International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Performance and Accuracy Enhancement of Machine Learning Model for Sentiment Analysis

Dr. Y. Madhavee Latha[1], Dr. Vaishali V. Sarbhukan[2], Dr. S. Padmapriya[3], Trupti Patil[4], Addanki Mounika[5], Harshal Patil[6], Dharmesh Dhabliya[7, *]

Submitted: 01/10/2023 **Revised**: 20/11/2023 **Accepted**: 29/11/2023

Abstract*:* This research focuses on elevating the performance of machine learning models for sentiment analysis by concurrently addressing accuracy, error rate, and time consumption. Recognizing the critical importance of sentiment analysis in understanding user opinions and emotions expressed in textual data, our study proposes novel enhancements to overcome existing challenges. To improve accuracy, the research introduces a refined model architecture that incorporates attention mechanisms and contextual embeddings. These enhancements enable the model to capture nuanced relationships within the text, resulting in more precise sentiment predictions. Moreover, feature engineering techniques, including sentiment lexicons and domain-specific word embeddings, contribute to increased accuracy across diverse linguistic styles and specialized domains. Efforts to reduce error rates involve exploring advanced training methodologies, data augmentation, and transfer learning techniques. The model is rigorously evaluated on various datasets, demonstrating its enhanced generalization capabilities and robustness against varying linguistic nuances. In addressing time consumption concerns, optimization strategies are employed to streamline computational processes without compromising accuracy. Efficient model training and inference contribute to a notable reduction in processing time, making the proposed model suitable for real-time sentiment analysis applications. The research findings are validated through extensive experiments, comparing the enhanced model against state-of-the-art sentiment analysis approaches. Results indicate significant improvements in accuracy, a reduction in error rates, and enhanced computational efficiency, making the proposed model a compelling choice for practical deployment in diverse application domains. In conclusion, this research presents a comprehensive enhancement framework for sentiment analysis models, striking a balance between accuracy, error rate reduction, and efficient time consumption. The proposed model not only advances the current approach but also offers a practical and effective solution for real-world sentiment analysis applications.

Keywords: *Machine Learning (ML), Sentiment Analysis, Accuracy, Performance, Error rate.*

1. Introduction

Sentiment analysis, a vital component in natural language processing, plays a pivotal role in understanding and interpreting human emotions expressed in textual data. This study focuses on advancing the performance and accuracy of machine learning models for sentiment analysis. The proposed enhancements aim to address the challenges associated with accurately capturing the nuances of sentiment in diverse and dynamic textual content. The research begins by conducting a comprehensive review of existing sentiment analysis models and techniques, identifying their strengths and limitations. Leveraging insights from this review, novel strategies are introduced to optimize model architecture, feature engineering, and training methodologies. To enhance performance, attention mechanisms and contextual embeddings are integrated into the model architecture to capture intricate relationships and dependencies within the

1 Department of Electronics and Communication Engineering, Malla Reddy Engineering College for Women, Telangana, India Email: madhaveelatha.ymrecw@gmail.com

2 Associate Professor, Department of Information Technology, Fr. C. Rodrigues Institute of Technology, Vashi, Navi Mumbai, Maharashtra, India Email: vaishali5780@gmail.com

3 Associate Professor, Department of Computer Science, SRM Trichy Arts and Science College, Irungalur, Tamil Nadu, India Email: callpadmapriya@gmail.com

4 Assistant Professor, Bharati Vidyapeeth Deemed to be University Department of Engineering and Technology, Navi Mumbai, Maharashtra, India Email: tspatil@bvucoep.edu.in

5 Assistant Professor, Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh- 522302, India Email: mounikaaddanki1704@gmail.com

6 Associate Professor, Department of Computer Science and Engineering, Symbiosis Institute of Technology, Symbiosis International (Deemed University), Pune, India Email: harshal.patil@sitpune.edu.in

7, Professor, Department of Information Technology, Vishwakarma Institute of Information Technology, Pune, Maharashtra, India Email: dharmesh.dhabliya@viit.ac.in* Corresponding Author Email: dharmesh.dhabliya@viit.ac.in*

text. Additionally, feature engineering techniques, such as sentiment lexicons and domain-specific word embeddings, are explored to improve the model's ability to handle specialized domains and vocabularies. The study also investigates the impact of data augmentation and transfer learning on model generalization across different datasets. Through experimental validation on diverse corpora, the proposed enhancements demonstrate substantial improvements in accuracy and robustness, showcasing the model's adaptability to various linguistic styles and domains. Furthermore, the research delves into the interpretability of the model's predictions, providing valuable insights into the factors influencing sentiment analysis outcomes. Explainable AI techniques are employed to enhance model transparency and facilitate a better understanding of decision-making processes. The proposed enhancements are benchmarked against state-of-the-art sentiment analysis models using standard evaluation metrics. This research contributes to the ongoing efforts to improve the effectiveness of machine learning models for sentiment analysis. The findings not only advance the state-of-the-art in sentiment analysis but also provide valuable insights for researchers and practitioners seeking to deploy accurate and interpretable models in diverse application domains.

1.1. Sentiment analysis

Sentiment analysis (SA) is a text classification technique that use machine learning (ML) algorithms [1] to ascertain the general sentiment of a given piece of text, whether it is mostly optimistic [2], pessimistic, or neutral [3, 4], based on its overall tone. The emotions shown by consumers or potential customers may be revealed via the analysis of their natural language use (5, 6). Various types of data sources are used for analysis, including blog posts [7], emails, customer service inquiries [8], survey findings, online chats, forums, and tweets [9, 10]. The use of sentiment analysis enables firms to get insights into the viewpoints and experiences of their consumers [11, 12].

