International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(1s) 2026 pp. 391-406
https://doi.org/10.70917/ijcisim-2026-2291
Article

Improving Sentiment Analysis in Social Media Using Deep Learning Techniques: A Comparison Study of Generative AI Model Approaches in Gujarati Language Context

Sharma Vishalkumar Sureshbhai[1], Tulsidas Nakrani[2], Sachin Sharma[3]

[1]Department of Computer Application (MCA), Sankalchand Patel University, Visnagar, Gujarat, India, phd.vishalsharma@gmail.com
[2]Department of Computer Application (MCA), Sankalchand Patel University, Visnagar, Gujarat, India, tvnakranimca_spce@spu.ac.in
[3]Associate Professor, Department of Computer Science and Engineering, Indrashil University, Mehsana, India, sharma.f@gmail.com

Abstract: Social media's explosive growth has drastically changed how people express their thoughts, feelings, and sentiments online. Low-resource languages like Gujarati are still not well studied, despite the significant advancements in sentiment analysis in widely spoken languages. In order to close this gap, this study compares generative AI and deep learning methods for sentiment analysis in the context of Gujarati Language. A dataset that captured a variety of linguistic nuances, colloquial expressions, and code-mixed content typical of user-generated posts was assembled from multiple social media platforms. To ascertain of their respective advantages and disadvantages, a number of models, including LSTM, BERT, RNN, CNN, GPT-2, Gemma, LLaMA, RoBERTa, and Word2Vec-based architectures, were put into practice and method assessed. We used cross validation and standard evaluation metrics to test how well the model performed. The results showed that transformer-based and generative models work better than traditional sequence and the embedding-based models when handling complex context and meaning. This study adds new data and shows that generative AI can make models more robust and easier to interpret, which helps advance sentiment analysis for supervised represented languages. Our findings show strong results for Gujarati and open the door for future research on sentiment analysis in other regional and low- resource languages used on social media.
Keywords: Sentiment Analysis, Generative AI, Deep Learning Algorithms, Social Media, Gujarati Language.

1. INTRODUCTION

1.1. Sentiment Analysis

Finding and extracting subjective information from text is the goal of Sentiment analysis, commonly referred to as analysis of opinions, which is a task in NLP. Utilize this method to ascertain whether the author exhibits a positive, negative, or neutral attitude or emotional disposition. Sentiment analysis has significantly progressed due to the emergence of deep learning and generative AI models, facilitating a more precise and intricate comprehension of human emotions conveyed in written language.

Sentiment analysis [1] is a prevalent and difficult subject in the field of artificial intelligence. Automated methods are utilized to recognize subjective information, such as ideas, attitudes, and sentiments that are represented in text. These tools are applied to various sources, including news, blogs, and social networks. In modern times [2], social media, blogs, and networks generate vast amounts of information. Individuals express their opinions on diverse occurrences and matters. This extensive data is utilized for forming opinions and making decisions within a corporate context. Sentiment research [3] is crucial in the financial markets as it assesses the emotions and beliefs of investors towards the market and individual stocks. Sentiment analysis [4], constructs ML and deep learning algorithms, is a widely used method for categorizing text into its appropriate classification.

Sentiment analysis is a Natural Language Processing (NLP) activity that ascertains the sentiment or emotional tone underlying a text. Deep learning methodologies have markedly enhanced the precision of sentiment analysis by utilizing neural networks to comprehend intricate patterns in textual input.

1.2. Generative AI Models

The limits of what may be achieved in sentiment analysis have been significantly expanded by recent developments in GAN artificial intelligent models, especially those built on the Transformer architecture. Generative models [5] have demonstrated exceptional skill in comprehending and producing text that resembles human language. Examples of such models are BERT [6], LSTM, GAN [7], LLM, CNN, and RNN. These models have already been trained on large datasets and can be adjusted for applications, such as sentiment analysis, using smaller datasets [8]. Much investigation and development has been prompted by the prospect of using generative AI models into sentiment analysis. Their effectiveness in sentiment analysis tasks like emotion identification, aspect-based sentiment analysis, and sentiment classification has been proven by studies to outperform both conventional and earlier deep learning methods. Various sectors are putting these sophisticated models to use to enhance consumer satisfaction, monitor company image, and even predict market conditions [3] by analyzing shifts in sentiment. Businesses are quickly adopting these innovations to make data-driven decisions and gain a deeper understanding of customer sentiment.

Although Gan AI models for Opinion analysis offer benefits, their deployment poses problems including high computing resource demands, potential pre-existing biases in training information, and the requirement to adjust [9] to suit individual fields. Future investigations will focus on improving the efficiency of these models, reducing assumptions, and increasing their understanding. The amalgamation of deep learning techniques with generative artificial intelligence models has markedly enhanced sentiment analysis, providing robust tools for the assessment and interpretation of individuals' emotions [7,10] and thoughts articulated in text. As technology advances, these algorithms will develop, offering more precise and sophisticated sentiment analysis capabilities.

Generative AI models [22], particularly those based on deep learning architectures like Transformers (e.g., GPT, BERT, and T5), have revolutionized sentiment analysis by enabling more context-aware, nuanced, and adaptive sentiment classification. Traditional sentiment analysis methods relied on lexicon-based or rule-based approaches, but generative AI models have significantly improved accuracy, generalization, and adaptability.

2. RELATIVE WORK

2.1 LLM (Large Language Model)

Stanislav Chumakova et al. [5] the research discusses the constraints of ASTE based on BERT. A generative approach for open domains based on the GPT utilizing few-shot and adjusting methodologies is proposed. The application of LLM on computer and food ratings, specifically Yelp, Google food service reviews, as well as the e-commerce and resulted in models being evaluated on a multi-domain dataset of language consequently tagged data, utilizing a comprehensive large language model with a few-shot technique, alongside exclusively consequently assigned the English tongue data.

Konstantinos I. Roumeliotis. et al. [11] have this paper introduces and describes a Large Language Model-based research study on e-commerce customer satisfaction. The highlights e-commerce expansion, automated customer service, and product review emotion assessment issues. They used deep Learning on e-commerce datasets. GPT-3.5 model accuracy improved sentiment categorization.

