Engineered Science
DOI: https://dx.doi.org/10.30919/es1391

Machine Learning Techniques to Classify Movies Based on Reviews

Vishva Dalela,[1,#] Kunal Gupta[1,#] and G. Poornalatha[1,*]

Abstract

Movie reviews help viewers decide whether to watch the movie based on detailed reviews with analysis by other viewers and experts. It is an important indicator for the audience as a guiding choice amidst vast entertainment options. This paper aims to analyze movie reviews and examine the accuracies of different machine learning algorithms. The review data is pre-processed to transform into a format suitable to the model. We feed this pre-processed data to various models for the best possible outcome. It was observed that the support vector machine yields good results for the datasets considered. Furthermore, k-fold cross-validation is carried out to compare and check the efficiency of various models.

Keywords: Classifier; Machine learning; Natural language processing; Sentiment analysis; Movie reviews.
Received: 03 July 2024; Revised: 12 November 2024; Accepted: 10 January 2025.
Article type: Research article.

1. Introduction

Sentiment analysis is the use of natural language processing (NLP), text analysis, computation linguistics, and biometrics to systematically identify, extract, quantify, and study affective states and subjective information. It can be categorized into the following: application-oriented, customer care, sentence-level, aspect-level, concept-level, multilingual, and linguistic features analysis. Applications of sentiment analysis include analyzing social network data to understand the views of a given group of people. Companies like Google and Facebook make the most of their revenues from advertisements. They implement advanced sentiment analysis to help them gain meaningful insights and display the most relevant advertisement to their customers, in which the respective persons might be interested. Sentiment analysis also comes in handy for advertisement clicks through the rate optimization. The algorithm analyzes and chooses the best commercial depending on the promotions that the customer clicks on and the responses given by the other users.

Sentiment analysis is gaining more attention with the fast proliferation of existing data comprising opinions, movie reviews, Twitter tweets, product reviews, discussions on social platforms, *etc*.[1] Opinion mining can be used to design recommender systems. It is used in advertisements to identify suitable places. It can be used to predict whether people prefer the same kind of product or switch to some other product. Organizations release beta versions of their products and make them accessible to a few customers who get the version in advance and must provide feedback on the product as they use it. The feedback can be positive or negative. All the feedback provided by the users over the given period is analyzed and placed in different classes based on the type of reviews. These organizations improve features that users did not find satisfactory before releasing them to the public. This helps them to release a better product and increase their overall profit. News channels during elections take opinions from the public on different political parties and forecast the outcome using sentiment analysis. A significant amount of data managed by companies helps them make real-time decisions when analyzed properly.

Anshul *et al*. explored the application of machine learning (ML) algorithms in human activity classification.[2] They focused on the performance evaluation of popularly used ML methods namely k-nearest neighbor (KNN), support vector machine (SVM), logistic regression, *etc*. They observed that the kernel SVM outperformed the other methods. Harsh *et al*. performed sentiment analysis of movie reviews by using ML techniques like logistic regression, Naïve Bayes (NB), and SVM. They obtained the highest accuracy for SVM.[3]

In general, movies are liked by all irrespective of the age. Whenever we watch a movie at a theater or a television series

[1] *Department of Information and Communication Technology, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, 576104, India*
[#] *All authors contributed to this work equally.*
**Email:* poornalatha.g@manipal.edu (G. Poornalatha)

at home, we tend to feel something about the show, good or bad. We can call this a sentiment of an individual attached to a particular item or thing. The present work focuses on inferring whether a movie is "good" or "bad" based on review comments given by viewers. This information could help other interested viewers decide whether to watch a particular movie or not.

Several efforts were made to analyze online opinions based on the review comments given by viewers. A modified Naïve Bayes algorithm was proposed to analyze Twitter sentiments.[4] A sentence-level approach is followed, and a sentiment score is assigned to the tweets for analysis.[5] Different opinions given by the customers on the service provided by the restaurant are analyzed based on feature-based opinion mining.[6] An overview of different approaches used for analyzing the sentiments is presented.[7] The information regarding sentiment analysis from the perception of opinion mining is discussed along with the open research questions.[1] Nagamma *et al*. attempted to analyze the correlation between the movie's revenue and online reviews.[8] They used the fuzzy clustering technique for testing. Wankhede *et al*. designed an approach to classify the movie into three different polarities by inspecting the sentiment in the textual document.[9] Singh *et al*. proposed a feature-based, aspect-level method to investigate movie reviews.[10-11] Tripathi *et al*. used different feature selection techniques to analyze reviews for Indian movies.[12] Sahu *et al*. presented an approach for movie rating that uses feature selection.[13] Yan *et al*. considered various review-related factors like timeliness, length, *etc*.[14] Sunil *et al*. used a neural network structure for sentiment analysis.[15] Wang used ML-based sentiment analysis of movie reviews by considering a case study of "Beyond the Clouds".[16] Suresh *et al*. analyzed movie reviews using ML techniques and observed the highest accuracy for linear SVC.[17]

Much of the work proposed by various authors, as mentioned above, predicts whether a particular movie will be a hit using different parameters. However, the present work aims to provide information regarding whether a movie is "good" or "bad," as determined by reviews.