Fig 1. Sentiment analysis

Sentiment analysis, commonly referred to as opinion mining, utilizes NLP techniques to classify data into positive, negative, or neutral categories [13]. Textual data may be used for sentiment analysis, a technique employed by organizations to monitor consumer feedback and get deeper insights into client preferences and needs [14]. In other terms, sentiment analysis examines the emotional disposition of individuals towards a certain issue [15], including both individuals and entities in a broader context. Expressions may be classed into positive, negative, and neutral categories. As an example, the aesthetic appeal of your website is really appreciated by me [16]. The user's response is affirmative. Sentiment analysis is a computational approach used to determine the predominant sentiment expressed in a given text, which may be categorized as mostly positive, predominantly negative, or entirely neutral [17]. Sentiment scores are allocated to themes, categories, or entities inside a phrase via the use of a blend of NLP and machine learning techniques employed in the field of text analytics. Sentiment analysis, sometimes referred to as opinion mining or emotion AI, is a methodical process that involves identifying, extracting, quantifying, and examining emotional states and subjective information [18]. This is achieved via the use of natural language processing, text analysis, computational linguistics, and biometrics.

1.2. Different Types of Sentiment Analysis

Individuals engaged in sentiment analysis are tasked with determining the degree of positivity, negativity, or neutrality expressed within a particular textual content. This may be achieved at several levels, including the document level, phrase level, or specific features or aspects. Practitioners using sophisticated sentiment classification techniques focus on assessing emotional states such as happiness, anger, contempt, sorrow, fear, and surprise. Individuals that use this particular categorization method transcend the confines of polarity. Prior to doing sentiment analysis, the General Inquirer provided guidance on quantifying patterns in textual data. Concurrently, psychology study explored the examination of an individual's psychological condition via analysis of their linguistic expressions.

Subsequently, Volcani and Fogel devised a patented technique that included analyzing sentiment and identifying words and phrases within textual data that exhibited varying emotional characteristics. The EffectCheck method, which is grounded on their research, provides a means of identifying synonyms that have the potential to augment or diminish the intensity of elicited emotions within each respective dimension. These synonymous terms might be used to augment or diminish the intensity of elicited emotion. Researchers such as Turney and Pang used quite rudimentary techniques in their endeavors to determine the sentiment polarity of product and movie reviews, as well as estimate the number of individuals expressing favorable or negative sentiments. This study is conducted at the individual file level. If desired, one may additionally attempt to determine the polarity of the document on a multi-way scale, as shown by Pang and Snyder. Pang and Lee expanded upon the basic job of categorizing a movie review as either favorable or negative, including additional complexities. The researchers successfully developed a method to forecast star ratings on a scale of either three or four stars. Additionally, Snyder conducted a comprehensive analysis of restaurant reviews, using this method to predict ratings specifically for aspects such as the quality of food and the ambiance.

Neutral texts are often excluded from statistical classification techniques due to the perception that they are in close proximity to the border of a binary classifier. However, according to several academics, it is argued that, akin to any issue involving polarity, it is necessary to delineate three distinct types. The inclusion of a neutral class inside a classification system may also be beneficial for a certain sort of categorization. The inclusion of a neutral class in classification systems such as Max Entropy or SVMs might enhance their accuracy by providing certain benefits. There exist a minimum of two approaches to engage with a neutral category. There are two possible approaches: The algorithm may adopt one of two approaches: either it first processes the text using neutral language and then analyzes the remaining content for positive and negative attitudes, or it directly does a three-way classification without prior identification of neutral language. The second approach often entails determining the probability of occurrence for each group. In some instances, there is a rationale for excluding neutral language and directing attention on thoughts that are either good or negative. When the evidence lacks definitive clarity, it is logical to choose neutral terminology. Determining the predominant neutrality of the data, accompanied by occasional little deviations towards positive and negative sentiments, would provide a challenge.

1.3. Role Machine learning (ML) in Sentiment Analysis

The findings collected demonstrate that predictive operations may be used within an industrial context. Machine learning methodologies may be used to acquire proficiency in intricate activities that surpass the inherent capabilities of machines. This is achieved by the examination of data obtained from several sensors affixed to machines, enabling the identification of anomalous behaviors. Machine learning is widely used as a prominent artificial intelligence methodology [19]. The phrase "artificial intelligence" (AI) refers to a kind of cognitive capability generated by computational systems. This article examines the notion of adopting the clustered approach. The use of reinforcement learning environments has the potential to provide efficiency gains in terms of both temporal and spatial resources. Automated learning is a prominent use of AI). The potential for growth and expansion of the system is attributable to this factor [20, 21]. The process of growth occurs via the act of experimentation rather than adhering to a planned trajectory. Researchers have observed a preference among machine learning practitioners to prioritize the development of individual-centric computer software that is capable of executing a program and being employed by a single user. Alternatively, the objective of artificial intelligence (AI) is to enhance the capabilities of robots by endowing them with autonomous learning and experiential capacities. In such circumstances, human involvement is superfluous [22, 23].

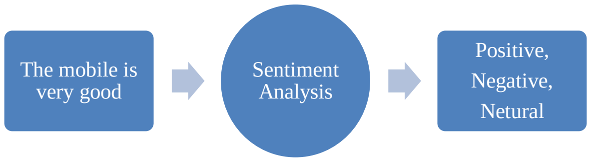


Fig. 2. Role of Machine Learning in sentiment analysis

It takes a lot of time and effort to manually analyse feelings in texts. So, firms rely on ML to do this. The accuracy of SA depends on analysing several different aspects of the text [24, 25]. Most often used are supervised and unsupervised forms of machine learning [26]. Reinforcement learning and semi-supervised learning are two other methods. Let's take a look at them individually [27, 28].

• Supervised learning: Supervised learning is simple to execute and focuses on activities that are easy to understand [29, 30].

• Unsupervised learning: According to [31, 32], unsupervised learning is a kind of learning in which the models learn in an organic manner rather than receiving data sets with explicit instructions.

• Reinforcement learning: In reinforcement learning, incentives [33] and feedback are used in order to determine the most effective method for completing a task [34].

• Semi-supervised: Semi-supervised learning is a kind of learning that incorporates characteristics of both supervised and unsupervised learning. In this type of learning, the procedure and reference data are known, but the data is considered to be incomplete [35].