Frans Mikael Sinaga et al. [12] the study examines twitter comments about ChatGPT's potential drawbacks. To evaluate Positive, Negative, and Neutral feelings. Fast RCNN employing LLM, CNN on Twitter datasets showed enhanced sentiment classification accuracy.

Xiang Deng, Vasilisaet al. [13] This research discusses social media market sentiment analysis issues and proposes a semi-supervised learning approach utilizing a big language model. The difficulties human rats have in understanding financial markets and social media platforms leaves standard supervised learning approaches without high-quality labeled data. They used LLM on Reddit datasets and found greater sentiment categorization accuracy.

Megharani Patil et al. [14] This study emphasizes the influence of fake updates on societal matters and the need of identifying and combating disinformation. LLM was used on Whatsapp bot and Twitter Data datasets, and it worked better to identify bogus news and reassure users.

2.2 Bidirectional Encoder Representations from Transformers (BERT)

Priyank Sonkiya et al. [3], the primary purpose of the analysis is to present a collection of techniques for forecasting stock prices. The main objective is to perform sentiment analysis on news and headlines concerning Apple

Inc. This will be done using a modified version of BERT, a natural language processing model. Additionally, a Generative Adversarial Network will be used to forecast iPhone Inc.'s stock price. The prediction will be based on technical measurements, stock indices, products, previous prices, and sentiment metrics. The researchers employed BERT and GAN algorithms on datasets of Apple Inc shares, showcasing exceptional performance. This enables investors and traders to make informed decisions by predicting the stock price for the next few days and determining whether to purchase or sell a specific stock.

Zhexing Sun et al. [6] to highlight information disparity between clients and manufacturers of agricultural commodities, especially on online platforms. An Opinion analysis algorithm based on an upgraded BERT model is proposed. The purpose is to analyze online agricultural product after-sales reviews, determine consumer emotions, and optimize and enhance agricultural products. BERT performed well on agricultural product reviews from social networking platforms, media outlets, online shopping sites, and offline polls. Improved BERT improves sentiment categorization accuracy.

Sin-Chun Ng et al. [15] the aim is to tackle the difficulties in opinion mining in Content by suggesting and assessing the efficacy of employing BERT with various deep neural network structures. The author employed models to analyze Tweets obtained from Kaggle datasets. Their explore indicate that the proposed planning exhibits higher performance, with an average accuracy that surpasses other approaches.

2.3 LSTM (Long Short-Term Memory)

Chase Cotton et al. [1] emphasize how sentiment analysis helps make sense of the increasing user-generated content online. Their research discusses the difficulties in detecting viewpoints, emotions, and sentiments in emails, tweets, feedback, and comments. They applied Long Short-Term Memory (LSTM) and Convolutional Neural Network (CNN) models to Amazon reviews and the IMDB dataset of 50,000 movie reviews. These models outperformed others, with Dataset2 reaching the highest accuracy in classifying sentiment.

Mouli Venkata Srinivas et al. [2] identify sentiment analysis as a significant research domain and underscore the necessity for robust methodologies to process extensive, interconnected datasets. Their application of Long Short-Term Memory (LSTM) networks to 1.6 million tweets resulted in superior accuracy for sentiment classification.

Abdullah as Sami et al. [9] investigate the objectives and challenges of sentiment analysis within Bangla natural language processing, particularly emphasizing the limitations posed by insufficient standardized labeled data. Their implementation of a CNN-BiLSTM model on social media comment datasets demonstrated enhanced accuracy in sentiment classification.

Yun Peng et al. [16] propose a BiLSTM model integrating TextCNN and Self-Attention mechanisms, optimized using an enhanced Particle Swarm Optimization technique. Evaluation on IMDb, English, SST-2, and Hotel Reviews datasets demonstrated superior accuracy in sentiment categorization.

Rashid Amin et al. [17] assert that sentiment analysis is essential for interpreting consumer reviews and social media content, particularly given the increasing volume of unstructured online data. Their research emphasizes the challenges associated with large-scale data management and the necessity for precise polarity detection. The study employed Long Short-Term Memory (LSTM) models on Amazon-Food-Feedback, Mobile and Accessories, and Amazon-Products datasets, with LSTM achieving the highest performance, especially on Amazon-Products.

Bilen B et al. [18] note that comprehending complex human emotions remains a significant challenge in sentiment analysis. Their study introduces a binary classification approach utilizing a singular-tag/plural-class strategy with an LSTM network alongside other machine learning models. Testing on electronic store and IMDB review datasets revealed strong performance and improved precision in extracting sentiment from electronic retail data.

Retno Kusumaningruma et al. [20] highlight the necessity for effective sentiment evaluation methods in Indonesian motel reviews, given the substantial volume of customer feedback. Their application of Long Short-Term Memory (LSTM) models to hotel review datasets from the Traveloka website resulted in high accuracy for sentiment classification.

Shanmukha Rao et al. [21] introduced automated sentiment analysis to overcome the limits of manually reviewing large amounts of text data. Their study focused on finding positive and negative opinions in datasets like electrical application feedback and merchandise evaluations. They used an LSTM model on IMDB and Amazon product datasets and achieved better accuracy in sentiment categorization.

2.4 GAN (Generative Adversarial Network)

Maryam Edalatiet al. [4] The COVID-19 days highlights the value of online student input. Online studymakes it important to realize students' thoughts reviews. The research focuses on training ML and DL models for sentiment categorization utilizing balanced and imbalanced datasets. To address imbalanced datasets, text generation models (CatGAN and SentiGAN) are used to improve online education sentiment classification. GAN-SentiGAN and

CatGAN performed better on Coursera's CR23k-21937 course reviews, Kaggle's CR100k-107016 reviews, and Amazon's AMR datasets. Results show CR100k has a greater average growth degree than CR23k.

Dr. G. Kavitha et al. [7] this research introduces an innovative method that utilizes Generative Adversarial Networks (GANs) to generate text with emotional content. The structure incorporates a punishment technique into the GAN to improve the model capacity to create content sequences that are sensitive to sentiment. This study examines the integration of sentiment restrictions in the process of generating text, with the goal of creating content that is more contextually appropriate and emotionally expressive. The researchers employed a Generative Adversarial Network and Bi-LSTM model to analyze client reviews of Amazon products using Kaggle datasets. Their study showed exceptional results. Proposed Task, the Bi-LSTM model demonstrates superior performance in sentiment categorization, achieving greater accuracy.

