Enhancing Twitter Sentiment Analysis using a Dual Perspective Robust Optimized BERT with Deep Learning

[1st]S. Prathi
Department of Computer Application
Vels Institute of Technology and advanced Studies (VISTAS)
Chennai, India
prathi.sundar1986@gmail.com

[2nd]J. Jebathangam
Department of Computer Application
Vels Institute of Technology and advanced Studies (VISTAS)
Chennai, India
jthangam.scs@vistas.ac.in

***Abstract*—Twitter is among the most common social media platforms on which the users can be seen sharing opinions, emotions and experiences, most of the times, using short texts. Twitter Sentiment Analysis (TSA) has gained relevance as the user-generated content has grown exponentially, and it is crucial to find out what users feel and classify their views as positive, negative, or neutral. This is important in most applications of Natural Language Processing (NLP) including opinion mining, social monitoring and market analysis. Nevertheless, most of the current TSA methods primarily depend on textual characteristics and usually encounter problems in the analysis of short, informal, and ambiguous tweets. Moreover, the application of conventional methods of sentiment analysis often omits emojis data despite the popularity of the latter in conveying feelings and contextual meaning in social media communication. These factors can be overlooked, which can decrease the precision of sentiment classification. In an effort to evade these constraints, this paper suggests a two-perspective robust optimized bert (DP-ROBERTa) that combines textual and emoji based data in enhancing the sentiment classification. Twitter sentiment data is retrieved in the Kaggle repository and processed with the Lexicon Text Terms Normalization (LT2N) feature in order to eliminate punctuations, redundant tags and noisy components. Aspect-based Emoji Text Affinity Rate (AETAR) method is then applied to calculate the relationship between emojis and textual aspects by giving affinity weights to them. The proposed DP-ROBERTa model will process these affinity-weighted representations of the tweets, as it will focus on local fine-grained sentiment patterns and global semantic representations. The experimental outcomes show that the proposed framework has a better performance than the baseline models in terms of accuracy, precision, recall, and F1-score to effectively classify the emotions of the Twitter posts.**

Keywords— Natural Language Processing (NLP), Twitter Sentiment Analysis (TSA), Lexicon Text Terms Normalization (LT2N), Aspect-based Emoji Text Affinity Rate (AETAR), Dual Perspective Robust Optimized BERT algorithm (DP-ROBERTa)

I. INTRODUCTION

Social media sites have emerged as a significant communication platform of opinion, emotions and experience, with Twitter being one of the most commonly used sites with its instant and short format communication. Twitter Sentiment Analysis (TSA) is valuable to understand the opinion of people and to assist in the application of the tool in marketing analysis, policy evaluation, disaster management, or social monitoring (Zhou et al., 2022). Nonetheless, the analysis of tweets has its own set of issues due to the error of briefness, informality, and slang, abbreviation, hashtags, and emojis that are inherent in a tweet. These features complicate sentiment and contextual meaning interpretation through the use of conventional Natural Language Processing (NLP) models. Previous TSA algorithms were primarily based on lexicon-based algorithms or classical machine learning algorithms which relied on the existence of handcrafted dictionaries and handcrafted features. These methods offer a baseline sentiment classification performance, but fail to address contextual associations and entailed emotional indicators expressed in tweets (Aljedaani et al., 2022). As deep learning progressed, transformer-based models, like Bidirectional Encoder Representations from Transformers (BERT), have contributed immensely to the performance of NLP, because they learn contextual associations of words in both directions in the form of bidirectional representations (Feng, 2025). RoBERTa is an optimized BERT that further purposes training datasets of a larger size and more effective pretraining approaches (He, 2024). Nevertheless, most of the models that operate on the basis of a transformer are primarily textual and do not consider non-verbal information sources like emojis, which is crucial in the communication format of social media (Li et al., 2021). Recent works highlight the advantages of dual-perspective sentiment analysis strategies that integrate textual elements of locality with global elements of semantics to get a more accurate sentiment recognition (Wang et al., 2023). The local viewpoint reveals subtle patterns of emotional feel of single tweets whereas the global viewpoint determines the existence of semantic connections among the entire data set. Moreover, emojis often enhance or alter the meaning of textual message, and give good contextual emotion indications. Based on these observations, this paper presents a Dual Perspective Robust Optimized BERT (DP-ROBERTa) model to enhance sentiment analysis in Twitter. The given solution will use Lexicon Text Terms Normalization (LT2N) as a preprocessing tool and Aspect-based Emoji Text Affinity Rate (AETAR) to measure the connection between emojis and textual aspects. Through the combination of textual and emoji sentiment-based information, the DP-ROBERTa model is effective in classifying sentiment of tweets based on positive, neutral, and negative with an enhanced performance on sentiment analysis as a whole.

A. Objectives

- To enhance Twitter sentiment analysis by effectively incorporating emoji-based emotional cues.
- To develop a dual-perspective framework that captures local hidden patterns in tweets.
- To extract global semantic representations for a more comprehensive understanding of tweet context.
- To improve classification accuracy for short and ambiguous tweets.
- To achieve robust sentiment prediction for emoji-rich tweets by integrating text

The Contribution of the work is DP-ROBERTa algorithm, a novel BERT-based model that integrates textual and emoji information for robust sentiment classification. It proposes Lexicon Text Terms Normalization (LT2N) for preprocessing and Aspect-based Emoji Text Affinity Rate (AETAR) for weighting the importance of emojis and text. The approach demonstrates superior performance compared to existing baseline and deep learning models across precision, recall, accuracy, and F1-score metrics.