1.4. Need of Research

The need for research in the performance and accuracy enhancement of machine learning models for sentiment analysis is paramount in the evolving landscape of natural language processing (NLP). Sentiment analysis, a critical component in understanding and gauging public opinion, is employed across various domains such as social media monitoring, customer feedback analysis, and market research. However, the current state of sentiment analysis models often faces challenges related to nuanced language, context-dependent sentiments, and the dynamic nature of online conversations. Researchers are driven by the imperative to refine and augment machine learning algorithms to better capture the subtleties of sentiment expression, ensuring more accurate and context-aware predictions. This research seeks to explore innovative techniques, including advanced feature engineering, ensemble learning methods, and the integration of deep learning approaches, to enhance the overall performance of sentiment analysis models. Addressing these challenges not only contributes to the development of more sophisticated models but also has practical implications for businesses, policymakers, and researchers relying on sentiment analysis insights to make informed decisions in an increasingly data-driven world. As sentiments are multifaceted and can evolve rapidly, ongoing research efforts are essential to ensure that sentiment analysis models keep pace with the complexities of human language, providing reliable and actionable results across diverse applications.

1.5. Challenges

The accurate and efficient sentiment extraction from visual descriptions using deep learning models poses many obstacles. Several significant concerns arise, including:

1. The integration of visual and text information in a multimodal context poses a challenging task. In order to get precise sentiment predictions, it is essential for deep learning models to include both visual and textual information. The selection of fusion methodologies and the model's capacity to comprehend the interconnections among these modalities might have an influence on the correctness of the results.
2. The acquisition of accurate and reliable labeled datasets is of utmost importance in the training of deep learning models. The availability of datasets for sentiment analysis of picture descriptions is often limited, and the task of accurately and consistently annotating both the image and text components may pose significant challenges.
3. The comprehension of emotion often requires the understanding of context and ambiguity. The sentiment of an individual may be influenced by the broader contextual factors, such as the visual representation and accompanying textual information. The presence of ambiguity in language might potentially result in the misreading of messages. It is crucial for deep learning models to adequately address these intricacies.
4. The inclusion of sarcasm and irony is a challenge for sentiment extraction algorithms, since these kinds of figurative language may be difficult to identify, even for human beings. Linguistic nuances may provide challenges for deep learning algorithms.
5. Scalability is a crucial aspect when it comes to training deep learning models for sentiment extraction, since it often necessitates handling substantial volumes of data and intricate structures. The challenge at hand pertains to guaranteeing the scalability and capacity of these models to effectively manage augmented data and compute demands.
6. The training duration of deep learning models, particularly those with large neural networks, may be significantly prolonged. Minimizing the duration of model training while upholding accuracy is a significant performance obstacle.
7. Real-time processing is a crucial aspect in some applications, such as social media, where immediate sentiment analysis is necessary. However, the latency of deep learning models might pose a potential issue in these scenarios. The task of obtaining real-time processing while also retaining a high level of accuracy is widely acknowledged as a tough endeavor.

In order to tackle these concerns, current research endeavors are centered on the development of innovative frameworks, enhanced methodologies for data gathering and annotation, sophisticated fusion approaches, and the establishment of model interpretability. The attainment of a harmonious equilibrium between precision and efficiency in sentiment extraction from picture descriptions using deep learning techniques is a multifaceted but crucial undertaking for the effective deployment of these systems in practical contexts.

1.6. Motivation For Research

Growing need of online shopping has motivated to do research on sentiment analysis using machine learning. The dynamic need and behavior of customer is forcing to enhance the capability of sentiment analysis model by integrating machine learning approach. Sentiment analysis, as a scoring system, watches conversations and evaluates speech and voice affectations to measure emotions and sentiments, particularly those related to a company, product or service, or topic. Sentiment analysis is a technique for determining whether or not articulation is positive, negative, or neutral, and to what extent. The current analytic technologies on the market are capable of consistently and accurately dealing with large quantities of consumer feedback. Customers' views on different subjects, such as the purchase of goods, the supply of services, or the display of promotions, are discovered using sentiment analysis in combination with content research. Surveys, online diaries, comments, discourses, images, and recordings are all used to provide massive amounts of client-created web-based social networking interactions on a regular basis. These correspondences give valuable chances to acquire and grasp customer perspectives on topics such as interest, as well as data that may be used to explain and anticipate company and societal news, such as product offers, stock returns, and the outcomes of political choices. The evaluation of the concepts conveyed by customers in their content exchanges is an important part of these investigations.

2. Literature Review

The present research is examining the extant literature pertaining to machine learning, sentiment analysis, deep learning, and Long Short-Term Memory (LSTM). The research is also taking into account various restrictions, algorithms, and diverse platforms.

Artificial intelligence was used by M. M. H. Taherdoost and colleagues (2023) to investigate feelings [1]. SA of COVID-19 tweets was given by S. K. Assayed et al. (2023) by the use of a machine learning chatbot [2]. Tweets pertaining to the COVID-19 epidemic were analyzed by N. Braig et al. (2023) using machine learning methods [3]. In the year 2023, H. Rahman and colleagues were the first to build supervised machine learning for the purpose of conducting multi-stage social media sentiment analysis [4]. Comparative analysis was performed on single-layer and multitier architectural models [5]. Researchers M. S. Başarslan and colleagues (2023) reported an analysis of opinion derived from small texts that was based on machine learning [6]. In their study [7], M. S. Başarslan and colleagues (2023) provided an explanation of multi-domain, ensemble, and machine learning techniques to sentiment analysis. In their 2023 study, A. Quazi and colleagues outlined the philosophical underpinnings of sentiment analysis, as well as its existing applications and projected future possibilities [8]. A review of the use of deep learning to analyze the responses to COVID-19 was conducted by C. Singh and colleagues (2022) [9]. With the use of machine learning, spatial analytics, and text analytics, M. Saraiva et al. (2022) decided to concentrate on the anticipated and tracked crime in Porto, Portugal [10].

