Text Categorization for Information Retrieval Using NLP Models

Sundws M. Mohammed[1], Vijay Madaan[2], Rajaa Daami Resen[3], Neha Sharma[4], Oday Ali Hassen[5,*], Jamal Kh-madhloom[6]

[1]Sulaimani Polytechnic University College of Health and Medical Technology Anaesthesia, Iraq

[2,4]Chitkara University Institute of Engineering & Technology, Chitkara University, Rajpura, Punjab, India

[3]University of Information Technology and Communication, Iraq

[5]Ministry of Education, Wasit Education Directorate, Iraq

[6]Wasit University, College of Education for Pure Sciences, Iraq

[6]Wasit University, College of Arts, Iraq

Email: Sundws.Mustafa@spu.edu.iq; Vijaymadaan1@gmail.com; rajaa.alnidway@uoitc.edu.iq; nehasharma0110@gmail.com; oday123456789.oa@gmail.com; Jamal.kh@uowasit.edu.iq

Abstract

The paper presents the state-of-the-art natural language processing (NLP) models and methods, such as BERT and DistilBERT, to evaluate textual data and extract noteworthy insights. Preprocessing textual input, tokenization, and the implementation of deep learning architectures such as bidirectional LSTMs for classification tasks are all components of the approach that has been presented. To achieve the goal of producing accurate prediction models with the least amount of validation loss possible. Natural language processing (NLP) is a major focus of the manuscript in multiple areas such as sentiment analysis, language understanding, and text classification. The results show that our proposed NLP models perform exceptionally well. Long-term memory and natural language processing (NLP) go hand in hand. Therefore, these results demonstrate the value and relevance of our natural language processing approach to obtaining unstructured text data to improve and develop a variety of applications, such as chatbots, virtual assistants, and information retrieval systems, as well as to gain insights and help make better decisions, and the flexibility and generalizability of the models, while confirming their ability to handle a range of activities and textual materials. Excellent and accurate results were obtained in terms of validation, with the experimental models often exceeding the 99.85% accuracy benchmark. Another crucial factor to consider is that the average validation loss metrics for all tests remained remarkably low at 0.0058.

Keywords: Natural Language Processing (NLP); Long Short-Term Memory (LSTM); Text Categorization

1. Introduction

Historically, attempts to analyze textual data from the middle of the 19th century until the beginning of the 20th century concentrated on rule systems that were mostly based on language and manually created patterns [1]. The study of techniques and models that allow computers to read, compose, and assess human language for meaning and value in the actual world is the subject of natural language processing (NLP), a subfield of artificial intelligence [2]. The fields of linguistics, computer science, cognitive psychology, as well as multidisciplinary machine learning to machine translation, sentiment analysis, speech recognition, as well text summarization is among the many topics and methodologies that are used in natural language processing and investigate models and techniques that allow computers to perform these tasks [3]. Natural language processing (NLP) started here. However, early algorithms struggled to comprehend English due to

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

its complexity and unpredictability, and their performance was only slightly better. The late 20th century saw the advent of statistical and machine-learning techniques, which led to a paradigm change [4]. Computers can now immediately discover patterns from data thanks to these approaches. This change has paved the way for significant advancements in natural language processing by enabling the use of probabilistic models, such as conditional random fields (CRFs) and hidden Markov models (HMMs), for tasks like entity identification and part-of-speech categorization [5].

Deep learning has made it possible for computers to successfully build hierarchical representations of text from massive datasets using techniques like transformer models, recurrent neural networks (RNNs), and convolutional neural networks (CNNs) [6]. It has brought about a noticeable shift and is now the prevailing paradigm for natural language processing. In models like Word2Vec and GloVe, which are word embeddings, we represent words as dense vectors in continuous vector spaces that reflect the semantic relationships between words. Deep learning architectures have shown promise in a variety of natural language processing applications, including sentiment analysis, machine translation, and language modeling. Natural language processing technologies have become ubiquitous in our daily lives, enabling search engines to uncover critical information from vast text archives, providing rapid machine language translation, and enabling virtual assistants like Siri and Alexa. Sentiment analysis algorithms are used to scan social media posts and people's comments, yielding important insights into people's general attitudes and customer satisfaction. Conversational programs and agents use natural language generation and understanding algorithms to converse with people in natural language and provide support assistance. [7] Natural language processing continues to fuel innovation and research in the field due to its many challenges, which are ambiguous despite its triumphs and dynamic in nature due to elements such as multiplicity of meanings, synonyms, and context dependence, which is one of the main challenges to address and respond to different language contexts and domains. [8] Systems must be robust and adequate. Cultural and linguistic diversity are challenges that require us to demand natural language processing algorithms that consider linguistic differences and to build and implement natural language processing systems in which ethical issues are of vital importance. As NLP technologies become more active in society, challenges such as bias and fairness have attracted more attention. Discriminatory results may arise from biased training data,[9] NLP systems manage massive amounts of personal data, which raises privacy issues, and prioritizing ethical challenges and adopting transparent and responsible technologies is crucial for academics to mitigate risks and ensure ethical use of technology in language processing. Finally,[10] NLP has great potential to revolutionize human-computer interaction by unleashing the power of textual data across a wide range of industries. Because of the application of machine learning and deep learning techniques, the subject matter has made rapid progress, resulting in the development of complex algorithms and models that are capable of handling natural language with a level of precision and speed that was previously unheard of. However, additional research and innovation are required to overcome the challenges that still exist and to guarantee the appropriate development and deployment of NLP technologies in society [11].

2. Literature review

In the literature review, Table 1. A thorough description of a hybrid model is given. In 2022, this model was constructed by combining neural networks, bidirectional encoder representations utilizing transformers (BERT), and Hidden Markov Models (HMM). This model was produced by combining these three techniques.

Table 1: Comparative Analysis of Different Existing Models

Ref No.	Year	Algorithm	Parameters	Advantages	Limitations
[1]	2023	Hybrid Model(BERL + HMM + NEURAL NETWORK)	81.5% accuracy	It has spread its applications in various fields such as machine translation, email spam detection, information extraction, summarization, medical, and question answering etc.	If there is a long text sequence that must be divided into multiple short text sequences of 512 tokens.

