in textual data more effectively, thereby enhancing the accuracy and reliability of sentiment analysis results. The paper by A. S. Safitri, et al. [23] is to improve the accuracy of sentiment analysis models by applying effective preprocessing techniques such as data cleaning, normalization, and feature refinement. The study focuses on how these preprocessing steps enhance the quality of textual data and lead to better classification performance. C. G. Maurya and S. K. Jha [24] is to develop a hybrid approach for sentiment analysis on Twitter data by combining multiple techniques to improve classification accuracy. The study aims to leverage the strengths of different methods to handle the complexity and variability of social media text, thereby enhancing overall performance in sentiment prediction. The study by A. Angelpreethi et al. [25] proposes a hybrid predictive modelling approach that integrates machine learning and deep learning techniques. The objective of the study is to enhance the performance of predictive models when dealing with complex and large-scale datasets. The results demonstrate improved accuracy and efficiency compared to traditional methods. This approach is highly relevant to the current research, as it supports multi-source sentiment analysis using datasets such as IMDb, Kaggle, and Twitter.

The study by S. Parimala et al. [26] is to develop an efficient and accurate method for colorectal cancer classification from biomedical images by integrating an improved Moth Flame Optimization (MFO) algorithm with a Deep Convolutional Neural Network (DCNN). The study aims to enhance feature selection and optimize model parameters using the modified MFO technique, thereby improving the performance of the DCNN in terms of classification accuracy, robustness, and computational efficiency, ultimately supporting early and reliable detection of colorectal cancer. The authors of [27] developed FUDL-SM, a hybrid fuzzy deep learning framework for sentiment analysis of social media data related to zoological and wildlife conservation. The study aims to improve classification accuracy by combining fuzzy logic with deep learning techniques to effectively handle uncertainty and domainspecific data.

The paper by P. Anitha [28] is to study a hybrid approach for Twitter opinion mining by combining lexicon-based and machine learning methods to improve sentiment classification accuracy. The research aims to leverage the strengths of both techniques for better analysis of social media data and more effective opinion extraction. The paper by A. Angelpreethi, [29] is to develop DeepPre-OM, an enhanced preprocessing framework for opinion classification of microblog data, aiming to improve data quality and boost sentiment classification accuracy by effectively handling noise, informal language, and irrelevant features in social media text. The paper by A. Angelpreethi [30, 32] is to propose OPINE_NEG, an approach for detecting negations and intensifiers in social media data to improve sentiment analysis accuracy and to analyze students' feedback on online education using advanced deep learning models such as BiLSTM with attention and Transformer architectures which aims to improve sentiment classification accuracy by effectively capturing contextual and sequential patterns in textual data. The paper by M. Soni, [31] is to perform a comparative study of different techniques used in movie review sentiment analysis. The study aims to evaluate and compare various methods to determine their effectiveness in accurately classifying sentiments and improving overall performance.

Though there have been several research studies on sentiment analysis with particular datasets, little work has been done on exploring genre-based sentiment analysis with multiple databases. Therefore, this study has been designed to perform a comparative analysis of genre-based sentiment classification with IMDb, Kaggle, and Twitter datasets.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15454

III. METHODOLOGY

The architecture for the genre-wise sentiment analysis system is given in the following figure 1.

