INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

OPEN ACCESS

Received: 21-06-2023

Accepted: 26-06-2023

Published: 03-08-2023

Citation: Mathur K, Jain P, Gupta S, Mathur P (2023) Sentiment Analysis Framework of Social Media Text by Feature Extraction and Machine Learning Model. Indian Journal of Science and Technology 16(29): 2233-2243. https://doi.org/10.17485/IJST/v16i29.1537

* **Corresponding author**.

info@icheindia.org

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Sentiment Analysis Framework of Social Media Text by Feature Extraction and Machine Learning Model

Kajal Mathur[1]*, Paresh Jain[2], Sunita Gupta[3], Puneet Mathur[4]

1 Research Scholar, Department of Computer Science and Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
2 Associate Professor, Department of Electronics and Communication Engineering, Suresh Gyan Vihar University, Jaipur, Rajasthan, India
3 Associate Professor, Department of information technology, Swami Keshvanand Institute of Technology, Jaipur, Rajasthan, India
4 Assistant Engineer (IT), RVUN, Jaipur, Rajasthan, India

Abstract

Objectives: This research paper aims to analyze sentiment and opinions in online resources like discussion forums, review sites, and blogs. It also compares the effectiveness of three feature extraction techniques (TF-IDF, Word2Vec, and WAM) and evaluates three machine learning algorithms (Naïve Bayes, SVM, and ANN) for sentiment classification to determine the most accurate algorithm. **Methods**: The study utilizes sentiment-rich datasets from IMDB movie reviews, Yelp reviews, and tweets. Three feature extraction techniques are applied to extract relevant features and patterns from the text. Three machine learning algorithms are implemented to classify sentiments into positive, negative, and neutral categories. Accuracy, precision, recall, and F-measure are used to assess algorithm performance. The model is updated and refined three times to ensure reliability. **Findings**: The Artificial Neural Network (ANN) algorithm outperforms Naïve Bayes and Support Vector Machines, achieving an impressive accuracy rate of 99.74% for sentiment classification. Precision, recall, and F-measure exceed 98.5% after model refinement, demonstrating the approach's robustness. The study highlights the potential of sentiment analysis in online resources and emphasizes the ANN's superior accuracy, providing valuable insights for future sentiment analysis studies. **Novelty**: This research combines three popular feature extraction techniques in sentiment analysis, compares three machine learning algorithms on multiple datasets, and achieves a remarkable accuracy rate of 99.74% with the ANN. The study demonstrates the robustness of the approach through model refinement and contributes insights into sentiment analysis in online resources.

Keywords: Dataset; Feature Extraction; Machine Learning; Sentiment Analysis; Accuracy and Precision

1 Introduction

Social media has emerged as a dominant platform for online communication, allowing individuals to express their thoughts and emotions in real-time. However, the informal nature of social media text poses challenges for accurate classification and information extraction. To address this, techniques such as TF-IDF weighting combined with a Word Article Matrix (WAM) have been proposed to categorize and analyze social media text effectively. Yet, determining the optimal iteration number for WAM updating remains an unexplored area [(1–3)].

Moreover, sentiment analysis techniques have been applied to movie reviews, with a focus on comparing supervised machine learning approaches like Support Vector Machines (SVM) and Naive Bayes. The findings indicate the superiority of Naive Bayes, particularly when dealing with a large number of reviews, achieving higher accuracy compared to other methods. With social media playing a vital role in public opinion on various topics, sentiment analysis enables businesses to gain valuable insights for informed decision-making [(4,5)].

Sentiment analysis involves predicting sentiments using classification algorithms and employing text pre-processing techniques. These techniques involve removing symbols, punctuation, and word stems, while also eliminating stop words. The construction of a vector space model using term frequencies and inverse document frequencies serves as the foundation for sentiment analysis [(6–8)].

While previous studies have explored sentiment analysis using various algorithms, there are still gaps in understanding algorithm performance across different datasets, including movie comments, political tweets, and drug-related tweets. Furthermore, research conducted on Turkish datasets highlights the significant role of data distribution in the success rate of classification algorithms. These gaps justify the need for further investigation and contribute to the advancements of sentiment analysis on social media text [(9,10)].