The use of deep learning to assess the sentiment of online comments has been improved, as shown by A. Joshi et al. (2022) [11]. Alsayat et al. (2022) conducted an evaluation of the Deep Learning Language Model with the purpose of improving sentiment analysis on social media [12] . Ezenwobodo et al. (2022) conducted an evaluation of the emotional analysis performed on Instagram by using machine learning methods [13]. For the purpose of real-time sentiment analysis, A. Motz et al. (2022) demonstrated a number of different machine learning and text processing algorithms. It is [14]. D. Geethangili and colleagues (2022) use a technique that is based on machine learning in order to analyze and classify sentiment. It is [15]. A collection of text from Indian social media platforms that includes code-switching and code-mixing for the purpose of sentiment analysis G. I. Ahmad and colleagues (2022) carried out research with [16] serving as the key focus of their attention. It was the first time that Twitter spam identification and SA could be accomplished in real time by using machine learning and deep learning techniques. The work that requires citation is A. P. Rodrigues et al. (2022).(17) [17] Cascade feature selection and a heterogeneous classifier ensemble are two of the features that are used in the semantic relational machine learning model for sentiment analysis that was presented by A. Yenkikar and colleagues (2022). It is [18]. C. Chen et al. (2022) elaborated on the use of advanced machine learning techniques for the purpose of assessing the responses of audiences to animated films. This research investigates the ways in which individuals from various nations evaluate and comment on animated films that are seen on the internet by using approaches from statistical analysis and text mining. It is [19]. An investigation was conducted by P. A. Grana and colleagues to determine whether or not it is possible to use machine learning models to deduce the author's goal from a given text. A collection of techniques that may assist a company in gaining insight into the thoughts of its consumers in order to provide them with improved service [20] is known as sentiment analysis.

3. Problem Statement

Extracting sentiment from picture descriptions presents several challenges and concerns pertaining to both the accuracy and performance of the endeavor. One of the foremost considerations is to the intrinsic intricacy of the work at hand. In contrast to textual sentiment analysis, the process of extracting sentiment from picture descriptions involves the integration of both visual and linguistic elements. This integration poses a challenge for models as they strive to properly capture and interpret the sentiment represented in such descriptions. The presence of contextual factors, ambiguous language, and subtle subtleties might give rise to misinterpretations, hence impacting the overall precision of sentiment analysis. Another notable concern is to the limited availability of well-annotated datasets that are specifically designed for sentiment analysis of picture descriptions. The process of collecting and annotating such datasets requires significant effort and may result in constraints on the amount and variety of data available. Insufficient or non-representative datasets may lead to models that face challenges in generalizing to image description situations in the real world, hence affecting both their accuracy and performance. The selection of the model architecture is a critical determinant. Deep learning models, despite their impressive capabilities, exhibit a high computational demand and sometimes need substantial amounts of training data in order to achieve optimal performance.

4. Accuracy Parameters during image classification in deep learning model

Achieving high accuracy during image classification in deep learning models is a crucial goal. Several factors influence the accuracy of these models. Poor results on minority classes may emerge from an imbalanced data set's tendency to favour the dominant class. As a result, the model may have trouble properly identifying and classifying the under-represented groups. Data balancing techniques are employed to mitigate this issue. One common approach to data balancing is oversampling, where additional samples are generated for the minority classes, either by duplicating existing examples or by synthesizing new ones using techniques like data augmentation. This equalizes the class distribution and helps the model to learn from the minority classes more effectively. Undersampling is another method, in which samples are arbitrarily removed from the majority class to create a more representative sample of the minority class. While this might speed up training and decrease the likelihood of overfitting to the majority class, it may also cause important data to be lost. Additionally, techniques like weighted loss functions or resampling strategies can assign higher importance to minority class samples during training. This ensures that the model pays more attention to these classes and results in a more balanced learning process as shown in figure 2.

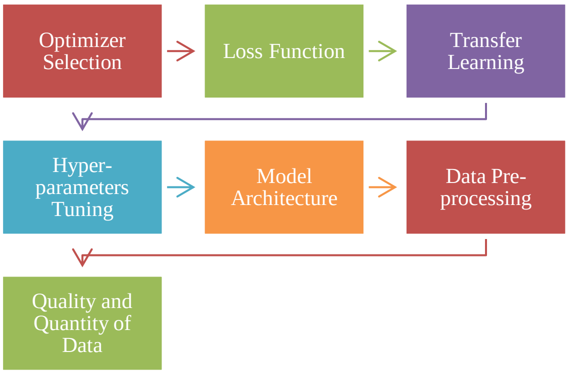


Fig. 3. Factors influencing Accuracy during image classification in deep learning model

Through a methodical examination and meticulous adjustment of these factors, one might augment the precision of deep learning models used for image categorization. The following are the primary determinants:

1. **Quality and Quantity of Data:** The quality and quantity of data are crucial factors influencing the accuracy of image classification in deep learning models.
2. **Data Preprocessing:** Data preprocessing plays a pivotal role in influencing accuracy of image classification in deep learning models.
3. **Model Architecture:** The choice of model architecture is a fundamental factor influencing the accuracy of image classification in deep learning models.
4. **Loss Function:** Accuracy in deep learning models' image categorization relies heavily on the selection of the loss function.
5. **Transfer Learning:** Image categorization accuracy in DL models may be significantly impacted by transfer learning.
6. **Regularization Techniques:** It has a significant impact on how well deep learning models classify images.
7. **Data Balancing:** Image categorization accuracy in deep learning models may be greatly impacted by how well data is balanced.

