Anuj Kumar
Rakesh Kumar
Shashi Shekhar

HYBRID FRAMEWORK FOR SENTIMENT ANALYSIS OF PATIENT REVIEWS USING LEXICON BASED BIO-BIDIRECTIONAL ENCODER REPRESENTATIONS FROM TRANSFORMERS

Abstract *This study integrates a hybrid approach to sentiment analysis in the area of medical texts (patient reviews) that merges lexical strategies with deep learning and machine learning models. Different methodologies are used, such as reviews annotated by SenticNet and Text Blob lexicons. and for extracting crucial features in which TF and TF-IDF are utilized. Lastly, classification tasks are performed using machine learning models and deep learning models, and models based on transformers such as Bio-BERT. Performance metrics are utilized to evaluate the effectiveness of this combined methodology. Experimental results demonstrate that hybridization of lexicon and a transformer-based medical learning model produces superior outcomes compared to using each method independently in sentiments finding. Text Blob exhibits impressive performance, achieving 97 percent accuracy with a hybrid of LSTM and CNN, and another medical transformer model is Bio-BERT, on a drug review dataset, with 95 percent accuracy in term frequency and the logistic regression model. TextBlob also attains 94 percent accuracy when paired with term frequency and the LSTM model, and 97 percent accuracy when combined with the Bio-BERT transformer-based model on a dataset sourced from tweets.*

Keywords convolutional neural network, TextBlob, sentiments, sentiment analysis, pharmaceutical reviews, hybrid lexicon–deep learning approach, Bio-bert, term frequency–inverse document frequency (tf-idf)

Citation Computer Science 27(1) 2026: 65–90

1. Introduction

The sentiment analysis tool is used by healthcare practitioners to rapidly evaluate the advantages and disadvantages of their practice. By analyzing patient feedback and reviews, these tools can generate actionable insights that help improve overall patient experience. Implementing such insights can lead to enhanced patient satisfaction rates, as providers can address specific areas of concern identified through sentiment analysis.

Two main applications are, first, that it offers a more lucid understanding of the feelings a patient has about different parts of the medical institution. This emotional data is crucial for understanding how patients perceive their care, the environment of the facility, and interactions with healthcare professionals. By focusing on both areas of improvement and excellence, healthcare providers can continuously refine their services to better meet patient needs and expectations. Therefore, an accurate medical model and lexicon are needed for sentiment analysis in the healthcare domain. An efficient medical learning model needs to develop to overcome these challenges. This research is aimed to solve some of the below listed research questions. The formulated research questions are as follows:

- **Research Question 1 (RQ1)** – How can a specialized medical lexicon be developed to accurately capture the sentiment expressed in medical texts?
- **Research Question 2 (RQ2)** – What methods can be used to enhance the accuracy of sentiment analysis models specifically trained on medical datasets?
- **Research Question 3 (RQ3)** – Which feature extraction techniques are most effective for capturing sentiment-relevant features from unstructured medical texts?
- **Research Question 4 (RQ4)** – How can the integration of structured and unstructured medical data improve the feature extraction process for sentiment analysis?

The primary objective of sentiment analysis using a medical lexicon and medical model is to accurately identify and interpret sentiments expressed in medical texts. This involves developing a comprehensive medical lexicon that encompasses the vast array of medical terminology, including diseases, treatments, symptoms, and procedural descriptions, ensuring that it captures context-dependent meanings specific to the medical domain. The medical model, trained specifically on annotated medical datasets, leverages this lexicon to parse and analyze the sentiment embedded in various forms of medical communication, such as clinical notes, patient reviews, and research articles. By focusing on context and the unique nuances of medical language, the goal is to enhance the model's ability to discern subtle emotional cues and sentiment shifts, which are critical for understanding patient experiences and clinical outcomes. This objective aims to provide more precise and actionable insights into the sentiments conveyed in medical documents, thereby supporting better patient care, enhancing doctor-patient communication, and informing clinical decision-making processes.

The contributions of this study are manifold:

- Investigation into the applicability of sentiment analysis to medical reviews and Twitter tweets.
- Assessment of sentiment lexicons available to the public in the medical field.
- Proposal of a hybrid framework integrating lexicons, machine learning, and deep learning algorithms for accurate sentiment analysis in healthcare.
- Utilization of TextBlob and SenticNet sentiment lexicons for categorizing medical reviews into sentiments.
- Evaluation of feature extraction methods, TF and TF-IDF, on medical reviews and Twitter datasets.
- Adoption of machine learning models and four deep learning models for sentiment classification.
- Conduct multiple experiments on publicly available datasets to assess the performance of the proposed approach.
- Comparison of the results of the proposed method with state-of-the-art approaches to validate its efficiency.

2. Literature review

In this study, the author categorizes attitude polarity into positive, negative, and neutral, demonstrating that deep learning architectures, particularly CNNs, outperform traditional methods. Using the enhanced Cohn Kanade (CK+) and FER-2013 datasets, our CNN-based approach surpasses previous designs in facial expression recognition, as confirmed by extensive studies [20].

The author presents a computationally efficient CNN for automatic MRI scan classification to detect brain tumors, using the Br35H benchmark data set. By augmenting the data, the author aims to reduce training time and improve accuracy, achieving high accuracy in identifying brain tumors in MRI images [19].

In this study the author utilized three datasets to discover attitudes. To verify the accuracy and consistency of the dataset annotations, the datasets were validated using Cohen's kappa and Krippendorff's alpha metrics. The preprocessed datasets were trained using a variety of methods, such as ensemble learning, deep learning, machine learning, transformers, and hybrid approaches [35].

This paper provides a unique hybrid-model architecture for Urdu sarcasm that combines multilingual BERT (mBERT) embeddings with multi-head attention (MHA) and BiLSTM. Word embeddings from fastText were used to train deep learning classifiers [14].

In this paper, reducing the necessary proportion of labeled data was the author's aim. Three publicly accessible datasets (OpenBookQA, SWAG, and HellaSWAG) and three CNLI models (BERT, LSTM, and ESIM), which reflect various tasks, are used to show the effectiveness of our technique [1]. The authors introduce a deep generative adversarial network for image de-hazing. Traditional de-hazing techniques rely on per-pixel loss, causing significant differences from minor pixel variations. To address this, the authors used a perceptual loss function with pretrained ImageNet models to extract high-level image features, overcoming per-pixel loss limitations [8].

This study investigates the effectiveness of e-learning through an analysis of social media attitudes on platforms such as Facebook, Instagram, and Twitter, recognizing social media's role as a major communication tool [26] to achieve better classification accuracy and employ a transformer-based BERT model with an appropriate preprocessing pipeline [41]. It is crucial to investigate sentiment analysis theory and its useful applications in many settings. This study explores the methods and strategies used, assessing the benefits and drawbacks of each.