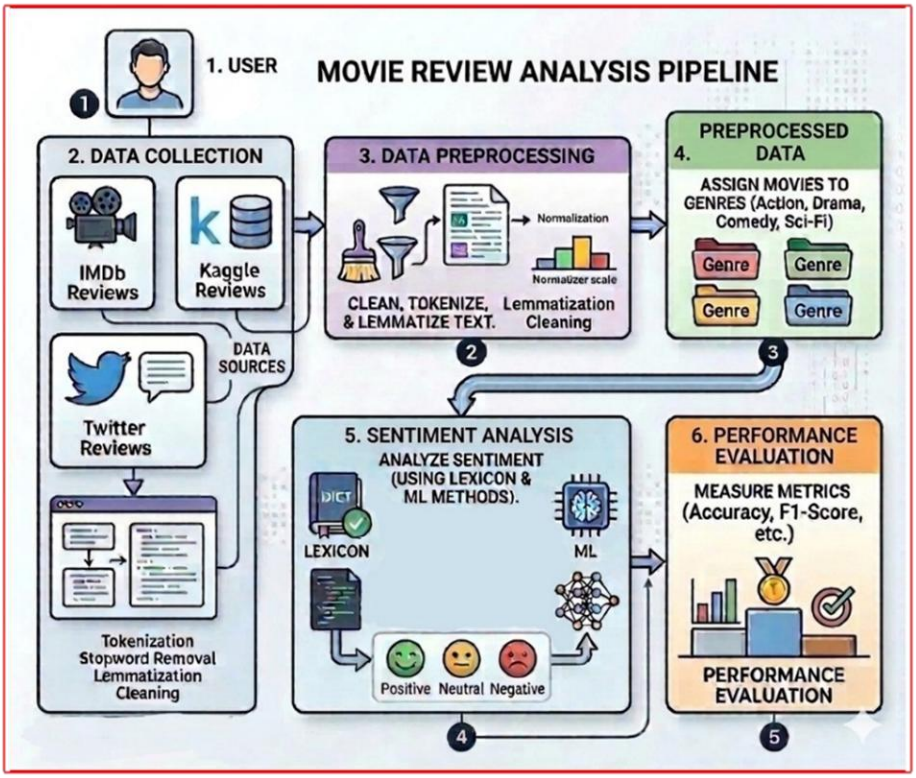

Figure 1. Architecture of the proposed genre-wise sentiment analysis system

1. DATASETS USED

Three distinct datasets are used in the proposed work to enable cross-platform comparison analysis.

1. **The IMDb Dataset -** The IMDb dataset is a compilation of official movie reviews that include user-submitted text comments and rating data. The reviews are usually lengthy and grammatically appropriate. Additionally, movie genre information is included, enabling genrebased classification. IMDb is a standard dataset for sentiment analysis tasks due to its wellorganised format and thorough reviews.
2. **The Kaggle Movie Review Dataset -**A collection of labelled movie reviews created especially for sentiment analysis tasks is known as the Kaggle movie review dataset. Both good and negative samples are equally distributed across the rather lengthy assessments.
3. **Twitter Dataset -**The Twitter dataset is a collection of short text movie-related tweets gathered by exploiting movie-related hashtags. In contrast to the IMDb and Kaggle datasets, the Twitter dataset is very informal, noisy, and contains abbreviations, emojis, slang, and hashtags. The Twitter dataset allows the assessment of the effectiveness of sentiment classification in a realtime social media setting.

Using the above three datasets, it is possible to compare the effectiveness of sentiment classification on Structured long text data (IMDb), Curated labeled data (Kaggle) and Informal short text social media data (Twitter) in the following figure 2.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15454

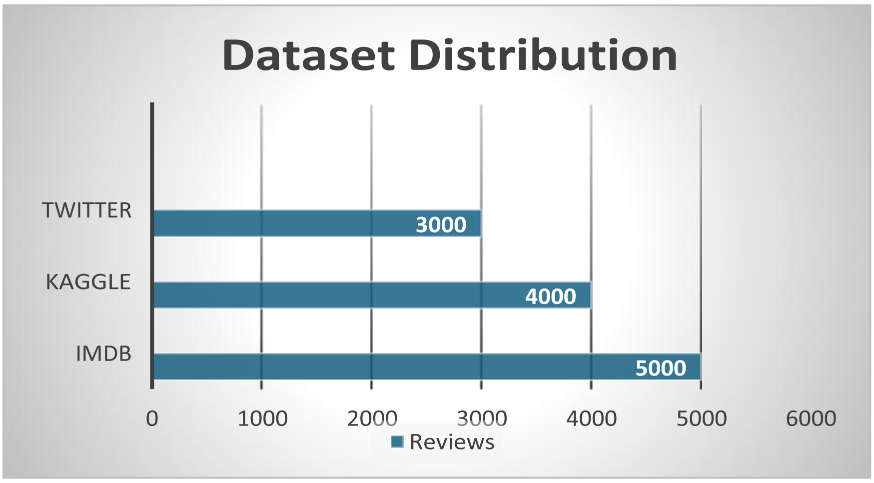

Figure 2. Distribution of movie reviews across datasets

B. GENRE CLASSIFICATION

For genre-level sentiment analysis, reviews were classified based on movie genres such as Action, Romance, Comedy, Horror and Thriller. Genre information was obtained either from the metadata of the dataset or from the corresponding movie information and each review was classified into its respective genre category.