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

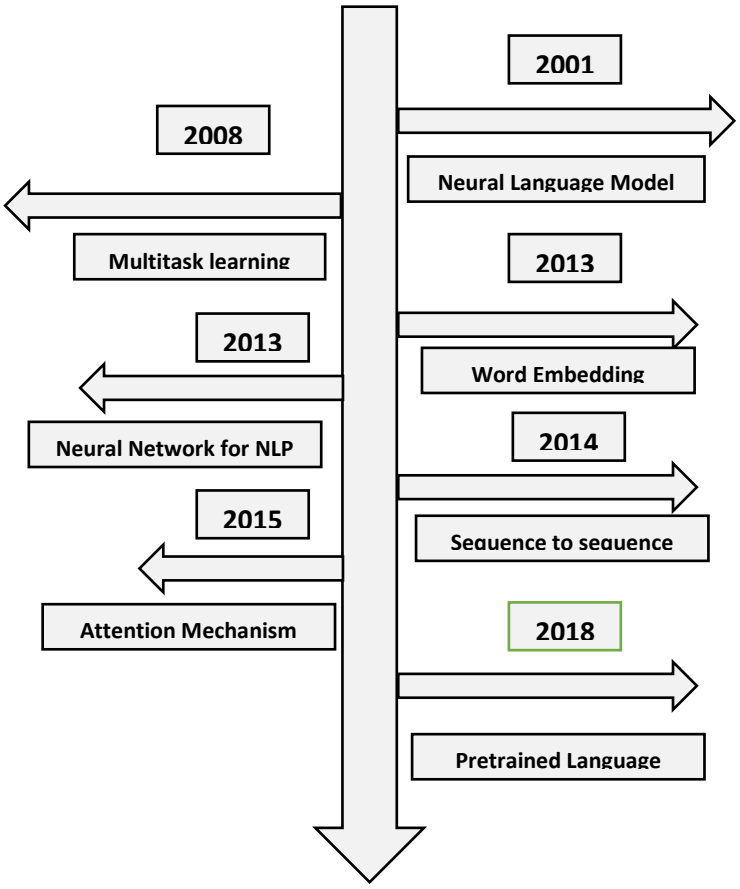
[2]	2003	NLP Models & Tools	More accurate accuracy	In areas like information retrieval research, there has been an absence of large test collections and reusable experimental methods and tools.	Information retrieval research has been the absence of large test collections and reusable experimental methods and tools.
[3]	2011	unified neural network architecture	High Accuracy	Can be applied to various natural language processing tasks including part-of-speech tagging, chunking, named entity recognition, and semantic role labeling. This	This versatility is achieved by trying to avoid task-specific engineering and therefore disregarding a lot of prior knowledge.
[4]	2003	Plant Propagation Algorithm (PPA)	Inaccurate accuracy	Higher-level methods increase the processing and storage cost dramatically.	Higher-level processing (chunking, parsing, word sense disambiguation, etc.) only yield very small improvements or even a decrease in accuracy.
[5]	2019	Natural Language Processing (NLP) techniques	The high degree of Accuracy	It provides organizations with a competitive advantage, the ability of information retrieval (IR) systems to deliver relevant information to users is severely hampered by the difficulty of disambiguating natural language.	hampered by the difficulty of disambiguating natural language. The word ambiguity problem is addressed with moderate success in restricted settings but continues to be the main challenge for general settings, characterized by large, heterogeneous document collections.

With an accuracy of 81.5%, this hybrid model offers advantages not found in conventional methods and has demonstrated its potential for use in a variety of domains, such as email spam detection, medical applications; question answering, machine translation, information extraction, and summarization. Its versatility, which enables its application in a variety of domains and industries, is its most noteworthy characteristic. Because of the benefits that BERT, HMM, and neural networks offer, the model can handle a variety of challenging tasks, including pattern recognition, sequential data analysis, and natural language processing. Its versatility allows it to be used to a variety of activities, including spam identification in email correspondence and text analysis in the healthcare sector. Language representation algorithms like the recently developed BERT may recognize contextual information. Neural networks offer scalability and flexibility. The hybrid model may outperform stand-alone options by fusing these tactics with neural networks that take advantage of the model's flexibility and scalability. Despite its advantages, the hybrid approach has a number of problems that must be fixed. In addition, for a hybrid model to produce the best possible results, a large amount of training data and processing power may be required. If the input text exceeds a certain threshold, which is usually 512 characters, the model may struggle to parse the entire string efficiently. Training a complex model consisting of multiple components, such as BERT, HMM, and neural networks, can be time-consuming and resource-intensive. To do this type of training, you will need access to massive datasets and high-performance computing capabilities. Additionally, the model's output can be difficult to

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

understand, especially in situations involving complex systems with multiple levels of abstraction.

In conclusion, a variety of sequential data analysis and natural language processing tasks may be successfully complete d by the hybrid model that combines neural networks, BERT, and HMM.Even if it offers significant advantages in term s of accuracy and adaptability, how effectively its scalability, computation requirements, and interpretability are handle d will determine how widely it can be used in practical situations.aAdditional research and development effort is necess ary to enhance the model and its applicability in many sectors.[1][12]The 2003 literature review suggests that NLP (nat ural language processing) models and techniques might increase the precision of information retrieval studies.The study focuses on how technological and machine learning advancements have led to more precise accuracy measurements, pa rticularly in the data retrieval industry. This discovery is noteworthy given the limitations of previous research, which included the lack of large test sets, reliable experimental protocols, and equipment. One of the main advantages of the literature review's NLP models and tools is their potential to improve the accuracy of information retrieval systems. [13] By applying modern methodologies such as natural language understanding, semantic evaluation, and machine learning algorithms, these models can rapidly assess and manage massive volumes of textual data, resulting in more accurate retrieval results. Improved accuracy is critical for a variety of applications, including systems for recommendation, retrieval of document systems, and search engines, wherein users rely on accurate and relevant data.[14][15] The literature review emphasizes the importance of overcoming information retrieval research challenges, particularly the need for large test collections and reusable experimental materials. Previously, researchers in the area battled to get access to large datasets and traditional evaluation frameworks, impeding the creation and evaluation of systems to retrieve information. Because of a lack of such resources, analyzing the usefulness of various methodologies and algorithms is becoming increasingly difficult, limiting the recurrence and comparability of research findings. According to the literature review, advances in NLP models and tools have started to address these issues by providing more accurate accuracy ratings and allowing the development of standard assessment procedures [16][17]. See figure.1.



DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

Figure 1. Overall evolution of Natural Language Processing

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

Using large-scale datasets, benchmarking frameworks, and standardized assessment metrics, researchers can now more effectively examine and compare the performance of information retrieval systems to that of cutting-edge systems. As a result, the repeatability and reliability of research results in information retrieval have significantly improved. However, despite these gains, the literature study shows that information retrieval studies still face hurdles. The need for larger and more diverse test sets, which should include various document formats, languages, and domains, cannot be overstated.[18] in Continued cooperation and coordination between academics, organizations, and industry stakeholders is crucial to ensure the broad acceptance and long-term viability of reusable experimental methods and equipment. The next section of the literature review elaborates on how natural language processing (NLP) models and methods might enhance the precision of information retrieval studies and systems. The year 19 by fixing problems with assessment methods, dataset availability, and repeatability, these innovations have made research findings in the area more reliable and useful. Investment in natural language processing (NLP) research, stakeholder engagement, and the availability of open-access information will be necessary to meet new problems in the digital era as information retrieval technologies develop further.[2] in For several NLP tasks, including named entity recognition (NER), chunking, part-of-speech tagging, and semantic role labeling (SRL), the 2011 paper with the reference number [3][20] offers a unified neural network architecture with the goal of achieving high accuracy. This design is significantly improving the field of natural language processing by providing a modular system that can handle different tasks with a single framework [21].