Table 1

Comparison of Latest Research Findings on Sentiment Analysis

S.No	Author name	Methodology used	Accuracy [%]	Dataset	Year
1	Meena [20]	CNN	79% and 95%	Cohn Kanade (CK+) and FER-2013	2024
2	Mohbey [19]	CNN	98.99%	Br35H	2023
3	MA Rahman [35]	XLM-RoBERTa	89.46%	3 Datasets	2024
4	M. Hasan [14]	mBERT-BiLSTM-MHA	79.51%	Urdu sarcasm dataset	2024
5	S.N. Aakur [1]	GPT, GPT-2, BERT	63%	Open Book QA, SWAG, and HellaSWAG	2023
6	Mujahid [27]	RoBERTa and BERT	96.02%	Tweets (27,780)	2023
7	Rustam [25]	BERT	96.49%	Tweets	2023
8	Sudheesh [21]	Transformer-based BERT	99%	Twitter's reviews	2023
9	Meena [22]	CNN based on InceptionV3 architecture	99.5%, 86% and 73%	CK+, FER2013, and JAFFE datasets	2023
10	Choudhary [8]	Deep Learning	PSNR (17.09), SSIM (0.14)	NYU Depth Dataset	2020
11	Mujahid [26]	Random Forest and Support Vector Machine	95%	Tweets (17,155)	2021

[43] adds to this conversation by contrasting and comparing various approaches, highlighting the holes and unexplored aspects that are essential in obtaining thorough sentiment comprehension. [34] expand on this investigation by imagining a path forward that takes unsolved issues into account. This paper summarizes sentiment analysis's applications and clarifies how it works. It carefully evaluates approaches to offer a fair assessment of their benefits and drawbacks. The intricacies of senti-

ment analysis are explored [6], opening the door to possible fixes that might increase precision and dependability.

Traveling or accommodation sector services are one practical arena in which sentiment analysis is examined [10, 15, 29]. The emergence of social media sites such as Twitter, where users express a wide range of thoughts and feelings, has led to a boom in interest in sentiment analysis, as noted [23, 24, 30].

Sentiment analysis procedures are dominated by two main strategies: machine learning techniques and the semantic orientation (SO) approach. These tactics are divided into a number of sub-methods and algorithms, each of which has advantages and disadvantages [9, 31]. Prominent sentiment lexicons such as Ml-Senticon, iSOL, SentiWordNet, and eSOL are essential for sentiment extraction using the SO technique. The contextual variety of words is difficult, though, which has led to the creation of domain-specific lexicons, as suggested in [13, 33, 38]. However, labeled datasets are necessary for model training in supervised machine learning techniques [5, 7, 40]. Conversely, supervised machine learning techniques, as explored in [5,7], analyze and evaluate technique-dependency features [2,3,32,37,39], POS-related features [28] and, in some cases, a combination of these factors [4].

To improve the efficiency of classification algorithms, feature extraction techniques were included. But, as noted [11, 12, 16, 17, 36], there are still issues, such as the requirement of substantial labeled datasets in supervised machine learning and the necessity for substantial language resources in SO techniques [26, 27].

Deep learning models hold significant promise for image-based diagnostics, including Covid-19 and cancer detection. To address diagnostic challenges, the author developed a transfer learning-based model using InceptionV3 to detect Monkey-pox, trained on a publicly available dataset, aiding medical professionals [18, 21].Author-enhanced image-classification capabilities used deep features and the InceptionV3 CNN architecture. By employing the CK+, FER2013, and JAFFE datasets, we successfully recognized and categorized emotions [22].

3. Proposed framework

3.1. Problem statement

Patient reviews on healthcare platforms aid in evaluating new medicines' effectiveness and identifying adverse drug reactions (ADRs). These insights are crucial for post-market monitoring and gauging patient satisfaction, helping healthcare stakeholders make informed decisions.

3.2. Dataset description

This research utilized two distinct data sets for analysis. The first set comprised medication reviews sourced from the UCI Machine Learning Repository, renowned for its benchmark data sets. The second data set consisted of tweets collected from Twitter.

3.2.1. Medicine review dataset

The data set consisted of medication reviews by patients along with information about their underlying conditions at the time of drug use. On a scale of 1 to 1, 1 denoting the highest degree of satisfaction, each medicine was assessed. With seven variables and 161,291 entries, the data set mostly focused on patient evaluations. Specifically, 3,436 drug records with user-expressed pleasure or discontent were included in the collection. These evaluations included user testimonies about side effects, adverse effects, and the overall influence of the medication, which expressed how satisfied the user was with it overall.

Table 2
Characterization of dataset features [36]

Series number	Variable	Description
1	Unique ID	User ID
2	Drug name	Evaluated drug brand name
3	Condition	Ailment name
4	Review	User-authored review
5	Rating	Rating (1-5 Stars)
6	Date	Review submission date
7	Useful count	Number of users finding the review informative

Tweets
The sentiment analysis data set comprises 32,047 tweets gathered from various Twitter accounts discussing diabetes, extracted using the Python code. The data set was structured with several distinct fields:

- **Date of the Tweet**
- **Tweet ID**
- **Tweet Text**
- **User handle**
- **User name**

Table 3
An Illustration of Ataset 2's Attributes for Testing Hypotheses

S.No	Variable	Description
1	Tweet. Date	Date of tweet posting
2	Tweet. ID	Unique user ID
3	Tweet. Content	User-written review
4	Tweet. User	Holder's name
5	Username	Username

3.3. Prospective fusion system for sentiment analysis

Both lexical-based and learning-based techniques are used by natural 'language processing (NLP) to analyze sentiments regarding medication evaluations. After preprocessing, the data set was divided into three sentiment classes using lexical methods, with pertinent features being extracted using TF and TF-IDF. These characteristics were used in the training and evaluation of learning model systems, which measure efficiency by F1 score, accuracy, precision, and recall.

To categorize the attitudes of drug reviews, a composite technique was employed, which combines opinion mining with a sentiment-vocabulary-based learning algorithm. Reviewers can be categorized as positive, negative, or neutral using the SenticNet and TextBlob sentiment vocabularies, which evaluate words and provide emotional values. To train and test certain classifiers, the labeled data is divided into practice and test sets (75% and 25%, respectively).

