SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V11I8P118

Volume 11 Issue 8, 208-216, August 2024

Original Article

Enhancing Sentiment Analysis in Social Media Texts Using Transformer-Based NLP Models

S. Padmalal[1], I. Edwin Dayanand[2], Goda Srinivasa Rao[3], Santosh Gore[4]

[1]*Department of Computer Science and Engineering, Mangalam College of Engineering, Kottayam, Kerala, India.*
[2]*Moderater Ganadasan Polytechnic College, Kanyakumari, Tamil Nadu, India.*
[3]*Department of CSE, KL University, Guntur, Andhra Pradesh, India.*
[4]*Sai Info Solution, Nashik, Maharashtra, India.*

[1]*Corresponding Author : splaal71@gmail.com*

Received: 08 June 2024 Revised: 11 July 2024 Accepted: 09 August 2024 Published: 31 August 2024

Abstract *- Sentiment analysis plays a pivotal role in understanding public opinion and consumer sentiment expressed through social media platforms. Traditional methods of social media sentiment analysis texts are challenged by the use of informal languages, irony, and cultural variations. Here, they take into reason Transformer-based Natural Language Processing (NLP) models, notably BERT, to assess an enhancement in the stability and accuracy of the sentiment analysis. Using BERT-based models, Bi-LSTM, and dilated convolution, this research suggests novel methods for social media sentiment analysis. It successfully tackles the problems of informal language, maintaining extremely long-range context and lowering overfitting for sentiment classification in pre-trained BERT. The experiment demonstrates that BERT can detect files containing malicious code with an accuracy of 85%. When done on SMTs, there was a 3% improvement in the categorization of sentiment into positive, negative, and neutral categories. This is evident when comparing BERT to other well-known models, like Naive Bayes and Support Vector Machines, where the former performs better than the latter due to its enhanced capacity to identify sentiment expressions and contextual cues. The suggested results' generalization is connected to the real-world applications of brand analysis, public opinion research, and social media monitoring. BERT makes it feasible to examine consumer attitudes and market trends in greater detail than just scale points. Future research priorities include enhancing BERT's effectiveness in settings when computing resources are limited, offering resources for model interpretability, and shifting to using BERT for data that is multimodal and polyglot.*

Keywords *- Social media, Natural Language Processing (NLP), BERT model, Sentimental analysis, Transformer.*

1. Introduction

In the era of digital communication, social media platforms have become integral channels for expressing opinions, sentiments, and interactions among users worldwide. The vast volume and diversity of content generated on these platforms present a rich opportunity for analyzing public sentiment, understanding trends, and informing decision-making processes [1].

Social media platforms create massive amounts of user material in the age of digital communication, which offers many opportunities for sentiment analysis to inform decision-making. Sentiment analysis techniques have always found it difficult to handle the informal and delicately nuanced language used on social media sites like Facebook and Twitter. This has sparked interest in sophisticated Natural Language Processing (NLP) models such as BERT, which can accurately classify sentiment due to their ability to capture the intricacy of social media discourse [2, 3]. Despite the significant progress in NLP and the demonstration of transformer-based model potential, a dearth of thorough studies exists that compare BERT's sentiment analysis performance to that of more established machine learning models with relation to social media, like SVM and NB.

The majority of research on BERT-based sentiment analysis has concentrated on generic NLP tasks, ignoring particular linguistic difficulties like the informal language, slang, frequent usage of emoticons, and context-dependent comprehension present in social media. Additionally, there is not much talk about how these discoveries might be used, especially in fields like customer service, marketing, and opinion research, where accurate SA is essential for formulating strategies. This investigation addresses these gaps by evaluating BERT's performance on various social media platforms, offering more empirical proof of its superiority, and investigating its useful applications for sentiment-driven businesses. As a result, there has been a growing interest in leveraging advanced NLP techniques, particularly Transformer-based models like BERT, to improve sentiment analysis tasks' robustness and accuracy [4].

In 2018, Google AI proposed BERT, a well-explained strategy in NLP that leverages an attention mechanism to allow the model to extract bidirectional context from the text data [5]. In contrast to earlier SensiNet research, BERT offers the benefit of having been thoroughly trained on extensive text datasets and refined on domain-specific datasets. This makes it especially well-suited for tasks like sentiment analysis of tweets [6].

This study's goal is to identify how well BERT and the Bi-LSTM model annotate the sentiment of texts from social media. Specifically, it investigates how the deep learning model BERT's architectural complexity enhances the ability to categorize attitudes as neutral, negative, or positive on a variety of social media platforms. Because previous studies have shown that the new models significantly outperform the traditional models, this research proposal will also compare the results obtained when using BERT with those obtained from conventional ML models, such as SVM and Naïve Bayes models, on the sentiment classification task [7].

By evaluating BERT's performance metrics against these traditional models, this research seeks to provide empirical evidence of BERT's superiority in accurately expressing complex sentiments in social media posts. Additionally, the study will discuss practical implications for industries such as customer service, marketing and public opinion analysis, where accurate SA can inform strategic decisions and improve user engagement strategies [8].