In this paper a framework of sentiment analysis framework for social media text is proposed by using enhancing advance feature extraction techniques and machine learning to obtain the accuracy, precision, sensitivity and F-measures of the proposed framework

The novelty of this study lies in the development of a sentiment analysis framework specifically designed for social media text in two-fold. Firstly, the framework focuses specifically on social media text, which presents distinct challenges compared to other types of text, such as news articles or product reviews. Social media text often contains informal language, abbreviations, emojis, and contextual references that require specialized techniques for accurate sentiment analysis.

Secondly, the framework integrates feature extraction and machine learning models. Feature extraction involves identifying relevant aspects of the text that can capture sentiment, such as keywords, linguistic patterns, syntactic structures, or contextual cues. By leveraging machine learning models, such as Support Vector Machines (SVM), Artificial Neural Network (ANN), and Naïve Bayes (NB), the framework can learn from the extracted features to accurately classify the sentiment of social media text.

Overall, the novelty of this topic lies in its targeted focus on sentiment analysis in the context of social media, as well as the integration of feature extraction techniques and machine learning models to achieve accurate sentiment classification. By addressing the unique characteristics of social media text, this framework contributes to advancing the field of sentiment analysis and enables deeper insights into public opinion, customer feedback, and social media trends.

1.1 Research Gaps

Based on the information provided in the research papers, here are some potential research gaps that could be explored:

- The findings of the previous research paper are limited to specific datasets, and there is a need for further research to examine the generalizability of the results across different types of online resources, including news articles, forum threads, and social media posts from various platforms.
- A comprehensive comparison of TF-IDF, Word2Vec, and Word Article Matrix methods in terms of effectiveness and performance is lacking. Future studies should conduct a more extensive evaluation to determine the most suitable feature extraction approach for sentiment analysis in different contexts.
- A performance comparison of SVM, NB and ANN using TF-IDF, Word2Vec, and Word Article Matrix feature extracting methods is lacking if they outperform in terms of accuracy.

2 Methodology

2.1 Data Collection

This paper uses three types of data is collected as [(11)]:

- Internet Movies Database (IMDB)
- Twitter Database
- Yelp Database

The Internet Movies Database (IMDB) movie review dataset. This data consists of unprocessed, unlabelled file. In this dataset 1400 processed text files are available.

The files of all three datasets are divided in two types with respect to their classification as 'Positive' and 'Negative', indicating the true classification (sentiment) of the component files.

2.2 Text Pre-processing

The initial step in this stage involves obtaining the actual text from the dataset, treating each review as a separate entity. To achieve this, the content of the file is split based on the end-of-line character, effectively separating individual reviews. Additionally, the reviews are converted to lowercase to facilitate matching with the AFINN data being utilized. Punctuation marks, numbers, and control characters are omitted during this process to enhance matching accuracy. In this research, feature extraction is performed using the following techniques[12–14]:

- Term Frequency-Inverse Document Frequency (TF-IDF)
- Word2Vec (W2V)
- Word Article Matrix (WAM)

These techniques are employed to extract meaningful features from the reviews, enabling further analysis and classification[13].

2.3 Classifications

The algorithms are employed to get the best results as given bellow:

- SVM
- ANN
- Naïve Bayes

2.4 Proposed Framework

This section provides an overview of the datasets used in the study, including Twitter, IMDB, and Yelp. Feature extraction techniques such as TF-IDF, Word2Vec, and WAM are employed to extract meaningful features from the data. The paper also utilizes various classification algorithms from the field of machine learning. The flowchart illustrating the methodology employed in this paper is presented in Figure 1.

2.5 Datasets

Twitter is a popular microblogging site that allows users, including Jack Dorsey, to share text, pictures, and videos instantly within a 280-character limit[10,15]. Users can follow other accounts, like tweets, and retweet them to share with their own followers.

In this research paper, a dataset of 4,500 health-related tweets was collected using the Twitter Application Programming Interface (API). These tweets were then pre-processed and assigned sentiment scores using a Python program. Out of the collected and labeled tweets, 1,680 were categorized as neutral, 1,220 as positive, and 1,600 as negative[16]. The attributes of the collected tweets obtained via the Python program are presented in Table 1.

In addition to the analysis of Twitter data, the same models were applied to two other datasets. The first dataset consisted of 500 positive and 500 negative opinions collected by[14] from IMDB movie reviews, as shown in Table 2. The second dataset, called Yelp, consisted of 200 neutral, 350 positive, and 300 negative reviews, as presented in Table 3.