Jo ao P. Papa et al. [8] the primary proposes of the analysis are is to provide a thorough examination of contemporary research and progress in text generation using GANs. This essay underscores the favorable aspect of adversarial learning in the context of text generation, placing emphasis on its capacity to produce language that appears "natural". Furthermore, it addresses the difficulties arising from theGANs that were first developed for sequential data, specifically photos, rather than individual data, such as text. They employed Generative Adversarial Networks (GAN) on various datasets including Amazon Feedback, Poems, COCO Captions, and WMT News. Their experiments revealed that the sentiment classification task performed exceptionally well on the WMT News dataset, achieving greater accuracy compared to other datasets.

YuliangMa et al. [10] the primary objective is to showcase the efficacy of the suggested model in acquiring the distribution of unprocessed data and producing superior synthetic samples, hence facilitating the training of a highly accurate emotional model. The researchers applied VAE-GAN to Emotion EEG datasets, showcasing the improvised performance of the suggested method. The average accuracy of the new proposed method is 5% bigger thanthe original dataset.

Shariq Imran et al. [19] in the textual of sentiment mining of user comments, the goal of the is to tackle the problem of disparity concerning educational response databases. In unstructured feedback, students are free to share their ideas in response to open-ended inquiries about the class, the material, and the instructor. The educational feedback datasets CR23K and CR100K were processed using CatGAN and SentiGAN. The final output showed that the CR100K dataset had improved sentiment classification performance and higher accuracy.

3. RESEARCH METHODOLOGY

This area allocated a clear and accurate explanation of the research findings, how they can be understood, and the conclusions that can be drawn from them.

3.1 Data Organizing & Data Preparation

Database was gathered from primary and secondary section where primary data was collected from different social media platform like Amazon, Flipkart, Instagram, Tweeter, and YouTube. The secondary data was collected from GitHub and finally I combine both of them into a one. As indicated in Table 1 below, the database consists of 5737 rows and 2 columns. After removing the null values, the mapping was finished. A positive score was 1, while a negative score was 0. As seen in Table 2 below, the dataset sentiment has 2843 positive and 2894 negative responses, with a 49.06% positive and 50.04% negative sentiment percentage. Figure 1 illustrates how the preprocessed raw Gujarati dataset was fed into various models.

Table1. Structure of the Dataset

Name of Dataset	Rows by Columns	Background
gujarati_domain_sentiment	5737 x 2	GitHub

Table 2. Number of Sentiments

Sentiment	Number of Sentiment	Percentage
Positive	2843	49.06%
Negative	2894	50.04%

3.2 All Model Architecture Design

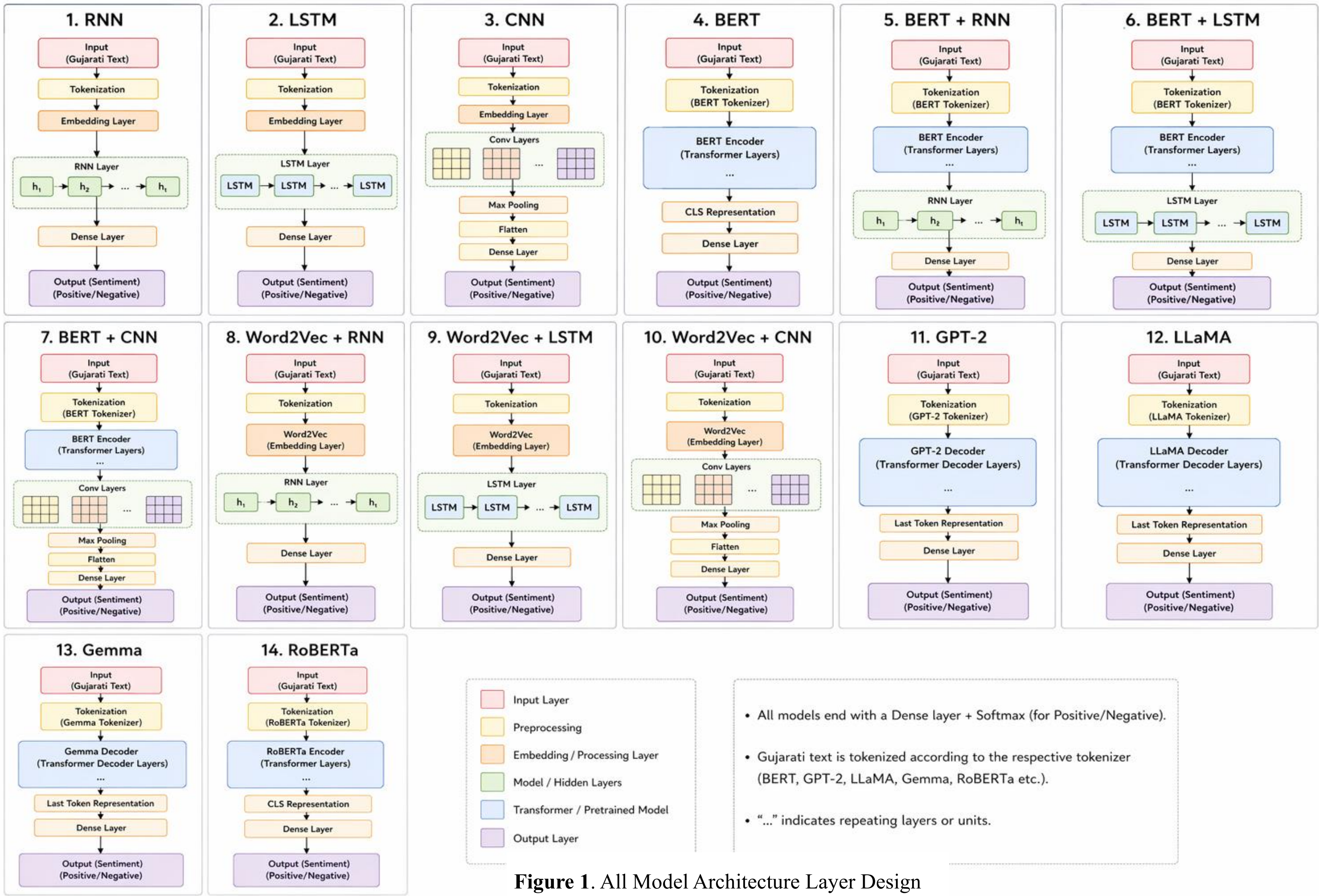


Figure 1. All Model Architecture Layer Design

3.3 Model Selection & Parameter setup

3.3.1 RNN

A particular kind of artificial neural network called an RNN uses patterns found in data to predict the most likely result that will occur next. In order to forecast the output of each layer, it works on the principle that the output of every layer is saved and feedback into the input of the structure.