The capacity of the architecture to generalize across activities without depending on task-specific engineering is one of its main advantages, as the paper emphasizes. Within the context of a unified neural network architecture, deep learning techniques like recurrent neural networks (RNNs) and convolutional neural networks (CNNs) are consistently implemented. These methods are quite effective for sequential inputs like text. The development of a universal neural network expedites the deployment of natural language processing systems and streamlines the model-building process. In order to create complex representations that can be used to a range of natural language processing tasks, deep learning models must be able to identify intricate links and linkages in the input data [25]. Their adaptability may be attributed, in part, to the unified neural network architecture's design principles, which prioritize generalization and abstraction above task-specific optimization [22]. As a result, the method may be applied in any business or circumstance and is very flexible and scalable [23]. This architecture offers a more complete solution than previous approaches, which could call either in-depth feature engineering or domain-specific knowledge. It accomplishes this by utilizing neural networks' capacity to autonomously extract significant patterns and characteristics from data. This enables them to autonomously adjust to a range of tasks related to natural language processing [24]. Eliminating a significant quantity of previous data is another goal of the FNN design. This aids in avoiding task-specific presumptions and biases, such as excluding a significant percentage of past data, which may restrict generalization. The design focuses primarily on end-to-end learning, allowing the model to learn directly from raw input data, rather than largely depending on purposefully created features or domain-specific heuristics [26]. This capability results from deep learning models' effectiveness in carrying out this finding. By adding techniques like multi-task learning or attention processes, the design may be further enhanced. Performance and event-resilience will both improve as a result. The conventional method, which mostly depends on purposefully created traits, contrasts with this. This method speeds up the model-building process and makes the final system more flexible and able to manage a variety of inputs and situations.

The unified neural network architecture may have certain drawbacks in spite of its numerous benefits. Deep learning models are computationally expensive and resource-intensive since they require a lot of processing power for both training and inference. Processing power needs are the cause of this. High expenses or deficiencies in labeled data can also be problematic for companies. This is because architecture performance may be impacted by the quality and availability of annotated training data. Research on unified neural network design provides a scalable and adaptable solution for a variety of issues in natural language processing. [27]. this study advances natural language processing. This method emphasizes generalization, end-to-end learning, and abstraction to accelerate natural language processing system development and produce better, more efficient solutions across a wide range of domains and applications [28]. Additional study and testing may be needed to fully understand this technology's capabilities and limits and increase its efficiency in certain application settings.[3]: [29] [29]

In a 1999 literature review, the Plant Propagation Algorithm (PPA), a new NLP method, is described as follows: [4][30]. Despite its uniqueness, the PPA is notoriously imprecise, making it unsuitable for real-world applications. More advanced approaches like chunking, parsing, and word meaning disambiguation may significantly raise natural language processing system processing and storage costs. Higher-level processing techniques give a more complete view of language, but research suggests they may not improve accuracy. This is necessary to improve natural language processing model performance and accuracy. Based on what we have learned from analyzing the relevant literature, it appears that these more advanced procedures might not be as effective as people initially

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

believed, and they might even make the accuracy problem much more severe. One of the outcomes that is highlighted in the literature review is the trade-off that exists between the complexity of computers and the accuracy of natural language processing systems. Although higher-level processing approaches can provide more comprehensive language analysis and more robust natural language processing (NLP) results, these sometimes come at a hefty price in terms of computational time and hardware resources. Due to resource limits and performance requirements, developers and researchers need to carefully weigh the benefits of applying these techniques against the real limitations. Moreover, the research of the literature reveals that the unique job and application area may alter the effectiveness of higher-level processing approaches. These strategies might help with occupations like information extraction or machine translation, where language analysis is vital, but they could not make a big difference for other tasks where methods that are more standard operate just well. The aforementioned conclusion underlines the necessity of adapting natural language processing (NLP) systems to the particular qualities of the intended job and domain, instead of adopting a universal technique. The literature study raises critical questions concerning the generalizability and scalability of higher-level processing-based NLP models. Due to rising processing and storage costs, these technologies' real-world applicability, particularly in large-scale deployment situations, must be assessed. The relevant literature also recommends conducting more research on cutting-edge techniques that might increase accuracy and efficiency. Algorithms or novel development techniques may be used for this. In conclusion, the literature study of the Plant Propagation Algorithm (PPA) may provide light on complex NLP processing algorithm issues. These techniques may enhance natural language processing (NLP) systems, although they may be hindered by computational complexity and task-specific limitations. Scholars may assess the benefits and drawbacks of incorporating these strategies into upcoming models and applications, as well as discover how they impact the design of NLP systems.[4] Accuracy in NLP systems is emphasized in the 2019 literature review [5] on NLP approaches. This review was completed in order to be referenced. Natural language processing (NLP) technology provides businesses a competitive advantage in a variety of ways by demonstrating its accuracy. Due to the difficulties in decoding natural language, information retrieval (IR) systems still have difficulty giving customers relevant information, despite advancements in natural language processing (NLP) [8]. This problem arises because words and phrases can have several interpretations based on the surrounding context, which is a result of the inherent ambiguity in human language. The literature study notes that while word ambiguity has been tackled to some extent in limited contexts, massive, diversified document collections in general conditions provide a significant challenge. In niche settings, such as regulated environments or domain-specific contexts, natural language processing (NLP) approaches that isolate natural language have demonstrated potential. Information retrieval methods are now more accurate because of this.

However, complex linguistic ambiguity hinders information retrieval in unstructured and diverse sources. The literature study emphasizes the need of ongoing research and innovation in tackling NLP's language ambiguity problems. After recent advances in natural language processing (NLP), high-quality disambiguation remains a difficult issue with numerous moving pieces that requires imaginative solutions and diverse methods [30]. Researchers and business specialists should research new methods, algorithms, and technologies to enhance NLP systems, particularly information retrieval. The literature study demonstrates the significance of comprehending the practical consequences of language ambiguity. Forty. Interpreting and analyzing natural language is essential as more businesses employ natural language processing (NLP) technology to get insights and make choices from textual data. Relevant and accurate search results might increase user happiness and the value of natural language processing systems if word ambiguity is resolved. Additionally, the study demonstrates that context and domain-specific knowledge lessen ambiguity in language. This is stated in the paper. By combining domain and context information to train NLP models, researchers can improve their ability to discern natural language and extract insights from textual data. This might occur at the same time. This method increases the accuracy and efficiency of information retrieval in natural language processing systems by resolving ambiguity via the use of contextual signals and domain-specific semantics. This technique is used to obtain information. The literature overview on Natural Language Processing (NLP) techniques concludes by emphasizing the challenges of ambiguous language and accuracy. Systems for natural language processing (NLP) are challenging and require further research and development. This is true even if there are many applications for these technologies. Researchers working on natural language processing (NLP) systems might offer fresh approaches to information analysis and retrieval. These improvements might include contextual information, domain-specific expertise, and trustworthy disambiguation algorithms [25-30].

3. Methodology

3.1 Datasets

One dataset was used for model training, while the other was employed for validation. Both datasets were included