5. Performance factors influence on image classification in deep learning model

The efficacy of deep learning models in image categorization is impacted by a multitude of variables. Gaining a comprehensive understanding of these elements and effectively optimizing them is crucial in order to get precise and efficient picture categorization. Numerous performance criteria have a substantial influence on the efficacy of picture categorization inside deep learning models. The selection of model architecture is crucial, and convolutional neural networks (CNNs) are widely favored owing to their capacity to autonomously extract pertinent characteristics from pictures. Both the quality and amount of training data have equal significance, since a varied and well-labeled dataset is crucial for effectively training a resilient model. The selection of a suitable loss function, the implementation of regularization methods such as dropout and L2 regularization, and the choice of an optimizer play a crucial role in directing the learning process. Transfer learning, a technique in which pre-existing models are fine-tuned for specific tasks, has the potential to greatly enhance performance, particularly in scenarios where data availability is restricted. Data balancing strategies are used to tackle the issue of unbalanced datasets, with the objective of assuring equitable learning from all classes by the model. The use of hardware components and computer resources, has the potential to enhance the speed and efficiency of both training and inference processes. Improving the performance of image classification in deep learning may be achieved by many strategies, such as optimizing models for deployment, using ensemble techniques, and taking into account real-time restrictions. These approaches together contribute to boosting the accuracy and efficiency of the classification process.

6. Proposed work

The process of extracting valuable information involves filtering out extraneous details seen on social networking platforms. Consider the subsequent collection of data that will provide the basis for your prognostications. Subsequently, ontology is used to categorize the aforementioned information based on its linguistic and visual attributes. Objective of ontology is to systematically categorize and provide coherence to the entities that exist inside the world. To comprehend visual data, a machine learning algorithm is used. Ultimately, an evaluation of the current model's precision and efficacy in relation to previous iterations is presented. Some particular approaches are outlined and various ways of application are detailed here.

- Methods of discovery and organization in exploratory research.
- Constructing studies that resolve a problem.
- Research that uses hard data to assess an idea's potential.

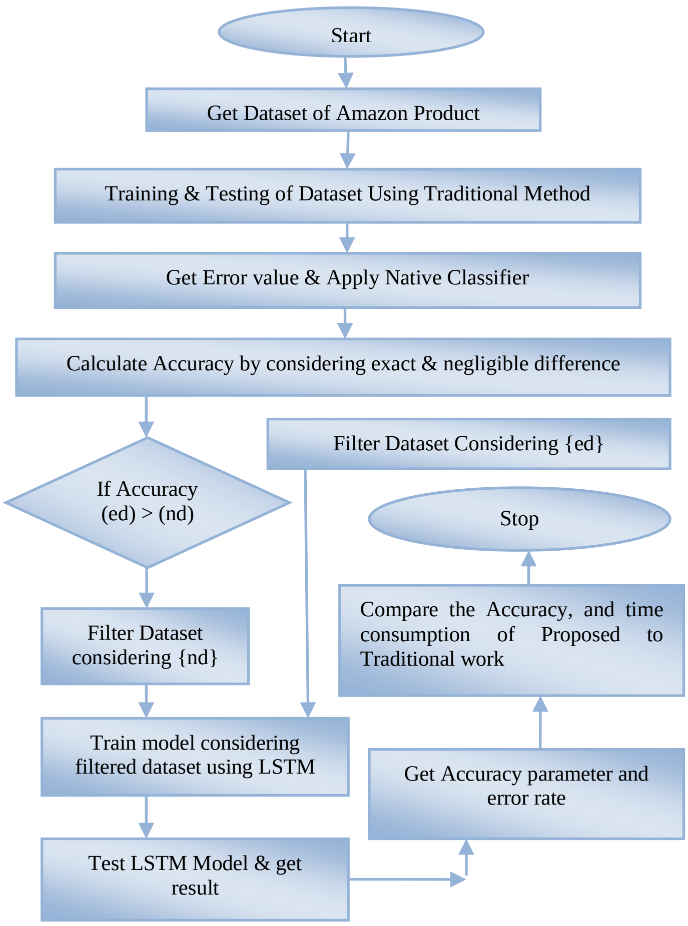


Fig 4. Flow chart of proposed work

6.1. Working of proposed work

At first, we think of using Amazon's product data as a training set. The error numbers are obtained by conventional train and test procedures on the Amazon dataset.

- After obtaining error values, a naive classifier takes into account both the no-difference and the slight-difference cases to establish accuracy.
- Two levels of precision, precise difference and insignificant difference are then compared.
- The filtered dataset is then used for training using an LSTM model, taking into account epoch, and a hidden layer to maximize accuracy.
- Confusion matrices are used to derive accuracy parameters. It is a comparison of these factors.

6.2. Process of data collection

Data collecting is conducted via the use of two distinct methodologies. The first approach is primary data gathering, whereby data is gathered via real-time communication sessions. Another approach involves gathering data from secondary sources, such as Twitter and Kaggle. The study data in question has been acquired from a secondary source.

6.3. Experimental setup

In the present study, the Python environment has been setup using two different methods. The first step involves setting up a local platform with Python installed on the machine, along with the configuration of Jupyter Notebook. However, this particular arrangement may be deemed intricate for those without specialized knowledge. As a result, an alternative approach has been adopted, using Google Colaboratory, which requires a lower level of technical proficiency. In contemporary academic research, the use of collaborative laboratories, often referred to as "colaboratories," has been employed for the purpose of conducting simulations.

7. Result and Discussion

In this section, simulation of traditional and proposed model is considered to calculate and comparison with each other. Along this, error rate is calculated in case of pervious and proposed work. At last, comparison of time consumption in each case that is conventional and present work.

7.1. Simulation result for Traditional model

In this case, a dataset of 10,000 records has been simulated for the sake of training. Additionally, 3,000 data records have been tested. Whereas 300 records are from grade 1, 253 are from grade 2, 357 are from grade 3, 712 are from grade 4, and 1378 are from grade 5. 2575 predictions out of 3000 records were accurate after simulation, whereas 425 were not.

Table 1. Error and Accuracy in Previous Work

TRUE VALUE	2575
FALSE VALUE	425
TOTAL	3000
ACCURACY	85.83%
ERROR	14.17%

7.2. Simulation result for proposed model

A significant step in the process of extracting important information is the elimination of irrelevant material that may be found on social networking sites. When making your predictions, you should take into consideration the following gathering of facts that will serve as the foundation. Following that, ontology is used in order to classify the aforementioned information according to the linguistic and visual characteristics that it has. The goal of ontology is to offer coherence to the entities that exist inside the world by methodically classifying them and providing them with a classification system. It is necessary to use a machine learning algorithm in order to grasp visual input. At the end of the process, an assessment of the accuracy and effectiveness of the current model in comparison to earlier iterations is offered. Prior model's output was used to choose just the most relevant rows from a 10,000-row dataset using a NB classifier.