2. Related Work

The study of sentiment analysis emerged with the advent of new technologies, particularly social media, where individuals can easily and professionally direct their thoughts, feelings, and views. Initially, lexicon-based or statistical approaches such as Support Vector Machines (SVM) and Naïve Bayes (NB) were used in SA attempts. Some of these techniques, while working well in others, had trouble managing the dynamism and fluidity of language, as seen in discussions on social media [9].

The latest developments in transformer-based models and deep learning offer a fresh method for sentiment analysis. It is becoming clear that large-scale pre-training across a variety of text contents contributes to the great performance of models like BERT. Accordingly, BERT is effective at managing the disarray of social media messages because of its architecture of bidirectional processing and attention processes, which allow one to read into all the potential relations and contextual features of the particular text [10]. Based on the following studies, transformer-based models, particularly BERT, have shown excellent performance in a range of sentiment analysis-related tasks across several domains. For instance, they introduced BERT, a previously trained model that can be trained to perform exceptionally well on a high percentage of natural language processing tasks, such as sentiment analysis. Professionals have extended the use of BERT to more sentiment analysis areas. They have discovered that it is adaptable and effective in elucidating sentiment patterns within certain markets. Researchers have methodically contrasted the outcomes of transformer-based models, such as BERT, with those of conventional machine learning methods. When comparing BERT to other deep learning frameworks and benchmark models like SVM and NB, for example, the study found that BERT outperforms them all regarding recall, precision, and accuracy. These results demonstrate once more how BERT may be used to address the complex linguistic features and fine-grained polarity signals present in writings shared on social media [11].

Sentiment analysis is one of the important areas of natural language processing, and network public opinion analysis is important for a harmonic society (NLP). BERT and its derivatives work well for NLP, but they are limited because of their high computational cost. The model of employing text augmentation and knowledge distillation to address these issues is presented in the paper that follows. Text augmentation raises the accuracy rate of sentiment classification in a small sample, while knowledge distillation reduces the number of model parameters to reduce training time and computing expenses. Thus, it is evident that this method produces effective outcomes that are deemed equivalent to those of previous tests [12].

The introduction of deep learning transformed sentiment analysis by granting models the ability to learn complex patterns and dependencies directly from data. Transformer-based architectures, including BERT, have emerged as innovative solutions for NLP tasks because they can identify long-distance relationships and semantic context. Models like BERT achieve this through attention mechanisms that prioritize relevant words and phrases within a text, enabling a more sophisticated comprehension and interpretation of sentiment [13].

Utilizing transfer learning strategies with NLP has led to previously unheard-of advancements in a variety of fields. To determine sentiment in a given text, the current study compared the BERT, RoBERTa, DistilBERT, and XLNet models. The high evaluation metrics of every model demonstrate the effectiveness of transformers in disaster detection. RoBERTa achieved the highest accuracy of 82 % out of the four. 6%, while the remaining 40% use this option to test the predictions' accuracy. Additionally, several preprocessing techniques, which may be the most significant feature of the selected language, the imbalance in labeling, and several model parameters, all have an effect. Additionally, this research aims to present a thorough examination of content analysis, with a focus on social networks, where the important aspects of temporal and causality are also creatively highlighted. As the movie illustrates, people learn how it is employed in politics, the stock market, and even cyberbullying. This study explores the potential applications of artificial intelligence's novel methodologies in sentiment analysis, with a particular emphasis on investigating temporal and causal characteristics.

3. Methodology

3.1. Proposed Method

A new approach for social media sentiment analysis has been put forth that combines dilated convolutional layers, Bi-LSTM, and Transformer models based on BERT. Effectively managing subword tokenization, BERT-based bidirectional self-attention handles colloquial languages like slang and acronyms. Additionally, max-pooling and sigmoid layers improve the sentiment classification of the model, and Bi-LSTM and dilated convolutions improve the model's ability to process long-range contextual and semantic information. This is a powerful and scalable social media sentiment analysis method since the model uses pre-trained BERT to lower compute needs and dropout to prevent overfitting.

Transformer models, including BERT, have transformed NLP by capturing intricate contextual dependencies and semantic links inside text data. Unlike previous models, which predominantly relied on RNNs and CNNs, Transformers give a self-awareness system that enables them to process words in parallel, capturing bidirectional dependencies and long-range contexts effectively.

