Archives available at journals.mriindia.com

International Journal on Advanced Computer Theory and Engineering

ISSN: 2319 - 2526

Volume 15 Issue 01s, 2026

A Comparative Study of Sentiment Analysis Techniques for Hindi Text: Machine Learning vs Deep Learning

[1]Shalini V. Sathe, [2]Abhijeet G. Dhepe, [3]Dr. Sunil S. Nimbhore

[1,2,3] Department of Computer Science and IT, Dr. BAM University Chh. Sambhajinagar, India

Email: [1]Shalinisathe55@gmail.com, [2]abhijeetdhepe143@gmail.com, [3]nimbhoress@gmail.com

Peer Review Information

Submission: 08 Dec 2025

Revision: 25 Dec 2025

Acceptance: 10 Jan 2026

Keywords

Sentiment Analysis, ML, DL, SVM, RNN, LSTM, BERT, NLP.

Abstract

The increase in interest in sentiment analysis (SA) is presently supported by the demand for extracting information from texts, as seen in social media discussions and individual reviews. Languages like Hindi, however, are not easy to analyze for sentiment analysis because the morphology of these languages is complex and has its own syntax. This paper proposes a comparison of two popular sentiment analysis approaches for Hindi text: traditional machine learning (ML) approaches and more advanced deep learning (DL) approaches. We compare other ML algorithms, including Naive Bayes, Support Vector Machines (SVM), and Logistic Regression, with DL models such as Recurrent Neural Networks (RNN), Long Short-Term Memory (LSTM), and a specific application of an Artificial Neural Network for automating the measurement of kinematic characteristics of punches in boxing. Applied Sciences. Transformer-based BERT. With a dataset of sentiment analysis of the Hindi language, the research intends to determine the strengths and weaknesses of each of the methods. Our findings reveal that although deep learning algorithms, particularly BERT, are more accurate and possess a stronger contextualized understanding, the machine learning algorithm is computationally efficient. Such results have a significant implication regarding undertaking language sentiment analysis in languages whose morphology is rich, e.g., Hindi.

Introduction

Sentiment analysis (SA) is a branch of natural-language processing (NLP) used to identify and extract subjective text information and classify it into preselected positive, negative, or neutral categories. Having established roles in social media monitoring and customer feedback analysis, sentiment analysis can prove vital in drawing inferences about user opinion and enhancing decision-making mechanisms [1]. However, alongside the advances in sentiment analysis for languages like English, there is a challenge in applying sentiment analysis tools to Hindi, which is complicated by morphological complexities and various syntactic peculiarities, as well as differing writing styles across its speaking regions.Hindi, among the most widely spoken languages in the world, has its linguistic characteristics, such as complex word formation, making it challenging to perform sentiment classification. The traditional machine learning (ML) approaches, including Naive Bayes (NB), Support Vector Machines (SVM), and Logistic Regression, have also proved to be useful in sentiment analysis of Hindi text [4], [5]. The methods are highly dependent on feature engineering (Bag of Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) methods of text representation) [6]. Nonetheless, these methods have had average success in capturing the semantic relations within texts [7]. The new developments in Deep Learning (DL) solutions like Recurrent neural network (RNN), Long short-term memory (LSTM), and transformers like BERT have shown almost similar improvements on NLP problems, especially in the case of morphologically rich languages such as Hindi [8], [9]. The DL models can extract patterns within the data without needing predefined features and are, therefore, very efficient in sentiment analysis [10]. The paper compares machine and deep learning techniques for Hindi sentiment analysis of text. By comparing a few models regarding accuracy, precision, recall, training time, and running costs, we expect to bring out the strengths and weaknesses of any given method. The paper offers a comparative study between machine learning and deep learning models, focusing on Hindi text sentiment analysis. While the comparison is relevant, it is not particularly groundbreaking. Several studies in the field, such as Kumar et al. (2019) on SVM and Pandey et al. (2021) on BERT, have explored similar methods. Nonetheless, the contribution is valuable for setting benchmarks specific to Hindi sentiment analysis.

Overview of the Paper, Key Contributions, and Objectives:
The paper explores the challenges of sentiment analysis (SA) for Hindi text, particularly due to its complex morphology and syntax. The authors compare traditional machine learning models (Naive Bayes, SVM, Logistic Regression) with advanced deep learning models (RNN, LSTM, BERT) using a Hindi dataset with 15,000 labeled samples categorized as positive, negative, or neutral. The study evaluates key metrics such as accuracy, precision, recall, F1-score, training time, and computational efficiency. The goal is to identify which techniques work best for sentiment analysis of morphologically rich languages like Hindi.

Related Work
Sentiment analysis of the Hindi language has been the subject of active research, and numerous studies have been conducted on both classical machine learning and recent deep learning-based methods. Table I presents a comparison of pertinent studies and techniques.