For the RNN, The model is a Bidirectional RNN (SimpleRNN layers) with dropout and batch normalization, trained with Adam optimizer**,** categorical crossentropy loss**,** ReLU + Softmax activations**,** using 15 epochs max**,** batch size 64**, and** learning rate scheduling with early stopping to optimize performance shown in following Table 3.

Table 3. Defining Constrains of RNN

Parameter's	**Constrains Values**
Learning-rate	0.001
Epoch	15
Optimizer	Adam
Batch-size	64
Activation	ReLU + Softmax
Loss Function	Categorical cross-entropy

3.3.2 CNN

CNN is made up of many neuronal networks. When a picture is entered, each network layer generates a some of processes that are then transfer to the next network, which is how this works best with images. Usually, basic features such as diagonal or horizontal edges are taken out in the initial network. This output is forward to the next network, which uses it to identify more complex features such as many edges or corners. As we delve further, the network may be able to identify ever-more-complex items, such as faces, objects, etc.

The CNN model was train for 5 epochs with the Adam optimizer with a batch size of 32, employing softmax activation and categorical crossentropy as the loss function, as illustrated in Table 4 below.

Table 4. Defining Constrains of CNN

Parameter's	**Constrains Values**
Learning-rate	0.001
Epoch	5
Optimizer	Adam
Batch-size	32
Activation	ReLU + Softmax
Loss Function	Categorical cross-entropy

3.3.3 LSTM

Since processing is done in order in this approach, a bidirectional LSTM was used. It consists of two LSTMs, one of which permits feedback in a particular way and the other in a different direction. Multiple all connected network that connect every neuron in network to every other network neuron make up a fully connected neural network.

For the LSTM, The model uses a BiLSTM + Dense architecture with dropout and batch normalization, trained using Adam (lr=0.001) for up to 15 epochs (early stopping enabled), batch size 64, optimizing categorical crossentropy, and dynamically adjusting learning rate if validation loss stalls. as illustrated in Table 5 below.

Table 5. Defining Constrains of LSTM.

Parameter's	Constrains Values
Learning-rate	0.001
Epoch	15
Optimizer	Adam
Batch-size	64
Loss Function	Categorical cross-entropy

3.3.4 Word2Vec

The Word2Vec embedding method provides only one independent embedding vector per word. Word2vec saves a single vector per word in the output model. As the model is maintained as a singular vector for each word,Word2vec is used without context in a downstream NLP job even if it was trained using contextual neighbors. stagnating in use as a result. This limits one's capacity to comprehend a word's meaning in two distinct settings and situations.For example, the word "કાંઠો" means the bank of a river, while the word "કાંઠો" is also used to indicate the edge or side of something. Similarly, the word "માઉસ" is used as a computer device, while "માઉસ" is also known as an animal.

The Word2vec model, utilizing CNN, RNN, and LSTM architectures, was trained for 20, 10, and 25 epochs, respectively, with a learning rate of 0.001, employing the Adam optimizer and a batch size of 32. The activation functions utilized were ReLU and softmax, while categorical crossentropy served as the loss function, as detailed in Table 6 below.

Table 6. Defining Constrains ofWord2vec with CNN, RNN, and LSTM

Parameter's	Constrains Values
Learning-rate	0.001
Epoch	20,10,25
Optimizer	Adam
Batch-size	32
Activation	ReLU + Softmax
Loss Function	Categorical cross-entropy

3.3.5 *BERT*

BERT was trained on two different assignments: masked language tasks, where the model is given sentences and other word's are masked or hidden for the model, and the model make attempts to expect these hidden word's; and unmasked language activity, where the model is given sentences and some word's are hiding or covered up for the model. The second task is sentence prediction, where the model is given a pair of sentences every round and must determine whether or not one sentence is followed by the other. For these two tasks, BERT has been train on a sizable database. A transformation model, a kind of neural network layer that accepts a statement as input, provides the encoder used by BERT. The BERT model is then fed the tokenized words after each word in the phrase has been tokenized. Each tokenized word is represented by a vector in the BERT output.

Transformer's encoders allow BERT to understand context better than conventional neural networks like LSTM or RNN because they process all inputs, or the entire sentence, at once. This means that when BERT builds a context for a word, it considers both the inputs before and after the word's, whereas LSTM and RNN method the implement by considering simply the earlier inputs, which is reflected in the word's output matrix value., so the word "કાંઠો" in the two sentences (નદીના કાંઠા પર ઘણા લોકો સવારના સમયે ફરવા આવે છે.) and (મેજના કાંઠા પર રાખેલી ફાઇલ જમીન પર પડી ગઇ.) would have two distinct vectors when using BERT, however when using LSTM or RNN, it would possess the same matrix value and hence the actual meaning. As a result, BERT will typically perform better than the conventional ML algorithms. There are main two types of the BERT: the BERT big and the BERT base.

The implementation of utilized the BERT base. For the BERT base alongside CNN, RNN and LSTM, the model was fine-tunes google/muril-base-cased for text classification using Stratified K-Fold cross-validation (5 folds). It trains with AdamW optimizer, learning rate 2e-5, 2e-4, 2e-5, batch size 16,32,16, and 5,15,5 epochs, while monitoring F1 score with early stopping. The classification head uses ReLU, ReLU + Softmax, Softmax + categorical cross-entropy loss.