The paper is organized as follows: Section 2 reviews related work on TSA and transformer-based models. Section 3 describes the proposed methodology, including preprocessing, AETAR, and DP-ROBERTa architecture. Section 4 presents experimental results and performance analysis. Section 5 concludes the study and discusses future research directions.

II. RELATED WORK

Wang et al. (2023) presented a dual-perspective fusion network (DPFN) for multimodal Aspect-Based Sentiment Analysis (ABMSA). They combined both text data and image data. They adopted Deep Neural Language Models (DNLMs) for extracting features from text data, while Convolutional Neural Network (CNN) features for extracting features from image data. They used a Dual Perspective Fusion (DPF) approach for obtaining the features of both Global Sentiment (GS) and Aspect Sentiment (AS).

Qu et al. (2025) introduced a dual attention-based graph convolutional neural network framework for carrying out multimodal sentiment analysis. This framework involves the use of a graph structure to represent the multimodal information, in which the nodes contain the textual, visual, and contextual information. The dual attention mechanism is employed to capture the inter-modal and intra-modal relationships.

Abrar et al. (2025) proposed a dual-modal credibility analysis framework for the detection of fake news using the combination of Interpretive Structural Modeling (ISM) and BERT. The proposed technique uses the BERT model for the extraction of text features, whereas ISM is applied for the modeling of structural relationships among credibility factors.

Abulfaraj et al. (2025) suggested a parallel CNN approach based on the BiLSTM model with an attention mechanism and an ensemble mechanism to handle Twitter sentiment analysis. This approach utilizes Bidirectional LSTM (BiLSTM) to learn dependences in contextual words and utilizes parallel CNN layers to achieve a better representation of local n-gram features. Finally, an Attention Mechanism (AM) highlights sentiment-relevant words within tweets.

Bello et al. (2023) presented a framework using BERT for analyzing Twitter sentiments. The approach employs BERT fine-tuning to obtain high-level contextual and semantic data from tweets. Preprocessing approaches were employed to cope with noisy data on Twitter, which contains many informal words like hashtags, emojis, and abbreviations. The proposed framework utilizes BERT's self-attention technique to tackle long-term dependencies.

Kannan et al. (2022) proposed an approach for fine-tuning BERT-based sentiment analysis on multi-classes of emotion datasets from Twitter. Deep contextual embeddings using BERT are utilized to capture nuances at the semantic and syntactic levels of tweets. Preprocessing cleans noisy social media text, with slang, hashtags, and emojis.

Albeladi et al. (2025) introduced a framework using the BERT model for the sentiment analysis of Twitter datasets related to smart city services. The framework incorporates the use of BERT Embeddings (BE), which enables the extraction of context and meaning by analyzing the tweets. The Preprocessing (PP) component deals with the noise in the text on the social media platforms.

Talaat et al. (2023) presented a hybrid sentiment analysis classification system based on hybrid models of BERT. This technique uses a combination of different versions of the BERT model in order to identify a range of context and semantic information present within the given text. Various preprocessing techniques are used for the cleaning of the given text, such as punctuation, stopwords, and special characters.

A heuristic-based approach for Twitter sentimental analysis using BERT was proposed by **Yenduri et al. (2021)**. It addresses the problem by combining the power of BERT, which captures contextual and semantic information in the context of features, with heuristic methods in order to improve the process for sentimental analysis.

Joloudari et al. (2023) introduced the use of the BERT-Deep CNN model for the sentiment analysis task of COVID-19-related tweets. This model utilizes the embeddings derived from the Bidirectional Encoder (BE) for the extraction of deep-level context and semantic information present within the tweet texts. Deep-level CNNs are used for the extraction of local n-gram information. This information is further propagated through the Deep CNN for the improvement of the representation capabilities of the model.

TABLE I. COMPARISON OF BERT-BASED APPROACHES FOR SENTIMENT ANALYSIS IN SOCIAL MEDIA AND CUSTOMER INSIGHTS

Ref	Application Domain	Technique	Methods	Contribution
Rahman et al. (2024)	Customer Satisfaction	BERT	Machine Learning (ML)	BERT-based sentiment analysis to extract customer insights and optimize satisfaction

Wu et al. (2024)	Social Media Sentiment	Deep Learning-based BERT	CNN / LSTM	Deep learning integration with BERT for improved sentiment classification
Prottasha et al. (2022)	Social media	BERT	Transfer Learning (Supervised Fine-Tuning)	Transfer learning approach to fine-tune BERT for sentiment analysis tasks
Mann et al. (2023)	Twitter	Enhanced BERT	–	Enhanced BERT model for improved Twitter sentiment analysis performance
Muneer et al. (2023)	Cyberbullying Detection	Enhanced BERT	Stacking Ensemble Learning	Combines BERT with ensemble learning to detect cyberbullying on social media
Phan et al. (2022)	Aspect-Level Sentiment	BERT-GCN	CNN	Uses BERT embeddings with Graph Convolutional Networks and CNN for aspect-level sentiment analysis
Nistor et al. (2021)	Twitter	Recurrent Neural Network (RNN)	Deep Learning-based Sentiment Analysis	Developed a Twitter sentiment analysis system using recurrent neural networks to classify tweet sentiments effectively

Table 1 provides a comparative summary of recent BERT-based methods used for sentiment analysis in social media and customer-oriented domains. The articles demonstrate the overall adaptability of BERT and its variants across various sentiment analysis tasks, including customer satisfaction analysis, Twitter emotion classification, cyberbullying detection, and aspect-level sentiment analysis. The vast majority of works enhance the performance of baseline BERT using additional deep learning methods such as CNNs, LSTMs, Graph Convolutional Networks, and ensemble or transfer learning approaches.