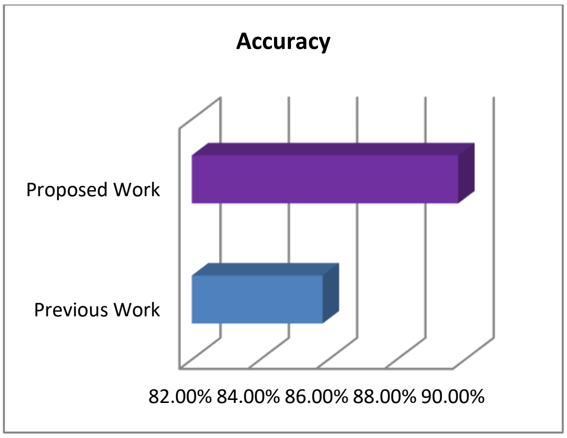
7.2.1. Applying LSTM Model on filtered dataset

When the naive classifier is taken into account, 425 entries are dropped. The LSTM model has been trained using a dataset including 9577 samples. The parameters used in this case are a gradient threshold of 2, a batch size of 16, and a total of 200 hidden units. This model has a tightly connected softmax and classification layer.

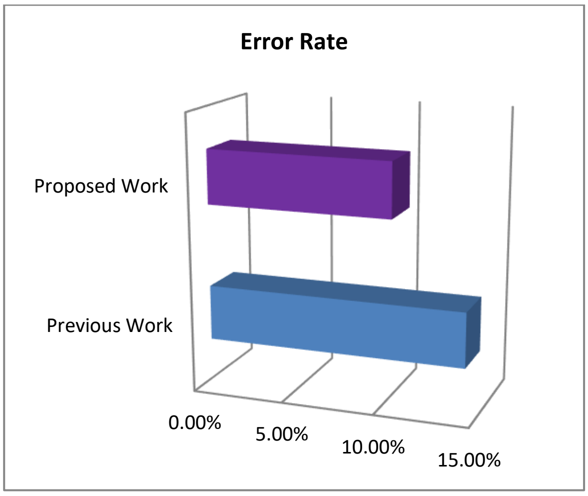
Table 2. Errors and Accuracy in Proposed Work

TRUE VALUE	2605
FALSE VALUE	395
TOTAL	3k
ACCURACY	89.8%
ERROR	10.2 %

Accuracy, precision, & recall value comparisons have been done at this stage.



(a)



(b)

Fig. 5. (a) Accuracy and (b) Error Rate

7.3. Performance Comparison

It has been noted that unfiltered data has increased the duration of earlier studies. It took longer to complete the training procedure. Additionally, the suggested model's design has shortened processing times. Traditional methods take 5 times as long as the suggested Hybrid method. The following diagram depicts the results of a simulation of the time spent in each of the three scenarios.

Table 3. Comparison of time consumption in three cases

Dataset	**Traditional**	**Proposed work**
100	0.51	0.34
200	1.02	0.63
300	1.55	0.94
400	2.03	1.26
500	2.61	1.52
600	3.02	1.90

700	3.57	2.12
800	4.08	2.47
900	4.55	2.77
1000	5.045	3.05

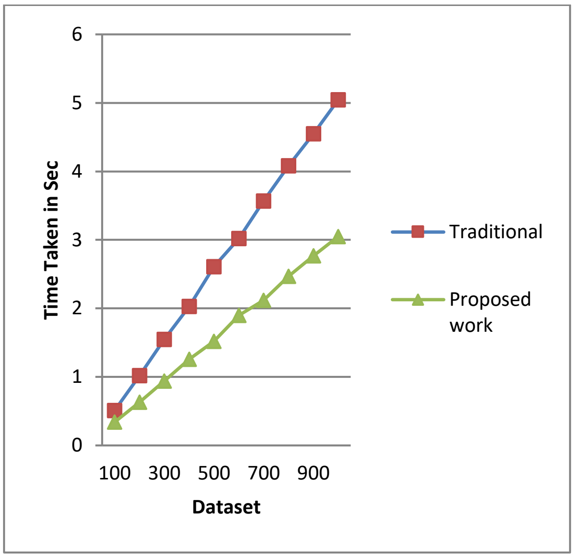


Fig. 6. Comparison of time consumption in three cases

Using machine learning, the study's results are computed. The end objective of research is to find a trustworthy solution that can be adjusted as required. Acknowledging the issue of accuracy, researchers resorted to sentiment analysis in textual data. Autonomous learning and development by automated systems may soon be within reach, thanks to machine learning. The end result is that the system can now identify and correct its own mistakes. This allows us to make well-informed decisions. One deep learning method for text categorization is neural networks, which give each text its own unique identity by assigning it a class and a label.

8. Conclusion

In conclusion, it can be seen that the development of machine learning models for sentiment analysis has undergone substantial advancements and remains very promising for a wide range of sectors and applications. In the contemporary era of digitalization, the significance of models that possess the capacity to analyze and comprehend thoughts derived from textual data is steadily growing. As we progress, it is expected that there will be advancements in models that possess more complexity in capturing the nuances of human language, effectively manage feelings in several languages and particular domains, and smoothly connect with various technology. Sentiment analysis models are expected to maintain their significant influence in several domains such as corporate decision-making, customer service, marketing, and public opinion monitoring. Organizations would be enabled to get useful insights from user-generated content and comments, therefore facilitating effective responses to consumer wants and attitudes. The accuracy of photo classification in deep learning models is contingent upon the quality and amount of the data. In order to develop models capable of accurately categorizing photographs in real-world situations, it is essential to use a well-balanced combination of both training data and real-world data.