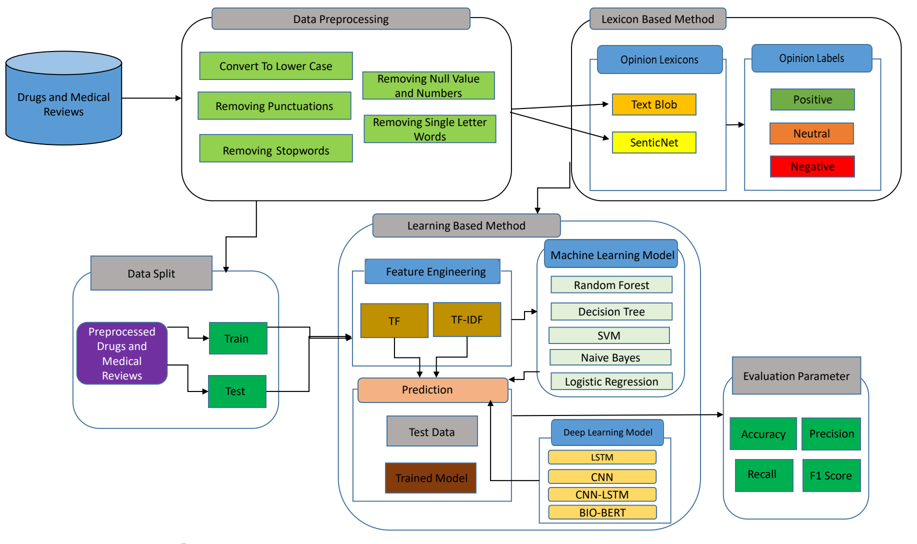


Figure 1. Architecture of the lexicon-based learning model with transformers for sentiment analysis

3.3.1. Pre-processing

The data in this study goes through a number of phases in preparation in order to get it ready for classification:

- **Handle negations:** Negative connotations in tweets need to be eliminated.
- **Convert to lower case:** Any tweets that start with an uppercase letter will be changed to lowercase.
- **Removing punctuation:** Punctuation has been stripped out of the text.

- **Removing single-letter words:** Single-letter terms that were deemed offensive in previous text have been deleted.
- **Removing blank spaces:** The text that had extra spaces between the words was deleted.
- **Removing numbers:** Text containing irrelevant or offensive figures have been deleted.
- **Removing stop words:** Text containing stop words has been deleted.

3.4. Word2Vec model

The Word2Vec model is trained by feeding it a sizable corpus of text. During training, the model learns to modify the word vectors so that words with comparable contexts or meanings are mapped closely together in the vector space. Semantic connections, such as synonyms or related concepts, are represented in the vector space as geometric closeness.

Furthermore, in addition to using the preprocessed text corpus to train a Word2Vec model and produce dense vector representations (embeddings) for every word, these Word2Vec embeddings extract contextual information and semantic connections from the text. Next, depending on their relevance to your job or subject, a subset of significant words or phrases from the Word2Vec embeddings is chosen to generate a lexicon.

3.5. Lexicon-based methods

Combining assumption dictionaries with a set of guidelines for categorizing phrases found in a text into positive, negative, and neutral categories is part of the lexicon technique [37]. Lexicon-based sentiment classification is based on the assessment of the sentiment-bearing phrases' intensity in a given text, allowing one to ascertain the overall sentiment of the text.

One common phrase for a sentiment lexicon is a lexicon that has word polarity values. Every word or phrase in the lexicon has its emotion polarity score annotated on it. Tuples in the sentiment lexicon database are frequently represented using the format (term, sentiment polarity score). An intensity score, which can be positive, negative, or neutral, is included for each word. The text in question is categorized based on this score.

3.5.1. TextBlob

TextBlob is a popular Python-based toolkit used for natural-language processing (NLP) activities on input text. For a range of natural language processing (NLP) operations, including sentiment analysis, entity identification, part-of-speech tagging, translation, and classification, it offers user-friendly APIs.

The authors of [4] use TextBlob as a sentiment analyzer to forecast the tweets' intensity. TextBlob facilitates sentiment analysis with machine learning and the natural language toolkit (NLTK) [12]. There are 2,918 terms in the sentiment lexicon of

TextBlob. You may categorize a text as factual information or a person's sentiment using TextBlob's sentiment analysis tool by grading the text's intensity and subjectivity. This is the format that TextBlob uses when it gives a tuple for a word or phrase:

$$\text{Sentiment} = (\text{polarity\_score}, \text{subjectivity\_score}) \tag{1}$$

where $Rp = [+1., 1.]$ and $Rp = [., 1.]$. Scores for subjectivity and polarity are represented by the float values.

3.5.2. SenticNet

ScenicNet is a language created especially for finding sentiments that are conceptual. Its foundation is sentic computing, a sophisticated, interdisciplinary technique of sentiment analysis. SenticNet, in contrast to other technologies, can integrate emotional and polarity information into complex ideas, such as reaching objectives or remembering important events. Sentiment ratings, ranging from -1 to 1, are currently available for over 14 popular ideas on SenticNet.

Table 4
Sentiment Analysis Parameter Key Value

Series number	Parameter name	Parameter value
1	Positive	Value > 0
2	Negative	Value < 0
3	Neutral	Value $= 0$

4. Feature engineering

Feature engineering is used to separate important characteristics from the preprocessed data in order to increase the efficacy of prediction models on unseen data [38]. By identifying attributes pertinent to the issue statement, this method improves the learning model's performance. The study's researchers [39] demonstrated how feature manipulation outperformed conventional models in terms of performance. There are several ways to manipulate features in textual data.

4.1. Term frequency (TF)

The assessment of a word's frequency inside a certain document is referred to as a "term feature" [16]. You may calculate it by dividing the quantity of words in Document D by the frequency with which each word w appears. It may be represented mathematically as:

$$tf(w, D) = \frac{\text{Occurrences of } w \text{ in } D}{\text{Total number of words in } D} \tag{2}$$

The count vectorizer-function from Python is used in this study to extract word frequencies from the data set.

4.2. TF-IDF (term frequency-inverse document frequency)

Using TF-IDF, a word in a given document is quantified. A word's weight is often calculated based on how relevant it is to the given material. The relevance of a word to a document will increase with its weight score, and vice versa [34]. This is accomplished by combining the metrics of inverse document frequency (IDF) and term frequency (TF), where IDF measures a word's frequency over the whole corpus of texts, while TF measures a word's relevance within a given document. The word "w" in document D has the following mathematical representation in IDF:

$$idf(w, D) = \log\left(\frac{N}{df(w)} + 1\right) \tag{3}$$

where N represents the total number of documents in the corpus and $df(w)$ is the number of documents that include the word w.

$$tf - idf(w, D) = tf(w, D) \times idf(w, D) \tag{4}$$

$$tf - idf(w, D) = tf(w, D) \times \log\left(\frac{N}{df(w)} + 1\right) \tag{5}$$

Where:

- The term frequency of the word w in document D ($TF(w, D)$) counts the frequency of the word w appearing in document D.

- The inverse document frequency (IDF) of the word w evaluates the significance of the term over the whole document collection.

5. Machine learning methods for sentiment categorization

5.1. Support vector machines (SVM)

A technique for binary classification called support vector machine (SVM) looks for the optimal hyperplane to divide data points into distinct groups. SVM may be trained on a feature representation of the text input in order to be used for sentiment analysis. The primary idea of support vector machines (SVM) is to maximize the margin by optimizing the decision boundary to attain the best possible separation between classes.

5.2. Logistic regression

A linear model used for binary classification issues is logistic regression. It expresses the likelihood that an instance falls into a given class using a logistic function.

5.3. Decision tree

In this method, the input data is recursively divided into subgroups according to the characteristics that most effectively distinguish between classes or forecast the target variable. The branches of the tree indicate the potential results of each node, which symbolize a choice made in response to a characteristic.