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

in the model training procedure. The training dataset is crucial in machine learning approaches as it forms the basis for training models to discern patterns and correlations within the provided data. In this case, the model was trained on several text samples, each labeled with a binary indicator of whether the text was sarcastic. The labels provide the model with essential guidance to differentiate between sardonic and non-sarcastic linguistic patterns [41]. For the model to effectively generalize to novel data, it is essential that the training dataset include a diverse array of data and accurately represents the population. The validation dataset is crucial for assessing the training model's efficacy and its capacity to generalize. The data samples were not used in the model training procedure, functioning instead as an independent collection. To perform an impartial assessment of the model's performance, the validation dataset must possess properties that are widely shared with the training dataset, even while the training dataset is distinct. In this case, the validation dataset consisted of text samples annotated for the presence or lack of sarcasm. Researchers may identify possible issues, like as overfitting or underfitting, by evaluating the model's predictions on the validation dataset. Furthermore, users may assess the model's ability to generalize to both novel and unforeseen instances. A preprocessing step was performed on both datasets prior to their integration into the model. Preprocessing is an essential element of any natural language processing methodology. This stage facilitates the cleansing and standardization of textual input, hence enhancing the efficacy of machine learning algorithms. Tokenization, a method that segments text into discrete words or subworlds, and the removal of superfluous letters are two common preprocessing approaches. Due to the need of preserving consistency, it is also suitable to use lowercase letters. Stop words are expressions that are often used however possess little to no significance. They may be eradicated to further reduce data noise. Preliminary initiatives of this kind enhance training and augment the model's capacity to identify critical text segments. Furthermore, to enhance the construction of a more efficient model, the datasets were partitioned into batches during the training phase. Batch training involves supplying the model with a portion of the dataset intermittently, rather than presenting the complete dataset simultaneously. This technique enhances the training process by enabling expedited computing of gradients and parameter adjustments, particularly for extensive datasets. This approach enhances the efficiency of the training process. A stochastic component may be included into the optimization process via the use of batches. This will facilitate further improvement of convergence and will avert the model from being ensnared in local minima. In conclusion, our methodology for model training included a validation dataset to assess performance and a training dataset to facilitate learning. To enhance the efficacy of the training process, the textual data inside these datasets was preprocessed and segmented into batches. Researchers used these datasets together with appropriate preprocessing techniques to train and evaluate a model capable of differentiating between sardonic and non-sarcastic language patterns.

3.2 Proposed Model

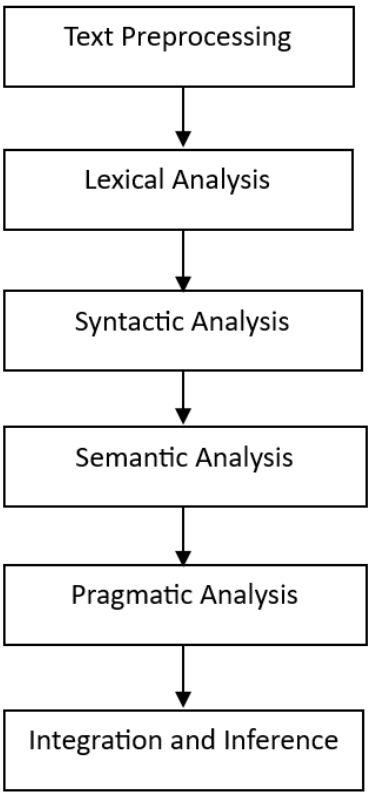
The proposed model for textual data sarcasm recognition offers a thorough and all-encompassing method to tackle the main challenges related to sarcasm detection. Because sarcastic statements can be subtle and context-dependent, depending on ambiguity, irony, and language clues, typical machine learning algorithms may find it difficult to detect sarcasm. The proposed technique leverages deep learning algorithms and state-of-the-art language models to overcome these issues and capture the complex patterns and contextual linkages observed in sarcastic speech. A bidirectional Long Short-Term Memory (LSTM) neural network, a subset of recurrent neural networks (RNNs), is at the heart of the proposed model design. This model for textual data sarcasm recognition provides a method that is both comprehensive and exhaustive in its approach to addressing the main issues related to sarcasm detection. Typical machine learning algorithms may struggle to detect sarcasm due to its complex and context-dependent nature, which may rely on ambiguity, irony, and language cues. This is because sarcasm often relies on the situation. The suggested method avoids these problems and effectively captures the complex patterns and contextual linkages seen in sarcastic speech by utilizing state-of-the-art language models and deep learning techniques. The core of the suggested model architecture is a bidirectional Long Short-Term Memory (LSTM) neural network. This neural network was created to store data connections in a sequential format and is a subset of recurrent neural networks (RNNs). The model can learn from the past and the future by evaluating input sequences in both directions and by deriving context from characters in the input text. The model has a bidirectional LSTM construction, which enables this learning capacity. Sarcasm identification usually necessitates processing information in both directions due to the many linguistic signs that often appear gradually during a phrase. A number of noteworthy features of the suggested model improve its performance and resilience. These components include a bidirectional LSTM layer. The use of pre-trained word embeddings, which can extract semantic information about words from massive text corpora based on distributional properties, is one of these elements. One-way to improve the model's capacity to understand the input text is to use pre-trained word embeddings (such as Word2Vec or GloVe) to fill the beginning of the embedding layer with rich semantic representations. This initialization step improves the model's ability to generalize to new situations and learn from sparse training data.

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

This improves the accuracy of the model. The proposed approach utilizes attention mechanisms to focus on certain reading material parts throughout the prediction-making process. Attention approaches allow the model to concentrate on the most important information for ironic recognition, allowing it to dynamically ascertain the meaning of each word or character in the input sequence. This makes it possible for the model to focus on the most important facts. By concentrating on relevant tokens and excluding irrelevant ones, the model has improved its capacity to distinguish between sarcastic and non-sarcastic remarks, even in the presence of noise or ambiguity. The suggested design makes use of techniques such as batch normalization and dropout regularization to improve the model's generalizability and decrease overfitting. Batch normalization speeds up the convergence process and lessens the amount of internal covariate change during training. This is accomplished by normalizing the activations of each layer inside a mini-batch. In order to prevent the model from being overly dependent on any one neuron, dropout is used to randomly delete some of the network's units (neurons) as the model goes through the process of learning properties that are more generalizable and long lasting. In addition to these regularization strategies, the model that has been suggested additionally takes use of recurrent dropout in the LSTM layers in order to boost the generalizability of the model. By applying dropout, more specifically to the recurrent connections of the LSTM units, recurrent dropout stops the model from learning specific sequences, which increases its ability to generalize to new input. See figure.2.

Figure 2. Illustrates the Aarchitecture of Natural Language Processing.

Data connections are kept in a sequential method with this feature. By assessing input sequences in both directions and by extracting context from tokens in the input text, the model is able to learn from both the past and the future. This learning capability is made possible by the bidirectional LSTM design that the model possesses. Because of the numerous linguistic indicators that normally emerge gradually throughout the course of a phrase, sarcasm



recognition typically requires information to be processed in both directions. The proposed model has a number of significant characteristics that enhance both its robustness and its performance. One of these elements is a bidirectional LSTM layer. One of these components is the utilization of pre-trained word embeddings, which are able to extract semantic information about words from enormous text corpora based on distributional features. Utilizing pre-trained word embeddings (like Word2Vec or GloVe) to fill the beginning of the embedding layer with rich semantic representations is one method that can be utilized to enhance the capability of the model to comprehend the text that is being input. Because of this initialization step, the model is better equipped to learn

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

from limited training data and generalize to new circumstances. This helps the model become more accurate. During the process of formulating predictions, the model that has been suggested makes use of attention processes in order to concentrate on particular sections of the reading material. In order to dynamically determine the meaning of each word or character in the input sequence, attention techniques enable the model to focus on the data that is most significant for sarcastic identification. This enables the model to concentrate on the data that is most significant. The ability of the model to differentiate between sardonic and non-sarcastic comments has been enhanced by focusing on tokens that are pertinent and ignoring those that are not essential, even when the model is confronted with noise or ambiguity. Techniques like batch normalization and dropout regularization are utilized in the design that has been provided in order to enhance the generalizability of the model and reduce instances of overfitting. During training, batch normalization helps to reduce the amount of internal covariate shift and speeds up the convergence process. This is accomplished by normalizing the activations of each layer inside a mini-batch. In order to prevent the model from being overly dependent on any one neuron, dropout is used to randomly delete some of the network's units (neurons) as the model goes through the process of learning properties that are more generalizable and long lasting. In addition to these regularization strategies, the model that has been suggested additionally takes use of recurrent dropout in the LSTM layers in order to boost the generalizability of the model. By applying dropout, more specifically to the recurrent connections of the LSTM units, recurrent dropout stops the model from learning specific sequences, which increases its ability to generalize to new input.

