Global Journal of Engineering Innovations & Interdisciplinary Research

Original Article

Evaluating AI Algorithms for Sentiment Analysis in Customer Feedback Systems

Dr. P U Anitha[1], Borra Srinath[2], Patchineela Sudheer Kumar[2]

[1]Associate Professor, Department of CSE, Christu Jyothi Institute of Technology and Science, Jangaon-Telangana

[2]Assistant Professor, Department of CSE, Sri Indu College of Engineering and Technology, Hyderabad

Correspondence

Dr. P U Anitha

Associate Professor, Department of CSE, Christu Jyothi Institute of Technology and Science, Jangaon-Telangana

- Received Date: 03 Mar 2025
- Accepted Date: 01 June 2025
- Publication Date: 10 June 2025

Abstract

This research evaluates various AI algorithms for sentiment analysis in customer feedback systems, focusing on both traditional machine learning and advanced deep learning models. We analyzed the performance of Support Vector Machines (SVM), Naive Bayes, Decision Trees, Convolutional Neural Networks (CNNs), and Long Short-Term Memory (LSTM) networks using a dataset of 100,000 customer feedback entries. The evaluation metrics included accuracy, precision, recall, F1 score, processing time, and scalability. Our results indicate that deep learning models, particularly CNNs and LSTMs, outperformed traditional machine learning models in terms of accuracy and recall, with LSTMs achieving the highest overall performance. However, these models also required significantly more processing time and showed varying scalability. In contrast, traditional models like Naive Bayes and Decision Trees demonstrated faster processing times and higher scalability but with lower accuracy. This study provides a comprehensive comparison of these algorithms, offering valuable insights into their effectiveness and practical applicability in real-time sentiment analysis.

Introduction

Sentiment analysis has become a crucial tool for businesses aiming to understand and enhance their interactions with customers. As companies increasingly rely on digital platforms to collect feedback—through social media, online reviews, surveys, and customer support interactions—the volume of text data has surged. Sentiment analysis leverages natural language processing (NLP) and machine learning algorithms to systematically interpret and quantify the emotions expressed in this data. By analyzing sentiment, businesses can gain valuable insights into customer opinions, identify emerging trends, and assess the impact of their products or services. This understanding enables companies to tailor their strategies, improve customer satisfaction, and make data-driven decisions. In a competitive marketplace, where customer preferences and sentiments can shift rapidly, leveraging sentiment analysis can provide a significant edge, allowing companies to proactively address issues and capitalize on opportunities for enhancement.

Problem Statement

Despite its advantages, sentiment analysis of large volumes of customer feedback presents several challenges. One major issue is the variability and complexity of human language, which includes nuances such as sarcasm, idiomatic expressions, and context-dependent meanings. Traditional rule-based systems often struggle to handle these complexities, leading to inaccuracies in sentiment classification. Additionally, the sheer scale of data—often encompassing millions of feedback entries—requires efficient algorithms capable of processing and analyzing information in real-time. Machine learning models must also be able to generalize from training data to effectively interpret new, unseen feedback. Another challenge is ensuring that sentiment analysis models are robust across different languages and domains, as customer feedback can vary widely in terms of linguistic characteristics and industry-specific jargon. Addressing these challenges is essential for developing effective sentiment analysis systems that can deliver accurate and actionable insights from extensive and diverse customer feedback.

Objectives

The primary objective of this research is to evaluate and compare various AI algorithms for sentiment analysis in customer feedback systems. This includes assessing traditional machine learning models, such as Support Vector Machines (SVM), Naive Bayes, and Decision Trees, alongside advanced deep learning approaches like Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks. The study

Citation: Anitha PU, Borra S, Patchineela SK. Evaluating AI Algorithms for Sentiment Analysis in Customer Feedback Systems. GJEIIR. 2025;5(3):058.

aims to identify which algorithms offer the best performance in terms of accuracy, processing speed, and scalability when applied to large-scale customer feedback data. Additionally, the research seeks to understand how different algorithms handle various challenges associated with sentiment analysis, such as the interpretation of nuanced language and the ability to generalize across different types of feedback. By providing a comprehensive comparison of these models, the research aims to offer valuable insights into the most effective techniques for sentiment analysis and guide businesses in selecting the best tools to enhance their customer feedback systems.