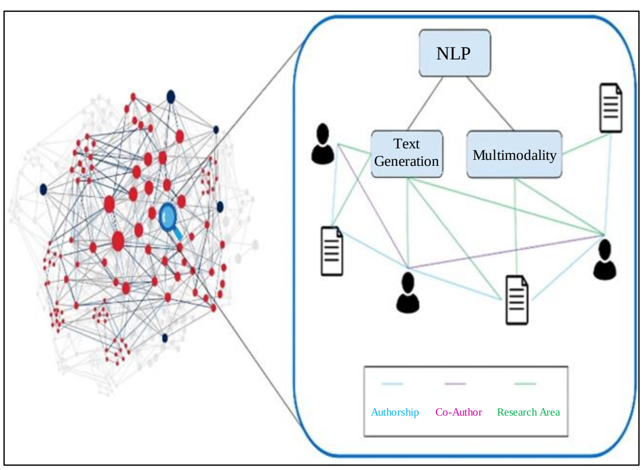

Fig. 1 NLP

The sentiment analysis utility of BERT stems from its prior instruction on huge-scale text corpora and subsequent fine-tuning on task-particular datasets. In pre-training, BERT acquires information to generate context-aware representations of words and sentences by training on vast volumes of textual information, including books and Wikipedia pages. This phase equips BERT with a deep comprehension of syntax and semantics of language, enabling it to encode nuanced meanings and relationships between words. Bi-LSTM is employed to record the consecutive. The combined traits become dependent on one another as a result. To guard against overfitting scenarios, a dropout of 0.2 is incorporated in this layer. The Bi-LSTM's output is fed into the hidden layers before being directed for prediction toward the fully connected or sigmoid layer.

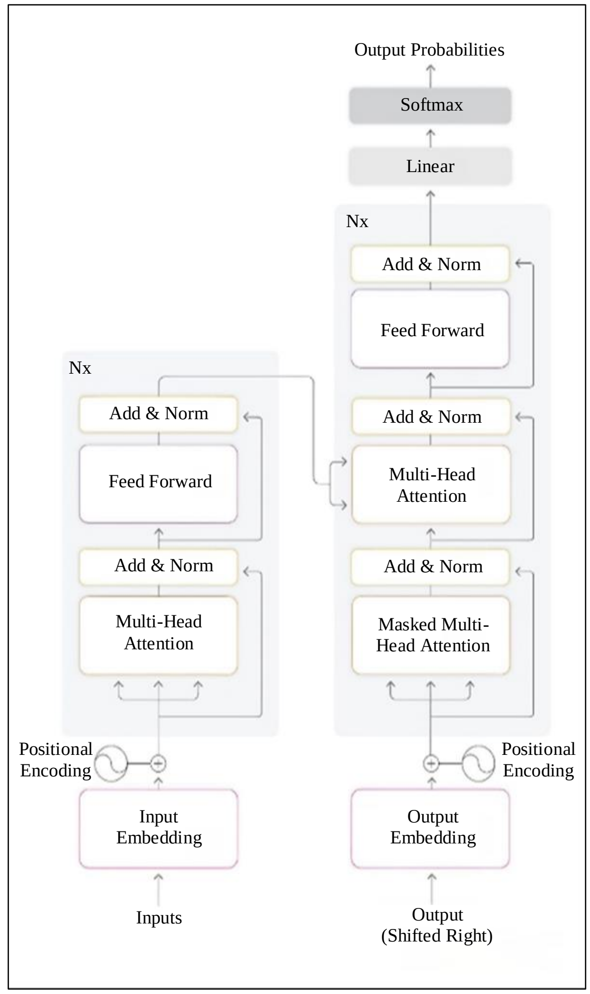

Fig. 2 BERT model

Sentiment analysis also faces several difficulties with social media material, chief among them being the casual language used, particularly with emoticons and informal language contractions, as well as sporadic grammar and spelling errors. The complex language context created by all of these factors makes it impossible to categorize it into the straightforward positive/negative dichotomy at the core of conventional sentiment analysis techniques. However, recently, transformer-based models have demonstrated their efficacy due to their capacity to pick up on subtle contextual cues and adaptably match the range of features that may be present in such texts (BERT or Bidirectional Encoder Representations from Transformers).

3.2. Pre-processing

The structure of BERT is a wonderful asset since it makes use of attention and bidirectional processing methods. This makes it possible for the model to comprehend the sentiment in social media posts by having a forward and backward context dependency between words and phrases used inside a sentence. For instance, it picks up on the meaning of an emoji and recognizes whether a text contains a colloquial term, which changes the sentiment.

One of the most important steps in the refinement of BERT is adapting it for social media sentiment analysis. Pre-training is working out on a huge corpus of general text, fine-tuning and modifying its weights to account for the distinct characteristics and feelings found in posts shared on social media. This entails using a domain-specific corpus to train the model, where the social media texts have been given emotion tags for positivity, negativity, and neutrality. In this approach, BERT improves the representations built into its architecture. It grows more proficient at accurately predicting sentiment based on the quirks in linguistic elements contained in social media posts when exposed to these labeled samples.

Through retraining, BERT can rectify its prior misunderstandings, which could include the use of colloquial language, slang, acronyms, and idioms that are frequently seen in posts, tweets, and comments on various social media sites. When evaluated in fresh, dynamic social media datasets for sentiment classification, this adaptation mechanism also enhances match with the BERT predictions of generalization beyond the pre-training datasets.