Table 7. Defining Constrains for BERT with CNN, RNN, and LSTM

Parameter's	Constrains Values
Learning-rate	2e-5, 2e-4, 2e-5
Epoch	5,15,5
Optimizer	AdamW
Batch-size	16,32,16
Activation	ReLU, ReLU + Softmax, Softmax
Loss Function	Categorical cross-entropy

3.3.6 GPT2

The model fine-tuning GPT-2 on a classification task with AdamW optimizer , batch size=8, epochs=5, cross-entropy loss, and GELU activations in the backbone.

Table 8. Defining Constrains for GPT2

Parameter's	Constrains Values
Learning-rate	5e-5
Epoch	5
Optimizer	AdamW
Batch-size	8
Activation	GELU
Loss Function	Cross-entropy loss

3.3.7 LLaMA

The model fine-tuning LLaMA on a classification task with AdamW optimizer , batch size=8, epochs=5, cross-entropy loss, and Softmax activations in the backbone.

Table 9. Defining Constrains for LLaMA

Parameter's	Constrains Values
Learning-rate	2e-5
Epoch	5
Optimizer	AdamW
Batch-size	8
Activation	Softmax
Loss Function	Cross-entropy loss

3.3.8 Gemma

The model fine-tuning Gemma on a classification task with AdamW optimizer , batch size=4, epochs=5, cross-entropy loss, and Softmax activations in the backbone.

Table 10. Defining Constrains for Gemma

Parameter's	**Constrains Values**
Learning-rate	2e-5
Epoch	5
Optimizer	AdamW
Batch-size	4
Activation	Softmax
Loss Function	Cross-entropy loss

3.3.9 RoBERTa

The model fine-tuning RoBERTa on a classification task with AdamW optimizer, batch size=16, epochs=5, cross-entropy loss, and Softmax activations in the backbone.

Table 11. Defining Constrains for RoBERTa

Parameter's	**Constrains Values**
Learning-rate	2e-5
Epoch	5
Optimizer	AdamW
Batch-size	16
Activation	Softmax
Loss Function	Cross-entropy loss

4 RESUTS AND DISCUSSING

4.1. Performance Evaluation

It's a performance metric that measures the recommendation of correct classifications out of all classifications made by a model. Accuracy is calculated as follows:

$$\text{Accuracy} = \frac{\text{True Positives} + \text{True Negatives}}{\text{Total Samples}} \tag{1}$$

The quantity of total positive prediction that is accurate is known as precision. This value is calculated by dividing the total amount of classified positives by the total number of anticipated positives. The precision of a well-performing model should be high. The following is a calculated of precision:

$$\text{Precision} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Positives}} \tag{2}$$

Here TP stands for true positive and FP stands for false positive.

Recall can be expressed as the ratio of all correctly detected favorably classed classes to all positively classified courses, or as the number of correctly anticipated classes with a good outcome. A strong model is indicated by a high recall rate. The calculated of recall as follows:

$$\text{Recall} = \frac{\text{True Positives}}{\text{True Positives} + \text{False Negatives}} \tag{3}$$

Here FN stands for false negative

F1 score is a machine learning evaluation metric that measures a model's accuracy. It combines the precision and recall scores of a model. It is calculated as follows:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4}$$

4.2. Experimental Results

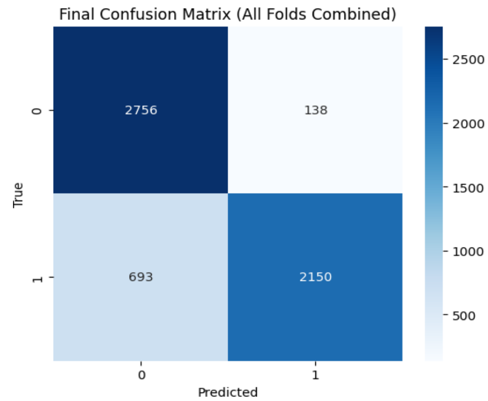
The key metrics of classification of every model taken into consideration in the study are summarized in Table 9.

Table 9. Overview of Model Performance

Sr. No	Model	Accuracy	Precision	Recall	F1 Score
1	RNN	85.50	93.97	75.62	83.80
2	LSTM	95.31	93.78	96.98	95.35
3	CNN	97.09	97.48	96.62	97.05
4	Word2Vec-RNN	96.29	96.70	95.78	96.24
5	Word2Vec-LSTM	96.69	96.99	96.31	96.65
6	Word2Vec-CNN	95.47	96.58	94.20	95.37
7	**BERT**	**98.31**	**97.97**	**98.63**	**98.30**
8	BERT-RNN	95.77	96.20	95.22	95.71
9	BERT-LSTM	98.22	98.27	98.14	98.21
10	BERT-CNN	98.13	98.17	98.07	98.12
11	GPT2	94.89	94.58	95.15	94.86
12	LLaMA with LoRA	94.17	94.19	94.17	94.17
13	Gemma with LoRA	95.76	95.77	95.76	95.76
14	RoBERTa with XLM	98.13	98.13	98.13	98.13

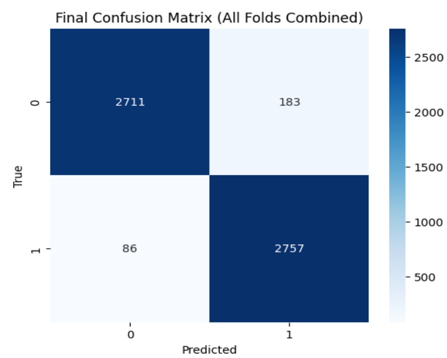
According to the RNN's confusion matrix, as seen in Figure 1 below, 2150 out of 2843 positive feelings and 2756 out of 2894 negative sentiments were accurately classified.

Figure 2. Confusion matrix of RNN



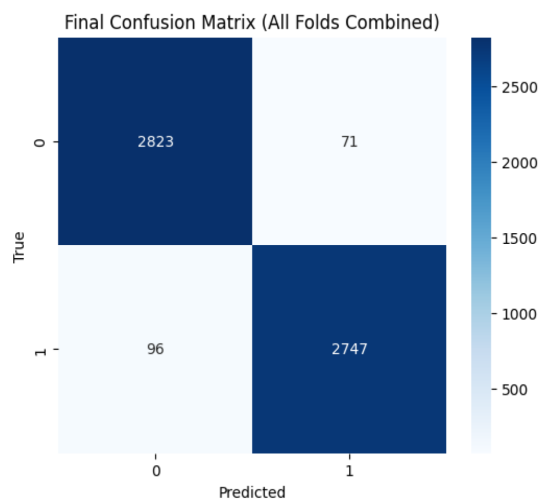
As seen in Figure 2 below, the LSTM confusion matrix shows that 2711 out of 2894 negative feelings and 2757 out of 2843 positive sentiments were properly identified.