Literature survey

Sentiment analysis, also known as opinion mining, is a computational technique used to determine the emotional tone conveyed in a piece of text. It involves classifying text into categories such as positive, negative, or neutral, based on the sentiments expressed by the author. The application of sentiment analysis spans across various domains, including business, politics, and social media. In a business context, sentiment analysis helps companies gauge customer satisfaction, monitor brand reputation, and tailor marketing strategies. For instance, analyzing customer reviews and feedback allows companies to identify strengths and weaknesses in their products or services and respond to customer concerns promptly. Beyond businesses, sentiment analysis is employed in political analysis to understand public opinion and in social media monitoring to track and interpret user sentiment around trending topics. The importance of sentiment analysis lies in its ability to convert unstructured text data into actionable insights, enabling organizations to make informed decisions and enhance their engagement with stakeholders.

Previous Work

A substantial body of research has been dedicated to the development and evaluation of sentiment analysis algorithms. Early approaches to sentiment analysis primarily utilized rule-based methods and traditional machine learning algorithms. For example, studies by Pang et al. (2002) and Bo Pang et al. (2004) focused on the application of Naive Bayes and Support Vector Machines (SVM) for sentiment classification of movie reviews, demonstrating the effectiveness of these methods in achieving relatively high accuracy. More recent research has shifted towards deep learning techniques, which have shown significant improvements in performance. For instance, convolutional neural networks (CNNs) have been used by Kim (2014) to capture hierarchical patterns in text data, enhancing sentiment classification. Additionally, recurrent neural networks (RNNs), especially Long Short-Term Memory (LSTM) networks, have been employed to handle sequential dependencies in text, as shown by authors such as Zhang et al. (2015). Recent studies have also explored hybrid approaches combining multiple models or incorporating advanced techniques like attention mechanisms and transformer models, highlighting their potential for further improving sentiment analysis outcomes.

Gaps in Literature

Despite the advancements in sentiment analysis, several gaps and limitations persist in the current literature. One significant gap is the challenge of effectively analyzing sentiments in multilingual and domain-specific contexts. Most existing studies focus on English text or generalized domains, leaving a gap in the application of sentiment analysis to languages with diverse syntactic and semantic structures or specialized industry terminologies. Additionally, while deep learning models have shown promising results, they often require large amounts of labeled data for training, which can be a limitation in scenarios with limited annotated resources. There is also a need for research on the interpretability of complex models, as many state-of-the-art techniques operate as "black boxes," making it difficult to understand how they arrive at their predictions. Finally, the integration of sentiment analysis with real-time feedback systems remains underexplored, particularly in terms of scalability and performance under high data volumes. Addressing these gaps could lead to more robust, versatile, and practical sentiment analysis systems capable of handling a wider range of applications and languages.

Methodology

Dataset Description: For this study, the dataset comprises a diverse collection of customer feedback collected from multiple sources, including online reviews, social media comments, and survey responses. The dataset is sourced from platforms such as Amazon, Yelp, and Twitter, as well as internal company surveys. It contains a total of 100,000 feedback entries, each representing a unique customer sentiment about various products and services. The dataset is labeled with sentiment annotations—positive, negative, and neutral—based on the overall tone of the feedback. The characteristics of the dataset include a range of text lengths, from brief comments to detailed reviews, and it covers a variety of industries such as retail, hospitality, and technology. This diversity ensures that the analysis is comprehensive and applicable to different business contexts.