The study's conclusions are done using machine learning. The ultimate goal of research is to identify a reliable answer that can be modified as needed. In order to address the accuracy problem, researchers turned to textual data sentiment analysis. Machine learning may soon make autonomous learning and development by automated systems possible. As a consequence, the system is now able to recognize and fix its own errors. This enables us to make knowledgeable judgments. Neural networks are one deep learning technique for text classification; they provide a class and a label to each text, giving it a distinct identity.

9. Future Scope

The future of sentiment analysis in machine learning holds immense promise and is poised for significant advancements. As machine learning models continue to evolve, we anticipate a deeper integration of natural language processing techniques, enabling systems to better grasp the intricacies of human sentiment expressed in text. The utilization of advanced deep learning architectures, such as transformers, may contribute to capturing more complex contextual relationships, thereby enhancing the accuracy of sentiment predictions. Additionally, the future is likely to see an increased emphasis on multimodal sentiment analysis, where models can effectively analyze sentiments from a combination of text, images, and possibly audio. As the demand for sentiment analysis across diverse languages and cultural contexts grows, models will likely become more adaptable and language-agnostic, fostering global applications. Ethical considerations, including bias mitigation and transparent decision-making processes, will also become paramount, ensuring responsible and fair deployment of sentiment analysis models. Overall, the future landscape of sentiment analysis in machine learning is marked by continuous innovation, expanding capabilities, and a broader scope of applications across various industries and domains.

Conflicts of interest

The authors declare no conflicts of interest.

References

[1] M. M. Hamed Taherdoost 1, “Artificial Intelligence and Sentiment Analysis : A Review in,” Comput. 2023, 12, 37., no. 0, 2023. https// doi.org/10.3390/computers12020037

[2] S. K. Assayed, K. Shaalan, M. Alkhatib, and S. Maghaydah, “Machine Learning Chatbot for Sentiment Analysis of Covid-19 Tweets,” pp. 41–55, 2023, doi: 10.5121/csit.2023.130404.

[3] N. Braig, A. Benz, S. Voth, J. Breitenbach, and R. Buettner, “Machine Learning Techniques for Sentiment Analysis of COVID-19-Related Twitter Data,” IEEE Access, vol. 11, no. February, pp. 14778–14803, 2023, doi: 10.1109/ACCESS.2023.3242234.

[4] H. Rahman, J. Tariq, M. A. Masood, A. F. Subahi, O. I. Khalaf, and Y. Alotaibi, “Multi-Tier Sentiment Analysis of Social Media Text Using Supervised Machine Learning,” Comput. Mater. Contin., vol. 74, no. 3, pp. 5527–5543, 2023, doi: 10.32604/cmc.2023.033190.

[5] L. H. B, Proceedings of the 2022 7th International Conference on Modern Management and Education Technology (MMET 2022). Atlantis Press SARL, 2023. doi: 10.2991/978-2-494069-51-0.

[6] M. S. BAŞARSLAN and F. KAYAALP, “Sentiment analysis with ensemble and machine learning methods in multi-domain datasets,” Turkish J. Eng., vol. 7, no. 2, pp. 141–148, 2023, doi: 10.31127/tuje.1079698.

[7] M. Bordoloi and S. K. Biswas, Sentiment analysis: A survey on design framework, applications and future scopes, no. 0123456789. Springer Netherlands, 2023. doi: 10.1007/s10462-023-10442-2.

[8] A. Quazi and M. K. Srivastava, “Twitter Sentiment Analysis Using Machine Learning,” Lect. Notes Electr. Eng., vol. 877, no. 05, pp. 379–389, 2023, doi: 10.1007/978-981-19-0312-0_38.

[9] C.- Reviews, C. Singh, T. Imam, and S. Wibowo, “applied sciences A Deep Learning Approach for Sentiment Analysis of,” 2022.

[10] M. Saraiva, I. Matijošaitienė, S. Mishra, and A. Amante, “Crime Prediction and Monitoring in Porto, Portugal, Using Machine Learning, Spatial and Text Analytics,” ISPRS Int. J. Geo-Information, vol. 11, no. 7, 2022, doi: 10.3390/ijgi11070400.

[11] A. Joshi, B. Akash, R. Bharathkhanna, and T. Srihari, “Improved Comment Sentiment Analysis Method Using Deep Learning,” vol. 10, no. 5, pp. 722–726, 2022.

[12] A. Alsayat, “Improving Sentiment Analysis for Social Media Applications Using an Ensemble Deep Learning Language Model,” Arab. J. Sci. Eng., vol. 47, no. 2, pp. 2499–2511, 2022, doi: 10.1007/s13369-021-06227-w.

[13] Ezenwobodo and S. Samuel, “International Journal of Research Publication and Reviews,” Int. J. Res. Publ. Rev., vol. 04, no. 01, pp. 1806–1812, 2022, doi: 10.55248/gengpi.2023.4149.

[14] A. Motz, E. Ranta, A. S. Calderon, Q. Adam, F. Alzhouri, and D. Ebrahimi, “Live Sentiment Analysis Using Multiple Machine Learning and Text Processing Algorithms,” Procedia Comput. Sci., vol. 203, pp. 165–172, 2022, doi: 10.1016/j.procs.2022.07.023.

[15] D. Geethangili, “Machine Learning Approach based Sentiment Analysis, Classification: An Application of Natural Language Processing,” vol. 71, no. 4, pp. 773–786, 2022, [Online]. Available: http://philstat.org.ph

[16] G. I. Ahmad, J. Singla, A. Ali, A. A. Reshi, and A. A. Salameh, “Machine Learning Techniques for Sentiment Analysis of Code-Mixed and Switched Indian Social Media Text Corpus: A Comprehensive Review,” Int. J. Adv. Comput. Sci. Appl., vol. 13, no. 2, pp. 455–467, 2022, doi: 10.14569/IJACSA.2022.0130254.

[17] A. P. Rodrigues et al., “Real-Time Twitter Spam Detection and Sentiment Analysis using Machine Learning and Deep Learning Techniques,” Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/5211949.