2. Methodology

Fig. 1 outlines the steps in the current study. Reviews are stored in a database, and are transformed to categorical variables by one-hot encoding. English stop words are removed to eliminate non-informative data, and the Porter algorithm[18] reduces words to their stems, decreasing vocabulary size and improving the results for small datasets.
The term frequency (Tf) denotes the count of a particular term in a document. Inverse document frequency (idf) is calculated by taking the logarithm of the ratio between the total documents in the dataset and the number of documents that contain the word.[19] Count vectorizer tokenizes text data and encodes it into a sparse matrix of word count. In contrast, the Tf-idf vectorizer assesses the importance of words by combining Tf and idf. Principal component analysis (PCA) is an unsupervised algorithm that reduces the number of dimensions by identifying principal components that preserve the required information for further processing. Various ML techniques are applied to the outcome of the PCA.

ML algorithms take a vast amount of labeled data as input and try to analyze it. ML algorithms allow the computer system to learn and improve continuously based on the data. We have chosen five distinct algorithms to analyze movie reviews, each of which is briefly described below. KNN is a classification algorithm that classifies data based on 'k' nearest neighbors. The ANN method is usually used to learn patterns from the given dataset. Random forest is a robust method and generates multiple decision trees during the training period and considers the mean value of individual trees to provide a final prediction. SVM uses hyperplanes to separate data into classes. Naïve Bayes is a well-known probabilistic-based text classifier model having 3 types – Gaussian, Multinomial, and Bernoulli - based on data distribution.

The analysis and findings of our experiment have been run on Lenovo Legion Y520 using Python 3.6.5. The major packages and libraries used are Scikit-learn, Natural Language ToolKit, Keras, Numpy, and Pandas. To evaluate the classifier accuracy, we followed an 80:20 train/test split on the datasets and performed the k-fold validation to verify the results further. K-fold validation is a statistic used to assess models in ML. The method splits the entire dataset into 'k' subsets, or folds. Each iteration uses 'k-1' folds for training, while the remaining fold is used for testing, ensuring that each fold serves as a test set exactly once. This repeated procedure provides a better performance metric and less dependence on the single train-test split, perfect for small datasets where any data usage needs to be maximized.

The k-fold validation is performed to assess the robustness of ML techniques, where k-values vary from 1 to 10. The final metric is the mean value of the k iterations, providing a more consistent and stable estimate of the model's generalization ability. This technique reduces the possibility of model overfitting and is applied to artificial neural networks (ANN), KNN, SVM, random forest, and Naïve Bayes, to ensure reliability. The data subsets of K-Fold Cross-Validation avoid overfitting by cross-validating across different subsets of data, which would prevent overfitting against a specific subset. Implementing this on models, such as ANN, KNN, SVM, random forest, and Naïve Bayes, K-cross-validation can ensure the generalizability of the respective models. Iterative training on different subsets may mean that every model is experimented with varying conditions, promoting generalizability and increasing the likelihood of a measure of reliability and performance as applied to unseen data.

3. Results and discussion

3.1 Data and hyperparameters

The current work considered 3 standard data sets. Dataset v1.0: movie review documents labeled by overall sentiment polarity, and sentences labeled by subjectivity status or polarity. It has