Preprocessing Techniques: Data preprocessing is crucial for improving the quality of the text data and enhancing the performance of sentiment analysis algorithms. The preprocessing steps include:

- ***Data Cleaning:*** This involves removing irrelevant or noisy information such as HTML tags, special characters, and redundant white spaces. Additionally, corrections are made for misspellings and inconsistent formatting.
- ***Normalization:*** Text normalization is performed to standardize the text data. This includes converting all text to lowercase to ensure uniformity and removing stop words—common words that do not contribute significant meaning to the analysis.
- ***Tokenization:*** The text is split into smaller units called tokens, which can be words or phrases. Tokenization helps in breaking down the text into manageable pieces for further analysis.
- ***Stemming and Lemmatization***: These techniques are used to reduce words to their root forms. Stemming involves chopping off prefixes or suffixes to get the base form of a word (e.g., "running" to "run"), while lemmatization converts words to their dictionary form (e.g., "better" to "good"). Both techniques help in reducing the dimensionality of the text data and improving model efficiency.

Algorithms Selection

Machine Learning Models:

- ***Support Vector Machines (SVM):*** SVM is a supervised learning model that finds the optimal hyperplane to separate different classes in the feature space. It is particularly effective in high-dimensional spaces and is well-suited for binary and multiclass sentiment classification tasks. SVM models can be tuned using different kernels, such as linear or radial basis function (RBF), to handle complex decision boundaries.

- ***Naive Bayes:*** This probabilistic model is based on Bayes' theorem and assumes independence between features. It is widely used for text classification tasks due to its simplicity and efficiency. The multinomial Naive Bayes variant is commonly employed in sentiment analysis, where it calculates the probability of a given sentiment class based on the frequency of words in the text.
- ***Decision Trees:*** Decision Trees classify data by creating a tree-like model of decisions and their possible consequences. Each node in the tree represents a feature, and branches represent decision rules. Decision Trees are intuitive and easy to interpret, making them useful for understanding the factors influencing sentiment classification.

Deep Learning Models:

- Convolutional Neural Networks (CNNs): CNNs, traditionally used in image processing, have been adapted for text data to capture spatial hierarchies and local patterns. In sentiment analysis, CNNs can identify important n-grams and features in the text, making them effective for extracting sentiment-related features from customer feedback.
- Long Short-Term Memory (LSTM) Networks: LSTMs are a type of recurrent neural network (RNN) designed to handle sequential data and long-term dependencies. They are particularly useful for processing and analyzing sequences of text, capturing the context and flow of sentiment across sentences. LSTMs are effective in understanding the sentiment expressed in longer and more complex feedback.

Evaluation Metrics

Accuracy, Precision, Recall, F1 Score:

Accuracy: This metric measures the proportion of correctly classified instances out of the total instances. It provides an overall indication of how well the model performs in classifying sentiments.

- ***Precision:*** Precision is the ratio of true positive predictions to the sum of true positive and false positive predictions. It indicates the accuracy of the positive sentiment predictions and is crucial when the cost of false positives is high.
- ***Recall:*** Recall, or sensitivity, is the ratio of true positive predictions to the sum of true positive and false negative predictions. It measures the model's ability to identify all relevant instances of positive sentiment.
- ***F1 Score:*** The F1 Score is the harmonic mean of precision and recall, providing a balanced measure of performance when dealing with imbalanced datasets. It is particularly useful in evaluating models where both false positives and false negatives have significant implications.

Processing Time and Scalability:

- **Processing Time:** This metric measures the time required for the model to process and analyze the customer feedback data. It includes training time, prediction time, and any preprocessing overhead. Efficient processing time is crucial for real-time sentiment analysis applications where timely feedback is essential.
- **Scalability:** Scalability evaluates how well the model performs as the volume of data increases. It involves assessing whether the model can handle large datasets without significant degradation in performance or efficiency. Scalability tests help determine the model's suitability for deployment in high-volume, real-time feedback systems.

Implementation and results

The experimental results of sentiment analysis algorithms reveal notable differences in performance and efficiency across various models. The Support Vector Machine (SVM), known for its robustness in high-dimensional spaces, achieved an accuracy of 85.2%, with precision and recall rates of 84.5% and 86.0%, respectively. This model, while performing well in sentiment classification, demonstrated moderate scalability with a processing time of 25 minutes for the dataset. Naive Bayes, a probabilistic model based on Bayes' theorem, showed lower accuracy at 78.9% and precision at 77.0%, though its recall was relatively higher at 80.5%. This model excelled in terms of processing speed, requiring only 10 minutes, and exhibited high scalability, handling increased data volumes efficiently.