Information gain: At each decision tree node, this measure aids in choosing the optimal split or attribute. It calculates the amount of entropy or impurity that is reduced when data is divided according to a particular feature.

6. Deep learning models for sentiment classification

Sentiment analysis (SA) uses four-deep learning models: long short-term memory (LSTM), CNN-long short-term memory (LSTM), generative pretrained transformer 2, and BIO bidirectional encoder representations from transformers to assess provided datasets.

6.1. LSTM model

An improved RNN design uses three control gates and a memory cell to efficiently store past data in lengthy sequences. This lessens the impact of training RNNs with several layers on the vanishing gradient issue and the loss of prior data. The LSTM's hierarchical structure is seen in Figure 2.

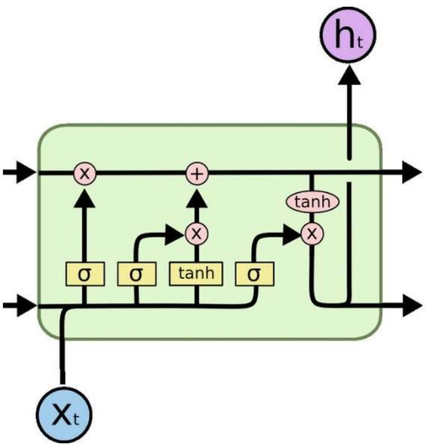


Figure 2. Architecture depiction of LSTM [36]

A memory cell is added to the framework to hold past data. The input gate, forget gate, and output gate are the three gates that control the input, update, and output of past data, respectively. The structure also has a memory cell that stores past data.

6.2. CNN-LSTM model

Initially, I used the CNN-LSTM model for my experiments. The first convolution layer in our CNN-LSTM model takes as input -(word missing?). Subsequently, the output is sent into an LSTM layer after being reduced in dimension. This strategy is justified by the fact that the convolution layer can identify local characteristics that the LSTM layer can use to decipher the input text's sequence. This is the fundamental idea behind this approach. However, in comparison with our alternative LSTM-CNN model, this model performs worse in real-world scenarios.

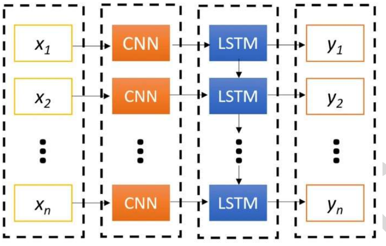


Figure 3. Architecture depiction of CNN LSTM [26]

6.3. Bio-BERT Model

The Bio-BERT Model is fine-tuned for biomedical and clinical text. Pretraining on biomedical literature enhances its performance in sentiment analysis.

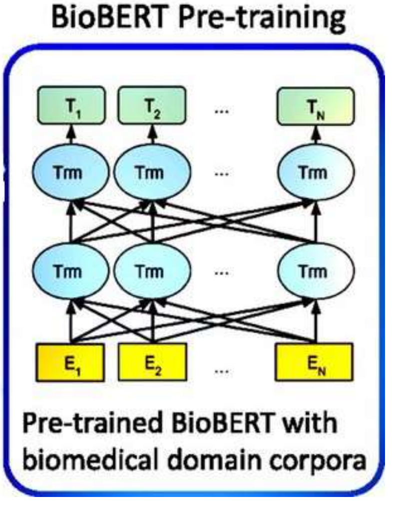


Figure 4. Architecture depiction of Bio-BERT [16]

BioBERT initially learns contextual relationships from biomedical data and is then fine-tuned on sentiment-labeled datasets specific to the domain. During inference, BioBERT generates contextualized word embeddings capturing word meanings within context. These embeddings are fed into a classifier for sentiment prediction. Leveraging domain-specific knowledge, BioBERT accelerates sentiment analysis work and offers nuanced understanding and interpretation of sentiment in biomedical texts.

6.4. GPT-2

Expanding on the transformational transformer design, OpenAI created the very advanced generative pretrained transformer 2 (GPT-2) architecture. It is made up of many stacked transformer blocks with feed-forward neural networks and built-in self-attention mechanisms. GPT-2 uses self-attention to assess each word's significance in relation to the complete input sequence and positional encoding to maintain the tokens' sequential order. These elements provide the model with the ability to represent complex semantic links and dependencies in language, which helps it perform well on a range of natural-language processing tasks. Furthermore, to improve training stability, information flow, and coherent text creation, GPT-2 incorporates layer normalization, residual connections, and softmax activation functions in its output layer.

7. Evaluation parameters

It is crucial to assess the model's performance following evaluation and training. Four different results can be obtained from a categorization model:

- **True Positives (TP)**: Situations in which a favorable conclusion is accurately predicted.
- **True Negatives (TN)**: Predicted negative cases that really fall under the negative category.
- **False Positives (FP)**: Events that are wrongly classified as positive when they are actually negative.
- **False Negatives (FN)**: Instances that are positively predicted but are mistakenly classified as negative.

Four measures are usually used to assess the performance of the classifier: accuracy, precision, recall, and F1-score.

7.1. Accuracy

The traditional metric for evaluating sentiment analysis performance is accuracy, which is the percentage of tweets that are properly classified out of all tweets.

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{True Positive} + \text{True Negative} + \text{False Positive} + \text{False Negative}} \tag{6}$$

7.2. Precision

The accuracy of positive predictions among all positive classifications is indicated by precision, which is the ratio of properly predicted positive tweets to the total number of tweets classified as positive.

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \tag{7}$$

7.3. Recall

Recall is computed by dividing the total number of tweets that are actually positive by the total number of tweets labeled as positive. Recall is also referred to as sensitivity or true positive rate. It assesses the model's accuracy in identifying positive cases among all real positive instances.

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \tag{8}$$

7.4. F1-score

The F1-score is frequently regarded as the most accurate indicator of a model's performance, as it incorporates recall and accuracy into a single statistic. The harmonic mean of recall and accuracy is used to calculate it. The performance of an improved model may be evaluated by looking at its F1 score. You may use the following formula to determine the F1 score:

$$F1 - \text{score} = \frac{2 \cdot (\text{Precision} \cdot \text{Recall})}{\text{Precision} + \text{Recall}} \tag{9}$$

In this study, we employed a hybrid approach to categorize sentiment in drug reviews. This method integrates a semantic lexicon that has proven crucial for sentiment analysis. The proposed technique is employed to develop a system that automatically interprets unlabeled data using a sentiment lexicon and categorizes emotions for drug reviews and patient feedback assessment using a trained model.

8. Fine-tuning phase

Many tests were run in order to maximize the model's performance throughout the fine-tuning phase. To get the best outcomes, several setups and hyperparameter adjustments were made. To improve training efficiency and model convergence, for example, the optimizer, number of epochs, batch size, and learning rate were adjusted. Furthermore, a variety of neural network topologies and layers were tested to find the best setup for sentiment analysis on medication evaluations. Additionally, cross-validation methods were used to verify the model's performance over many dataset folds, guaranteeing its generalizability and resilience. In addition, the best combination was found by methodically exploring the hyper-parameter space using grid search and random search techniques.