These datasets serve as valuable resources for examining sentiment analysis techniques and evaluating the performance of the models applied in the study. The attributes of the collected tweets and reviews provide insights into the data used for analysis and classification.

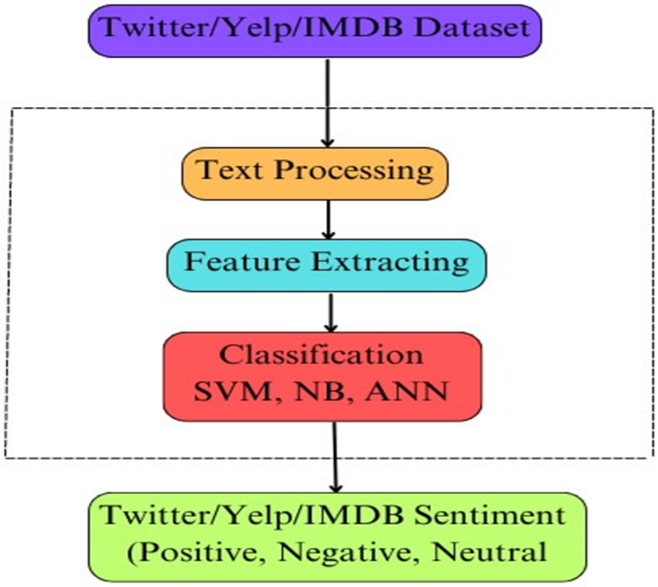

Fig 1. Flowchart of Proposed Methodology

Table 1. Twitter dataset

Dataset attribute	Explanation of Attribute
Id	Order of tweet dataframe
Text	Tweet
Created_at	Data and time the Tweet was posted
Retweeted	Tweet rerun status (bool)
Retweet_count	Number of retweets
User_screen_name	Username
User_followers_count	Number of followers
User_location	Followers location
Hastags	Tweet tag
Sentiment_score	Sentiment score
Sentiment_class	Positive, negative, neutral

Table 2. IDMB dataset of Kotzias

Dataset attribute	Explanation of Attribute
Text	Reviews from IDMB
Year	Year of release
Name_movie	Name of the movie
Genre	Genre of the movie
Runtime	Total runtime of the movie
Sentiment class	Positive, negative

Table 3. Yelp dataset of JSON

Dataset attribute	Explanation of Attribute
Name	Name of Business
User_id	Customer ID number
Review_id	ID of Reviewer
Business_id	Business number
Stars	Rating of business service and quality
Review_count	Number of reviews received
Sentiment class	Positive, negative, neutral

2.6 Feature Extracting Techniques

2.6.1 Term Frequency-Inverse Document Frequency (TF-IDF)

The TF-IDF (Term Frequency-Inverse Document Frequency) technique is used to calculate term weights in a document. The TF component calculates the frequency of a term in a document, as shown in Equation (1). The IDF component determines the significance of a term by considering its occurrence in multiple documents and distinguishing it from stop words. It is calculated by taking the logarithm of the ratio between the total number of documents and the number of documents containing the term, as shown in Equation (2) [5,8].

$$TF(t,d) = \frac{Term\ t\ frequency\ in\ document\ d}{Total\ words\ in\ document\ d} \tag{1}$$

$$IDF\ (t) = log\left(\frac{Total\ documents}{Documents\ with\ term\ t}\right) \tag{2}$$

t=Term, d=Documents

the TF-IDF formula is defined as (3):

$$\mathrm{TF-IDF(t) = TF(t,d) \times IDF(t)} \tag{3}$$

In Equation (1), "TF (t, d)" represents the term frequency of term "t" in document "d" divided by the total number of words in document "d". In Equation (2), "IDF(t)" is calculated as the logarithm of the ratio between the total number of documents and the number of documents containing term "t". "t" represents the term, and "d" represents the documents. The TF-IDF formula, given in Equation (3), combines these components to determine the importance of a term in a document based on its frequency and occurrence in the document collection.

2.6.2 Word2vec

Word2vec is a natural language processing tool that operates on unsupervised learning principles and is based on the artificial neural network structure developed by [3,9,15]. It functions by taking text input and representing each word in the text as a vector. The primary objective of word2vec is to cluster words with similar meanings close to each other in vector space. This is achieved through two different learning architectures: continuous bag of words (CBOW) and skip-gram (SG).