[18] A. Yenkikar, C. N. Babu, and D. J. Hemanth, “Semantic relational machine learning model for sentiment analysis using cascade feature selection and heterogeneous classifier ensemble,” PeerJ Comput. Sci., vol. 8, pp. 1–34, 2022, doi: 10.7717/PEERJ-CS.1100.

[19] C. Chen, B. Xu, J. H. Yang, and M. Liu, “Sentiment Analysis of Animated Film Reviews Using Intelligent Machine Learning,” Comput. Intell. Neurosci., vol. 2022, 2022, doi: 10.1155/2022/8517205.

[20] P. A. Grana, “Sentiment Analysis of Text Using Machine Learning Models,” Int. Res. J. Mod. Eng. Technol. Sci., no. 05, pp. 2582–5208, 2022, [Online]. Available: www.irjmets.com

[21] D. Panchal, M. Mehta, A. Mishra, S. Ghole, and M. S. Dandge, “Sentiment Analysis Using Natural Language Processing,” Int. J. Res. Appl. Sci. Eng. Technol., vol. 10, no. 5, pp. 2262–2266, 2022, doi: 10.22214/ijraset.2022.42711.

[22] K. L. Tan, C. P. Lee, K. M. Lim, and K. S. M. Anbananthen, “Sentiment Analysis With Ensemble Hybrid Deep Learning Model,” IEEE Access, vol. 10, no. September, pp. 103694–103704, 2022, doi: 10.1109/ACCESS.2022.3210182.

[23] S. Kumar, K. Sharma, D. Veragi, and A. Juyal, “Sentimental Analysis of Movie Reviews Using Machine Learning Algorithms,” 2022 Int. Conf. Mach. Learn. Big Data, Cloud Parallel Comput. COM-IT-CON 2022, vol. 02006, pp. 526–529, 2022, doi: 10.1109/COM-IT-CON54601.2022.9850878.

[24] S. Sahu, K. Divya, D. N. Rastogi, P. K. Yadav, and D. Y. Perwej, “Sentimental Analysis on Web Scraping Using Machine Learning Method,” J.

Inf. Comput. Sci., vol. 12, no. 8, 2022, doi: 10.12733/JICS.2022/V12I08.535569.67004.
[25] J. R. Monalisha Sahoo, “Survey on Sentiment Analysis to Predict Twitter Data using Machine Learning and Deep Learning,” Int. J. Eng. Res. Technol., vol. 11, no. 7, pp. 506–512, 2022.
[26] V. Talukdar, D. Dhabliya, B. Kumar, S. B. Talukdar, S. Ahamad, and A. Gupta, “Suspicious Activity Detection and Classification in IoT Environment Using Machine Learning Approach,” 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC). IEEE, Nov. 25, 2022. doi: 10.1109/pdgc56933.2022.10053312.
[27] P. R. Kshirsagar, D. H. Reddy, M. Dhingra, D. Dhabliya, and A. Gupta, “A Scalable Platform to Collect, Store, Visualize and Analyze Big Data in Real- Time,” 2023 3rd International Conference on Innovative Practices in Technology and Management (ICIPTM). IEEE, Feb. 22, 2023. doi: 10.1109/iciptm57143.2023.10118183.
[28] M. Dhingra, D. Dhabliya, M. K. Dubey, A. Gupta, and D. H. Reddy, “A Review on Comparison of Machine Learning Algorithms for Text Classification,” 2022 5th International Conference on Contemporary Computing and Informatics (IC3I). IEEE, Dec. 14, 2022. doi: 10.1109/ic3i56241.2022.10072502.
[29] D. Mandal, K. A. Shukla, A. Ghosh, A. Gupta, and D. Dhabliya, “Molecular Dynamics Simulation for Serial and Parallel Computation Using Leaf Frog Algorithm,” 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC). IEEE, Nov. 25, 2022. doi: 10.1109/pdgc56933.2022.10053161.
[30] P. R. Kshirsagar, D. H. Reddy, M. Dhingra, D. Dhabliya, and A. Gupta, “A Review on Application of Deep Learning in Natural Language Processing,” 2022 5th International Conference on Contemporary Computing and Informatics (IC3I). IEEE, Dec. 14, 2022. doi: 10.1109/ic3i56241.2022.10073309.
[31] P. R. Kshirsagar, D. H. Reddy, M. Dhingra, D. Dhabliya, and A. Gupta, “Detection of Liver Disease Using Machine Learning Approach,” 2022 5th International Conference on Contemporary Computing and Informatics (IC3I). IEEE, Dec. 14, 2022. doi: 10.1109/ic3i56241.2022.10073425.
[32] V. V. Chellam, S. Praveenkumar, S. B. Talukdar, V. Talukdar, S. K. Jain, and A. Gupta, “Development of a Blockchain-based Platform to Simplify the Sharing of Patient Data,” 2023 3rd International Conference on Innovative Practices in Technology and Management (ICIPTM). IEEE, Feb. 22, 2023. doi: 10.1109/iciptm57143.2023.10118194.
[33] P. Lalitha Kumari et al., “Methodology for Classifying Objects in High-Resolution Optical Images Using Deep Learning Techniques,” Lecture Notes in Electrical Engineering. Springer Nature Singapore, pp. 619–629, 2023. doi: 10.1007/978-981-19-8865-3_55.
[34] N. Sindhwani et al., “Comparative Analysis of Optimization Algorithms for Antenna Selection in MIMO Systems,” Lecture Notes in Electrical Engineering. Springer Nature Singapore, pp. 607–617, 2023. doi: 10.1007/978-981-19-8865-3_54.
[35] V. Jain, S. M. Beram, V. Talukdar, T. Patil, D. Dhabliya, and A. Gupta, “Accuracy Enhancement in Machine Learning During Blockchain Based Transaction Classification,” 2022 Seventh International Conference on Parallel, Distributed and Grid Computing (PDGC). IEEE, Nov. 25, 2022. doi: 10.1109/pdgc56933.2022.10053213.