9. Result and evaluation

This section provides the straightforward test components, including the hyper-parameters of the unique classification computations and a depiction of the testing system. The Intel Core i7 8th generation computer, which has a quad-core CPU and 8GB of random access memory running Windows 1, is used for the tests. The Python programming language is utilized in conjunction with various tools, including the Boa Constrictor 3 application, Spyder notebook, Google Colab, and others.

In this section, using two different datasets, we discuss the outcomes of the hybrid strategy for each feature building technique. Findings are presented when we use patient reviews as a source of information.

9.1. Patient reviews

The initial study delves into employing a lexicon-driven technique to assign sentiment labels to 161,291 unmarked medical reviews. Based on the sentiment score, patient reviews are categorized as traditional sentiment. Table 5 illustrates the distribution of reviews across each sentiment category.

Table 5
Number of Examples for Each Sentiment Lexicon Under Each Label Distribution

Method	Positive	Neutral	Negative
TextBlob	98255	8426	54610
SenticNet	97602	1820	61869

9.1.1. Evaluation of machine learning models using TextBlob reviews

Table 6
Performance Metrics for Different Machine Models Using TF

Machine Model	Accuracy	Precision	Recall	F1 Score
TB+LR	95	96	97	96
TB+DT	89	92	91	92
TB+SVM	89	93	92	92

Table 7
Performance Metrics for Different Machine Models Using TF-IDF

Machine model	Accuracy	Precision	Recall	F1 Score
TB+LR	93	92	97	95
TB+DT	89	90	92	91
TB+SVM	88	89	94	92

9.1.2. Assessment of machine learning models using SenticNet

The suggested combined feature engineering approach performs noticeably better on average, according to the SenticNet sentiment analysis. TF outperforms other feature engineering procedures in terms of classification accuracy, as seen in Figure 8. To be more precise, SenticNet with LR results in greater accuracy, with TF at 93% and TF-IDF at 91% precision. Table 7 shows that SenticNet obtains the greatest accuracy of 93% with TF when paired with LR. On the other hand, TF-IDF characteristics also demonstrate a high accuracy of 91% when combined with SenticNet and LR. The maximum F1 score of 95 is obtained when SenticNet, LR, and TF characteristics are combined.

Table 8

Performance Evaluation of Classification Models Using SenticNet Sentiments (Using TF)

Machine model	Accuracy	Precision	Recall	F1 Score
TF SN + LR	93	94	95	95
SN + DT	81	85	85	85
SN + SVM	87	91	89	90

Table 9

Performance Evaluation of Classification Models Using SenticNet Sentiments (Using TF-IDF)

Machine model	Accuracy	Precision	Recall	F1 Score
TF-IDF SN + LR	91	91	94	93
SN + DT	82	85	85	85
SN + SVM	86	88	90	89

Table 10

Comparison of Machine Learning Classification Models Using SenticNet and TextBlob Sentiments

Machine models	TF	TF-IDF		
	TextBlob	SenticNet	TextBlob	SenticNet
LR	95	93	93	91
DT	89	81	89	82
SVM	89	87	88	86

9.1.3. Exploring deep learning models using SenticNet and TextBlob sentiments

Table 11
Performance Analysis of Deep-Learning Classification Models Using SenticNet Sentiments

Deep Learning Model	Accuracy	Precision	Recall	F1 Score
SN+LSTM	91	91	95	93
SN+GPT-2	91	88	89	88
SNt+CNN+LSTM	93	93	93	93
SN+Bio Bert	95	96	94	96

Table 12
Performance Analysis of Deep-Learning Classification Models Using TextBlob Sentiments

Deep Learning Model	Accuracy	Precision	Recall	F1 Score
TB+LSTM	96	97	97	97
TB+GPT-2	96	94	93	93
TB+CNN+LSTM	97	96	96	96
TB+Bio Bert	97	97	96	96

9.1.4. Comparative analysis of learning models using SenticNet and TextBlob sentiments

Tables 8, 9, and 10 show that TextBlob combined with Bio-BERT yields the greatest sentiment classification accuracy, of 97%, followed by TextBlob, combined with CNN+LSTM at 97% and TextBlob employing LR with TF, at 95%. This demonstrates efficacy learning. TextBlob outperforms in terms of its pattern analyzer features and improves word sense disambiguation more accurately when combined with any of the two-feature engineering methodologies. On the other hand, SenticNet has not kept up with other sentiment lexicons because of its inability to handle brief sentences, which are typical of micro blogs.

Due learning models are surpassed by LR due to its resilience. However, tree-based models like DT perform poorly compared to LR. Tree-based models are prone to overfitting and sampling mistakes, which can lead to poor results if there is a significant discrepancy between the training and testing data. Furthermore, there is a relationship between the size of the feature set and the models' effectiveness. When the feature set size is greater than the training examples, LR models can be trained more effectively than tree-based models. As a result, TF has performed better, as it has a larger set of features.

9.2. Twitter tweets

9.2.1. Exploration of medication surveys using a lexicon-based approach

The first study looks at classifying sentiment in 3,247 untagged medical reviews using a lexicon-driven technique. Based on the sentiment rating—a score of ?? (needs to be added) denotes a neutral sentiment—drugs are evaluated and classified as either positive or negative. Table 11 shows the distribution of the reviews assigned to each sentiment group.

Table 13
Category-wise Count of Sentiment for Each Lexicon

Approach	Positive	Neutral	Negative
TextBlob	16102	5278	10667
SenticNet	20232	10395	1420

9.2.2. Performance evaluation of machine learning models using TextBlob sentiments

Table 14
Performance Analysis of Classification Models Using Text Blob Sentiments

Machine model	Accuracy	Precision	Recall	F1 Score
TF				
TB+LR	89	93	92	93
TB+DT	90	91	92	91
TB+SVM	92	95	93	94
TF-IDF				
TB+LR	85	87	91	89
TB+DT	78	85	81	83
TB+SVM	90	92	92	92

9.2.3. Evaluation of machine learning models using SenticNet opinions

Based on SenticNet's exploratory results, the hybrid feature engineering strategy that has been proposed performs better overall. Table 16 shows that the best accuracy of 92 for TF is obtained when SenticNet and SVM are coupled. On the other hand, TF-IDF features show the greatest accuracy rate of 93 when combined with SenticNet and SVM. The greatest F1 score of 91 is displayed by TF features combined with SenticNet and SVM.