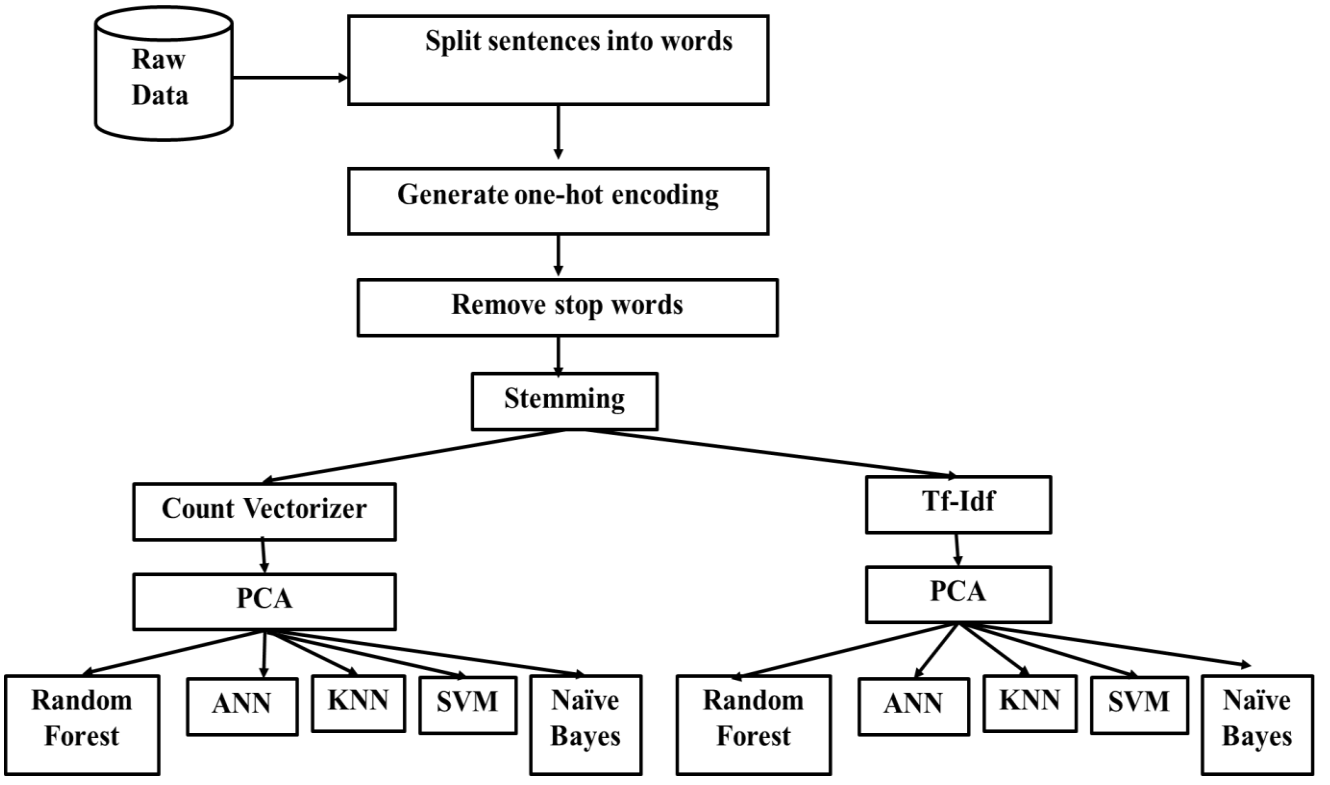

Fig. 1: Flowchart depicting workflow.

an equal number (5331) of positive and negative sentences.[20] Dataset v2.0 is a movie review document with 1000 reviews for each positive and negative, respectively.[19] Dataset v3.0 is a dataset for binary sentiment classification, a set of 25,000 movie reviews for training and 25,000 for testing.[21]

Entropy is used as the criteria for the random classifier. Radial basis function (RBF) is used for kernel SVM with random state set as 1 and degree as 2. The number of neighbors is set to 12, with Minkowski distance as the metric for KNN. 64 neurons (Relu) are set as hidden layers with Adam as an optimizer. The number of epochs is set to 10 with a batch size of 32. ANN kernel initializer is set to "glorot uniform" with Relu as an activation function.

3.2 Data pre-processing

The dataset is initially loaded, and the contents of each file are appended into a list called dataset followed by storing the label of each file in a separate list – labels. The English stop words are removed by using the NLTK library. These words are non-informative and are probably overheads. The Porter stemming algorithm removes all the suffixes from the words. Words that have a common stem usually have the same meanings. It helps to reduce the vocabulary size and thus improves the results for small datasets.

3.3 Scores for count vectorizer

The results of various classification algorithms on the dataset using a count vectorizer are shown in Table 1, Table 2, and Table 3. The settings slightly differ from one another in various versions: version 2 limits the max features to 1,500, while for version 1, the features are 2,500 with n-grams (0, 4) and with binary set to True. Version 3 uses a balanced approach but performs fine-tuning of parameters.

It is observed that Kernel SVM generates the best result compared to other methods, raising the mean accuracy to 85% and 86%, respectively, for dataset v2.0 and Sentence Polarity v3.0. Surprisingly enough, the performance of ANN and NB is as good as Kernel SVM for v1.0. However, it is observed that Kernel SVM is consistent with a better performance for all data sets. Fig. 2, Fig. 3, and Fig. 4 depict the precision, recall, F1 score, accuracy with 80:20 train and test ratio, and mean accuracy by considering k-fold validation where k is 10 for v3.0, v2.0, and v1.0 data sets in order.

3.4 Scores for Tf-idf vectorizer

The Tf-idf vectorizer plays a critical role in ranking the feature importance in the corpus across the versions of the Sentence Polarity dataset. In v1.0, we limited the maximum number of features to 2500 and utilized L2 normalization, while limiting the number of features to 1500 and using L1 normalization for the second version. For v3.0, the same feature restriction and normalization method as v2.0 were applied.

As we can observe from Table 4, ANN and Kernel SVM give the best result with a mean accuracy of 86 % and 85%, respectively. 4 layers of the neural network are used with glorot uniform as the kernel initializer and Relu as the activation function. NB and the Kernal SVM outperformed the other methods for v1.0 as shown in Table 5. The Kernel SVM and ANN have superior performance on v3.0 as shown in Table 6.

Table 1: Analysis of various algorithms on sentence polarity V2.0 using count vectorizer.

Algorithm	Precision	Recall	F1 Score	Accuracy (80:20)	K-fold Mean Accuracy
Random Forest	0.57	0.76	0.65	0.76	0.68
ANN	0.74	0.84	0.79	0.78	0.84
Gaussian NB	0.69	0.81	0.74	0.72	0.72
Multinomial NB	0.80	0.82	0.81	0.81	0.80
Bernoulli NB	0.70	0.78	0.74	0.73	0.76
KNN	0.54	0.86	0.67	0.58	0.55
SVM	0.86	0.77	0.81	0.82	0.80
Kernel SVM	0.86	0.77	0.81	0.82	0.85

Table 2: Analysis of various algorithms on sentence polarity V1.0 using count vectorizer.