In the CBOW architecture, the tool examines the neighboring words (both to the right and left) of a given word within a specific window size and performs word estimation based on these neighboring words. On the other hand, the skip-gram architecture estimates neighboring words by considering the target word in reverse, focusing on predicting the surrounding words given the target word.

By employing these learning architectures, word2vec can effectively capture semantic relationships between words and represent them as vectors, enabling various downstream natural language processing tasks such as sentiment analysis, text classification, and word similarity calculations.

2.6.3 Word Article Matrix (WAM)

WAM is a significant data structure [3,5,17]. It represents a large matrix that captures the weighted relationships between documents and keywords. The rows of the matrix correspond to document names (articles), while the columns correspond to words or keywords extracted from the documents. The WAM is filled in by counting the occurrences of keywords within each document, resulting in a table structure as shown in Table 4.

To generate the initial WAM (i-WAM), the term frequency (TF) value of each word is utilized. For example, considering a training set of 10 documents with a total of 100 words, the i-WAM will be constructed using the TF values, as depicted in Table 5. In this representation, documents and words are represented as vectors. Each row in the matrix represents a document, and the values within the row correspond to the vector of words that represent that particular document.

Suppose there is a query, such as "Microsoft stock got a small boost from the launch of Windows 10". This query is transformed into a model of word vectors, as illustrated in Table 6.

Table 4. An example of WAM

Article Word (Category)	Stock	Windows 10	Golf
Economic	5	2	2
IT	2	10	1
Sports		1	7
Entertainment	5	4	4
Foreign	3	5	6
Politics	4	7	2
Regional	1	6	4

Table 5. An example of the i-WAM

Article Word (Category)	Stock	Windows 10	Golf
Economic	0.05	0.02	0.02
IT	0.02	0.10	0.01
Sports		0.01	0.07
Entertainment	0.05	0.04	0.04
Foreign	0.03	0.05	0.06
Politics	0.04	0.07	0.02
Regional	0.01	0.06	0.04

Table 6. A sample query with word count

Query Word (Category)	Stock	Windows 10	Golf
Query	1	1	0

In the context of a corpus, the collection of documents can be seen as a set of vectors in a vector space, with each term representing a unique axis. The similarity between any two documents can be determined using the cosine similarity technique[4,18], which measures the similarity between their respective vectors.

The cosine similarity (d1, d2) is calculated as the dot product of the document vectors d1 and d2, divided by the product of their magnitudes (||d1|| and ||d2||), as shown in Equation (4):

$$\text{Cosine Similarity}\,(d1,\ d2) = (d1 \cdot d2)/(\|d1\| * \|\ d2\|) \tag{4}$$

Here, the dot product represents the similarity between the vectors, while the magnitude represents the length of the vectors.

Using the cosine similarity values, we can calculate the similarity between documents. For example, when applying this technique to an example query, the cosine similarity scores are computed and presented in Figure 2. In this table, the word "Stock" has a high weight of 0.5 in the economic category. The operation results indicate that the query is more likely related to the economic document, as it produces the highest cosine similarity score of 0.861.

2.7 Classification Algorithms

This research focuses on document-level sentiment analysis, which involves classifying the sentiment of entire documents rather than individual sentences or specific attributes. Two supervised machine learning models, Naive Bayes (NB) and support