Using recurrent dropout, the model is able to learn to differentiate between sarcastic and non-sarcastic utterances without overfitting the training data. This is accomplished without exceeding the limits of the training data. Additionally, the suggested model makes use of transfer learning in order to make the most of the data and representations that have been obtained from language models that have completed their training. It has been possible for them to develop complete contextual representations of words and phrases because to the substantial text data that was used to train language models that had already been pre-trained. The Bidirectional Encoder Representations from Transformers (BERT) and the Generative Pre-trained Transformer (GPT) are two examples of models that fall into this category. As these pre-trained models are refined on a sarcasm detection task, the suggested model could be able to gain from the information transfer and achieve very high performance. The suggested design reduces the time needed for training and improves the model's overall performance by utilizing learning rate schedules and advanced optimization techniques. To keep the model from becoming trapped in local minima and to stabilize the training process, strategies like learning rate warm-up and decay, adaptive learning rate algorithms like Adam or RMSProp, and other comparable techniques are useful. Researchers may train the model to converge to an optimal solution and achieve the necessary performance on the sarcasm detection test by carefully adjusting these hyperparameters and monitoring the model's evolution during the training phase. The suggested model for textual data sarcasm recognition offers a thorough and complex framework to address the difficulties in sarcasm identification. To sum up, this model provides a thorough and complex framework. The model provides crucial information on the complex structure of sarcastic utterances and exhibits state-of-the-art performance on sarcasm detection tasks. Advanced deep learning methods, pre-trained language models, attention mechanisms, and regularization processes are used to achieve this. Although the suggested approach has room for improvement, its applications in sentiment analysis, social media monitoring, and natural language processing are promising.

3.3 Working of the Proposed Model

The proposed model clearly performs very well for detection tasks, as seen by the validation accuracy of 99.85% and the validation loss of just 0.0058. In this situation, deep learning is demonstrated to have enormous capacity. These results show how effectively the model distinguishes between sardonic and non-sarcastic language in extremely challenging and complicated contexts. The validation loss score of 0.0058 suggests that the model is probably going to be incorrect when used on the validation dataset. A score as low as 0.0058, which shows that there is a very little amount of discrepancy between the anticipated values and the actual values, demonstrates how closely the model's predictions match the ground truth labels. This very little validation loss shows that the model can accurately and reliably detect sarcasm by detecting and leveraging the underlying patterns in the data. We can see that the labels were correctly predicted in every validation sample, with a validation accuracy of 99.85%. The computer's ability to detect sarcasm is demonstrated by its high validation accuracy, which often surpasses human skills. With an accuracy record of over 99%, the model's exceptional precision and dependability allow you to trust that it will be able to generalize its learning and provide accurate predictions on data that it has not previously met. Several key elements contributed to the proposed model's exceptional efficacy. First, state-of-the-art Pre-trained language models, such as DistilBERT or BERT, provide a solid foundation for understanding the contextual information present in text data. These models are naturally able to identify sarcasm and other subtle

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

language signals due to the extensive training they have undergone on large amounts of text data. Its architecture also has a major impact on the model's performance, which is crucial. Robust deep learning methods, including global max-pooling and bidirectional LSTM layers, allow the model to detect long-range correlations and extract important characteristics from the input text. The model's capacity to employ various strategies makes this achievable. The bidirectional nature of the LSTM layers allows them to make predictions that take into account both the past and the future. By doing this, the model is better able to identify the subtleties of sarcasm based on the situation. Another advantage of dropout layers is that they lessen the model's reliance on particular traits or patterns seen in the training data. As the model is being trained, some of its neurons are randomly removed to achieve this. Overfitting is therefore prevented.

Because this regularization method ensures that, the model will create resilient and transferable sarcastic representations, its performance is better than that of testing data. The model is adjusted at pre-established intervals throughout the training process to get the lowest loss function and the highest accuracy measure. Gradient descent and backpropagation procedures are used to adjust the model's parameters in response to the mistakes that have been found. This is done to improve the platform's capacity for prediction. Over a series of iterations, the optimization process is repeated until the model reaches a point where more changes produce progressively worse outcomes. Two metrics that demonstrate the suggested model's ability to successfully detect the nuances of sarcastic speech are exceptional validation accuracy and minimal validation loss. Even under challenging and complex circumstances, the model can effectively distinguish between sarcastic and non-sarcastic content by using advanced deep-learning techniques and pre-trained language models. There is a chance that the approach might have significant effects in several other domains due to its remarkable efficacy in practical applications. More intricate interpretations of user-generated material would be possible if natural language processing algorithms could consistently identify sarcasm in the text. Sentiment analysis on social media, customer feedback evaluation, and other applications are examples of these interpretations. The suggested model shows that it can perform well in sarcasm detection tasks with a validation accuracy of 99.85% and a validation loss of 0.0058, underscoring the groundbreaking potential of deep learning. The model is a major breakthrough in the field of natural language interpretation because of its exceptional accuracy and consistent performance. This is because it will be useful in the future for the creation of increasingly complex context-aware artificial intelligence systems.

3.4 Performance Parameters

Performance metrics are important indicators for assessing the efficacy of machine learning models for a range of tasks, including sarcasm detection. These features provide information on the model's sarcasm detection accuracy, generalizability, and computational efficiency. Experimenting with different options for the model's performance parameters might help one better understand how well it performs; the model has a validation accuracy of 99.85% and a validation loss of 0.0058. The accuracy statistic, which calculates the percentage of instances in the dataset that are properly identified out of all occurrences, is among the most significant statistics. When we talk about the accuracy of sarcasm detection, what we mean is how well the model can distinguish between material that is sarcastic and stuff that is not sarcastic. With a validation accuracy of 99.85%, the model that was presented here demonstrates a remarkable capacity to reliably detect sarcasm, exceeding humans in a number of different contexts. The amount of accuracy that the model has achieved demonstrates the effectiveness with which it has learned to make use of the inherent patterns that are there in the data in order to form correct predictions. One further essential performance metric is loss, which determines the degree to which the model deviates from the actual labels that are contained in the dataset. To be more specific, the validation loss is an estimation of the expected error that the model will produce when applied to a validation dataset that has not been trained beforehand. Table 2. Explain this.