Algorithm	Precision	Recall	F1 Score	Accuracy (80:20)	K-fold Mean Accuracy
Random Forest	0.72	0.71	0.71	0.70	0.61
ANN	0.74	0.80	0.77	0.75	0.73
Gaussian NB	0.78	0.65	0.71	0.72	0.70
Multinomial NB	0.74	0.74	0.74	0.74	0.74
Bernoulli NB	0.63	0.63	0.63	0.63	0.74
KNN	0.75	0.51	0.35	0.51	0.63
SVM	0.70	0.69	0.69	0.68	0.70
Kernel SVM	0.76	0.74	0.75	0.74	0.73

Table 3: Analysis of various algorithms on sentence polarity V3.0 using count vectorizer.

Algorithm	Precision	Recall	F1 Score	Accuracy (80:20)	K-fold Mean Accuracy
Random Forest	0.78	0.77	0.77	0.77	0.78
ANN	0.84	0.84	0.84	0.84	0.84
Gaussian NB	0.76	0.75	0.75	0.75	0.76
Multinomial NB	0.83	0.83	0.83	0.83	0.84
Bernoulli NB	0.84	0.84	0.84	0.83	0.84
KNN	0.68	0.62	0.58	0.61	0.65
SVM	0.85	0.84	0.84	0.84	0.82
Kernel SVM	0.85	0.83	0.84	0.84	0.85

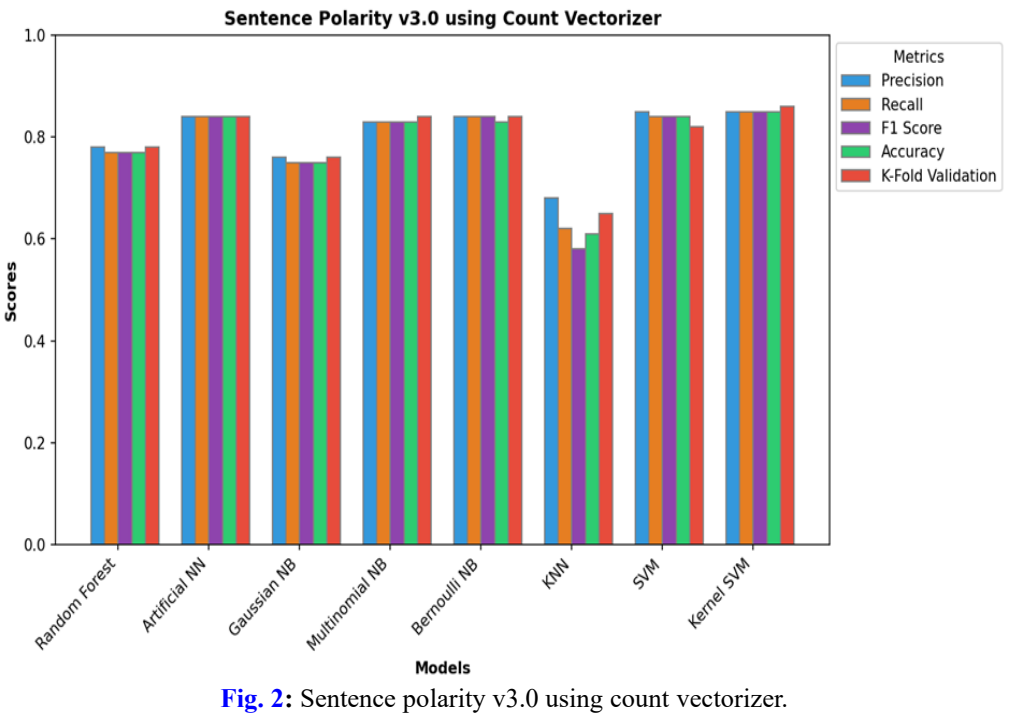

Fig. 2: Sentence polarity v3.0 using count vectorizer.