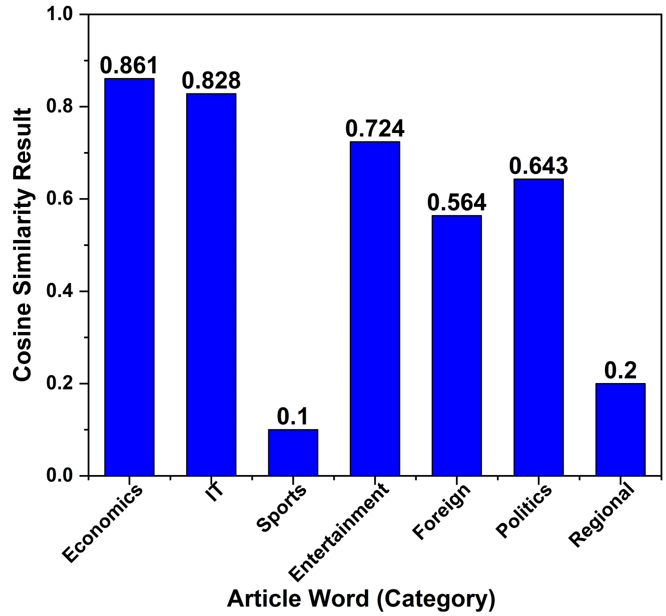

Fig 2. Cosine similarity result

vector machines, were utilized for sentiment classification of selected movie reviews. To represent the documents in a machine-readable format, a predefined set of features (f1, f2, ..., fm) was established, where ni(d) represents the frequency of feature fi in document 'd'. Consequently, each document 'd' was transformed into a document vector d := (n1(d), n2(d), ..., nm(d)).

The chosen machine learning algorithms, namely SVM, ANN, and NB, are widely recognized for their effectiveness in sentiment analysis tasks. This study contributes by evaluating the performance of these algorithms in comparison to traditional frequency-based text representation (TF-IDF) and prediction-based text representation (W2V) methods. Experimental analysis was conducted on datasets including IMDB, Yelp, and tweets that were collected and labeled by researchers based on their sentiments. The results indicated that the model created using W2V and ANN demonstrated superior performance compared to other approaches[1–4,19].

2.7.1 Naïve Bayes

The Naïve Bayes (NB) algorithm, named after the mathematician Thomas Bayes, belongs to the family of Bayesian algorithms and is based on the statistical Bayesian theorem. It is a statistical classification technique that utilizes the predictive power of Bayesian models. The Bayes classifier, which is relatively straightforward to apply, is a predictive model.

In the context of the algorithm, let's consider a sample set d = d1, d2, d3, ..., dn, and a class set c1, c2, c3, ..., cm. To classify a given sample, the probability is calculated using Equation (5):

$$\mathrm{P(c/d) = (P(c) * P(d/c))/P(d)} \tag{5}$$

Here, the probability of each class given the sample is determined. The class with the highest probability for the data sample is considered the classification result.

Although the role of P(d) in selecting c is negligible, it is important to note that the conditional independence assumption made by the Naive Bayes classifier does not hold in real-world situations. Nevertheless, Naive Bayes-based text classification tends to perform well, as it is a simple probabilistic classifier based on Bayesian probability. The classifier assumes that the probabilities of individual features in a document are independent of each other. It treats a document as a collection of words and assumes that the presence and position of each word in the document are independent of other words. The Naive Bayes classifier is derived from Bayes' rule[4,20].

2.7.2 Support Vector Machine

The Support Vector Machine (SVM) is a data mining method that operates in a vector space and aims to find a decision boundary between two classes that is farthest from a random point on the training data. It follows the principle of structural risk minimization in statistical learning theory, which is one of its key characteristics [3,4].

SVMs have proven to be efficient for document classification and are known as large margin classifiers. The fundamental concept behind SVM classification is to identify a hyperplane with the maximum margin that effectively separates the document vectors of one class from those of the other class. Unlike Naïve Bayes, SVMs are large-margin classifiers rather than probabilistic classifiers. The objective is to find a solution represented by the vector W:

$$W = \sum_j \propto_j c_j d_j, \qquad \propto_j \geq 0 \tag{6}$$

The αj values, obtained by solving a problem of dual optimization, play a crucial role in determining the support vectors. Only the document vectors with αj greater than zero contribute to the construction of the vector w. These support vectors are essential for the classification process, as they define which side of the hyperplane created by w an instance falls on.

2.7.3 Artificial Neural Network

Artificial neural networks are computational models inspired by the structure and functioning of the human brain. They are composed of interconnected processing elements, referred to as neurons, which have their own memory and communicate through weighted connections. These networks emulate the behavior of biological neural networks and are implemented as computer programs [2,3].

The structure of an artificial neural network comprises three main components: neurons, connections, and a learning algorithm. Neurons serve as the fundamental processing units within the network. They receive input from various sources, representing the factors that influence the problem, and produce output based on the desired outcome. Through the connections between neurons, an interconnected network is formed, resembling the biological neural connections. In most artificial neural network systems, neurons are organized into layers, with each layer processing information in a specific direction [15].