III. PROPOSED METHODOLOGY

The suggested methodology presents a Dual Perspective Robust Optimized BERT (DP-ROBERTa) pipeline that allows improving Twitter Sentiment Analysis by providing a combined text- and emoji-based information representation. The first step is to first take a benchmark TSA dataset obtained in the Kaggle repository. The preprocessing of the raw tweets is then done with the help of Lexicon Text Terms Normalization (LT2N) that filters out any noise like punctuations, URLs, undesired tags, and other irrelevant symbols and ensures the textual input is clean and standardized. An Aspect-based Emoji Text Affinity Rate (AETAR) mechanism is used to capture the emotional semantics contained in emojis; it gives textual terms and emojis in their contextual relevance the appropriate weights. Such affinity-weighted representations of tweets are then inputted into the presented DP-ROBERTa model, which works in two perspectives. The local viewpoint has the model deriving fine-grained contextual pattern and concealed linguistic features on small and ambiguous tweets, whereas the global viewpoint has the model deriving long-range dependencies and semantically rich representations on the entire tweet. Lastly, this is done using the integrated features to classify the tweets into positive, neutral, and negative sentiment. Extensive experiments prove the efficacy of the suggested methodology, the best performance in comparison with the baseline models in terms of precision, recall, accuracy, and F1-score.

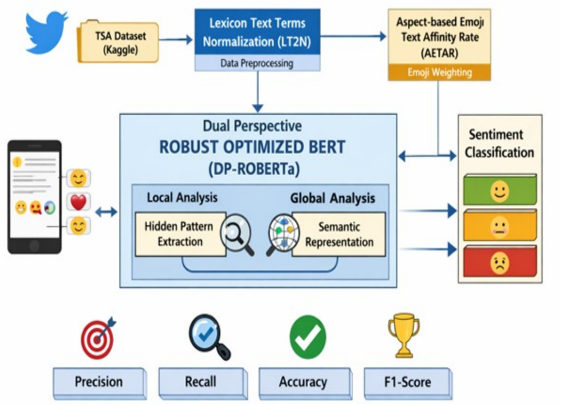

Fig. 1. Proposed Methodology for Tweet Sentiment Analysis Using DP-ROBERTa

Figure 1 is used to show the general process of the proposed Twitter Sentiment Analysis framework. The TSA data that can be found in the Kaggle repository is initially preprocessed by the Lexicon Text Terms Normalization (LT2N) tool that eliminates noise and normalizes tweet texts. Then, Aspect-based Emoji Text Affinity Rate (AETAR) method is used to provide affinity weights based on emojis and textual content relationship. The affinity-weighted tweets are subsequently fed through the Dual Perspective Robust Optimized BERT (DP-ROBERTa) model where both local analysis and hidden pattern extraction are carried out as well as global analysis and semantic representation learning. Lastly, the model categorizes the tweets into positive, neutral, and negative sentiment groups and its accuracy is tested in terms of precision, recall, accuracy, and F1-score.

A. *Lexicon Text Terms Normalization (LT2N)*

Lexicon Text Terms Normalization (LT2N) is a method for standardizing raw textual data and enhancing its quality prior to conducting an analysis. This method removes any symbols that do not add to the semantic meaning of the text, such as punctuation, Web addresses (URLs), social media icons (emojis), and unrelated tags. LT2N also normalizes all text into a consistent format by performing actions such as case normalization, removal of additional blank spaces, and correction of all forms of lexicon inconsistency (such as informal writing/slang or shortened last names). Creating a clean and normalized text through LT2N generates less interference, increases the consistency of the lexical content in the text input, and creates a more readable and accurate textual input which will assist with the success and accuracy of the future Natural Language Processing (NLP) and Sentiment Analysis processes.

The analysis of LT2N, AETAR, and the two-perspective DP-ROBERTa architecture is performed in terms of the contribution to the sentiment classification performance. LT2N method enhances the quality of the tweets used by normalizing the text and removing noise whereas AETAR mechanism assigns affinity scores to encode the connection between emojis and the text. The DP-ROBERTa model is then used to extract global and local sentiment features in the form of these weighted representations. The performance of the proposed framework is measured based on performance measures like accuracy, precision, recall and the F1-score.

Let the raw text corpus be represented as

$$T = \{t_1, t_2, \dots, t_n\} \quad (1)$$

According to equation (1) , $T = \{t_1, t_2, \dots, t_n$depicts the set of all objects or entities suo, hence t_i represents some specific member out of those many members; t_1 is the first member and t_2 is the second member of the collection, whilst t_n indicates that it is the final member of the collection - allowing for (T) to grow infinitely large by virtue of being limited by size (n). The notation represents a way in which to store groups of common actions/events/points of data within a single variable within the realm of analysis (Mathematics, Computer science) or can represent what may be a continuum of an ever-expanding set of related objects that provide a means of consolidating them into the smaller number of variables necessary to perform analysis between very large numbers of objects or groups. Initially, hyperlinks and HTML tags are removed to eliminate non-linguistic noise.