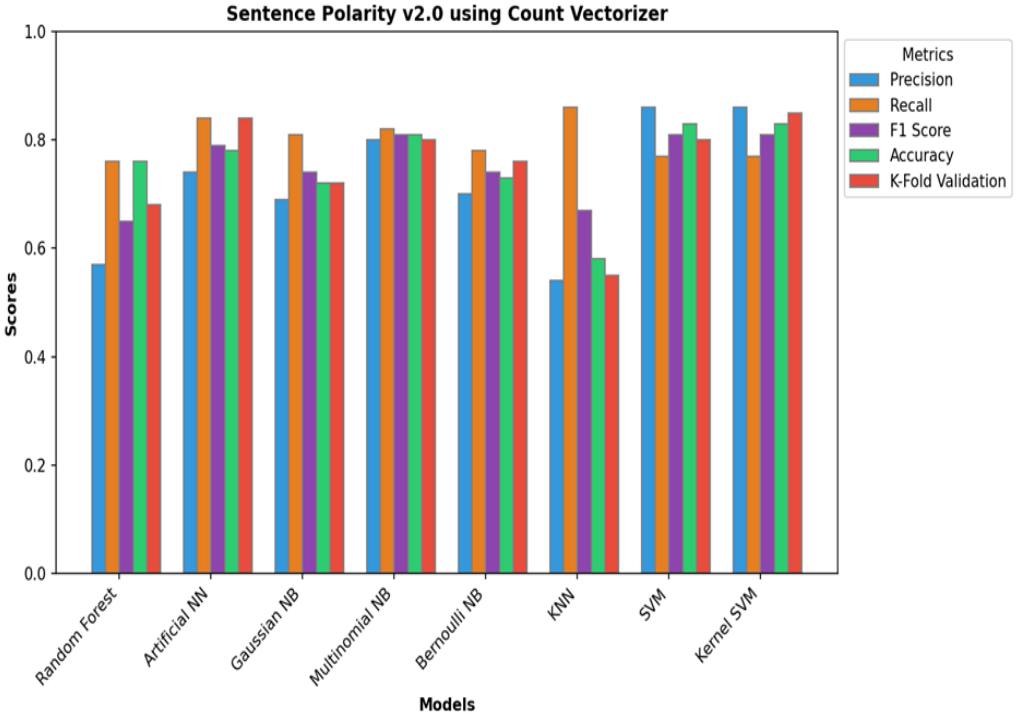

Fig. 3: Sentence polarity v2.0 using count vectorizer.

Table 4: Analysis of various algorithms on sentence polarity V2.0 using Tf-idf vectorizer.

Algorithm	Precision	Recall	F1 Score	Accuracy (80:20)	K-fold Mean Accuracy
Random Forest	0.68	0.67	0.67	0.67	0.68
ANN	0.94	0.89	0.91	0.92	0.86
Gaussian NB	0.77	0.77	0.76	0.77	0.74
Multinomial NB	0.79	0.79	0.79	0.79	0.78
Bernoulli NB	0.79	0.79	0.79	0.79	0.80
KNN	0.71	0.71	0.70	0.71	0.67
SVM	0.85	0.83	0.84	0.84	0.80
Kernel SVM	0.85	0.83	0.84	0.84	0.85

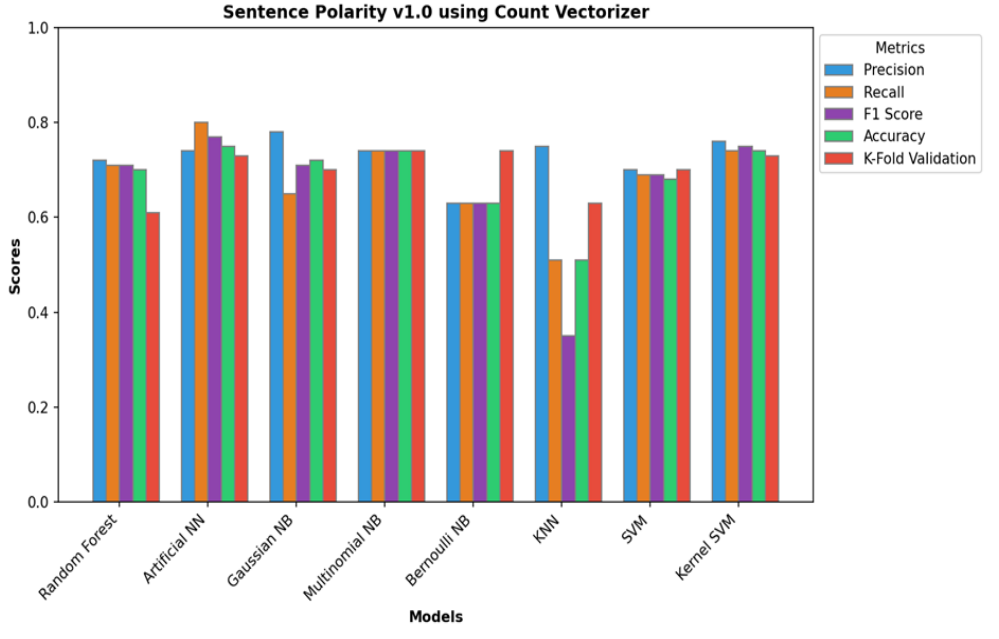

Fig. 4: Sentence polarity v1.0 using count vectorizer.

Table 5: Analysis of various algorithms on sentence polarity V1.0 using Tf-idf vectorizer.

Algorithm	Precision	Recall	F1 Score	Accuracy (80:20)	K-fold Mean Accuracy
Random Forest	0.71	0.71	0.71	0.70	0.59
ANN	0.75	0.75	0.75	0.75	0.72
Gaussian NB	0.75	0.69	0.72	0.72	0.70
Multinomial NB	0.75	0.75	0.76	0.75	0.74
Bernoulli NB	0.75	0.75	0.75	0.75	0.74
KNN	0.75	0.51	0.35	0.51	0.53
SVM	0.71	0.71	0.70	0.69	0.70
Kernel SVM	0.77	0.73	0.75	0.74	0.73

Table 6: Analysis of various algorithms on sentence polarity V3.0 using Tf-idf vectorizer

Algorithm	Precision	Recall	F1 Score	Accuracy (80:20)	K-fold Mean Accuracy
Random Forest	0.78	0.78	0.78	0.77	0.78
ANN	0.86	0.86	0.86	0.85	0.86
Gaussian NB	0.80	0.80	0.80	0.79	0.79
Multinomial NB	0.84	0.86	0.85	0.84	0.85
Bernoulli NB	0.85	0.85	0.85	0.84	0.84
KNN	0.69	0.68	0.67	0.68	0.67
SVM	0.85	0.85	0.85	0.85	0.82
Kernel SVM	0.86	0.86	0.86	0.86	0.86

Fig. 5, Fig. 6, and Fig. 7 depict the precision, recall, f1-score, accuracy with 80:20 train and test ratio, and mean accuracy by considering k-fold validation where k is 10 by considering v3.0, v2.0, and v1.0 data sets, respectively.