Genre classification is an important task because sentiment analysis varies from genre to genre. For instance:

- The word “dark” or “violent” may be used to express negative sentiments in general environments, but it may be used to express positive sentiments in horror or thriller genres.
- The romance and comedy genres may contain explicit expressions of positive or negative sentiments.

C. DATA PREPROCESSING

Text preprocessing is an important task in the overall sentiment analysis task that helps in the removal of noise and improvement of quality of the features. The following text processing tasks were performed on all the datasets equally:

1. **Removal of URLs and HTML Tags:** The removal of URLs, HTML tags, and links was performed to remove unnecessary tokens.
2. **Remove special characters and punctuation**: To make sure the text data was all the same, we got rid of the special characters, emojis (if they weren't processed separately), and punctuation marks
3. **Making it lowercase:** To make sure that there are no duplicate tokens based on case sensitivity (like "Good" and "good"), all the texts were changed to lowercase.
4. **Tokenisation:** The texts were broken up into separate words or tokens.
   For example, "The movie was very good" would become ["movie", "extremely", "good"].
5. **Deleting Stop Words:** The stop words "is," "the," "and," and "was" were taken out because they don't matter as much to the sentiment's polarity.
6. **Lemmatisation:** Changing each word to its base or root word. "Loved" and "Loving" both mean "Love." This is done to make sure that dimensionality reduction and feature
7. **Negation Handling:** The negation words “not” were handled carefully to ensure polarity.
   Example: “not good” should not be classified as positive.

These steps can potentially help in noise reduction and enhancement of feature quality.

D. SENTIMENT ANALYSIS TECHNIQUES

In the research study, two major approaches were employed to achieve the task of genre-based sentiment classification: (1) Lexicon-Based Approach, and (2) Machine Learning-Based Approach. The approaches were individually applied on each of the datasets (IMDb, Kaggle, and Twitter).

1) Lexicon Based

The lexicon-based approach identifies the polarity of the sentiment using a pre-defined sentiment dictionary where words are labeled with fixed polarity scores that denote positive, negative, or neutral sentiments. In this research study, after the application of preprocessing operations like tokenization, stop word removal, and lemmatization, the review was broken down into individual tokens. The tokens were then searched against the sentiment lexicon to fetch their respective polarity scores. The net sentiment score of a review is calculated by aggregating the polarity scores of all the sentiment-bearing words that are present in the text. The net sentiment score of a review is mathematically calculated as:

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026
DOI: 10.17148/IJARCCE.2026.15454

$$Sentiment\ Score = \sum_{i=1}^{n} Score\ (w_i)$$

where $w_i$denotes each word in the review and $Score(w_i)$represents the assigned polarity value of that word. The final sentiment classification was determined based on the cumulative score: if the total score was greater than zero, the review was classified as positive; if less than zero, it was classified as negative; and if equal to zero, it was considered neutral.

To make the classification more accurate, the negation handling feature was introduced to reverse the meaning of words that appear after the negation words "not," "never," or "no." This is to make sure that the classification is correct for phrases like "not good" or "not bad." The lexicon-based approach doesn't need any training data and can be used as a baseline for comparison. However, it only looks at static polarity assignment and doesn't take into account changes in genre or context. In genre-based sentiment analysis, the meaning of words may vary based on the type of movie. In horror or thriller movies, the words "dark" or "intense" could mean something good, but in other types of movies, they could mean something bad. Fixed lexical scoring is useful in analyzing how well it works on platforms such as IMDb, Kaggle, and Twitter.

2) Machine Learning-Based Approach

To overcome the drawbacks of context in the lexicon-based approach, supervised machine learning algorithms were employed for sentiment classification. The TF-IDF feature extraction technique was used to turn the text data from the reviews into numbers. This technique gives more weight to words that are more common in the document but less common in the corpus. This helps the algorithms that sort things out find the traits that are related to the sentiment's polarity.