2.8 Performance Criteria

In this paper, the models developed using classification algorithms were evaluated using a confusion matrix [35]. Four statistical measures were employed for performance evaluation: accuracy (ACC), sensitivity (SENS), precision (PREC), and F-measure (F). Sensitivity represents the probability of correctly identifying the True Positive (TP) class (where 'Y' means 'Yes'), while specificity represents the probability of correctly identifying the True Negative (TN) class (where 'Y' means 'No'). False Negative (FN) refers to the situation where the model predicts a negative class while the actual class is positive, while False Positive (FP) refers to the scenario where the model predicts a positive class while the actual class is negative. Accuracy reflects the overall probability of correctly detecting the true class. The F-measure is a harmonic mean of precision and recall, ranging from 0 (worst) to 1 (perfect PREC and SENS) [3,15].

The accuracy value is calculated using Equation (7):

$$Accuracy = \frac{T_P + T_N}{T_P + T_N + F_P + F_N} \tag{7}$$

The sensitivity value is calculated using Equation (8):

$$Sensitivity = \frac{T_P}{T_P + F_N} \tag{8}$$

Precision is calculated using Equation (9):

$$Precision = \frac{T_P}{T_P + F_P} \tag{9}$$

The F-measure value is calculated using Equation (10):

$$F - measure = \frac{2 * Precision * Sensitivity}{Precision + Sensitivity} \tag{10}$$

To establish the models using classifier algorithms and evaluate their performance, the dataset was divided into training and test sets.

3 Results and Discussion

In this paper, a dataset of 4500 tweets was used to perform sentiment analysis using the Naïve Bayes (NB), Support Vector Machine (SVM), and Artificial Neural Network (ANN) classification algorithms. The tweets underwent text pre-processing and vector space modelling. To evaluate the performance of the algorithms, a 5-fold cross-validation approach was applied to split the data into training and test sets. The evaluation metrics used were accuracy (AC), precision (PR), sensitivity (S), and F-measure (F), and the results are presented in Tables 8, 9 and 10.

Furthermore, the performance of the classification algorithms on the IMDB dataset, which contained labeled polarities provided by Kotzias, was also assessed and presented inTables 8, 9 and 10.

Table 7. Results with TF-IDF on Twitter, IMDB and Yelp datasets

Dataset	Algorithm	Accuracy	Precision	Sensitivity	F-Measure
Twitter	SVM	82%	83%	82%	81%
	NB	72%	73%	72%	76%
	ANN	86%	87%	84%	85%
IMDB	SVM	83%	84%	84%	84%
	NB	82%	82%	83%	82%
	ANN	89%	88%	88%	89%
Yelp	SVM	81%	82%	81%	80%
	NB	70%	72%	71%	74%
	ANN	85%	86%	82%	84%

Table 8. Results with TF-IDF on Twitter, IMDB and Yelp datasets

Dataset	Algorithm	Accuracy	Precision	Sensitivity	F-Measure
Twitter	SVM	82%	83%	82%	81%
	NB	72%	73%	72%	76%
	ANN	86%	87%	84%	85%
IMDB	SVM	83%	84%	84%	84%
	NB	82%	82%	83%	82%
	ANN	89%	88%	88%	89%
Yelp	SVM	81%	82%	81%	80%
	NB	70%	72%	71%	74%
	ANN	85%	86%	82%	84%

Table 9. Result with W2V on the Twitter, IMDB and Yelp datasets

Dataset	Algorithm	Accuracy	Precision	Sensitivity	F-Measure
Twitter	SVM	84%	80%	84%	82%
	NB	72%	76%	76%	77%
	ANN	**87%**	**84%**	**86%**	**85%**
IMDB	SVM	84%	84%	84%	84%
	NB	83%	84%	85%	84%
	ANN	**90%**	**91%**	**90%**	**96%**
Yelp	SVM	83%	79%	83%	81%
	NB	71%	75%	75%	75%
	ANN	**86%**	**83%**	**85%**	**84%**

To validate the performance results of the classifiers on the IMDB dataset, the same algorithms were applied to the Twitter and Yelp datasets using the TF-IDF, Word2Vec (W2V), and Word Article Matrix (WAM) methods for vector modelling. The performance of the algorithms on the three datasets is compared and presented in Tables 8, 9 and 10. It was observed that