$$t_i^{(1)} = t_i \setminus \{\text{URLs,HTML tags}\} \quad (2)$$

The removal of structural and web content from t_i represented in equation (2) as shows that this removal occurs in the first step of the data-processing pipeline (superscripted '1'). Removing these noise components allows cleaner input, allowing for subsequent transformations of the data. Next, emojis and special symbols that do not contribute to semantic meaning are filtered out using

$$t_i^{(2)} = t_i^{(1)} \setminus \{\text{emojis,special symbols}\}$$
(3)

Element t_i is processed a second time with the equation (3). The $t_i^{(2)}$element is created by removing 'emojis' and special characters from the first stage processed text, $t_i^{(1)}$ The superscript $t_i^{(2)}$ identifies it as the second transformation stage within the sequence of processing steps, as well as how much further the preprocessing will normalise the underlying data to increase its potential consistency and applicability for analytics, or machine learning.

$$t_i^{(3)} = t_i^{(2)} \setminus \{\text{punctuation}\} \quad (4)$$

The formula (4) states that $t_i^{(3)} = t_i^{(2)} \setminus \{\text{punctuation}\}$ indicating that all punctuation symbols have been removed from the text processed in$t_i^{(3)} = t_i^{(2)}$ Additionally, the superscript number $t_i^{(3)}$denotes the third transformation step in the preprocessing process. The text is then converted to lowercase to ensure lexical uniformity across all terms, formulated as

$$t_i^{(4)} = \text{Lowercase}t_i^{((3)})\quad (5)$$

The equation (5) represented $t_i^{(4)} = \text{Lowercase}t_i^{((3)})$as the fourth step of preprocessing that is done on ti to ensure that there is no case sensitivity in the representation used to conduct analysis or modelling. Extra whitespaces and redundant gaps are removed to improve text consistency, defined as

$$t_i^{(5)} = \text{TrimSpaces}\left(t_i^{(4)}\right) \quad (6)$$

The equation (6) represented $t_i^{(5)} = \text{TrimSpaces}(t_i^{(4)})$. It means that the lower-case text (distinctly $t_i^{(5)}$superscript $(t_i^{(4)})$ will be stripped of unnecessary leading, trailing and intermediary spaces in the preprocessing pipeline. The result of such operation is a small and formatted text that can be subjected to efficient processing and analysis. Stop words that do not contribute to sentiment or semantic meaning are removed using a predefined stop-word set S.

$$t_i^{(6)} = t_i^{(5)} \setminus S \quad (7)$$

The processing that is used on the element ti is represented as equation (7) $t_i^{(6)} = t_i^{(5)} \setminus S$. It shows that the set of elements (S) of the text, which are usually stop words or other irrelevant tokens, are eliminated out of the text $t_i^{(5)}$ The $t_i^{(6)}$ in the preprocessing sequence indicates the sixth transformation level. This action will minimize noise and dimensionality and make future text analysis or learning models more effective. Lexical normalization is then applied by mapping slang, abbreviations, and informal terms using a lexicon L, expressed as

$$t_i^{(7)} = \text{LexMap}\left(t_i^{(6)}, L\right) \quad (8)$$

The text processing within the element t_i is the equation (8) $t_i^{(7)} = \text{LexMap}(t_i^{(6)}, L)$ which means that the text is cleaned $(t_i^{(6)}, L)$and it is processed by a lexical resource or dictionary (L). The superscript denotes the seventh processing pipeline transformation step. It is an operation that transforms the tokens to standardised lexical forms or semantic representations, so that meaningful feature extraction and analysis is possible.

$$T_{\text{LT2N}} = \{, t_2^{(7)}, \dots, t_n^{(7)}\} \quad (9)$$

The equation (9) presents a last derived transformed set is the after all the preprocessing and lexical mapping are done. Each $t_1^{(7)}$represents the completely processed variant of the initial element t_i the suffix $LT2N$ represents this set as the result of the $LT2N$ transformation pipeline. This set is an input normalized and semantically enhanced to use in further analytic or machine learning activities.

B. Aspect-Based Emoji Text Affinity Rate (AETAR)

Aspect-Based Emoji Text Affinity Rate (AETAR) is a metric used to measure the semantic and emotional consistency between emojis and aspect-specific textual content within a sentence or document. It focuses on identifying how effectively emojis complement, emphasize, or modify the sentiment expressed toward particular aspects, rather than the overall text sentiment. By aligning emoji polarity, intensity, and contextual meaning with extracted aspects, AETAR captures the strength of emoji–text interaction at a fine-grained level. A higher AETAR score indicates stronger coherence between emojis and aspect-level sentiments, leading to more accurate and interpretable sentiment analysis outcomes.

The DP-ROBERTa is also trained with Adam optimizer, a popular deep learning model optimizer, as it has an efficient parameter updating method and a higher convergence rate. A proper learning rate is used in order to maintain steady training and to avoid a significant change in the performance of the model during the optimization process. The training process is handled with the help of Python-based deep learning framework, according to which the model is trained on the affinity-weighted representations of tweets that have been post-processed to LT2N and then calculated with AETAR. Its experiments are run in a typical computational setup with adequate processing power and memory availability onboard, such that it is possible to efficiently train and test the sentiment classification model.

$T = \{w1, w2, \dots, wm\}$ (10)

The equation (10) T= {w1, w2…, Wm} is used to represent the input text as an ordered sequence of mm tokens and w_i denotes each word or symbol that is a part of the text. The given formulation enables the writing to be analysed at the token level. It is the underlying representation of the extraction of aspects, sentiments and emojis. There is a high level of efficiency in semantic and contextual information being captured by modelling of text as a sequence of tokens using computational methods.