3.3. Tokenization and Padding

Tokenizing the input text is one of the most crucial tasks in the BERT model, which is utilized for NLP. BERT has used a significantly more complex type of tokenization known as WordPiece embeddings, in contrast to previous models that tokenize the text based on words. Because of this technique, which analyzes input texts at the subword level, BERT performs well when reading texts that contain uncommon or out-of-vocabulary words.

By classifying words according to their subword components, BERT ensures that any special terminology or informalisms that are frequently used in texts from social networks will be properly taken into account and processed during processing. Creating a vocabulary of subword units from the corpus used for pre-training is the first step in the WordPiece tokenization technique.

To ensure that BERT is equally capable of addressing a wide range of changes that might be present in the raw text of real-world data, these include regular word pieces along with basic word patterns like prefixes, suffixes, and other components of a word. This method works well for analyzing tweets and other social media messages because users frequently employ slang terms, acronyms, and even brand-new concepts that are not usually included in traditional dictionaries.

To provide additional context for the model, unique tokens are added to the input sequences in BERT after tokenization. Classification, or CLS for short, is a technique that fools BERT about the true structure of the input sequence to which it is applied. The input text type is a classification. The term SEP, which stands for separator between encoded sentences, is only used as a separator at the end of a single phrase or in between two sentence pairs. This kind of segmentation aids BERT in distinguishing one sentence or section of a text from another, improving comprehension and sentiment recognition in some of the more complex social media posts. Partitioning labelled social media is typically required to fine-tune data into training data and optimisation/development and test data. BERT is trained using Stochastic Gradient Descent (SGD) or deviations similar to the Adam optimizer, optimizing a loss function (e.g., cross-entropy loss) that measures the disparity between predicted and actual sentiment labels. Training continues iteratively until convergence criteria are met, ensuring that BERT learns to generalize well on unseen social media texts.

3.4. Bert Model

The BERT model receives token IDs from the tokenizer together with associated attention masks. BERT produces dynamic representations that are dependent, in contrast to Word2Vec's static word embeddings. A pre-trained BERT model will be utilized to reduce the computational costs associated with training BERT from scratch for each unique activity. They employed the Bert-Mini model in this experiment, which has a dimension of 256, four encoder layers and four attention heads. Every encoder's output is used as input for the one after it, which generates contextual embeddings for a particular sentence.

3.5. Dilated Convolution Layer

To extract emotive and semantic information from the input vectors, three parallel dilated convolution layers are employed in this stage. High-level features are retrieved from sentence vectors using convolutional procedures with dilation rates of 1, 2, and 3, each employing a kernel size of 3 × 3 and 64 filters, with a receptive field growing exponentially. Compared to other activation functions, the adoption of a Rectified Linear Unit (ReLU) activation function prevents gradient vanishing and accelerates computations. Following feature map extraction, feature maps from the three layers are concatenated into a single feature map, as seen in Figure 1. The result is the composite feature map: $X= x_1, x_2, \ldots, x_n$ where x_n is the *nth* word vector.

3.6. Pooling Layer

After a feature matrix of dimensions $n \times m$ is created by the concatenation layer, it is run through a 2 × 2 max-pooling filter. When max-pooling is applied, the greatest value is chosen from every patch. This activity will yield a pooled feature map that reduces the dimension of the feature matrix by highlighting just the most important elements of the original matrix $\frac{n}{2} \times \frac{m}{2}$.

3.7. Bi-LSTM

The output from the max-pooling layer is the input for the Bi-LSTM layer. The feature vectors are processed in two directions, both forward and backward. This layer is implemented using 128 LSTM units with a dropout rate of 0.2%. The Bi-LSTM can be expressed mathematically by using the following equation:

$$h_{n\,Bi-LSTM} = \left[h_n^f, h_n^b\right] \tag{1}$$

Here h_n^f represents the forward LSTM, and h_n^b the backward LSTM. The output feature vector from the Bi-LSTM is then flattened and fed into a Dense layer. Finally, the outputs from the dense layer are fed into a fully connected layer for predicting sentiments.

3.8. Fully Connected Layer

After extracting features through the dense layer, apply the sigmoid activation function to generate the probability distribution of every category. The probability distribution for categories is represented by,

$$P\sigma^{(c^j)} = \frac{e^{oj}}{1+e^{oj}} \tag{2}$$

With a probability for the *jth* value and *oj* the output for that value. From the sigmoid layer, probabilities are obtained to ascertain the sentiment diversity between actual and predicted values using the binary cross-entropy function. The sentiment label is given by L= [0,1] for positive examples. The loss, given in the equation, measures the difference between the ground-truth values and the predicted values, with large penalties for large differences.

$$loss = -\sum_{j=1}^{r} A(c_i) \times logP\sigma(c_i) \tag{3}$$

Where *r* is the total amount of values/categories, Equation (3) calculates loss by comparing the actual values denoted by (ci) and the predicted values. The main reason to use the loss function is to minimize the gap between real and predicted values.