3.5 Comparison and validation

The performance of SVM is better than that of other ML methods used in this study as it is suitable for small data sets and robust against overfitting. The Naïve Bayes is usually preferred for text classification. It performs equally well compared to SVM, and its results are closer to SVM. ANN also gives promising results. However, it is computationally extensive. The random forest is an ensemble method and hence slow compared to other methods. Its strength lies in handling large data sets, and it might not have produced better results due to the smaller data sets considered in this study. The recall obtained in KNN is good, but its low precision results in a low accuracy.

We considered a few recent works that used the same data sets and similar methods to validate our results. Maulana *et al*. tried to improve the accuracy of sentiment analysis for movie reviews with SVM based on information gain.[22] They used RBF kernel and 10-fold cross-validation. They attained an accuracy of 83% for v2.0, and 86% for v3.0. The accuracy of NB was 81% for both the data sets while KNN showed 56% for v2.0 and 61% for v3.0.

Sharma *et al*. used ML to analyze movie review sentiments and observed the highest accuracy of 73% for SVM, and 71% for NB by considering v3.0.[3] Table 7 and Table 8 show the accuracy obtained in this paper and the accuracies observed by Maulana *et al*.[22] and Sharma *et al*.[3] by considering the common method. This highlights that the results observed in our methods are better than the existing work.

Table 7: Result comparison.

ML method	Maulana *et al*.[22]	Our results
RBF Kernel SVM	0.83 (v2.0)	0.85 (v2.0)
	0.86 (v3.0)	0.86 (v3.0)
KNN	0.56 (v2.0)	0.67 (v2.0)
	0.61 (v3.0)	0.67 (v3.0)

Table 8: Result comparison.

ML method	Harsh Sharma *et al*.[3]	Our results
SVM	0.73 (v3.0)	0.85 (v3.0)

4. Conclusions

Many researchers have proposed various models by considering reviews. The present work analyzes the traditional ML techniques on the movie review dataset. The data was pre-processed, and sentence-level polarity was determined. The reviews were categorized as positive and negative polarity, and the movies were classified as good or bad by referring to these polarities. The kernel SVM method gives a better accuracy than other techniques for determining polarity using count vectorizer and Tf-idf. The accuracy of KNN is not as expected as it depends on the value of k. There is a scope to improve the KNN results by analyzing various values of k and different distance measures. Also, people have different ways of writing while sharing their reviews. This might have

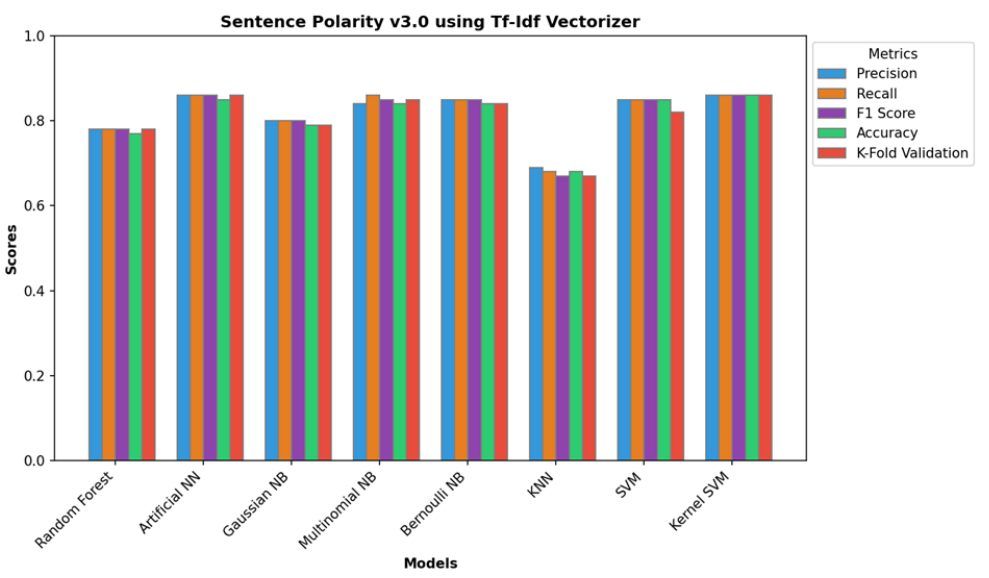

Fig. 5: Sentence polarity v3.0 using Tf-idf vectorizer.