$A = \{a1, a2, \dots, an\}$ (11)

The set of aspects extracted in the input text is represented by the equation (11) $A = \{a1, a2, \dots, an\}$. These a_i are attributes, features or topics of which sentiment is given. The given representation allows one to examine the situation at a small level, paying attention to the details, not to the general mood of the text. It lays the foundation of sentiment and emoji affinity computation at an aspect level.

$E = \{e1, e2, \dots, ek\}$ (12)

The formula (12) $E = \{e1, e2, \dots, ek\}$ is the collection of emojis within the text in question. All ei are single emojis that relay information that is emotional or expressive. The parameter k shows the amount of the emojis in the text. This representation allows performing a systematic study of emoji sentiment and its relations with the textual dimensions.

$S(ai) \in [-1,1], i = 1, \dots, n$ (13)

Sentiment polarity score in each of the extracted aspects a i is given by the equation (13) $S(ai) \in [-1,1], i = 1, \dots, n$ where $\in [-1,1]$,is the negative, zero, and positive sentiment respectively. The score nearer to -1 is strong negative sentiment and a score nearer to +1 is the strong positive sentiment. Using this formulation, it is possible to do a quantitative comparison of the strength of sentiments in different aspects.

$P(ej) \in [-1,1], j = 1, \dots, k$ (14)

The sentence (14) $P(ej) \in [-1,1], j = 1, \dots, k$ is the sentiment polarity or emotional intensity score of each emotion e_j. The negative emotional connotation of emojis is reflected in the range $[-1,1]$of the meanings. A negative value nearer to -1 represents negative emotions whereas a positive value nearer to $+1$ represents positive emotions. This expression enables sentiment and affinity analysis to quantitatively incorporate emojis.

$M = \{(a_i, e_j)\ e_j\ is\ contextually\ linked\ to\ ai\}$
(15)

The set of equations (15) Contextually connected with ai is the mapping set of aspects and emojis within the text. Every pair (ai, ej) means that the emoji (e_j) is semantically or positionally related to the aspect (a_i). This association is discovered according to contextual proximity or semantic relevance. The set (M) allows modeling the effects of emojis at the aspect level systematically.

$C(ai, ej) = cos(ai, ej)$ (16)

The contextual relevance between an aspect ai and an emoji e_j by the equation (16) $C(a_i, e_j) = cos(a_i, e_j)$ Computed It is calculated by the similarity of their vectors in an embedding space through cosine similarity. The larger the value of the cosine, the greater the semantic congruence between the aspect and the emoji. This measure represents the extent to which an emoji is contextually relevant to a particular aspect.

$Aff(ai, ej) = S(ai) \times P(ej) \times C(ai, ej)$ (17)

The equation (17) represented *is* the affinity of a given aspect and an emoji extracted to it. In this case, $S\ (a_i)$ represents the polarity of the sentiment of the aspect, and $P(e_j)$ represents the strength of emotion of the emoji. Their contextual relevance in the text is measured by the term $C\ (a_i\ e_j)$. A combination of these elements results in a composite score, which measures the extent of emoji-aspect sentiment congruence.

$AETAR = \frac{1}{|M|} \sum_{(a_i.e_j) \epsilon M} Aff(a_i, e_j)$ (18)

The equation (18) described an Aspect-Based Emoji Text Affinity Rate by averaging the affinity scores of all aspect emoji pairs in the mapping set M. Here, $\mid M \mid$ is the total number of aspects emoji pairs, and Aff $(a_i e_j)$ Aff (a_i, e_j) represents the individual affinity between an aspect and its linked emoji. This aggregation provides a single metric reflecting the overall coherence between emojis and aspect-level sentiments in the text. A higher $AETAR$ indicates stronger alignment between textual aspects and emoji. The combination of the AETAR weighting mechanism and the DP-ROBERTa architecture is carried out in the stage of feature preparation and training. Following the LT2N preprocessing, the affinity scores calculated by the AETAR technique reflect the contextual connection of emojis and the textual elements in every tweet. The weights to the respective tokens are assigned based on these scores, which indicates information in the twitter representation that is sentiment-relevant. The textual embeddings are affinity-weighted and presented as a form of input to the DP-ROBERTa model, and in the course of the training, the architecture acquires the global semantic characteristics as well as the local sentiment patterns.

C. Dual Perspective Feature Learning (DP-ROBERTa)

Dual Perspective Feature Learning (DP-ROBERTa) is a more advanced approach that builds on the existing ROBERTa models and adds two additional directions of textual information as complementary to each other. One is more concerned with global contextual knowledge, which encodes the general semantic and syntactic framework of the text, the other is more concerned with local or aspect-level, retrieving information of fine-grain, concerning a particular object or sentiment or topic. When combining the two views, DP-ROBERTa produces more discriminative feature representations and thus performs better when it comes to sentiment analysis, aspect-based opinion mining, and emotion detection. This method uses the ready-trained power of ROBERTa and adds to it a specific learning of features to gain macro and micro-level knowledge of the text.

$$T = \{w1, w2, \dots, wm\} \quad (19)$$

The equation (19) represents a T, which is an ordered set of tokens, where w_i is a single word or symbol. This tokenisation allows the model to break down the text in a fine level. It lays the basis to project every token into a vector space to continue with further processing. With the representation of text in such a structured manner, the semantic context and sequential relationships may be properly represented.