Table 15
Performance Analysis of Classification Models Using SenticNet Sentiments

Machine Model	Accuracy	Precision	Recall	F1 Score
TF				
SN+LR	87	90	92	91
SN+DT	72	79	78	79
SN+SVM	87	92	90	91
TF-IDF				
SN+LR	83	82	95	88
SN+DT	69	77	75	76
SN+SVM	85	86	93	89

Table 16
Result Analysis of Categorization by Models with SenticNet and TextBlob Sentiments

Machine Model	TF (TextBlob)	TF (SenticNet)	TF-IDF (TextBlob)	TF-IDF (SenticNet)
LR	89	87	85	83
DT	90	72	78	69
SVM	92	87	90	85

9.2.4. Exploring deep learning models using SenticNet and TextBlob sentiments

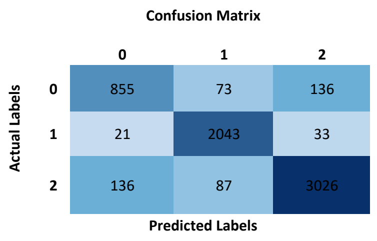
Table 17
Performance Analysis Deep Learning Classification Models Using SenticNet Sentiments

Deep Learning Model	Accuracy	Precision	Recall	F1 Score
SN+LSTM	82	84	93	91
SN+GPT-2	84	84	86	85
SN+CNN+LSTM	85	86	85	85
SN+Bio BERT	87	84	83	84

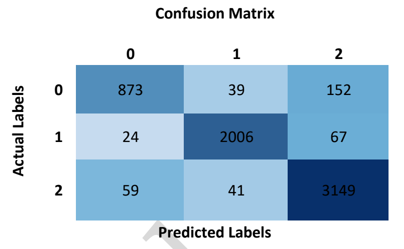
Table 18
Performance Analysis of Deep Learning Classification Models Using TextBlob Sentiments

Deep Learning Model	Accuracy	Precision	Recall	F1 Score
TB+LSTM	94	94	97	95
TB+GPT-2	93	91	93	93
TB+CNN+LSTM	92	92	92	92
TB+Bio BERT	97	96	97	96

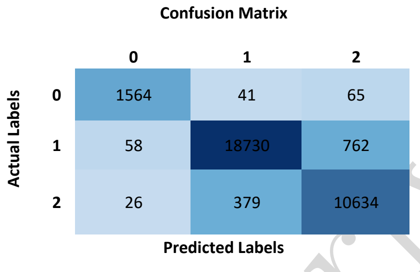
9.2.5. Comparative study learning models using SenticNet and TextBlob opinions



(a))**TextBlob+SVM (Tweets)**

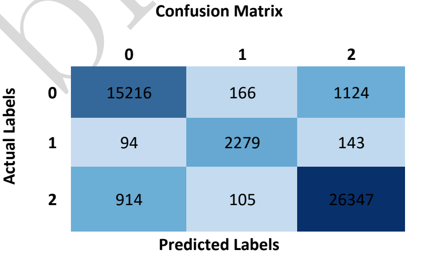


(b))**TextBlob+LSTM (Tweets)**

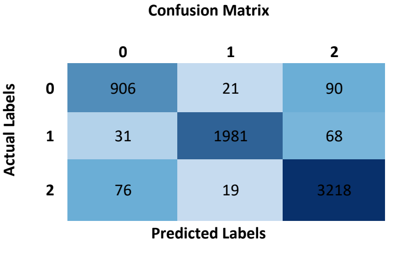


(c))**TextBlob+CNN+LSTM (Drug Reviews)**

0



(d))**TextBlob+LR (Drug Reviews)**



(e))**TextBlob+BioBERT (Tweets)**

Figure 5. Confusion matrix for the most accurate classification models to analyze positive, neutral, and negative predictions for drug reviews and tweets

According to Tables 14, 15, and 16, TextBlob with Bio Bert achieves the highest sentiment classification accuracy of 97, followed by TextBlob with LSTM 94 and TextBlob with SVM and TF 92. It illustrates how well a hybrid approach combining lexicon- and learning-based techniques (machine and deep learning) works. TextBlob performs better in terms of its pattern analyzer capabilities and more precisely uses word sense clarification when paired with any of the two-feature engineering methodologies. Conversely, SenticNet has fared worse than other sentiment lexicons when it comes to being limited to sentences that are the length of micro blogs.

9.2.6. Comparative study of a proposed framework with exiting work

By contrasting the hybrid approach's sentiment categorization outcomes with those of other cutting-edge techniques, its efficacy is assessed. This comparison is mostly concerned with sentiment analysis of drug reviews. The comparative effectiveness of several sentiment analysis methodologies is shown in Table 17. The authors of [42] suggest an adaption technique utilizing bag-of-words characteristics with little preprocessing to improve classification accuracy. In order to analyze domain-specific sentiment expressions they examine word weighting schemes; they find that the delta TF-IDF score works very well. In [12], the feature extraction techniques of count vectorizer and TF-IDF are coupled to analyze the sentiment of medication reviews. There are several classifiers used, such as ANN, LSTM, GRU, SVM, LR, and RF. With the integration of both generic and domain-specific lexicons, (word mossong) uses a hybrid approach for sentiment categorization and shows encouraging results. Furthermore, [14] proposes a fusion method that uses classifiers such as LR, ADA, MLP, and ETC to combine the lexicons of AFFIN, VADAR, and TextBlob. Table 17 shows that the present method outperforms the previous in terms of effectiveness.

Table 19

Comparative Study of Proposed Framework with Existing Work

S.No	References	Dataset	Method	Feature	A	P	R	F
1	[42]		ANN, LSTM, GRU, SVM, LR, RF	TF-IDF, Count vectorizer	93	95	95	90
2	[12]	Drug Review (50000)	Lexicon combination and information gain	Uni and bigram	91	76	53	62
3	[17]	Drug Review (5600)	RF, SVM, NB, RBFN	SWN and Position encoding	65	58	62	–
4	[36]	Drug Review (26060)	Corpus based sentiment classification	Uni, bigram and trigram	89	79	83	–
5	Proposed	r	TextBlob+BioBERT, TextBlob+CNN-LSTM	TF and TF-IDF	97, 97	97, 96	96, 97	98, 96
6	Proposed	Medical Review (32047)	TextBlob+BioBERT, TextBlob+LSTM	TF and TF-IDF	97, 94	97, 93	97, 94	97, 93

10. Conclusion and future work

Research addresses the problems of domain specificity of sentiment lexicons and the requirement for manual annotation in a learning-based approach to assess users' emotions. It does this by combining learning-based and lexicon-based approaches for labeling and categorization. TextBlob and SenticNet are evaluated as data annotation techniques for two datasets: medication reviews and patient tweets. Experimental investigation shows that TextBlob frequently yields better results when annotating drug reviews.

TF and TF-IDF are the two recognized techniques for feature engineering. Three machine-learning models (LSTM, SVM, DT) and four deep-learning models (LSTM, GPT-2, CNN-LSTM, and Bio BERT) are used to classify sentiment. Many featuring engineering and annotation procedures are used to evaluate their efficacy. When compared to all other machine- and deep-learning models in a deep learning model, CNN-LSTM and Bio-BERT yield the most accurate results. The results show that SVM and LR machine learning models perform well when performed on term frequency and term frequency-inverse document frequency with TextBlob lexicons.