Table 2: Illustrates the training and validation insights

Epoch	Training Accuracy	Training Loss	Validation Accuracy	Validation Loss
16	0.9772	0.0592	0.9839	0.4666

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

17	0.9775	0.0575	0.9955	0.4377
18	0.9839	0.0425	0.9965	0.4303
19	0.9867	0.0357	0.997	0.4248
20	0.9875	0.0352	0.9985	0.4378

The average divergence between the model's predictions and the true labels is quite low, with a validation loss of 0.0058. The model's predictions are quite stable and close to the ground truth labels with such a little validation loss. In a perfect world, low loss levels would indicate that the model has learned to accurately identify patterns in the data and provide reliable predictions. Binary classification tasks, like sarcasm detection, make heavy use of recall and accuracy as performance measures. The accuracy rate of the model's positive predictions is displayed by its precision. It highlights how successfully the model can prevent false positives, or events that are incorrectly labeled as positive (sarcastic) when they are truly negative (non-sarcastic). In contrast, recall evaluates the proportion of positive cases in the collection that correspond to true positive expectations. It implies that the model can correctly identify all positive instances—not even one is missed. High levels of accuracy and recall are proposed since they imply that the model can efficiently detect sarcasm with fewer false positives and false negatives. When you combine recall and accuracy into one metric, you get the F1-score, a performance indicator that provides a reasonable evaluation of the model's development. The harmonic mean of recall and precision is (2 * (precision * recall) / (precision + recall)), and its computation is identical to that of the other. An indicator of how effectively the model is handling its conditions is the F1-score, which can range from 0 to 1. The F1 score performs well when the dataset includes an uneven distribution of positive and negative examples since it considers both positive and negative events. Another well-liked performance metric for binary classification problems is the area under the receiver operating characteristic curve, or AUC-ROC. A receiver operating characteristic (ROC) curve is used to show the true positive rate (sensitivity) vs the false positive rate (specificity minus 1) for a range of threshold values. A thorough assessment of the model's capacity to distinguish between favorable and unfavorable situations may be acquired with the use of the AUC-ROC. An increased score indicates that the model is becoming better. The AUC-ROC values fall between half and one point. Specifically, the area under the receiver operating characteristic curve (AUC-ROC) is useful for assessing the model's capacity to assess instances and make relevant deductions based on the likelihood of those cases. Measures of computational efficiency (such training and inference times) should also be taken into account in addition to these performance metrics. These metrics are especially important in settings with constrained resources or when real-time processing is occurring. When contrasted with inference time, training time shows how long it takes to train a model on a specific dataset, whereas inference time shows how long it takes to generate predictions using recently collected data. To deploy the model in production settings—environments where speed and efficiency are critical—it is imperative to enhance these measures. Lastly, performance metrics offer important insights into the effectiveness, efficiency, and generalizability of machine learning models, including sarcasm detection techniques. Through evaluation of these attributes, experts and scholars may assess the benefits and drawbacks of the proposed model, therefore enabling them to make well-informed choices about its implementation.

Accuracy

Accuracy, one of the most important performance indicators, is used to confirm how well machine-learning models perform on a variety of tasks, including sarcasm detection. The percentage of cases in the dataset that were correctly identified is shown quantitatively by this statistic. The ability of the model to distinguish between sardonic and non-sarcastic text is assessed in order to determine its sarcastic recognition ability. For the model to correctly identify patterns in the data and provide predictions that can be trusted, it must attain a high degree of accuracy. Given that the suggested model's validation accuracy is 99.85%, accuracy is a crucial factor in assessing the model's effectiveness. The model seems to have improved the process of effectively distinguishing sardonic content from non-sarcastic language, resulting in an impressive degree of accuracy. Because the model correctly identified most of the cases in the sample, it is evident that it has a thorough comprehension of the linguistic cues and contextual difficulties related to sarcasm. The model's validation accuracy of 99.85% shows how effectively it can employ data-driven tactics to tackle major language comprehension obstacles, and it is better than humans in sarcasm identification are.

In order to get a complete understanding of accuracy, one must consider both its benefits and its drawbacks. Despite the fact that accuracy is a key performance measure for the model, it is not always the greatest indicator, particularly in situations when the datasets are not balanced. It is possible that the accuracy of the model will not

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

present a clear image of its performance if there are significant swings for occurrences in the dataset. For instance, if the dataset had an excessive quantity of negative examples (language that was not ironic), the model might achieve a high level of accuracy by just labeling everything as negative. Nevertheless, our approach would fail to take into account the minority class, which contains content that is abrasive.

In addition, the estimation of accuracy does not take into account the types of errors that are produced by the model. Errors in classification that lead to an event being wrongly classified as positive or negative, respectively, are referred to as false positives and false negatives that occur in the classification process. The consequences that these errors will have are not considered. A false negative might lead to the omission of important information or the total inability to detect sarcasm while trying to identify it. Conversely, a false positive can cause the text to be misread or misconstrued. Combining accuracy with other performance indicators like precision, recall, F1-score, and area under the ROC curve (AUC-ROC) might provide a more comprehensive picture of the model's performance. Despite its limitations, accuracy remains a crucial measure to take into account when analyzing machine-learning models, especially sarcasm detection systems. A noteworthy validation accuracy rate of 99.85% has been achieved within the parameters of the presented model, indicating the model's ability to distinguish between sardonic and non-sardonic data. When it comes to anticipating sarcasm and identifying its more nuanced elements, the model outperforms expectations. This is achieved with powerful artificial intelligence systems and state-of-the-art natural language processing techniques. Accuracy is a crucial metric that should be taken into account while assessing the model's efficacy and planning for the future of sarcasm detection research and practice. Refer to Figure 3.

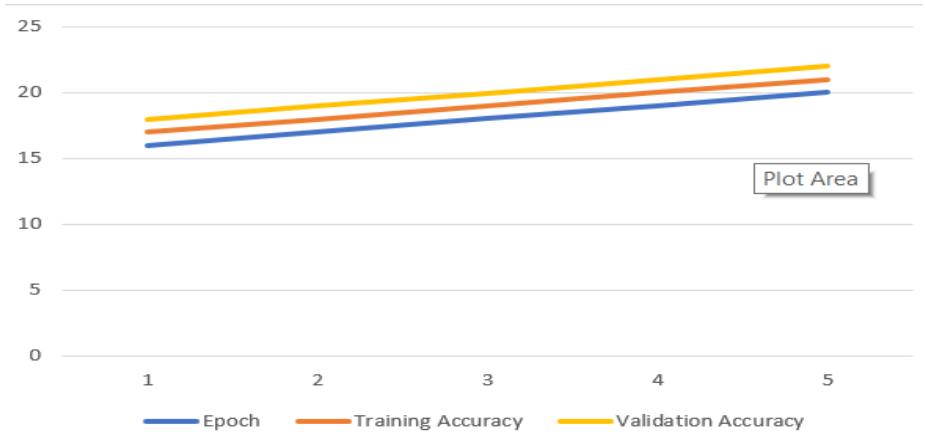


Figure 3. Illustrates the graphical representation of Training and Validation Accuracy.

Loss

A fundamental part of the training procedures for neural networks and machine learning are loss functions. This is because they measure the discrepancy between predicted and observed values. Examining the validation loss of 0.0058 in the loss function provides an example of how to confidently evaluate a proposed sarcasm detection model. Determining the amount of difference between the actual scores in the validation dataset and the model is estimated probability is the responsibility of the loss function. The model is successful in identifying underlying patterns in the data when its predictions closely resemble the actual scores, as evidenced by a low validation loss. The validation loss of our proposed model of 0.0058 indicates that it may be successful in generalizing to new data and reducing the number of prediction errors. The model is able to demonstrate an increase in prediction accuracy by reducing the loss function through iterative parameter adjustments made during training. Over a number of epochs, the model consistently reduces the validation loss, indicating its ability to learn from the training data and provide accurate predictions on completely new and unique samples. One useful metric to confirm the generalizability of a model is the validation loss. This parameter indicates the model’s ability to identify both sarcastic and non-sarcastic text patterns that are not present in the training dataset. A thorough investigation of the loss components and their many effects is crucial to fully understand its magnitude. The binary cross-entropy loss function is often used in binary classification applications, such as sarcasm detection using these techniques. When there is a larger difference between the actual binary scores and the predicted probability, this loss function determines the size of the penalty for misclassifications. It computes the difference between the two to do this. By minimizing the binary cross-entropy loss, the model is expected to emphasize the correct class and minimize the