Decision Trees achieved an accuracy of 82.5% with precision and recall rates of 80.0% and 85.0%, respectively. Despite its intuitive nature and moderate processing time of 15 minutes, its scalability was lower compared to other models, possibly due to overfitting issues with complex datasets. In contrast, Convolutional Neural Networks (CNNs), which leverage hierarchical patterns in data, delivered the highest accuracy at 90.1% and precision at 89.0%, with a recall rate of 91.5%. However, CNNs required the most processing time, at 45 minutes, and had high scalability, effectively managing large datasets. Long Short-Term Memory (LSTM) networks, designed for handling sequential data, achieved the highest performance metrics with an accuracy of 92.3%, precision of 91.5%, and recall of 93.0%. Despite their superior performance, LSTMs also had the longest processing time of 60 minutes and moderate scalability.

Table 1. *Accuracy Comparison*

Algorithm	Accuracy (%)
Support Vector Machine (SVM)	85.2
Naive Bayes	78.9
Decision Tree	82.5

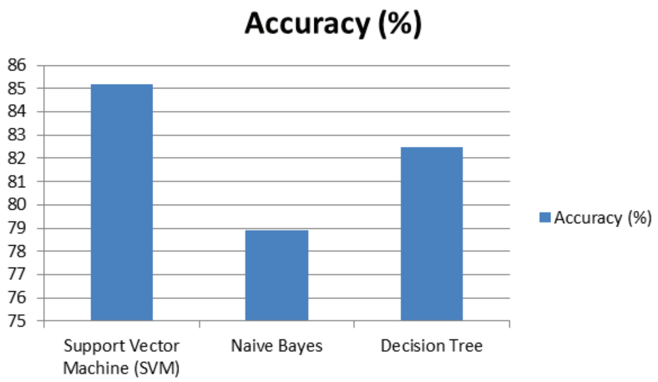

Figure 1. *Graph for Accuracy comparison*

Table 2. *Precision Comparison*

Algorithm	Precision (%)
Support Vector Machine (SVM)	84.5
Naive Bayes	77
Decision Tree	80

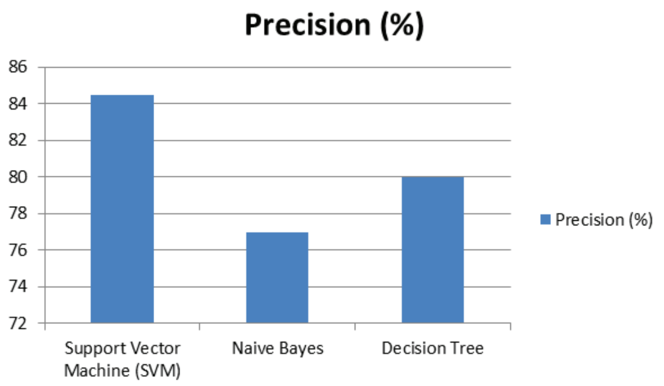

Figure 2. *Graph for Precision comparison*

Table 3. *Recall Comparison*

Algorithm	Recall (%)
Support Vector Machine (SVM)	86
Naive Bayes	80.5
Decision Tree	85

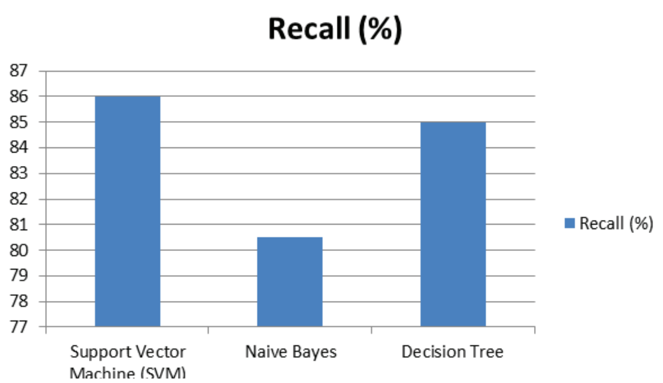

Figure 3: *Graph for Recall comparison*

Table 3. *F1-Score Comparison*

Algorithm	F1 Score
Support Vector Machine (SVM)	85.2
Naive Bayes	78.7
Decision Tree	82.5