$$X = ROBERTa\_Embed(T) = \{x_1, x_2, \dots, x_m\} \quad (20)$$

The equation (20) used to denote the process of encoding each token w_i in the input text T as a dense vector x_i by $ROBERT$ a embedding model. These embeddings are used to distil the semantic and contextual meaning of each of the tokens in the sequence. The output set X maintains the sequence of tokens whilst coded rich linguistic characteristics. This representation is used as the input to further global and local feature extraction in the model.

$$F_{global} = GlobalEncoder(X) \quad (21)$$

The equation (21) represents $F_{global} = GlobalEncoder(X)$ is the function of deriving global features of the token embeddings (X). The GlobalEncoder takes into account the general semantic and contextual detail of the whole text chain. This produces the feature vector F_{global} that represents the macro-level knowledge about the text. These global aspects are subsequently joined to local/aspect-level features to have an all-inclusive representation learning.

$$F_{local} = LocalEncoder(X, A) \quad (22)$$

The equations (22) denotes the extraction of local or aspect-level features of the token embeddings (X) under the influence of the set of extracted aspects (A). The LocalEncoder is concerned with the fine-grained data about every aspect, sentiment, context, or particular features. The output feature vector F_{local} stresses micro-level text of an object which might be lost to global encoding. Such local features are then combined with world features to provide a complete image of the text.

$$F_{global} = W_g F_{global} + b_g, F'_{local} = Wl F_{local} + b_l \quad (23)$$

The equations (23) represented is the linear projection of global and local feature vectors into a shared feature space. In this case $W_g$and W_l are learnable weight matrices, b_g and bl are the bias terms. This change is such that the global and local features are compatible in terms of their dimensions towards fusion. It also enables the model to highlight the significant elements of each type of feature and then merge them. Linear projection of both global and local features into a common feature space.

$$FDP = \alpha\, F'_{global} + (1 - \alpha) F_{local} \quad (24)$$

The equation (24) described a F local or local, the following step involves a non-linear transformation $1 - \alpha)F_{local}$ This activation function, e.g. ReLU or tanh, is what adds non-linearity to the model to allow the model to learn complex semantic relationships. The features are made more expressive by the transformed vector H. It is used as the input to the last task-specific prediction layers. Combines global and local features with a weighting factor α∈ [0,1] alpha in [0,1] α∈ [0,1].

$$H = \sigma(F_{DP}) \quad (25)$$

The equations (25) indicates that this is a non-linear activation function of the combined dual perspective feature vector. In this case, (F_{DP}) will include the global and local characteristics of DP-ROBERTa. The activation function σ (e.g. ReLU or tanh) allows the model to learn non-linear and complicated patterns. The resulting H will give a more expressive version to follow up the prediction tasks.

$$O = Softmax(W_o H + b_o) \quad (26)$$

The last stage of model prediction can be shown as its equation (26) $O = Softmax(W_o H + b_o)$ In this case, W o and b o are parameters of the learning process that maps the hidden feature vector H to the output space. The $SoftMax$ function transforms these values to normalized probability scores of each of the classes. O therefore gives the confidence of the model on each prediction label. Produces predictions in either classification or regression problems.

$$L = CrossEntropy(O, Y) \quad (27)$$

The equation (27) presents a loss function to train the model. In this case, O is the predicted distribution of probabilities that were given by the SoftMax layer, and Y is a true target distribution. $CrossEntropy$ is a measure of the difference between the actual and predicted distributions. Minimization of this loss directs the model to enhance accuracy of prediction in the training process. Trains a model depending on the ground truth labels Y.Optimizes model parameters based on the ground truth labels Y.

IV. RESULTS AND DISCUSSION

Table 1: Simulation Parameter

Parameters	Values
Dataset Name	Twitter Sentiment Analysis
Operating System	Windows 11
Total Number of Data	74682
Training Data	30682
Testing Data	40,000
Programming Language	Python 3.10 models

Table 2 shows the experimental design in assessing the proposed Twitter Sentiment Analysis model. The Twitter data has 74,682 tweets in which 30,682 are used as sample during training and 40,000 as sample during testing. The experiments are performed on the Windows 11 operating system with python 3.10. It is a conducive environment to preprocess data, train models and conduct sentiment classification analysis.

The efficiency of the suggested Dual Perspective Robust Optimized BERT (DP-ROBERTa) as a tool to improve the performance of Twitter Sentiment Analysis (TSA) compared to the existing baseline and deep learning models. Preprocessing stage; LT2N is added to make sure that the noise is removed and that the text of the tweets is standardized so as to enhance the quality of the textual features used in analysis. In AETAR mechanism is another instance that can be added to the model to capture the contextual interaction between the emojis and the text aspects in the text to allow the system to comprehend emotional messages in tweets. The DP-ROBERTa model is successful in training the local hidden sentiment patterns, as well as global semantic representations, on the affinity-weighted tweet embeddings. As the experimental findings on the benchmark Twitter dataset show, the given model has a better accuracy, precision, recall, and F1-score than the standard BERT, RoBERTa, and other deep learning networks. It is also demonstrated in the findings that the emoji information can be effectively utilized to enhance the classification of short and ambiguous tweets. Additionally, the dual-perspective learning approach guarantees the superior feature extraction and semantic consistency throughout the dataset, which is why the proposed DP-ROBERTa framework is an efficient and trustworthy method of sentiment classification and social media opinion mining.