Table 10. Result with WAM on the Twitter, IMDB and Yelp datasets

Dataset	Algorithm	Accuracy	Precision	Sensitivity	F-Measure
Twitter	SVM	99.68%	99.76%	99.11%	99.65%
	NB	99.60%	99.64%	99.83%	99.50%
	ANN	**99.72%**	**100%**	**100%**	**99.72%**
IMDB	SVM	99.68%	99.78%	99.21%	99.61%
	NB	99.62%	99.60%	99.15%	99.58%
	ANN	**99.74%**	**100%**	**100%**	**100%**
Yelp	SVM	99.62%%	99.73%	99.08%	99.64%
	NB	99.58%%	99.61%	99.81%	99.50%
	ANN	**99.70%**	**100%**	**100%**	**99.71%**

the ANN algorithm achieved the best performance across all three datasets, while the NB algorithm exhibited the worst performance.Table 9 demonstrates better performance results compared to Table 8, and similarly, Table 10 demonstrates improved performance compared to Table 9. As per the experiment result of ANN on different datasets using different feature extracting techniques, it is observed that the accuracy outperformed 99.74% on the IMDB dataset for WAM technique as depicted in Table 10.

3.1 Comparison with other Methods

Results of comparison between the proposed model using ANN classifier and those reported by others is shown in Table 11. It is revealed that our proposed method is superior to other methods in respect of accuracy, precision, sensitivity and F-measure. It is therefore apparent the method proposed by us is superior to the existing methods. This tends to authenticate the novelty of our proposition to use ANN classifier and therefore inherits its merit over other techniques advocated by a number of previous researchers

Table 11. Comparison between the proposed method and the methods suggested by previous workers

Ref.	Classifier	Feature Extraction Method	Dataset	Accuracy	Precision	Sensitivity	F-Measure
(1)	SVM	TF-IDF	Twitter	83%	83%	82%	81%
			IMDB	83%	84%	84%	84%
			Yelp	81%	82%	81%	81%
		W2V	Twitter	89%	88%	86%	87%
			IMDB	84%	84%	86%	85%
			Yelp	83%	84%	85%	84%
	NB	TF-IDF	Twitter	72%	73%	73%	76%
			IMDB	82%	82%	83%	82%
			Yelp	76%	77%	77%	77%
		W2V	Twitter	72%	76%	75%	76%
			IMDB	83%	84%	85%	84%
			Yelp	78%	78%	78%	81%
Proposed Method	ANN	WAM	Twitter	**99.72%**	**100%**	**100%**	**99.72%**
			IMDB	**99.74**	**100%**	**100%**	**100%**
			Yelp	**99.70%**	**100%**	**100%**	**99.71%**

4 Conclusions

The study aimed to evaluate the effectiveness of classifiers on three diverse datasets: IMDB, Twitter, and Yelp, using various text representation techniques. By leveraging existing categorization of online news categories, the study achieved human-like categorization of social media text. Classification algorithms employed were Artificial Neural Network (ANN), Support

Vector Machine (SVM), and Naïve Bayes. Results showed consistent performance across all datasets, with ANN outperforming other algorithms. Naïve Bayes had the lowest performance. Future studies should explore advanced neural network models for classification. These findings highlight the potential for accurate social media text categorization and suggest avenues for further research and improvement in classification techniques.