Figure 3. Confusion matrix of LSTM



According to CNN's confusion matrix, as seen in Figure 3 below, 2747 out of 2843 positive sentiments and 2823 out of 2894 negative sentiments were accurately classified.

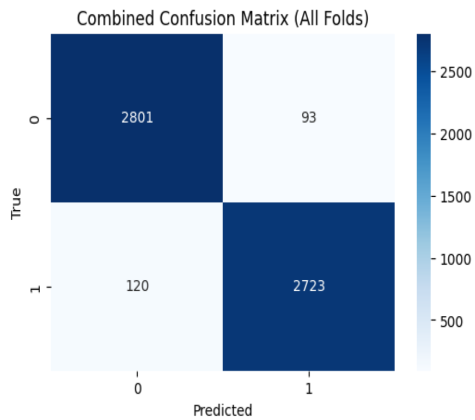
Figure 4. Confusion matrix of CNN



7

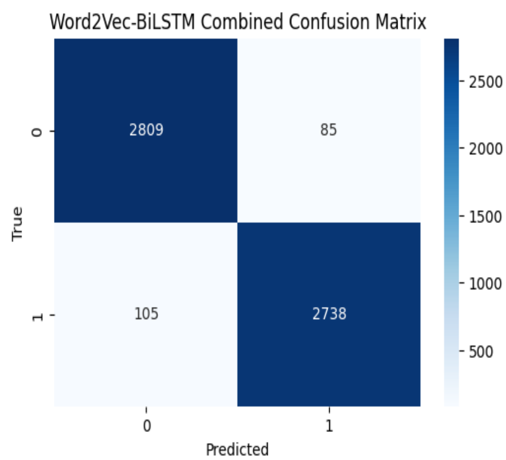
As seen in Figure 4 below, the Word2Ve-RNN confusion matrix shows that 2723 out of 2843 positive feelings and 2801 out of 2894 negative sentiments were accurately identified.

Figure 5. Confusion matrix of Word2Ve-RNN



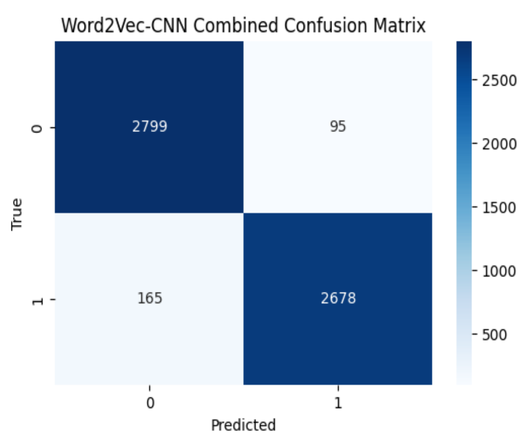
As seen in Figure 5 below, the Word2Ve-LSTM confusion matrix shows that 2738 out of 2843 positive feelings and 2809 out of 2894 negative sentiments were properly identified.

Figure 6. Confusion matrix of Word2Ve-LSTM



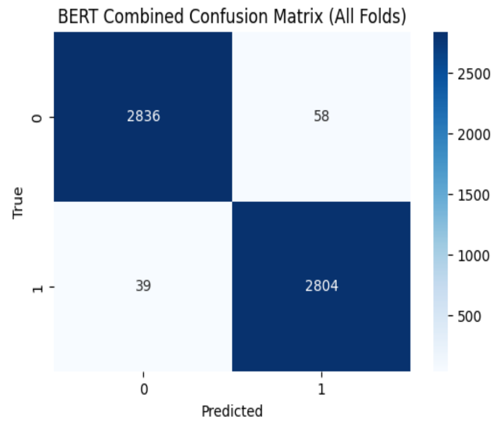
As seen in Figure 6 below, the Word2Ve-CNN confusion matrix shows that 2799 out of 2894 negative feelings and 2678 out of 2843 positive sentiments were properly identified.

Figure 7. Confusion matrix of Word2Ve-CNN



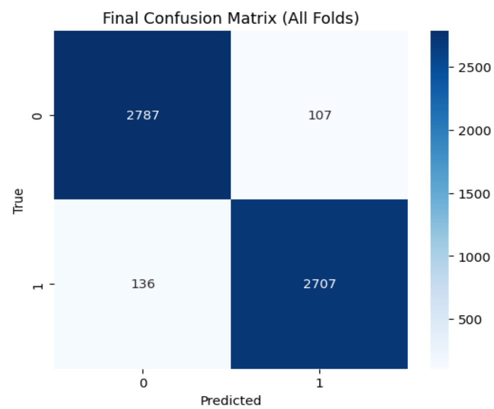
As seen in Figure 7 below, the BERT confusion matrix shows that 2836 out of 2894 negative sentiments and 2804 out of 2843 positive sentiments were accurately identified.

Figure 8. Confusion matrix of BERT



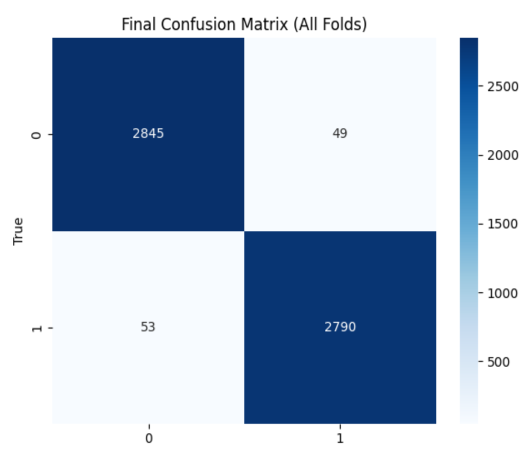
As seen in Figure 8 below, the BERT-RNN confusion matrix shows that 2787 out of 2894 negative feelings and 2707 out of 2843 positive sentiments were accurately identified.