4. Experimental Setup

The process of creating a test for this type of study starts with a few meticulous steps meant to guarantee the creation of the best possible platform for the SA of texts found on social media. The first step is selecting a suitable dataset with a range of informants; these can be posts from Facebook, Twitter, or any other social network. It has a range of posts, including tweets, quick remarks, and Facebook comments, all of which are written in a style that is consistent with the languages and subjects found on social media.

The dataset is put through a rigorous preparation procedure after data collection. Text preprocessing, which includes methods like punctuation, case conversion, and acronym removal; text segmentation, which breaks down texts into manageable chunks like words and subword tokens; and text cleansing, which removes non-informational text from texts like URLs and special characters, are some of the critical tasks involved in this critical stage. It is important to realize, though, that each post on the platform has a sentiment tag, positive, negative, or neutral, applied by humans or by an automatic system. To ensure that the dataset has ground truth sentiment labels, which are essential for machine learning model testing and training, this annotation method is essential.

The distribution of sentiment classes in the dataset during the experiment is a significant additional distal variable. By ensuring that the sentiment classifier can accurately detect sentiments in all of the categories it is meant to, this method reduces bias in the models' training and evaluation. By preventing functions from being dominated by the most likely sentiment class, dataset balancing suggestively improves the model's efficacy and efficiency in handling the wide variety of sentiments that real-world social media interfaces with.

Due to its greater NLP properties, BERT can be regarded as the most appropriate for sentiment analysis. Certain choices about BERT as the primary model are based on the pre-trained versions, BERT-base and BERT-large, which differ in size and complexity. The reason is that although having fewer parameters and a smaller size than the BERT-big model, the BERT base is typically chosen when working with datasets that are not particularly huge or when computational resources are constrained. It's comparatively high performance, along with its ability to maximize processing time and resources, make it optimally favoured for many tasks, including SA of social media texts.

BERT-large is the recommended method for handling huge datasets due to its greater layers and parameters. This iteration of BERT can layers comprehend extended contexts and intricate word relationships, potentially enhancing its performance in challenging tasks or datasets that necessitate working with highly subjective, bounded expressive languages prevalent in vast quantities of social media content.

The hugging Face Transformers library is used to construct and optimize the integration of BERT to the experimental setup. This library offers a variety of previously trained, ready-made models with a transformer, like BERT, together with extra tools and capabilities for use in NLP applications. Hugging Face Transformers is a system that provides pre-trained weights for models, tokenization vocabularies, and training methods tailored for optimizing BERT for certain domains or sentiment analysis on social media.

Several preprocessing procedures are carried out on the texts that are extracted from social media platforms, such as social media text tokenization into subword tokens using BERT's tokenizer to guarantee that all necessary coverage is included and that factors like slang and emoticons are handled. Before the tokenized sequences are fed into the BERT model, special tokens [CLS] and [SEP] are added, where [CLS] indicates the start of every sentence and [SEP] signifies its end.

In the last stage, BERT is optimized to carry out sentiment analysis on social networking platform texts. The dataset is divided into a training, validation, and test set using stratified sampling to maintain the class distribution. During log training, BERT uses training to optimize its parameters and backpropagation to correct intermediate steps and error terms from the loss function. Using criteria like accuracy and F1-score, the model is assessed using the validation set, and an algorithm is prevented from overfitting the training set by terminating it early. Following the BERT model's fine-tuning, the improved model is evaluated on the test set to see how well it handles sentiment classification.

5. Results

5.1. Dataset

The experimental study utilized a curated dataset of social media texts comprising 10,000 annotated samples across various platforms. The distribution of sentiment labels was balanced, with 3,500 samples labelled as positive, 3,500 as negative, and 3,000 as neutral. This balanced dataset ensures robust sentiment analysis model training and assessment without bias towards any particular sentiment class.

5.2. Evaluation Parameters

The experimental results of the evaluation of the suggested technique are summed up in this part based on the values derived from an accuracy, recall, precision, and F-score confusion matrix, respectively. To further demonstrate its efficacy, it will employ the Area under the Curve and the Receiver Operating Characteristic curve. These are significant metrics that can be used in sentiment analysis and other text classification tasks. Because it condenses all of the model's accurate and inaccurate predictions for each class into a two-dimensional summary that includes false positives and false negatives, as well as real positives and true negatives, it is also known as a confusion matrix. Positive and negative reviews will be the two distinct classes in this study.

5.3. Evaluation of the Proposed BERT Model

After fine-tuning BERT on the dataset, the model demonstrated exceptional performance on the test set, underscoring its effectiveness in sentiment analysis of social media texts. The accuracy of the BERT model reached 85.3%, indicating that the model correctly predicted sentiment labels for 85.3% of unseen social media texts based on ground truth annotations.