The data was divided into two groups: one for training and one for testing. The machine learning algorithms used the training set to learn and the testing set to test. Unlike the lexicon-based approach, machine learning algorithms possess the capability to learn from the context and patterns in the data. Three algorithms were employed in this research

i. Naïve Bayes

Naïve Bayes is a probabilistic classifier based on Bayes' Theorem:

$$P(\ C \mid X\ ) = \frac{P(\ X \mid C\ )P(C)}{P(X)}$$

where C represents the sentiment class and X denotes the feature vector corresponding to a review. The classifier computes the posterior probability of each class based on the features it sees and chooses the class with the highest probability. Naïve Bayes classifiers handle high-dimensional text data effectively despite the fact that it assumes the features are conditionally independent. It is also a good starting point for sentiment analysis tasks.

ii. Support Vector Machine (SVM)

Support Vector Machine is a type of classifier that finds the best hyperplane that separates positive and negative classes in a feature space with many dimensions. The decision function is defined as

$$f(x) = w^T x + b$$

where w denotes the weight vector, x represents the feature vector, and $b$is the bias term. SVM aims to maximize the margin between classes, thus improving the generalization capability. Because of its success in dealing with sparse and high-dimensional text features, it can be used on structured movie reviews such as IMDb and well-designed datasets such as Kaggle.

iii. Random Forest

Random Forest is an ensemble learning algorithm that grows a forest of decision trees on random samples of the training data and features. The overall sentiment classification result is obtained by majority voting of the individual decision trees. Random Forest works very well on noisy and unstructured datasets such as Twitter, where linguistic and short-text features may influence the correctness of the result of the classification task.

E. PERFORMANCE EVALUATION

The performance of both the lexicon-based model and the machine learning-based model for sentiment classification was measured using the standard classification performance metrics to ensure a fair comparison. As the problem under

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15454

consideration is a binary classification problem (positive and negative sentiments), the performance metrics were calculated using the confusion matrix, which holds True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). These performance metrics help in a systematic evaluation of the performance of the classification model in identifying the polarity of the sentiment correctly for the IMDb, Kaggle, and Twitter datasets.

Accuracy was the main way to measure performance to make sure the classification model was correct. Accuracy is the ratio of the number of correctly classified instances to the total number of instances predicted by the classification model. The formula for calculating accuracy is given as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

However, accuracy is used to verify the performance of the classification model, but it is not the bestperforming measure to be used when the classification problem is imbalanced.

Precision is the measure of the proportion of correctly classified positive instances out of all
instances classified as positive. Precision is a measure of the correctness of the positive classifications made by the classifier and is calculated as:

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

A high precision value indicates that the classifier makes fewer errors of the type false positive, which is an important aspect of genre-specific sentiment analysis, where incorrect polarity assignments can affect comparative analysis.

Recall, also known as sensitivity, is the measure of the proportion of actual positive instances that were correctly identified by the classifier. Recall is a function of the ability of the model to detect instances of relevant sentiment and is calculated as:

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

A high recall value will indicate that the model is able to detect most of the positive reviews in a genre with fewer errors of type false negative.

To address the problem of precision and recall, the F1-score was employed. The F1-score is the harmonic mean of precision and recall. It is a comprehensive measure of performance that takes into account both false positives and false negatives. It is calculated as:

$$F1 - Score = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The F1-score is very helpful in measuring the performance of sentiment classification on different genres, as there might be an unequal distribution of sentiments or context variations in some genres.

In the present study, the performance analysis is conducted at three different levels: dataset level, genre level, and algorithm level. The performance analysis conducted at the dataset level compares the performance of the sentiment classifiers on structured data (IMDb), curated data (Kaggle), and unstructured data (Twitter). Genre-level analysis analyzes the differences in the interpretation of sentiments on different movie genres like Action, Romance, Horror, Comedy, and Thriller. Algorithmlevel analysis determines the best-performing classification algorithm on all datasets and genres.

The multi-level performance analysis approach adopted in the present study enables a comprehensive insight into the effect of the nature of the datasets and genre-specific vocabulary on the performance of the sentiment classification task. The present study enables a more accurate and statistically valid comparison of the sentiment analysis approaches by adopting multiple performance measures instead of accuracy.

IJARCCE
ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15454

IV. RESULTS

This section shows the findings that resulted from applying the lexicon-based and machine learning-based approaches to IMDb, Kaggle, and Twitter movie review dataset. The effectiveness of these approaches is analyzed genre-wise based on various metrics such as accuracy, precision, recall, and F1-score. A comparative study is conducted to see how genre and other attributes of a dataset affect the overall performance.