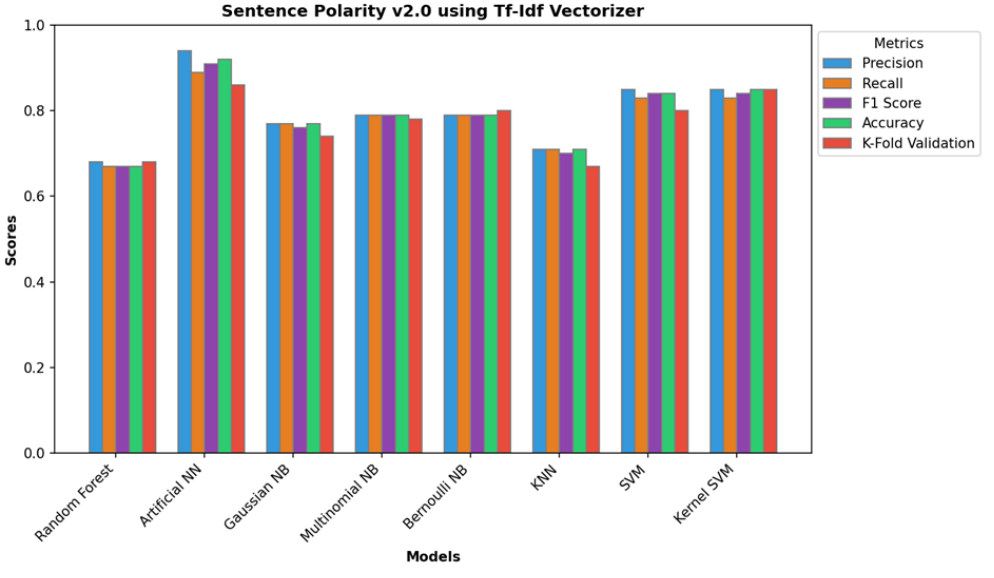

Fig. 6: Sentence polarity v2.0 using Tf-idf vectorizer.

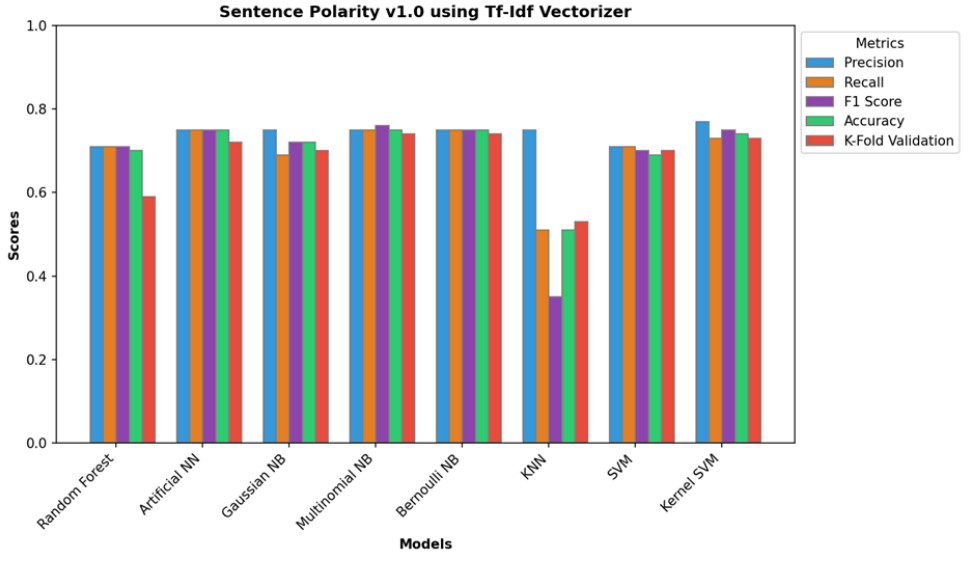

Fig. 7: Sentence polarity v1.0 using Tf-idf vectorizer.

introduced variance in the data and KNN may not be suitable for data with variance. The traditional ML methods are less prone to overfitting and are better generalized, given a small data set, like movie reviews considered in this paper. We observed good results for data sets v2.0 and v3.0. However, v1.0 results are not encouraging compared to the other 2 data sets though class distribution is symmetric. A hybrid model with various ML techniques may be regarded to check for possible accuracy improvement. Also, other NLP techniques can be explored to design a better classifier model as part of future work. A generalized model may be developed that considers cross-domain reviews like product reviews, tourist place reviews, hotel reviews, *etc*., and is not limited to only movie reviews. A larger corpus may be considered and recent transformer-based models like BERT and its variants can be used to analyze sentiments.

Conflict of Interest

There is no conflict of interest.

Supporting Information

Not applicable.

References

[1] C. Clavel, Z. Callejas, Sentiment analysis: from opinion mining to human-agent interaction, *IEEE Transactions on Affective Computing*, 2016, **7**, 74-93, doi: 10.1109/TAFFC.2015.2444846.

[2] A. Sheoran, R. Boora, M. Jangra, C. E. Valderrama, Performance analysis of machine learning models for human activity classification, *Engineered Science*, 2024, **31**, 1207, doi: 10.30919/es1207.

[3] H. Sharma, S. Pangaonkar, R. Gunjan, P. Rokade, Sentimental analysis of movie reviews using machine learning, *ITM Web of Conferences*, 2023, **53**, 02006, doi: 10.1051/itmconf/20235302006.