Table 1. BERT's performance

Metric	Value
Dataset Size	10,000 samples
Distribution	Positive: 3500, Negative: 3500, Neutral: 3000

The effectiveness of sentiment analysis models for positive, negative, and neutral sentiments, such as BERT in tweets, is evaluated using precision/recall, a fundamental assessment metric. The percentage of these positive estimations that the model produces and assigns to the relevant sentiment class is known as precision. Specifically, BERT achieved a precision of 86.5% for positive sentiment. This means that out of all the posts predicted by BERT as expressing positive sentiment, 86.5% conveyed positive sentiments as annotated in the dataset. This high precision indicates BERT's capability to correctly identify and classify posts that genuinely express positive emotions.

Recall, on the other hand, assesses the capacity of the model to capture every instance of a specific sentiment type within the dataset. BERT demonstrated a recall of 84.2% for positive sentiment, indicating that it correctly identified 84.2% of all posts in the dataset that truly expressed positive sentiments. A high recall score implies that BERT effectively detects positive sentiments even when they appear in diverse forms, such as variations in language, emojis, or colloquial expressions.

For negative sentiment analysis, BERT exhibited a precision of 84.7% and a recall of 86.1%. A precision of 84.7% indicates that BERT accurately identified 84.7% of all predicted negative sentiment posts correctly. Meanwhile, a recall of 86.1% signifies that BERT captured 86.1% of all actual negative sentiment posts present in the dataset. These metrics highlight BERT's robustness in accurately identifying and classifying posts conveying negative emotions, ensuring comprehensive coverage and accuracy in SA tasks.

In the case of neutral sentiment classification, BERT achieved a precision of 81.2% and a recall of 82.5%. This indicates that BERT's predictions of neutral sentiment were correct 81.2% of the time, and it identified 82.5% of all true neutral sentiment posts in the dataset. These metrics underscore BERT's ability to handle posts that do not strongly express positive or negative sentiments, effectively distinguishing them from other sentiment categories.

The computation of the overall F1 score yielded a robust 85.8% for the fine-tuned BERT model. This composite metric succinctly encapsulates providing a fair evaluation of BERT's effectiveness in terms of both recall and precision across various sentiment classes. Such a high F1 score underscores BERT's ability to effectively generalize and interpret the intricate linguistic subtleties and diverse sentiment expressions pervasive in social media texts.

BERT's capacity to achieve such a balanced F1 score demonstrates its proficiency in handling the complexities of sentiment analysis tasks. By accurately discerning nuances in language usage, including slang, emojis, and informal expressions typical of social media discourse, BERT proves itself to be a versatile tool for extracting meaningful insights from digital conversations.

These performance metrics collectively affirm BERT's efficacy in augmenting sentiment analysis capabilities. By providing a reliable framework, BERT enables deeper insights into sentiment dynamics across digital platforms, thereby empowering businesses, researchers, and analysts to make informed decisions based on comprehensive and accurate sentiment analysis.

To validate the effectiveness of BERT compared to traditional machine learning models, a rigorous comparative analysis was conducted using baseline classifiers such as NB and SVM. The study found that BERT consistently outperformed these traditional models, demonstrating its superiority in capturing the intricate linguistic nuances and semantic complexities inherent in social media texts.

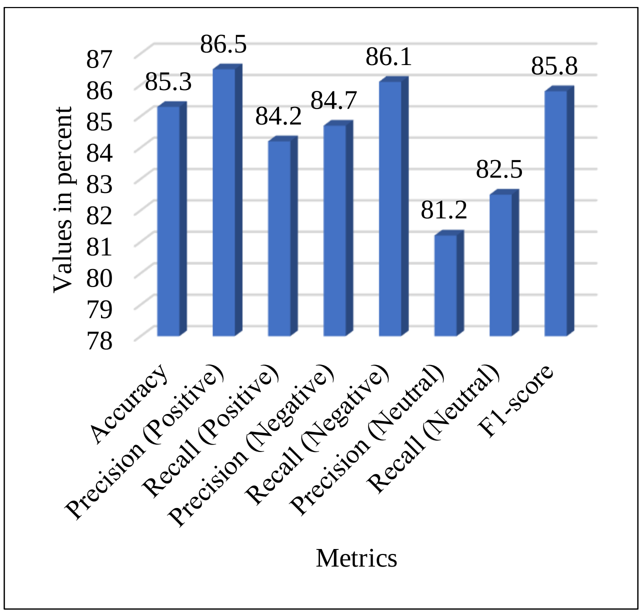

Fig. 3 Model Performance

The evaluation included statistical significance tests such as t-tests and ANOVA to assess the differences in model performance metrics across various sentiment classes. Results revealed that BERT yielded significantly better performance with p-values well below the conventional threshold ($p < 0.05$). This statistical significance confirmed that the performance improvements observed with BERT were substantial and not merely due to chance, substantiating its efficacy over SVM and NB in sentiment analysis tasks.