A. Dataset Description

The Twitter entity sentiment dataset will be partitioned into two training and testing datasets, with the former being utilized to acquire the model parameters and the latter to gauge the performance of the proposed DP-ROBERTa model. The data is made up of three sentiment classes, positive, negative, and neutral, which represent various opinions within twitter posts and the proportion of class distribution is kept in both subsets to get an equal evaluation. The tweet data is pre-processed by LT2N before the training, this includes the removal of URLs, punctuations, special characters and unnecessary tokens as well as normalization of the text. Notably, the irrelevant messages to the entity are also considered as Neutral, which means that sentiment analysis will be conducted with respect to the target content only. It creates a possible opportunity to analyze public opinion more accurately, which can be implemented in such applications as brand monitoring, social media analytics, and customer feedback evaluation.

https://www.kaggle.com/datasets/jp797498e/twitter-entity-sentiment-analysis?select=twitter_training.csv

B. Comparison Performance

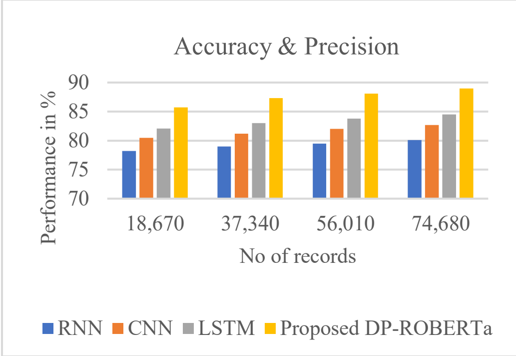

Fig. 2. Accuracy & Precision Comparison of RNN, CNN, LSTM, and Proposed DP-ROBERTa Across Different Dataset Sizes

Figure 2 Figure shows the accuracy performance comparison between RNN, CNN, LSTM and the proposed DP-RoBERTa model with various numbers of records of tweets. All models increase their accuracy slowly with the increase in the size of the dataset. Nevertheless, the proposed DP-RoBERTa has the greatest accuracy of approximately 85.7, 87.3, 88.1, and 89, as well as approximately 84, 85.2, 86, and 87 in various experiments. This enhancement demonstrates a stronger performance of the model in capturing more contextual and semantic information of tweets. The addition of semantic processing of LT2N and AETAR emoji-text affinity weighting, further increases the capability of the model to comprehend the semantics of tweets. Therefore, the suggested DP-RoBERTa exhibits better sentiment classification results than those of the classical deep learning models including RNN, CNN, and LSTM, which prove to be efficient when analyzing the sentiment of many tweets.

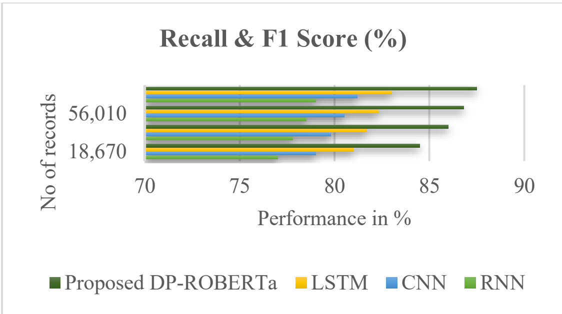

Fig. 4. Comparison of Recall (%) Across Models with Varying Dataset Sizes

Figure 4 shows the performance results of recall and F1-score of RNN, CNN, LSTM, and the proposed DP-RoBERTa model with varying tweet record counts. As one can see, the recall and the F1-score values of all the models grow slowly as the size of the datasets increases. Nevertheless, the best performance is obtained with the proposed DP-RoBERTa that has 84.5%, 86, 86.8, and 87.5% recall and 84.3, 85.6, 86.4, and 87.2% F1-scores respectively. This increase indicates that the model is good in capturing more contextual and semantic relationships in twetts. Therefore, DP-RoBERTa model performs better than the conventional deep learning models like RNN, CNN and LSTM in general performance in sentiment classification.

Fig. 5. F1-Score Comparison of RNN, CNN, LSTM, and Proposed DP-ROBERTa Models Across Different Dataset Sizes

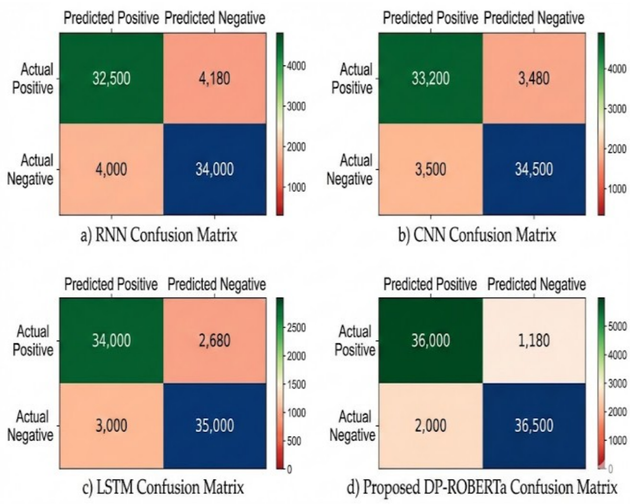

Fig. 6. Confusion Matrix Comparison Across Models

Figure 6 presents the The figure is used to compare the results of the confusion matrix of RNN, CNN, LSTM, and proposed DP-RoBERTa-based sentiment classification of tweets. RNN model accurately predicts 32,500 positive and 34,000 negative tweets but has bigger misclassification with 4,180 false negatives and 4,000 false positives. At CNN, the results are enhanced to 33,200 true positives and 34,500 true negatives, making the number of errors 3,480 and 3,500. LSTM also enhances results as it achieves an accuracy of 34,000 and 35,000 correct decisions. The optimal performance of the proposed DP-RoBERTa is 36, 000 true positives and 36,500 true negatives and reduced misclassifications.