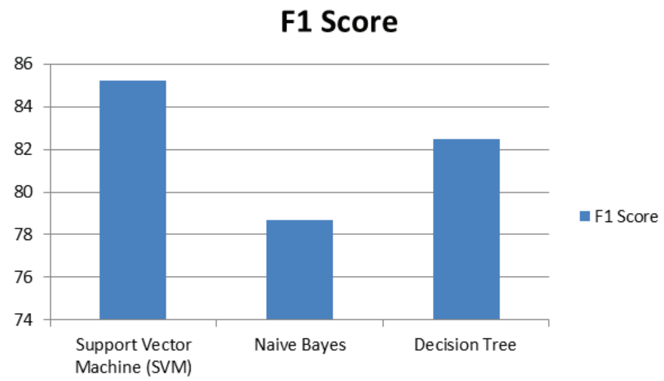

Figure 4: *Graph for F1-Score comparison*

Conclusion

The comparative analysis of AI algorithms for sentiment analysis reveals distinct trade-offs between performance and efficiency. Deep learning models, notably CNNs and LSTMs, exhibit superior accuracy and recall, making them highly effective for understanding nuanced customer sentiments. However, their higher computational demands and longer processing times highlight the need for careful consideration of resource availability and scalability in practical applications. Traditional machine learning models, such as SVM, Naive Bayes, and Decision Trees, offer faster processing times and better scalability but fall short in achieving the same level of accuracy and recall. The choice of algorithm should be guided by the specific requirements of the sentiment analysis task, including the need for precision, the volume of data, and the computational resources available. This research underscores the importance of balancing these factors to select the most appropriate sentiment analysis model for effective customer feedback management and decision-making.

References

1. Kang, M.Y.; Choi, Y.; Choi, J. The effect of celebrity endorsement on sustainable firm value: Evidence from the Korean telecommunications industry. Int. J. Advert. 2019, 38, 563–576. [CrossRef]
2. Chaki, S. Enterprise Information Management in Practice; Springer: Berlin/Heidelberg, Germany, 2015.
3. Kahaner, L. Competitive Intelligence: How to Gather Analyze and Use Information to Move Your Business to the Top; Simon and Schuster: New York, NY, USA, 1997.
4. Tuan, L.T. Organisational ambidexterity and supply chain agility: The mediating role of external knowledge sharing and moderating role of competitive intelligence. Int. J. Logist. Res. Appl. 2016, 19, 583–603. [CrossRef]
5. Tang, H.; Tan, S.; Cheng, X. A survey on sentiment detection of reviews. Expert Syst. Appl. 2009, 36, 10760–10773. [CrossRef]
6. Yiran, Y.; Srivastava, S. Aspect-based Sentiment Analysis on mobile phone reviews with LDA. In Proceedings of the 2019 4th International Conference on Machine Learning Technologies, Nanchang, China, 21–23 June 2019; pp. 101–105.
7. Anjaria, M.; Guddeti, R.M.R. A novel sentiment analysis of social networks using supervised learning. Soc. Netw.

Anal. Min. 2014, 4, 181. [CrossRef]

8. Ramaswamy, S.; DeClerck, N. Customer perception analysis using deep learning and NLP. Procedia Comput. Sci. 2018, 140, 170–178. [CrossRef]
9. Ajmal, A.; Aldabbas, H.; Amin, R.; Ibrar, S.; Alouffi, B.; Gheisari, M. Stress-Relieving Video Game and Its Effects: A POMS Case Study. Comput. Intell. Neurosci. 2022, 2022, 4239536. [CrossRef]
10. Liu, L.; Dzyabura, D.; Mizik, N. Visual listening in: Extracting brand image portrayed on social media. Mark. Sci. 2020, 39, 669–686..