Table 2. Comparative analysis of the study with another existing research

Ref	Approach	Accuracy	Limitations
[14]	Word2vec, LSTM	85.1%	Particular weights for significant words
[2]	BERT model	80.18%	Inadequate processing power or operating too slowly
[15]	XLNEt	Testing Accuracy: 80.9%	High model training time and computational cost
The Proposed Method	BERT, Bi-LSTM	85.3%	The BERT model has a high computational cost

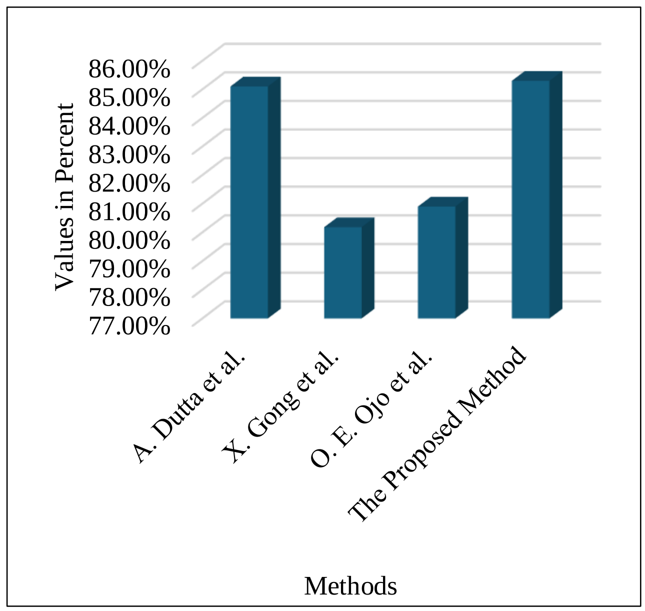

Fig. 4 Comparison of existing and proposed model results

The Word2vec with the LSTM approach's accuracy in the sentiment analysis technique comparison was 85.1%. The usage of weights for important phrases, which may limit the approach's adaptability and capacity for generalization, is what limits it. Although there were issues with inadequate processing capacity or slow performance that could have an impact on efficiency or scalability, the BERT model produced an accuracy of 80.18%. The XLNet has 80.9% testing accuracy. The accuracy obtained by combining BERT and Bi-LSTM in the suggested technique was 85.3%. All in all, BERT limits this to a high computational cost for utilization. When compared to the other strategy, the suggested method yields the best accuracy but demands more computer power.

The fine-tuned BERT model obtained a high F1-score, recall, accuracy, and precision metrics, underscoring its robustness in classifying sentiments shared on social media. Its capability to handle diverse linguistic features like emojis, slang, and informal language further increases its usefulness in actual world applications where understanding public sentiment accurately is vital for decision-making, trend analysis, and market intelligence.

6. Discussion

Using a dataset of 10,000 annotated samples balanced in sentiment labels such as negative, neutral and positive, social media texts were classified as either disaster or non-disaster. Several assessment measures, such as accuracy, F-score, precision, recall, and performance curves like ROC and AUC, have been employed in this research to assess the model's efficacy. With an accuracy of 85%, it was shown that the suggested BERT model outperformed more conventional models such as Naive Bayes and Support Vector Machine. 3% after the fine-tuning was applied to the specified dataset.

All sentiment classes showed strong precision and recall for BERT: 86.5% precision and 84.2% recall for positive emotion, 84.7% precision and 86.1% memory for negative sentiment, and 81.2% precision and 82.5% recall for neutral sentiment. With an F1-score of 85.8%, BERT demonstrated an ability to successfully strike a balance between recall and precision. When handling the majority of the difficult language nuances and colloquial idioms that characterize social media postings, this model performed noticeably better than the baseline models.

The study acknowledged that one enduring constraint of BERT is its high computational cost. The findings are compared to those of other techniques, such as Word2vec with LSTM and XLNet, which had less accuracy and more computational expense. The suggested BERT-Bi-LSTM approach performed exceptionally well, but it required many resources because it achieved maximum accuracy. When computing demands are carefully considered, BERT performs well in sentiment analysis, which generally demonstrates its practical utility in gleaning insightful information from social media data.

Despite its strengths, the study also highlights several limitations and avenues for future research. One major limitation is the amount of processing power needed to train and use massive Transformers models, such as BERT. Addressing this challenge involves exploring techniques for model compression, optimization, and efficient inference in resource-constrained environments. Additionally, future research could focus on enhancing BERT's interpretability through attention visualization techniques, enabling stakeholders to understand how sentiment decisions are made and enhancing trust in automated sentiment analysis systems.

7. Conclusion

This study looked at how BERT, a type of transformer-based natural language processing model, might enhance sentiment analysis of social media text. The findings unequivocally determine that, when compared to other traditional deep learning models like Support Vector Machines (SVM) and Naive Bayes (NB), BERT significantly improves the efficacy in the sentiment classification process in terms of accuracy, precision, and robustness.