A. DATASET-WISE PERFORMANCE ANALYSIS

Generally speaking, the performance of machine learning-based classifiers surpasses the lexiconbased method in all three datasets. Since IMDb reviews are of a well-structured and long nature, they provide the highest accuracy, while tweets have relatively lower accuracy because of informal language, abbreviations, and noise. Kaggle datasets provide moderately improved performance due to the presence of curated and labeled data. The datasetwise accuracy comparison is given in the table 1 and figure 3.

Table 1: Dataset-wise Accuracy Comparison

Dataset	Naïve Bayes (%)	SVM (%)	Random Forest (%)
IMDb	90	93	91
Kaggle	88	90	89
Twitter	84	86	85

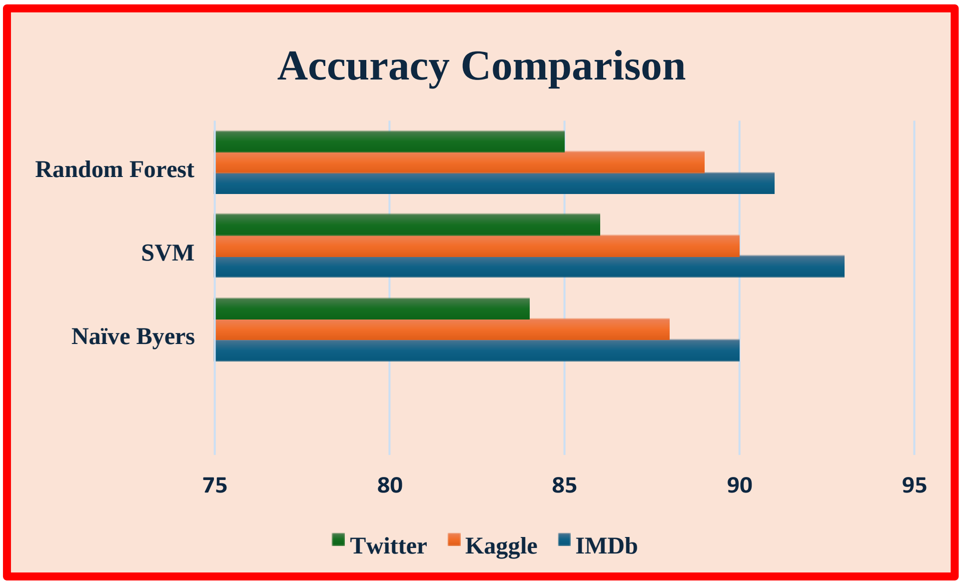

Figure 3. Accuracy comparison of sentiment analysis techniques across datasets

B. GENRE-WISE PERFORMANCE ANALYSIS

Genre-wise analysis of sentiment classification accuracy reveals that the accuracy of sentiment classification varies depending upon different genres of films. Romance and Comedy genres attain relatively greater accuracy owing to increased clarity in sentiment expressions. Horror and Thriller movies attain lower accuracy through techniques like Lexical Analysis, where words that reflect fear/violence do not necessarily reflect negative sentiment.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.471 ⁙ Peer-reviewed & Refereed journal ⁙ Vol. 15, Issue 4, April 2026

DOI: 10.17148/IJARCCE.2026.15454

Table 2: Genre-wise Best Performing Algorithm

Genre	IMDb	Kaggle	Twitter
Action	SVM	SVM	RF
Romance	SVM	NB	SVM
Comedy	RF	SVM	NB
Horror	RF	RF	SVM
Thriller	SVM	RF	RF

C. COMPARATIVE DISCUSSION

While it is true that lexicon-based methods do a reasonably good job on structured datasets, they fail when sentiment interpretation is genre-dependent or the content in social media is quite informal. Machine learning classifiers are more adaptable towards both dataset and genre variations; among them, Support Vector Machine achieved higher accuracy continuously in most genres and datasets. Twitter remains the most challenging data because it's noisy and context-dependent, making traditional sentiment analysis not perform well on runtime data. Clearly, the performance of sentiment classification can be significantly improved in different datasets using genre information and machine learning-based methods.