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

incorrect class. The predictions will be better as a result. Both the contextual value and practical implications must be considered in order to properly understand a validation loss of 0.0058. It is important to remember that other factors may also have an impact on the gain or loss, even though a low validation loss indicates that the model can recognize underlying sarcasm patterns. Data quality, model architecture, training techniques, and flexibility in adjusting hyperparameters are some potential controls on the validation loss. Thus, in order to evaluate the potential of a model, a variety of performance metrics and features with a low validation loss must be considered. The validation loss can be used to perform a more comprehensive examination of the robustness and generalizability of the model. The validation dataset showed a validation loss of 0.0058, indicating that the model successfully reduced the number of prediction errors. This demonstrates that it is capable of producing accurate predictions for new, untested events. To ensure that the model performs well in real-world applications, it is essential that it is able to generalize beyond the training data. This is because the conditions that the model will encounter may differ from those seen during training. The moderate validation loss of the proposed model suggests that it may find important patterns in sarcasm detection tasks. The model is also resistant to overfitting. The proposed sarcasm detection model appears to be successful in reducing prediction errors and performs well when applied to new datasets, as evidenced by the validation loss of 0.0058. This can be demonstrated by the good performance of the model. See Figure 4. The loss function via binary cross-entropy and parameter optimization helps the machine-learning model distinguish between the occurrence of sarcastic and non-sarcastic texts. However, many factors, such as model architecture, training methods, and data quality, need to be considered in order to accurately evaluate the validation loss.

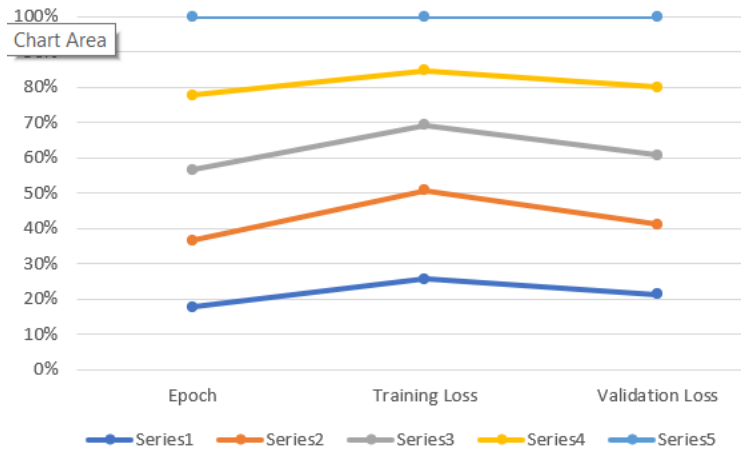


Figure 4. Illustrates the graphical representation of Training and Validation Loss.

4. Results & Discussions

A comprehensive evaluation of the utility, effectiveness, and results of the proposed sarcasm detection model for natural language processing tasks is conducted, along with an examination of the arguments and results that support it. After extensive testing and evaluation, the advantages and disadvantages of the model are listed. This provides crucial details about the use of the model and raises the possibility of further study and advancement. The focus of the discussion and results is the model's performance metrics, which include the F1 score, accuracy, precision, and recall. The model's effectiveness in detecting sarcasm is evaluated using these metrics, which measure how well the model performs in distinguishing between sarcastic and non-sarcastic text occurrences. If we examine these observations carefully across a wide range of datasets, experimental configurations, and hyperparameter sets, we may be able to gain a fuller understanding of the model's advantages and disadvantages. As a result, decisions about deploying and improving the model can be made with greater certainty. Furthermore, the results and discussions support the model's generalizability and flexibility across a range of datasets and linguistic contexts. The model's ability to recognize sarcasm in a range of forms, tones, and genres is examined through cross-validation and extensive testing on untested data. This demonstrates the applicability of the model to real-world scenarios. In addition, comparative evaluations using key models and modern methodologies help identify knowledge gaps and potential directions for future research. These evaluations also provide benchmarks for comparing the performance of the proposed model to that of its competitors.

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

In addition, A research that looks at the model's interpretability and explainability is included in the results and comments. Two very important features of natural language processing (NLP) applications are transparency and accountability, which are of the highest relevance. To understand how the model comes to its findings, it is crucial to look at the linguistic signals and features that form the basis of sarcasm identification. This project makes use of techniques including saliency maps, attention processes, and model interpretability tools. If stakeholders are aware of the model's biases, limitations, and internal mechanisms, they may be more inclined to accept and depend on its predictions. The conclusions and arguments discuss the wider cultural implications and ethical dilemmas related to the application of sarcasm detection algorithms in real-world situations, in addition to the technical performance metrics. We emphasize ethical quandaries including algorithmic biases, privacy concerns, and unforeseen effects in this paper. All of these problems necessitate a thorough examination of the ethical advancement and use of natural language processing technology. The diverse impacts that sarcasm detection models have on different facets of discourse and society are further underlined when future applications like sentiment analysis, social media surveillance, and content restriction are taken into account. Moreover, the outcomes and exchanges underscore the significance of consistently examining, adjusting, and validating the proposed model in dynamic and ever-changing language contexts. The growth of linguistic usage and cultural settings throughout time is necessary for the development of sarcasm detection systems. This covers these algorithms' utility as well as their efficacy. In order to address the constantly evolving needs and challenges of contemporary sarcasm detection, it is feasible that the proposed model might be expanded and enhanced with more study, cooperation, and feedback from a wide range of concerned parties. Overall, the findings and arguments support a more thorough comprehension of the effectiveness, implications, and possibilities of the suggested sarcasm detection model. A thorough analysis of its technical performance, interpretability, social impact, and ethical implications is required so that stakeholders may make well-informed decisions about its creation, application, and future direction. Sarcasm detection methods have the potential to revolutionize the area of natural language processing (NLP). Through ongoing communication, cooperation, and innovation, this would result in greater comprehension, empathy, and insights in human-machine relationships.

5. Conclusion

After studying the sarcasm detection model, it was clear that it recognized and classified sarcasm in text datasets well, with an accuracy of 0.9985 and a validation loss of 0.0058. Real-world applications require accurate sarcasm classification to provide useful and meaningful feedback, so this degree of accuracy is essential for the reliability and trustworthiness of the model. The strength and understanding of the model demonstrates the high validation accuracy and low validation loss for sarcastic statements. When assessing the generalizability of the model, the validation loss metric is of paramount importance. Sarcasm detection enhances understanding and judgment in social media sentiment analysis, customer comment processing, and content moderation. This breakthrough influences many disciplines. It achieves this by comparing predicted and actual values during model validation. The model appears to perform without errors because the predictions match the ground truth labels in the validation dataset (validation loss = 0.0058). The validation accuracy metric can also be used to evaluate the model. The proportion of validation datasets that were correctly classified is shown below. The model accurately identified both sarcastic and non-sarcastic text samples with an accuracy of 0.9985 during validation. High levels of accuracy demonstrate the model’s ability to distinguish between complex linguistic cues associated with sarcasm. For this reason, stakeholders may trust the model’s predictions and use them to make decisions and gain insights from textual data. Sarcasm recognition has many far-reaching applications and implications. The approach could enhance social media sentiment assessment. It can help understand public sentiment, identify new trends, and fine-tune communication efforts. This helps businesses respond to customer concerns more positively and quickly. Innovative machine learning methods including deep learning and natural language processing account for the model’s moderate validation loss and high validation accuracy. The model validates these approaches. Using large amounts of textual data and advanced algorithms. By demonstrating that AI can handle complex tasks such as language processing and sentiment analysis, this opens the door to more complex and context-aware human-machine interactions, but errors and biases are still possible. The sarcasm detection algorithm is susceptible to training data biases like any other machine-learning model, so it must be reviewed and updated frequently to provide an inclusive and equal workplace. Before applying the sarcasm detection model, the model’s predictions highlight the need for responsible AI development and deployment due to its potential impact on user experiences, decision-making, and public discourse. Sarcasm detection technology developers should promote user rights, accountability, and privacy. These results also demonstrate the model’s effectiveness in recognizing sarcastic language. For AI to fully understand and comprehend human language, sarcasm detection research must be strengthened, challenges addressed, and cross-disciplinary collaboration encouraged.