The suggested approach in the empirical results of this study produced 4% better accuracy results for drug review sentiment analysis than previous approaches. Furthermore, experiments on datasets from other domains show the utility of the proposed approach. Future study along these lines might greatly improve the system's functionality and usefulness. More specifically, improving the hybrid framework and investigating other methods will be essential in enhancing medical text sentiment analysis. The present hybrid architecture, which combines several learning models, has already demonstrated benefits in terms of increasing accuracy and capturing intricate emotion patterns.

Nevertheless, to improve sentiment detection's resilience across a variety of datasets, more advanced methods, such as ensemble learning (using majority voting, stacking, or boosting), may be the subject of future research. Additionally, transfer learning using pretrained models customized for medical contexts, such as BERT or GPT, may improve the system's ability to comprehend a complex language. The framework may be further optimized with improvements in feature engineering, especially by using medical lexicons and domain expertise.

Researching other approaches, such as multitask learning or graph-based models such as graph neural networks (GNNs), might help illustrate the connections between medical items and attitudes more accurately. Furthermore, an interactive method for constantly optimizing categorization based on feedback mechanisms may be introduced using reinforcement learning. Improving contextual comprehension using context-aware models augmented by attention mechanisms or transformers is another important line of inquiry. These models would be better able to manage complicated sentiment relationships and long-term dependencies seen in medical texts. As multilingual medical datasets increase, these systems will also need to be cross-lingually adapted.

References

[1] Aakur S.N., Sarkar S.: Leveraging symbolic knowledge bases for commonsense natural language inference using pattern theory, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45(11), pp. 13185–13202, 2023. doi: 10.1109/tpami.2023.3287837.

[2] Agarwal B., Mittal N.: Prominent feature extraction for review analysis: an empirical study, *Journal of Experimental & Theoretical Artificial Intelligence*, vol. 28(3), pp. 485–498, 2016.

[3] Agarwal B., Poria S., Mittal N., Gelbukh A., Hussain A.: Concept-level sentiment analysis with dependency-based semantic parsing: a novel approach, *Cognitive Computation*, vol. 7, pp. 487–499, 2015. doi: 10.1007/s12559-014-9316-6.

[4] Alharbi N.M., Alghamdi N.S., Alkhammash E.H., Al Amri J.F.: Evaluation of sentiment analysis via word embedding and RNN variants for Amazon online reviews, *Mathematical Problems in Engineering*, vol. 2021(1), 5536560, 2021. doi: 10.1155/2021/5536560.

[5] Ali T., Schramm D., Sokolova M., Inkpen D.: Can i hear you? sentiment analysis on medical forums. In: *Proceedings of the sixth international joint conference on natural language processing*, pp. 667–673, 2013.

[6] Birjali M., Kasri M., Beni-Hssane A.: A comprehensive survey on sentiment analysis: Approaches, challenges and trends, *Knowledge-Based Systems*, vol. 226, 107134, 2021. doi: 10.1016/j.knosys.2021.107134.

[7] Biyani P., Caragea C., Mitra P., Zhou C., Yen J., Greer G.E., Portier K.: Co-training over domain-independent and domain-dependent features for sentiment analysis of an online cancer support community. In: *Proceedings of the 2013 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining*, pp. 413–417, 2013. doi: 10.1145/2492517.2492606.

[8] Choudhary R.R., Jisnu K., Meena G.: Image dehazing using deep learning techniques, *Procedia Computer Science*, vol. 167, pp. 1110–1119, 2020. doi: 10.1016/j.procs.2020.03.413.

[9] Deng S., Sinha A.P., Zhao H.: Adapting sentiment lexicons to domain-specific social media texts, *Decision Support Systems*, vol. 94, pp. 65–76, 2017. doi: 10.1016/j.dss.2016.11.001.

[10] Fernández-Gavilanes M., Álvarez-López T., Juncal-Martínez J., Costa-Montenegro E., González-Castaño F.J.: Unsupervised method for sentiment analysis in online texts, *Expert Systems with Applications*, vol. 58, pp. 57–75, 2016. doi: 10.1016/j.eswa.2016.03.031.

[11] Fu X.L., Wu J., Chen J., Liu S.: Attribute-Sentiment Pair Correlation Model Based on Online User Reviews, *Journal of Sensors*, vol. 2019(1), 2456752, 2019. doi: 10.1155/2019/2456752.

[12] Goeuriot L., Na J.C., Min Kyaing W.Y., Khoo C., Chang Y.K., Theng Y.L., Kim J.J.: Sentiment lexicons for health-related opinion mining. In: *Proceedings of the 2nd ACM SIGHIT International Health Informatics Symposium*, pp. 219–226, 2012. doi: 10.1145/2110363.2110390.

[13] Habernal I., Ptáček T., Steinberger J.: Supervised sentiment analysis in Czech social media, *Information Processing & Management*, vol. 50(5), pp. 693–707, 2014. doi: 10.1016/j.ipm.2014.05.001.

[14] Hassan M.E., Hussain M., Maab I., Habib U., Khan M.A., Masood A.: Detection of sarcasm in Urdu tweets using deep learning and transformer based hybrid approaches, *IEEE Access*, vol. 12, pp. 61542–61555, 2024. doi: 10.1109/access.2024.3393856.

[15] Jain P.K., Pamula R., Srivastava G.: A systematic literature review on machine learning applications for consumer sentiment analysis using online reviews, *Computer science review*, vol. 41, 100413, 2021. doi: 10.1016/j.cosrev.2021.100413.

[16] Lee J., Yoon W., Kim S., Kim D., Kim S., So C.H., Kang J.: BioBERT: a pre-trained biomedical language representation model for biomedical text mining, *Bioinformatics*, vol. 36(4), pp. 1234–1240, 2020.

[17] Liu S., Lee I.: Extracting features with medical sentiment lexicon and position encoding for drug reviews, *Health information science and systems*, vol. 7, pp. 1–10, 2019. doi: 10.1007/s13755-019-0072-6.

[18] Meena G., Mohbey K.K.: Sentiment analysis on images using different transfer learning models, *Procedia Computer Science*, vol. 218, pp. 1640–1649, 2023. doi: 10.1016/j.procs.2023.01.142.

[19] Meena G., Mohbey K.K., Acharya M., Lokesh K.: An improved convolutional neural network-based model for detecting brain tumors from augmented MRI images, *Journal of Autonomous Intelligence*, vol. 6(1), 2023. doi: 10.32629/jai.v6i1.561.

[20] Meena G., Mohbey K.K., Indian A., Khan M.Z., Kumar S.: Identifying emotions from facial expressions using a deep convolutional neural network-based approach, *Multimedia Tools and Applications*, vol. 83(6), pp. 15711–15732, 2024.

[21] Meena G., Mohbey K.K., Kumar S.: Monkeypox recognition and prediction from visuals using deep transfer learning-based neural networks, *Multimedia Tools and Applications*, vol. 83(28), pp. 71695–71719, 2024. doi: 10.1007/s11042-024-18437-z.