Figure 9. Confusion matrix of BERT-RNN



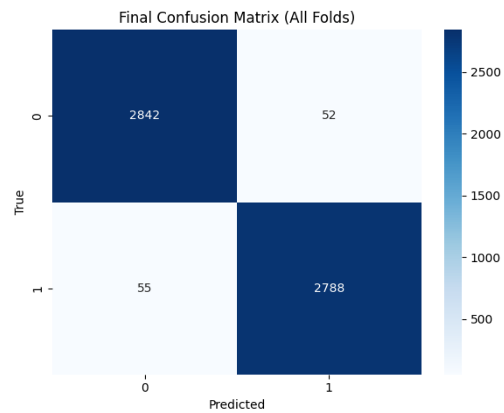
As seen in Figure 9 below, the BERT-LSTM confusion matrix shows that 2790 out of 2843 positive feelings and 2845 out of 2894 negative sentiments were accurately identified.

Figure 10. Confusion matrix of BERT-LSTM



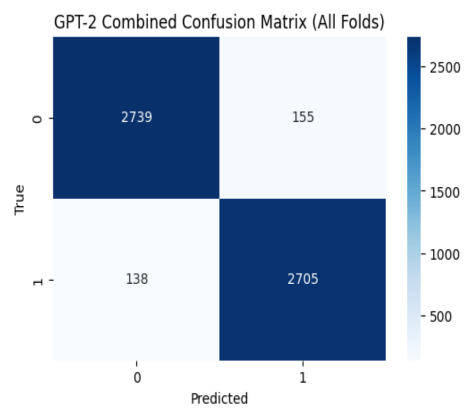
As seen in Figure 10 below, the BERT-CNN confusion matrix shows that 2788 out of 2843 positive feelings and 2842 out of 2894 negative sentiments were properly identified.

Figure 11. Confusion matrix of BERT-CNN



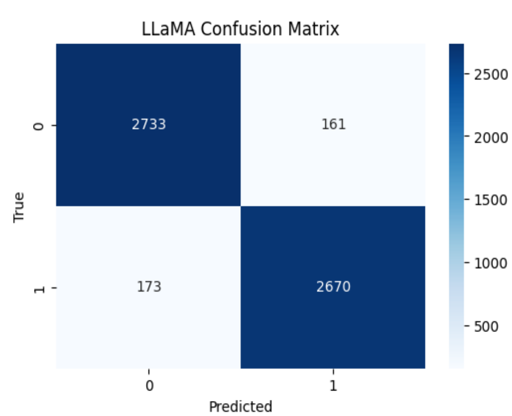
As seen in Figure 11 below, the confusion matrix of GPT2 shows that 2739 out of 2894 negative feelings and 2705 out of 2843 positive sentiments were accurately identified.

Figure 12. Confusion matrix of GPT2



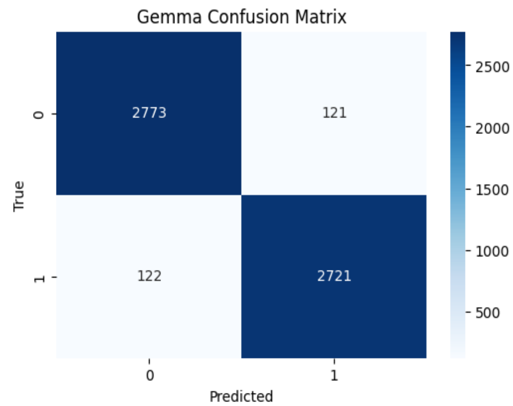
As seen in Figure 12 below, the confusion matrix of LLaMA shows that 2733 out of 2894 negative feelings and 2670 out of 2843 positive sentiments were accurately identified.

Figure 13. Confusion matrix of LLaMA



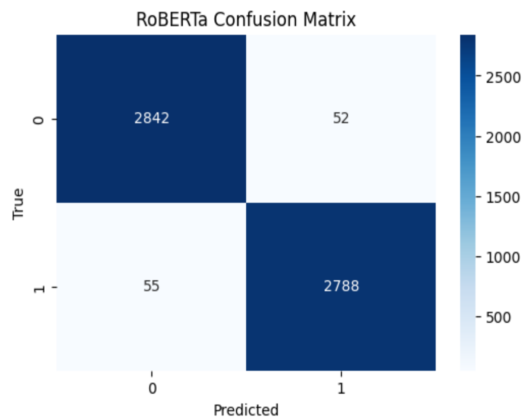
As seen in Figure 13 below, the confusion matrix of Gemma shows that 2773 out of 2894 negative feelings and 2721 out of 2843 positive sentiments were accurately identified.

Figure 14. Confusion matrix of Gemma



As seen in Figure 14 below, the confusion matrix of Gemma shows that 2842 out of 2894 negative feelings and 2788 out of 2843 positive sentiments were accurately identified.

Figure 15. Confusion matrix of RoBERTa



5. CONCLUSION

There are some disadvantages to the traditional approach to natural language processing (NLP), which uses CNN, RNN, LSTM, and Word2Vec. One of these drawbacks is that it fails to convey the word's deeper meaning. Word2Vec limitations the capacity to comprehend a words signifying in both different settings and scenarios, which will be represented in the word's output matrix value. On the other hand, because the encoder evaluates all inputs—that is, the full sentence—at once, the BERT has greater comprehension than conventional techniques. As a result, BERT will consider both the inputs that proceeded and followed a word when creating a context for it.

The transformer model BERT is used to deliver an outstanding performance in terms of key metrics of classification. With an accuracy of 98.31%, the BERT model produced my greatest results.

Future research may extend this work by developing a multimodal sentiment analysis framework for the Gujarati language that integrates textual data with additional modalities, including images, audio, and emojis frequently used in social media communication. This methodology would enable models to more effectively capture the comprehensive context of user opinions and emotions. Furthermore, real-time sentiment monitoring systems could be implemented using streaming data from social media platforms to dynamically assess public opinion. The incorporation of knowledge graphs and large language models may further enhance semantic understanding and contextual reasoning in Gujarati content. These advancements would contribute to the development of more robust and practical sentiment analysis systems for low-resource languages.