[4] M. Masrani, P. G, Twitter sentiment analysis using a modified Naïve Bayes algorithm, Information Systems Architecture and Technology: Proceedings of 38th International Conference on Information Systems Architecture and Technology – ISAT 2017. Cham: Springer International Publishing, 2018, 171-181.

[5] S. Bhat, S. Garg, G. Poornalatha, Assigning sentiment score for twitter tweets, 2018 International Conference on Advances in Computing, Communications and Informatics (ICACCI). September 19-22, 2018, Bangalore, India. IEEE, 2018, 934-937.

[6] N. Y.R, P. G, Feature based opinion mining for restaurant reviews, Advances in Signal Processing and Intelligent Recognition Systems. Cham: Springer International Publishing, 2018, 305-318.

[7] S. Mukherjee, Sentiment Analysis of Reviews, Encyclopedia of Social Network Analysis and Mining, 2017, 1-10.

[8] P. Nagamma, H. R. Pruthvi, K. K. Nisha, N. H. Shwetha, An improved sentiment analysis of online movie reviews based on clustering for box-office prediction, *International Conference on Computing, Communication & Automation*. May 15-16, 2015, Greater Noida, India. IEEE, 2015, 933-937, doi: 10.1109/CCAA.2015.7148530.

[9] R. Wankhede, A. N. Thakare, Design approach for accuracy in movies reviews using sentiment analysis, 2017 International conference of Electronics, Communication and Aerospace Technology (ICECA). April 20-22, 2017, Coimbatore, India. IEEE, 2017, 6-11.

[10] V. K. Singh, R. Piryani, A. Uddin, P. Waila, Sentiment analysis of movie reviews: a new feature-based heuristic for aspect-level sentiment classification, 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s). March 22-23, 2013, Kottayam, India, IEEE, 2013, 712-717.

[11] V. K. Singh, R. Piryani, A. Uddin, P. Waila, Sentiment analysis of movie reviews and blog posts, 2013 3rd IEEE International Advance Computing Conference (IACC). February 22-23, 2013, Ghaziabad, India. IEEE, 2013, 893-898.

[12] A. Tripathi, S. K. Trivedi, Sentiment analyis of Indian movie review with various feature selection techniques, 2016 IEEE International Conference on Advances in Computer Applications (ICACA). October 24-24, 2016, Coimbatore, India. IEEE, 2016, 181-185.

[13] T. P. Sahu, S. Ahuja, Sentiment analysis of movie reviews: a study on feature selection & classification algorithms, 2016 International Conference on Microelectronics, Computing and Communications (MicroCom). January 23-25, 2016, Durgapur, India. IEEE, 2016, 1-6.

[14] Y. Ou, A. Y. K. Chua, An investigation of factors that contribute to movie review helpfulness in China, International Multiconference of Engineers and Computer Scientists (IMECS), IAENG, 2021.

[15] S. Thimmaiah, R. Jayaram, A new gate control unit-recurrent neural network structure for audio-based sentiment analysis, *Engineered Science*, 2024, **30**, 1180, DOI: 10.30919/es1180.

[16] Y. Wang, Machine Learning Based Sentiment Analysis of Movie Reviews - A case study of Beyond the Clouds, CACML '24: Proceedings of the 2024 3rd Asia Conference on Algorithms, Computing and Machine Learning, 2024, 13-17.

[17] S. Kumar, K. Sharma, D. Veragi, A. Juyal, Sentimental analysis of movie reviews using machine learning algorithms, 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON). May 26-27, 2022, Faridabad, India. IEEE, 2022, 526-529.

[18] M. F. Porter, An algorithm for suffix stripping, *Program*, 1980, **14**, 130-137, doi: 10.1108/eb046814.

[19] B. Pang, L. Lee, A sentimental education: sentiment analysis using subjectivity summarization based on minimum cuts, Proceedings of the 42nd Annual Meeting on Association for Computational Linguistics - ACL '04. July 21-26, 2004. Barcelona, Spain. Morristown, NJ, USAACL, 2004.

[20] B. Pang, L. Lee, Seeing stars: exploiting class relationships for sentiment categorization with respect to rating scales, Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics - ACL '05. June 25-30, 2005. Ann

Arbor, Michigan. Morristown, NJ, USAACL, 2005.
[21] T. Zhou, Q. Hu, J. Li, L. Wu, Design and Development of a Sentiment Analysis System for Chinese Online Comment Texts. In 2023 3rd International Conference on Electronic Information Engineering and Computer (EIECT) 2023 Nov 17 (pp. 312-317). IEEE.
[22] R. Maulana, P. A. Rahayuningsih, W. Irmayani, D. Saputra, W. E. Jayanti, Improved accuracy of sentiment analysis movie review using support vector machine-based information gain, *Journal of Physics: Conference Series*, 2020, **1641**, 012060, doi: 10.1088/1742-6596/1641/1/01206.

Publisher's Note: Engineered Science Publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access
This article is licensed under a Creative Commons Attribution 4.0 International License, which permits the use, sharing, adaptation, distribution and reproduction in any medium or format, as long as appropriate credit to the original author(s) and the source is given by providing a link to the Creative Commons licence and changes need to be indicated if there are any. The images or other third-party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/ licenses/by/4.0/.

©The Author(s) 2025