V. CONCLUSION

The study presents a comprehensive comparative analysis of genre-wise sentiment classification using movie review datasets from IMDb, Kaggle, and Twitter. Both lexicon-based and machine learningbased approaches were systematically evaluated across multiple genres, revealing that the performance of sentiment classification models is significantly influenced by the nature of the dataset as well as genrespecific characteristics. The results demonstrate that machine learning classifiers consistently outperform lexicon-based methods, particularly when applied to well-structured datasets such as IMDb, where the textual content is more formal and less noisy. In contrast, sentiment analysis on Twitter data remains more challenging due to the presence of informal language, abbreviations, slang, and high levels of noise, which negatively impact classification accuracy. Furthermore, the findings highlight that different movie genres exhibit distinct linguistic patterns and emotional expressions, which affect the effectiveness of sentiment classification models. Therefore, incorporating genre-specific features and adopting hybrid or advanced learning approaches can further enhance performance. Overall, this study underscores the importance of considering both dataset type and genre variations to improve the reliability and accuracy of sentiment analysis systems.

REFERENCES

[1] Bo Pang, & Lillian Lee. (2008). Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, *2*(1–2), 1–135.

[2] Alec Go, Richa Bhayani, & Lei Huang. (2009). *Twitter sentiment classification using distant supervision*. Stanford University.

[3] Efthymios Kouloumpis, Theresa Wilson, & Johanna D. Moore. (2011). Twitter sentiment analysis: The good, the bad and the OMG! In *Proceedings of the International AAAI Conference on Weblogs and Social Media* (pp. 538–541).

[4] Angelpreethi, A., & Ramesh Kumar, S. B. (2018). Dictionary based approach for improving the accuracy of opinion mining on big data. *International Journal of Research and Analytical Reviews, 5(4), 1836–1844.* https://ijrar.com/upload_issue/ijrar_issue_20542657.pdf

DOI: 10.17148/IJARCCE.2026.15454

[5] Olga Kolchyna, Tiago T. P. Souza, Philip Treleaven, & Tomaso Aste. (2015). *Twitter sentiment analysis: Lexicon method, machine learning method and their combination*. arXiv. https://arxiv.org/abs/1507.00955

[6] V. Kharde, & S. S. Sonawane. (2016). Sentiment analysis of Twitter data: A survey of techniques. *International Journal of Computer Applications, 139*(11).

[7] A. Angelpreethi, & S. Britto Ramesh Kumar. (2016). An enhanced architecture for feature-based opinion mining from product reviews. In *Proceedings of the IEEE World Conference on Computing and Communication Technologies (WCCCT)*. 10.1109/WCCCT.2016.30

[8] T. P. Sahu, & S. Ahuja. (2016). Sentiment analysis of movie reviews: A study on feature selection & classification algorithms. In *International Conference on Microelectronics, Computing and Communications*.

[9] A. Angelpreethi, & S. B. Ramesh Kumar. (2016). Visualizing big data mining: Issues, challenges and opportunities. *International Journal of Control Theory and Applications, 9*(27), 455–460. https://doi.org/10.5281/zenodo.18215005

[10] R. Merlin Packiam, & Sinthu Janita Prakash. (2017). Multilevel sparse dimension selection approach for improved big data processing using taxonomy.

[11] A. Angelpreethi, & S. Britto Ramesh Kumar. (2018). Opinion mining using hybrid approach on big data: A perspective. *International Journal of Scientific Research in Computer Science Applications and Management Studies, 7*(5). https://doi.org/10.5281/zenodo.17399553

[12] R. Merlin Packiam, & V. Sinthu Janita Prakash. (2019). Relevance-based feature selection algorithm for efficient preprocessing of textual data using HMM. *International Journal of Computer Sciences and Engineering, 7*(1), 15–21.

[13] Angelpreethi, A., & Ramesh Kumar, S. B. (2019). NIC_LBA: Negations and intensifier classification of microblog data using lexicon based approach. *Journal of Emerging Technologies and Innovative Research, 6(6), 753–759.* https://www.jetir.org/papers/JETIR1906R05.pdf

[14] M. Yasen, & S. Tedmori. (2019). Movies reviews sentiment analysis and classification. In *IEEE Jordan International Joint Conference on Electrical Engineering and Information Technology (JEEIT)* (pp. 860–865).

[15] A. Angelpreethi, & S. B. Ramesh Kumar. (2019). Dom_Classi: An enhanced weighting mechanism for domain-specific words using frequency-based probability. *International Journal of Applied Engineering Research, 14*(1), 140–148. https://doi.org/10.5281/zenodo.17280663

[16] Noura O. Aljehane. (2020). Twitter sentiment analysis using LSTM neural network. *International Journal of Advanced Computer Science and Applications, 11*(1), 345–351.