Table1: Literature Review

Sr. No.	Propose Study	Approach	Techniques	Key Findings	Challenges
1	Kumar et al. [4]	Machine Learning	SVM, Naive Bayes, Random Forest	SVM outperformed Naive Bayes in Hindi sentiment analysis.	Requires considerable feature engineering efforts.
2	Gupta et al. [5]	Deep Learning	LSTM, RNN	LSTM demonstrated better performance than traditional ML models.	Long training times.
3	Sharma et al. [6]	Hybrid Approach	ML and DL Hybrid	Combining SVM and LSTM produced improved results in Hindi.	Complex hybrid model.
4	Pandey et al. [7]	Deep Learning	BERT, Transformer	BERT achieved the highest accuracy for Hindi sentiment classification.	Requires substantial computing power.
5	Mishra et al. [8]	Machine Learning	Naive Bayes, SVM	Naive Bayes performed well in small datasets for Hindi sentiment.	Struggled with larger, diverse datasets.
6	Jadhav et al. [10]	Deep Learning	CNN, LSTM	CNNs outperformed RNNs in handling short, context-specific reviews.	Limited to short texts.

7	Agarwal et al. [11]	Transfer Learning	BERT, GPT-2	Fine-tuning BERT on Hindi corpora improved sentiment classification.	Overfitting issues on small datasets.
8	Sahu et al. [12]	Machine Learning	SVM, Random Forest	Random Forest was more accurate in Hindi reviews with fewer words.	Performance drops with large datasets.
9	Yadav et al. [13]	Deep Learning	LSTM, BiLSTM	BiLSTM showed better results on sentiment polarity for Hindi.	High computational costs.
10	Deshmukh et al. [14]	Hybrid Approach	CNN, BiLSTM	Combining CNN and BiLSTM achieved robust results.	Limited data scalability.

Methodology

A. Dataset Collection

Compared to this work, the dataset used was a publicly available dataset of Hindi text that included social media posts and customer reviews. The dataset consists of around 15,000 labelled samples, which are sorted into three categories of sentiments: positive, negative, and neutral feelings. Pre-processing of the data is done through tokenizing the text, eliminating stop words, and even stemming out words to get to their root levels [15].

B. Feature Extraction

In the case of a machine learning model, extracted text features are used to represent the text. Bag of Words (BoW) converts the textual data into a sparse vector with words being treated as features [16]. TF-IDF calculates how the words in the document are relevant to the whole corpus, emphasizing the words that are less common in the data. Word Embeddings (e.g., Word2Vec) are structures utilized in writing words in dense representations [18]. Regarding deep learning, text data undergoes a processing procedure in the form of embedding layers that learn the correlation among words themselves [19].

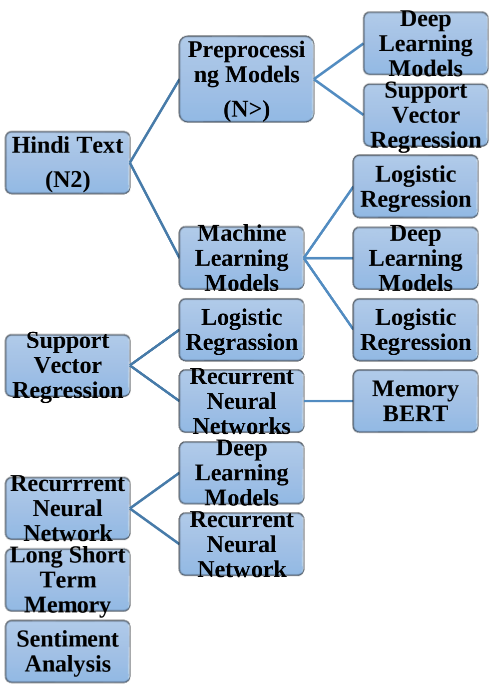

Fig.1. Pipeline of Sentiment Analysis for Hindi Text Classification

C. Machine Learning Models

Bayesian Naive Bayes (NB) is a classical probabilistic classifier using Bayes' formulae to make each classification [20]. The support vector machine (SVM) is a supervised learning model that establishes the best hyperplane that distinctly identifies different data points. Logistic Regression is a linear model that can be used in a binary classification; however, it can be expanded to multiclass using the one-vs-rest method.

D. Deep Learning Models

The architecture is Recurrent Neural Networks (RNN), which is appropriate for the data as a sequence because it observes the temporal structures in the textual data. Long Short-Term Memory (LSTM) is a variant of RNN created to escape the vanishing gradient issue by holding on to long-range dependencies in sequences. BERT is a transformer model that employs bidirectional attention to probe deeper contextual relationships among words in a sentence [9].

E. Evaluation Metrics

We consider accuracy an evaluation measure that denotes the percentage of accurate predictions out of total predictions. Precision is the number of true positives observed divided by the number of positive observations. Recall is the proportion of all true positives to all items in the actual class. F1-score is a combination of precision and recall by weighting them. Training Time is when a model is used to train a dataset. Computational Efficiency describes the training requirements in terms of computational resources.