References

1. Ishaani Priyadarshini, Chase Cotton, A novel LSTM–CNN–grid search based deep neural network for sentiment analysis, The Journal of Supercomputing (2021) 77:13911–13932 https://doi.org/10.1007/s11227-021-03838-w
2. Akana Chandra Mouli Venkata Srinivas, Ch. Satyanarayana, Ch.Divakar,KatikireddyPhani Sirisha, Sentiment Analysis using Neural Network and LSTM, IOP Conf. Series: Materials Science and Engineering 1074 (2021) 012007 https://doi:10.1088/1757-899X/1074/1/012007
3. Priyank Sonkiya, Vikas Bajpai, Anukriti Bansal, Stock price prediction using BERT and GAN, arXiv:2107.09055v1 [q-fin.ST] 18 Jul 2021
4. Ru Yang, Maryam Edalati, Using GAN-based models to sentimental analysis on imbalanced datasets in education domain, arXiv:2108.12061v1 [cs.CL] 26 Aug 2021
5. Stanislav Chumakova, Anton Kovantseva, Anatoliy Surikov, Generative approach to Aspect Based Sentiment Analysis with GPT Language Models, Procedia Computer Science 229 (2023) 284–293
6. Ying Cao, ZhexingSun, Ling Li and Weinan Mo, A Study of Sentiment Analysis Algorithms for Agricultural Product Reviews Based on Improved BERT Model, Symmetry 2022, 14, 1604. https://doi.org/10.3390/sym14081604
7. Ms. K. Chitra, Dr. G. Kavitha, Dr. P. Latchoumy, Penalty based Sentimental Text Generation Framework using Generative Adversarial Networks, International Conference on Automation, Computing and Renewable Systems (ICACRS), 978-1-6654-6084-2/22/$31.00 ©2022 IEEE https:109/ICACRS55517.2022.10029135/Doi: 10.1
8. Gustavo H. de Rosa, Jo ao P. Papa, A Survey on Text Generation using Generative Adversarial Networks, arXiv:2212.11119v1 [cs.CL] 20 Dec 2022
9. Nusrat Jahan Prottasha, Abdullah asSami, Md Kowsher, Saydul Akbar Murad, Anupam Kumar Bairagi, Mehedi Masud and Mohammed Baz, Transfer Learning for Sentiment Analysis Using BERT Based Supervised Fine-Tuning, Sensors 2022, 22, 4157.
10. Chenxi Tian, YuliangMa, Jared Cammon, Feng Fang, Yingchun Zhang, and Ming Meng, Dual-Encoder VAE-GAN With Spatiotemporal Features for Emotional EEG Data Augmentation, IEEE TRANSACTIONS ON NEURAL SYSTEMS AND REHABILITATION ENGINEERING, VOL. 31, 2023
11. Konstantinos I. Roumeliotis, Nikolaos D. Tselikas, Dimitrios K. Nasiopoulos, LLMs in e-commerce: A comparative analysis of GPT and LLaMA models in product review evaluation, Natural Language Processing Journal 6 (2024) 100056
12. Sio Jurnalis Pipin1, Frans Mikael Sinaga, SunaryoWinardi, Muhammad Noor Hakim, Sentiment Analysis Classification of ChatGPT on Twitter Big Data in Indonesia Using Fast R-CNN, JURNAL MEDIA INFORMATIKA BUDIDARMA Volume 7, Nomor 4, Oktober 2023, Page 2137-2148 https://DOI: 10.30865/mib. v7i4.6816
13. XiangDeng, VasilisaBashlovkina, FengHan, SimonBaumgartner, MichaelBendersky, LLMs to the Moon? Reddit Market Sentiment Analysis with Large Language Models, WWW '23 Companion, April 30–May 04, 2023, Austin, TX, USA
14. Megharani Patil, Hrishikesh Yadav, MahendraGawali, Jaya Suryawanshi, Jaikumar Patil, Anjali Yeole, Prathik Shetty, Jayesh Potlabattini, A Novel Approach to Fake News Detection Using Systems and Applications in Engineering IJISAE, 2024, 12(4s), 343–354
15. Abayomi Bello, Sin-Chun Ng and Man-Fai Leung, A BERT Framework to Sentiment Analysis of Tweets, Sensors 2023, 23, 506. https://doi.org/10.3390/s23010506
16. Yaowei Yue, Yun Peng, and Duancheng Wang, Deep Learning Short Text Sentiment Analysis Based on Improved Particle Swarm Optimization, Electronics 2023, 12, 4119. https://doi.org/10.3390/electronics12194119
17. Amjad Iqbal, Rashid Amin, Javed Iqbal,RoobaeaAlroobaea, Ahmed Binmahfoudh and Mudassar Hussain, Sentiment Analysis of Consumer Reviews Using Deep Learning, Sustainability 2022, 14, 10844. https://doi.org/10.3390/su141710844
18. Bilen B., Horasan F., "LSTM network-based sentiment analysis for customer reviews", PoliteknikDergisi, 25(3): 959-966, (2022). DOI: 10.2339/politeknik.844019
19. Shariq Imran, Ru Yang,ZenunKastrati, Sher Muhammad Daudpota, Sarang Shaikh, the impact of synthetic text generation for sentiment analysis using GAN based models, Egyptian Informatics Journal 23 (2022) 547–557
20. Putra Fissabil Muhammada, RetnoKusumaningruma, Adi Wibowoa, Sentiment Analysis Using Word2vec and Long Short-Term Memory (LSTM) For Indonesian Hotel Reviews, Procedia Computer Science 179 (2021) 728–735
21. Dr. G. S. N. Murthy, Shanmukha Rao Allu, Bhargavi Andhavarapu, Mounika Bagadi, Mounika Belusonti, Text based Sentiment Analysis using LSTM, International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-0181, Vol. 9 Issue 05, May-2020
22. Sharma Vishalkumar Sureshbhai, Dr. Tulsidas Nakrani, A Literature Review : Enhancing Sentiment Analysis of Deep Learning Techniques Using Generative AI Model, International Journal of Scientific Research in Computer Science, Engineering and Information Technology, ISSN : 2456-3307, Volume 10, Issue 3 , Page Number 530-540, May-June-2024 , doi : https://doi.org/10.32628/CSEIT24103204