[22] Meena G., Mohbey K.K., Kumar S., Chawda R.K., Gaikwad S.V.: Image-based sentiment analysis using InceptionV3 transfer learning approach, *SN Computer Science*, vol. 4(3), 242, 2023. doi: 10.1007/s42979-023-01695-3.

[23] Molina-González M.D., Martínez-Cámara E., Martín-Valdivia M.T., Perea-Ortega J.M.: Semantic orientation for polarity classification in Spanish reviews, *Expert systems with applications*, vol. 40(18), pp. 7250–7257, 2013. doi: 10.1016/j.eswa.2013.06.076.

[24] Montejo-Ráez A., Martínez-Cámara E., Martín-Valdivia M.T., Ureña-López L.A.: Ranked wordnet graph for sentiment polarity classification in twitter, *Computer Speech & Language*, vol. 28(1), pp. 93–107, 2014. doi: 10.1016/j.csl.2013.04.001.

[25] Mujahid M., Kanwal K., Rustam F., Aljedaani W., Ashraf I.: Arabic ChatGPT tweets classification using RoBERTa and BERT ensemble model, *ACM Transactions on Asian and Low-Resource Language Information Processing*, vol. 22(8), pp. 1–23, 2023. doi: 10.1145/3605889.

[26] Mujahid M., Lee E., Rustam F., Washington P.B., Ullah S., Reshi A.A., Ashraf I.: Sentiment analysis and topic modeling on tweets about online education during COVID-19, *Applied Sciences*, vol. 11(18), 8438, 2021. doi: 10.3390/app11188438.

[27] Mujahid M., Rustam F., Shafique R., Chunduri V., Villar M.G., Ballester J.B., Diez I.d.l.T., Ashraf I.: Analyzing sentiments regarding ChatGPT using novel BERT: A machine learning approach, *Information*, vol. 14(9), p. 474, 2023. doi: 10.3390/info14090474.

[28] Na J.C., Kyaing W.Y.M., Khoo C.S., Foo S., Chang Y.K., Theng Y.L.: Sentiment classification of drug reviews using a rule-based linguistic approach. In: *The Outreach of Digital Libraries: A Globalized Resource Network: 14th International Conference on Asia-Pacific Digital Libraries, ICADL 2012, Taipei, Taiwan, November 12-15, 2012, Proceedings 14*, pp. 189–198, Springer, 2012. doi: 10.1007/978-3-642-34752-8_25.

[29] Nakov P., Rosenthal S., Kiritchenko S., Mohammad S.M., Kozareva Z., Ritter A., Stoyanov V., Zhu X.: Developing a successful SemEval task in sentiment analysis of Twitter and other social media texts, *Language Resources and Evaluation*, vol. 50, pp. 35–65, 2016. doi: 10.1007/s10579-015-9328-1.

[30] Ofek N., Caragea C., Rokach L., Biyani P., Mitra P., Yen J., Portier K., Greer G.: Improving sentiment analysis in an online cancer survivor community using dynamic sentiment lexicon. In: *2013 international conference on social intelligence and technology*, pp. 109–113, IEEE, 2013. doi: 10.1109/society.2013.20.

[31] Park S., Lee W., Moon I.C.: Efficient extraction of domain specific sentiment lexicon with active learning, *Pattern Recognition Letters*, vol. 56, pp. 38–44, 2015. doi: 10.1016/j.patrec.2015.01.004.

[32] del Pilar Salas-Zárate M., López-López E., Valencia-García R., Aussenac-Gilles N., Almela Á., Alor-Hernández G.: A study on LIWC categories for opinion mining in Spanish reviews, *Journal of Information Science*, vol. 40(6), pp. 749–760, 2014. doi: 10.1177/0165551514547842.

[33] del Pilar Salas-Zárate M., Paredes-Valverde M.A., Limon J., Tlapa D.A., Báez Y.A.: Sentiment Classification of Spanish Reviews: An Approach based on Feature Selection and Machine Learning Methods., *J Univers Comput Sci*, vol. 22(5), pp. 691–708, 2016.

[34] Poria S., Hazarika D., Majumder N., Mihalcea R.: Beneath the tip of the iceberg: Current challenges and new directions in sentiment analysis research, *IEEE transactions on affective computing*, vol. 14(1), pp. 108–132, 2020. doi: 10.1109/taffc.2020.3038167.

[35] Rahman M.A., Begum M., Mahmud T., Hossain M.S., Andersson K.: Analyzing sentiments in elearning: A comparative study of bangla and romanized bangla text using transformers, *IEEE Access*, vol. 12, pp. 89144–89162, 2024. doi: 10.1109/access.2024.3419024.

[36] Saad E., Din S., Jamil R., Rustam F., Mehmood A., Ashraf I., Choi G.S.: Determining the efficiency of drugs under special conditions from users' reviews on healthcare web forums, *IEEE Access*, vol. 9, pp. 85721–85737, 2021. doi: 10.1109/access.2021.3088838.

[37] Saif H., He Y., Alani H.: Alleviating data sparsity for twitter sentiment analysis. In: *Making Sense of Microposts*, CEUR Workshop Proceedings (CEUR-WS. org), 2012.

[38] Saleh M.R., Martín-Valdivia M.T., Montejo-Ráez A., Ureña-López L.: Experiments with SVM to classify opinions in different domains, *Expert Systems with Applications*, vol. 38(12), pp. 14799–14804, 2011. doi: 10.1016/j.eswa.2011.05.070.

[39] Sharif H., Zaffar F., Abbasi A., Zimbra D.: Detecting adverse drug reactions using a sentiment classification framework, *Computer Science, Medicine*, 2014.

[40] Smith P., Lee M.: Cross-discourse development of supervised sentiment analysis in the clinical domain. In: *Proceedings of the 3rd workshop in computational approaches to subjectivity and sentiment analysis*, pp. 79–83, 2012.

[41] Sudheesh R., Mujahid M., Rustam F., Mallampati B., Chunduri V., de la Torre Díez I., Ashraf I.: Bidirectional encoder representations from transformers and deep learning model for analyzing smartphone-related tweets, *PeerJ Computer Science*, vol. 9, e1432, 2023. doi: 10.7717/peerj-cs.1432.

[42] Vijayaraghavan S., Basu D.: Sentiment analysis in drug reviews using supervised machine learning algorithms, *arXiv preprint arXiv:200311643*, 2020.

[43] Wankhade M., Rao A.C.S., Kulkarni C.: A survey on sentiment analysis methods, applications, and challenges, *Artificial Intelligence Review*, vol. 55(7), pp. 5731–5780, 2022. doi: 10.1007/s10462-022-10144-1.

Affiliations

Anuj Kumar
GLA University Mathura, India, anujkumar.gla@gla.ac.in

Rakesh Kumar
GLA University Mathura, India, rakesh.kumar@gla.ac.in

Shashi Shekhar
AMITY University Patna, India, sshekhar1@ptn.amity.edu

Received: 19.07.2024
Revised: 22.09.2024
Accepted: 15.08.2025