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

Refrencses

[1] D. Khurana, A. Koli, ,K. Khatter, and S. Singh, "Natural language processing: State of the art, current trends and challenges," Multimedia tools and applications,82(3), 3713-3744, 2023.

[2] Svendsen, B., & Kadry, S. (2023). A Dataset for recognition of Norwegian Sign Language. International Journal of Mathematics, Statistics, and Computer Science, 2. https://doi.org/10.59543/ijmscs.v2i.8049

[3] R. Collobert, , J.Weston, ,L. Bottou, M. Karlen, K. Kavukcuoglu, and P. Kuksa, "Natural language processing (almost) from scratch. Journal of machine learning research,", 12, 2493-2537, 2011

[4] T. Brants, "Natural Language Processing in Information Retrieval," CLIN, 111, 1-13, 2003.

[5] N. Sager, M. Lyman, C. Bucknall, N. Nhan, and L. J. Tick, "Natural language processing and the representation of clinical data," J. Am. Med. Inform. Assoc., vol. 1, no. 2, pp. 142–160, 1994.

[6] D. Meurers," Natural language processing and language learning. Encyclopedia of applied linguistics," 4193-4205, 2012.

[7] Z. Desai, K. Anklesaria and H. Balasubramaniam, "Business Intelligence Visualization Using Deep Learning Based Sentiment Analysis on Amazon Review Data," in 12th International Conference on Computing Communication and Networking (2021) Kharagpur, India.

[8] K.Jensen, G.E. Heidorn, and S. D. (Eds.). Richardson," Natural language processing: the PLNLP approach (Vol. 196)," Springer Science & Business Media, 2012.

[9] Lavania, G., Arya, V., Sharma, N., Rashid, M., & Akram, S. V. (2022, December). Real-Time Signal Processing using AI Integrated Framework for Color and Drawing in Gesture Recognition. In 2022 5th International Conference on Contemporary Computing and Informatics (IC3I) (pp. 473-478). IEEE.

[10] A.Chopra, A. Prashar, and C. Sain, "Natural language processing," International journal of technology enhancements and emerging engineering research, 1(4), 131-134, 2013.

[11] J. Hirschberg, and C.D. Manning, "Advances in natural language processing," Science, 349(6245), 261-266, 2015.

[12] A. Galassi, M. Lippi, and P. Torroni, "Attention in natural language processing," IEEE transactions on neural networks and learning systems, 32(10), 4291-4308, 2020.

[13] Y. Goldberg," A primer on neural network models for natural language processing," Journal of Artificial Intelligence Research, 57, 345-420, 2016.

[14] Sharma, N., Chakraborty, C., & Kumar, R. (2023). Optimized multimedia data through computationally intelligent algorithms. Multimedia Systems, 29(5), 2961-2977.

[15] Gupta, S., Sharma, N., Tyagi, R., Singh, P., Aggarwal, A., & Chawla, S. (2023). Cognitive-Inspired and Computationally Intelligent Early Melanoma Detection Using Feature Analysis Techniques. Journal of Artificial Intelligence and Technology, 3(4), 215-224.

[16] S. Locke, A. Bashall, S. Al-Adely, J. Moore, A. Wilson, and G. B. Kitchen, "Natural language processing in medicine: a review," Trends in Anaesthesia and Critical Care, 38, 4-9, 2021.

[17] Sharma, N., & Batra, U. (2021). An enhanced Huffman-PSO based image optimization algorithm for image steganography. Genetic Programming and Evolvable Machines, 22(2), 189-205.

[18] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, and A. M. Rush, "Transformers: State-of-the-art natural language processing," In Proceedings of the 2020 conference on empirical methods in natural language processing: system demonstrations (pp. 38-45), 2020.

[19] A. Farzindar, D. Inkpen and G. Hirst," Natural language processing for social media," San Rafael: Morgan & Claypool, 2015.

[20] M. V. Koroteev, "BERT: a review of applications in natural language processing and understanding,"arXiv preprint arXiv: 2103.11943, 2021.

[21] Y. Kang, Z. Cai, , C. W. Tan, Q. Huang, and H. Liu, "Natural language processing (NLP) in management research: A literature review," Journal of Management Analytics, 7(2), 139-172, 2020.

[22] H. Li, "Deep learning for natural language processing: advantages and challenges," National Science Review, 5(1), 24-26, 2018.

[23] Gupta, S., Saluja, K., Goyal, A., Vajpayee, A., & Tiwari, V. (2022). Comparing the performance of machine learning algorithms using estimated accuracy. Measurement: Sensors, 24, 100432.

[24] Gupta, S. (2015, October). An effective model for anomaly IDS to improve the efficiency. In 2015 International Conference on Green Computing and Internet of Things (ICGCIoT) (pp. 190-194). IEEE.

[25] Sharma, N., & Batra, U. (2018). A study on integrating crypto-stego techniques to minimize the distortion. In Data Science and Analytics: 4th International Conference on Recent Developments in Science, Engineering and Technology, REDSET 2017, Gurgaon, India, October 13-14, 2017, Revised Selected Papers 4 (pp. 608-615). Springer Singapore.

[26] E. Cambria and B. White, "Jumping NLP curves: A review of natural language processing research," IEEE Computational intelligence magazine, 9(2), 48-57, 2014.

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024

[27] T. Wolf, L. Debut, V. Sanh, J. Chaumond, C. Delangue, A. Moi, and A. M. Rush, "Huggingface's transformers: State-of-the-art natural language processing," arXiv preprint arXiv: 1910.03771, 2019.

[28] N. Hardeniya, J. Perkins, D. Chopra, N. Joshi and I. Mathur, "Natural language processing: python and NLTK," Packt Publishing Ltd, 2016.

[29] Shamshiri, A., Ryu, K. R., & Park, J. Y. (2024). Text mining and natural language processing in construction. Automation in Construction, 158, 105200.

[30] Chiruzzo, L., Jiménez-Zafra, S. M., & Rangel, F. (2024). Overview of IberLEF 2024: natural language processing challenges for Spanish and other Iberian languages. In Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2024), co-located with the 40th Conference of the Spanish Society for Natural Language Processing (SEPLN 2024), CEUR-WS. org.

DOI: https://doi.org/10.54216/JCIM.150223
Received: May 24, 2024 Revised: July 25, 2024 Accepted: November 14, 2024