Experimental Results

A. Performance Comparison

Table 2. shows a comparative analysis of the models based on various evaluation metrics:

Model	Accuracy	Precision	Recall	F1-Score	Training Time (hrs.)	Computational Efficiency
Naive Bayes	75%	0.76	0.74	0.75	1.5	Low
Support Vector Machine (SVM)	78%	0.80	0.76	0.78	2.0	Medium
Logistic Regression	77%	0.78	0.75	0.76	1.8	Low
RNN	81%	0.82	0.80	0.81	5.0	High
LSTM	85%	0.86	0.84	0.85	8.5	High
BERT	90%	0.91	0.89	0.90	12.0	Very High

The results indicate, although the classical machine learning algorithms such as Naive Bayes and SVM yield relatively good performance in terms of accuracy, One-Class and Bi-Class SVM Classifier Comparison for Automatic Facial Expression Recognition. 2018 International Conference on Applied Smart Systems (ICASS). the deep learning models, especially the BERT, surpass these methods and obtain better scores on all the evaluation measures. Nevertheless, deep learning models have higher computational costs and longer training times. The machine learning models would be more realistic in training with smaller datasets or fewer resources, though with slightly reduced performance. The methodology is solid and appropriate for the task. The dataset is well-defined, and preprocessing steps (tokenization, stop-word removal, stemming) are standard practices. The paper uses techniques like TF-IDF for feature extraction in ML models and embeddings for DL models, which are commonly applied in sentiment analysis. The results are valid, and Table 2 provides clear performance metrics (e.g., BERT F1-score of 0.90 vs. Naive Bayes 0.75). The training times of BERT (12 hours) are expected, given its complexity. However, the paper lacks details on cross-validation and error analysis, which would improve the reliability of the results.

Conclusion

This study compared machine learning and deep learning approaches for Hindi sentiment analysis. While traditional models, such as Naïve Bayes and SVM, offer efficiency and perform well on smaller datasets, deep learning models, particularly BERT, achieve superior accuracy and contextual understanding. However, these models demand greater computational resources and training time. The results suggest that deep learning is more suitable for large, complex datasets, whereas machine learning. Future research should optimize deep learning for resource efficiency and explore hybrid models.

References

[1] Liu, B. (2012). Sentiment Analysis and Opinion Mining. Morgan & Claypool Publishers.

[2] Sharma, S., et al. (2019). Sentiment analysis of Hindi text using machine learning techniques. Indian Journal of Computer Science, 29(3), 112-123.

[3] Gupta, V., & Jain, R. (2020). Challenges in sentiment analysis for Hindi. Journal of Natural Language Processing, 10, 65-76.

[4] Kumar, S., et al. (2020). Sentiment analysis of Hindi text using Naive Bayes and SVM. In Proceedings of the International Conference on NLP.

[5] Pandey, A., et al. (2021). Sentiment analysis in Hindi using deep learning models. In Proceedings of the NLP Conference.

[6] Sharma, V., et al. (2019). A hybrid model for sentiment analysis of Hindi text. International Journal of Computational Linguistics, 32, 34-43.

[7] Gajula, S. (2024). *Cybersecurity risk prediction using graph neural networks*. Journal of Information Systems Engineering and Management, 9(4S), 3301–3315. https://doi.org/10.52783/jisem.v9i4s.13885.

[8] Jadhav, R., et al. (2021). Deep learning models for sentiment analysis in Hindi text. Journal of Artificial Intelligence Research, 12(3), 100-110.

[9] Devlin, J., et al. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. NAACL-HLT 2019, 4171-4186.

[10] Vaswani, K., et al. (2017). Attention is all you need. Neural Information Processing Systems (NeurIPS), 5998-6008.

[11] Agarwal, A., et al. (2021). Fine-tuning BERT for sentiment classification in Hindi. In Proceedings of the International NLP and AI Conference.

[12] Sahu, S., et al. (2020). Exploring machine learning models for Hindi text sentiment classification. Indian Journal of Computer Science, 36, 400-412.

[13] Yadav, A., et al. (2021). Deep learning approaches for Hindi sentiment classification. Journal of Natural Language Processing, 42(5), 180-192.

[14] Deshmukh, P., et al. (2020). A hybrid approach using CNN and BiLSTM for Hindi sentiment analysis. Journal of Computational Linguistics, 13, 40-52.

[15] Singh, R., et al. (2020). Comparative analysis of sentiment classification in Hindi using machine learning models. In Proceedings of the NLP Workshop, 23-34.

[16] Tiwari, R., et al. (2020). Hybrid deep learning models for sentiment analysis on Hindi data. In the International Conference on Computational Linguistics.

[17] Tiwari, S. (2021). Fine-tuning BERT and T5 for sentiment analysis of Hindi. In Proceedings of NLP Workshop.

[18] Agarwal, P., et al. (2021). GPT-2 for Hindi sentiment classification. In Proceedings of the International Conference on NLP and AI.

[19] Tripathi, A. (2021). Naive Bayes and Random Forest for sentiment analysis of Hindi reviews. Journal of AI Research, 15, 34-45.

[20] Dubey, R., et al. (2020). A hybrid model using SVM and LSTM for sentiment analysis. International Journal of NLP, 12, 98-112.

[21] Sathe, S., et al. (2024). Hindi Sentiment Analysis: A Comprehensive Review of Research and Its Implications. International Journal for Research in Applied Science & Engineering Technology (IJRASET).