[17] M. Gayathri, S. Shajun Nisha, & M. Mohamad Sathik. (2020). Twitter sentiment analysis: Survey. *International Journal of Engineering Research & Technology, 8*(3). https://doi.org/10.17577/IJERTCONV8IS03029

[18] D. O. Ratmana, G. Fajar Shidik, A. Z. Fanani, Muljono, & R. A. Pramunendar. (2020). Evaluation of feature selections on movie reviews sentiment. In *International Seminar on Application for Technology of Information and Communication (iSemantic)* (pp. 567–571).

[19] S. Tripathi, R. Mehrotra, V. Bansal, & S. Upadhyay. (2020). Analyzing sentiment using IMDb dataset. In *International Conference on Computational Intelligence and Communication Networks (CICN)* (pp. 30–33).

[20] Angelpreethi A. (2023) Lexicon and Machine Learning Based Comparative Analysis to Classify the Students Opinions on Covid-19 Pandemic. *IJIEMR Transactions. 2023;12(2):380–387.*https://ijiemr.org/public/uploads/paper/590451676720775.pdf. 10.48047/IJIEMR/V12/ISSUE02/60.

[21] F. Sayeedunnisa, N. P. Hegde, & K.-U.-R. Khan. (2020). Using slang and emoticon for sentiment analysis of social media data. *International Journal of Engineering Research & Technology, 8*(15).

[22] A. Angelpreethi. (2023). Fuzzy-based sentiment classification using fuzzy linguistic hedges for decision making. *Mapana Journal of Sciences, 22*(Special Issue 2), 63–79. https://doi.org/10.12723/mjs.sp2.4

[23] A. S. Safitri, I. Wijayanto, & S. Hadiyoso. (2023). Improving classification accuracy with preprocessing techniques for sentiment analysis. In *International Conference on Data Science and Its Applications*.

[24] C. G. Maurya, & S. K. Jha. (2024). Sentiment analysis: A hybrid approach on Twitter data. *Procedia Computer Science, 235*, 429–436. https://doi.org/10.1016/j.procs.2024.04.094

[25] Dr A Angelpreethi , Subash K, Lekka P (2025) A Hybrid Predictive Modelling of Environmental Pollutant Concentrations using Machine and Deep Learning Techniques. *Indian Journal of Science and Technology 18(48): 3851-3862*. https://doi.org/ 10.17485/IJST/v18i48.1889

[26] S. Parimala, R. Kumutha, F. Iram, M. V. Sunena Rose, N. R., & A. Angelpreethi. (2024). Improved moth flame optimization with deep convolutional neural network for colorectal cancer classification using biomedical images. In *Second International Conference on Advances in Information*. 10.1109/ICAIT61638.2024.10690631

[27] A. Angelpreethi, P. Anitha, & R. Sivagami. (2025). FuDL-SM: A hybrid fuzzy deep learning framework for sentiment analysis of zoological and wildlife conservation based social media data. *Indian Journal of Science and Technology, 18*(43), 3422–3431. https://doi.org/10.17485/IJST/v18i43.1709

[28] P. Anitha. (2026). A study on hybrid Twitter opinion mining using lexicon and machine learning approaches. *International Journal of Research and Analytical Reviews, 13*(1), 761–766. https://doi.org/10.56975/ijrar.v13i1.326887

[29] A. Angelpreethi, M. L. Priya, & R. Kavitha. (2025). DeepPre-OM: An enhanced preprocessing framework for opinion classification of microblog data. *The Scientific Temper, 16*(12), 5302–5311. https://doi.org/10.58414/SCIENTIFICTEMPER.2025.16.12.17

[30] A. Angelpreethi. (2025). OPINE_NEG: An approach to detecting negations and intensifiers using social media data. *International Multidisciplinary Research Journal Reviews*, *2*(8), 13–17. IMRJR.2025.020803.pdf

[31] M. Soni, N. Kumari, P. Singh, & V. Kashtariya. (2024). Analysis on movie reviews: A comparative study. In *Eighth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 252–257).

[32] A. Angelpreethi, & Karan Marshal J. (2026). Opinion mining of students' feedback on online education using BiLSTM–attention and transformer models. *International Journal of Research and Analytical Reviews, 13*(1), 317–322. IJRAR26A1648.pdf