V.CONCLUSION

Finally, this work suggested a Dual Perspective Robust Optimized BERT (DP-ROBERTa) model that can enhance the Twitter Sentiment Analysis with the help of textual and emoji data. Conventional sentiment analysis techniques primarily use textual information and tend to be incapable of comprehending brief and ambiguous tweets in an appropriate manner. To overcome this drawback, the proposed framework uses Lexicon Text Terms Normalization (LT2N) to clean and standardize texts of tweets and Aspect-based Emoji Text Affinity Rate (AETAR) to remove the semantic connection between emojis and textual aspects. The DP-ROBERTa architecture then processes these affinity-weighted representations of tweets and learns local fine-grained representations of sentiment and global contextual representations. The experimental assessment of the Twitter data shows that the proposed model has better performance than the baseline models in terms of precision, recall, accuracy and F1-score. The framework may be further expanded in future work to multilingual sentiment analysis, multimodal data, e.g. images and videos, and real-time sentiment monitoring systems that are easier to interpret with explainable AI techniques.

REFERENCE

[1] Feng, Jiawen. "Improving sentiment analysis in literary texts through bidirectional encoder representations: A BERT-based approach." Journal of Computational Methods in Sciences and Engineering (2025): 14727978251338004.

[2] He, Luoyao. "Enhanced twitter sentiment analysis with dual joint classifier integrating RoBERTa and BERT architectures." Frontiers in Physics 12 (2024): 1477714.

[3] Zhou, Bing, et al. "VictimFinder: Harvesting rescue requests in disaster response from social media with BERT." Computers, Environment and Urban Systems 95 (2022): 101824.

[4] Li, Hongchan, et al. "Weibo text sentiment analysis based on bert and deep learning." Applied Sciences 11.22 (2021): 10774.

[5] Aljedaani, Wajdi, et al. "Sentiment analysis on Twitter data integrating TextBlob and deep learning models: The case of US airline industry." Knowledge-Based Systems 255 (2022): 109780.

[6] Wang, Di, et al. "Dual-perspective fusion network for aspect-based multimodal sentiment analysis." IEEE Transactions on Multimedia 26 (2023): 4028-4038.

[7] Qu, Na, et al. "Dual attention-based graph convolutional neural network for multimodal sentiment analysis: N. Qu et al." The Journal of Supercomputing 81.10 (2025): 1154.

[8] Abrar, Muhammad Faisal, et al. "A dual-modal analysis of credibility in integrating interpretive structural modeling (ISM) and BERT for enhanced fake news detection." Multimedia Systems 31.5 (2025): 317.

[9] Abulfaraj, Anas W. "BiLSTM-Based Parallel CNN Models With Attention and Ensemble Mechanism for Twitter Sentiment Analysis." IEEE Access (2025).

[10] Bello, Abayomi, Sin-Chun Ng, and Man-Fai Leung. "A BERT framework to sentiment analysis of tweets." Sensors 23.1 (2023): 506.

[11] Kannan, Eswariah, and Lakshmi Anusha Kothamasu. "Fine-Tuning BERT Based Approach for Multi-Class Sentiment Analysis on Twitter Emotion Data." Ingénierie des Systèmes d'Information 27.1 (2022).

[12] Albeladi, Abrar, and Turki Alghamdi. "Sentiment Analysis of Twitter Data on Smart City Services using BERT Model." International Journal of Computer Science & Network Security (2025): 64-80.

[13] Talaat, Amira Samy. "Sentiment analysis classification system using hybrid BERT models." Journal of Big Data 10.1 (2023): 110.

[14] Yenduri, Gokul, et al. "Heuristic-assisted bert for twitter sentiment analysis." International Journal of Computational Intelligence and Applications 20.03 (2021): 2150015.

[15] Joloudari, Javad Hassannataj, et al. "BERT-deep CNN: State of the art for sentiment analysis of COVID-19 tweets." Social Network Analysis and Mining 13.1 (2023): 99.

[16] Rahman, Ben. "Optimizing customer satisfaction through sentiment analysis: A BERT-based machine learning approach to extract insights." IEEE Access (2024).

[17] Wu, Yichao, et al. "Research on the application of deep learning-based BERT model in sentiment analysis." arXiv preprint arXiv:2403.08217 (2024).

[18] Prottasha, N. J., Sami, A. A., Kowsher, M., Murad, S. A., Bairagi, A. K., Masud, M., & Baz, M. (2022). Transfer learning for sentiment analysis using BERT based supervised fine-tuning. Sensors, 22(11), 4157.

[19] Mann, Suman, et al. "Twitter sentiment analysis using enhanced bert." Intelligent Systems and Applications: Select Proceedings of ICISA 2022. Singapore: Springer Nature Singapore, 2023. 263-271.

[20] Muneer, Amgad, et al. "Cyberbullying detection on social media using stacking ensemble learning and enhanced BERT." Information 14.8 (2023): 467.

[21] Phan, Huyen Trang, Ngoc Thanh Nguyen, and Dosam Hwang. "Aspect-level sentiment analysis using CNN over BERT-GCN." Ieee Access 10 (2022): 110402-110409.

[22] Nistor, Sergiu Cosmin, et al. "Building a Twitter sentiment analysis system with recurrent neural networks." *Sensors* 21.7 (2021): 2266.