References

1) Muhammet SB, Fatih K. Sentiment Analysis on Social Media Reviews Datasets with Deep Learning Approach Article. *Sakarya University Journal of Computer and Information Sciences·*. 2021;4(1). Available from: https://doi.org/10.35377/saucis.04.01.833026.
2) Bordoloi M, Biswas SK. Sentiment analysis: A survey on design framework, applications and future scopes. *Artificial Intelligence Review*. 2023. Available from: https://doi.org/10.1007/s10462-023-10442-2.
3) Alantari HJ, Currim IS, Deng Y, Singh S. An empirical comparison of machine learning methods for text-based sentiment analysis of online consumer reviews. *International Journal of Research in Marketing*. 2022;39(1):1–19. Available from: https://doi.org/10.1016/j.ijresmar.2021.10.011.
4) Bhuvaneshwari P, Rao AN, Robinson YH, Thippeswamy MN. Sentiment analysis for user reviews using Bi-LSTM self-attention based CNN model. *Multimedia Tools and Applications*. 2022;81(9):12405–12419. Available from: https://doi.org/10.1007/s11042-022-12410-4.
5) Ping W, Li J, Jingrui H. S2SAN: A sentence-to-sentence attention network for sentiment analysis of online reviews. 2021. Available from: https://doi.org/10.1016/j.dss.2021.113603.
6) Li L, Goh TTT, Jin D. How textual quality of online reviews affect classification performance: a case of deep learning sentiment analysis. *Neural Computing and Applications*. 2020;32(9):4387–4415. Available from: https://doi.org/10.1007/s00521-018-3865-7.
7) Li W, Qi F, Tang M, Yu Z. Bidirectional LSTM with self-attention mechanism and multi-channel features for sentiment classification. *Neurocomputing*. 2020;387:63–77. Available from: https://doi.org/10.1016/j.neucom.2020.01.006.
8) Zhiqiang G, Guofei C, Yongming H, Lu G, Li F. Semantic relation extraction using sequential and tree-structured LSTM with attention. *Information Sciences*. 2020;509:183–192. Available from: https://doi.org/10.1016/j.ins.2019.09.006.
9) Jain R, Kumar A, Nayyar A, Dewan K, Garg R, Raman S, et al. Explaining sentiment analysis results on social media texts through visualization. *Multimedia Tools and Applications*. 2023;82(15):22613–22629. Available from: https://doi.org/10.1007/s11042-023-14432-y.
10) Nandwani P, Verma R. A review on sentiment analysis and emotion detection from text. *Social Network Analysis and Mining*. 2021;11(1). Available from: https://doi.org/10.1007/s13278-021-00776-6.
11) Rahman H, Tariq J, Masood MA, Subahi AF, Khalaf OI, Alotaibi Y. Multi-Tier Sentiment Analysis of Social Media Text Using Supervised Machine Learning. *Computers, Materials & Continua*. 2023;74(3):5527–5543. Available from: https://doi.org/10.32604/cmc.2023.033190.
12) Budhi GS, Chiong R, Pranata I, Hu Z. Using Machine Learning to Predict the Sentiment of Online Reviews: A New Framework for Comparative Analysis. *Archives of Computational Methods in Engineering*. 2021;28(4):2543–2566. Available from: https://doi.org/10.1007/s11831-020-09464-8.
13) Fan FLL, Xiong J, Li M, Wang G. On Interpretability of Artificial Neural Networks: A Survey. *IEEE Transactions on Radiation and Plasma Medical Sciences*. 2021;5(6):741–760. Available from: https://doi.org/10.1109/TRPMS.2021.3066428.
14) Kaur G, Sharma A. A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis. *Journal of Big Data*. 2023;10(1). Available from: https://doi.org/10.1186/s40537-022-00680-6.
15) Sayyida TK, Sohail A, Shehneela N. Transformer-based deep learning models for the sentiment analysis of social media data. *Array*. 2022;14:100157. Available from: https://doi.org/10.1016/j.array.2022.100157.
16) Qianwen AX, Victor C, Chrisina J. A systematic review of social media-based sentiment analysis: Emerging trends and challenges. *Decision Analytics Journal*. 2022;3:100073. Available from: https://doi.org/10.1016/j.dajour.2022.100073.
17) Dimple T, Bharti N, Bhoopesh SB, Ashutosh M, Manoj K. A systematic review of social network sentiment analysis with comparative study of ensemble-based techniques. *Artif Intell Rev*. 2023;12:1–55. Available from: https://doi.org/10.1007/s10462-023-10472-w.
18) Muhammet SB, Fatih K. Sentiment analysis with machine learning methods on social media. 2020. Available from: https://doi.org/10.14201/ADCAIJ202093515.
19) Wang H, Wang X. Sentiment analysis of tweets and government translations: Assessing China's post-COVID-19 landscape for signs of withering or booming. *Global Media and China*. 2023;8(2):213–233. Available from: https://doi.org/10.1177/20594364231181745.
20) Yin Z, Shao J, Hussain MJ, Hao Y, Chen Y, Zhang X, et al. DPG-LSTM: An Enhanced LSTM Framework for Sentiment Analysis in Social Media Text Based on Dependency Parsing and GCN. *Applied Sciences*. 2022;13(1):354. Available from: https://doi.org/10.3390/app13010354.