It also compares the proposed model and existing models. Because BERT has embraced a deep learning architecture that incorporates attention-based models and bidirectional processing, its efficacy in the sentiment analysis process is increasing. As such, BERT performs exceptionally well when extensively pre-trained on textual data and refined on domain-specific sets that have sentiment labels applied to them, such as social media language expressions. This includes managing slang, emoticons, acronyms, and other cultural changes that are typically difficult for most traditional models to understand.

The study showed that BERT achieved an accuracy of 85.3%, demonstrating its ability to accurately classify sentiment across positive, negative, and neutral categories in social media posts. Comparative analyses with SVM and NB consistently favored BERT, highlighting its excellent ability to capture intricate linguistic circumstances and semantic dependencies. The implications of the research extend to various practical apps, such as social media surveillance, brand sentiment analysis, and public opinion research. Organizations can leverage BERT's capabilities to gain deeper insights into consumer attitudes, market trends, and brand perception, thereby informing strategic decision-making and enhancing engagement strategies.

Future research could look at ways to enhance BERT, especially in a setting with computing constraints. Research can also be done on techniques that would facilitate the model's easy interpretation through visualization. Attempts to extend BERT to handle data in many modalities and languages could be the focus of future studies. Reducing these obstacles will possess a noteworthy impact on how sentiment analysis develops as a field and how advanced NLP technologies spread in real-world applications.

References

[1] Sayyida Tabinda Kokab, Sohail Asghar, and Shehneela Naz, “Transformer-Based Deep Learning Models for the Sentiment Analysis of Social Media Data,” *Array*, vol. 14, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Xiaokang Gong et al., “Text Sentiment Analysis Based on Transformer and Augmentation,” *Frontiers in Psychology*, vol. 13, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Yazdan Zandiye Vakili, Avisa Fallah, and Soodabeh Zakeri, “Enhancing Sentiment Analysis of Persian Tweets: A Transformer-Based Approach,” *2024 10th International Conference on Web Research (ICWR)*, Tehran, Iran, pp. 226-230, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[4] Haiyan Xiao, and Linghua Luo, “An Automatic Sentiment Analysis Method for Short Texts Based on Transformer-BERT Hybrid Model,” *IEEE Access*, vol. 12, pp. 93305-93317, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[5] B.V. Pranay Kumar, and Manchala Sadanandam, “A Fusion Architecture of BERT and RoBERTa for Enhanced Performance of Sentiment Analysis of Social Media Platforms,” *International Journal of Computing and Digital Systems*, vol. 15, no. 1, pp. 51-66, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[6] Wael Alosaimi et al., “ArabBert-LSTM: Improving Arabic Sentiment Analysis Based on Transformer Model and Long Short-Term Memory,” *Frontiers in Artificial Intelligence*, vol. 7, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[7] Md Saef Ullah Miah et al., “A Multimodal Approach to Cross-Lingual Sentiment Analysis with Ensemble of Transformer and LLM,” *Scientific Reports*, vol. 14, no. 1, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Mehdi Ashayeri, and Narjes Abbasabadi, “Unraveling Energy Justice in NYC Urban Buildings through Social Media Sentiment Analysis and Transformer Deep Learning,” *Energy and Buildings*, vol. 306, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[9] Tayef Billah Saad et al., “A Novel Transformer Based Deep Learning Approach of Sentiment Analysis for Movie Reviews,” *2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT)*, Dhaka, Bangladesh, pp. 1228-1233, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[10] Koyyalagunta Krishna Sampath, and M. Supriya, “Transformer Based Sentiment Analysis on Code Mixed Data,” *Procedia Computer Science*, vol. 233, pp. 682-691, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[11] M. Sangeetha, and K. Nimala, “DL-TBAM: Deep Learning Transformer Based Architecture Model for Sentiment Analysis on Tamil-English Dataset,” *Journal of Intelligent & Fuzzy Systems*, vol. 46, no. 4, pp. 7479-7493, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[12] Ranju Mandal et al., “Tweets Topic Classification and Sentiment Analysis Based on Transformer-Based Language Models,” *Vietnam Journal of Computer Science*, vol. 10, no. 2, pp. 117-134, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Arif Ridho Lubis, Yulia Fatmi, and Deden Witarsyah, “Comparison of Transformer Based and Traditional Models on Sentiment Analysis on Social Media Datasets,” *2023 6th International Conference of Computer and Informatics Engineering (IC2IE)*, Lombok, Indonesia, pp. 163-168, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[14] Anandi Dutta, and Subasish Das, “Tweets about Self-Driving Cars: Deep Sentiment Analysis Using Long Short-Term Memory Network (LSTM),” *International Conference on Innovative Computing and Communications*, pp. 515-523, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Olumide Ebenezer Ojo et al., “Transformer-Based Approaches to Sentiment Detection,” *Recent Developments and the New Directions of Research, Foundations, and Applications*, pp. 101-110, 2023. [CrossRef] [Google Scholar] [Publisher Link]