© 2026 IJRTI | Volume 11, Issue 4 April 2026 | ISSN: 2456-3315

Deep Learning-Based Multi-Class Sentiment Analysis of Social Media Tweet

Authors
[1]Ashutosh Narayan Sharma , [2]Priyanshu Sahu , [3]Raj Choubey , [4]Raj Rajeshwar Yadu , [5]Rashmi Banchhor
[1]Team Leader , [2]Coder , [3]Coder , [4]Paper Writer , [5]Assistant Professor
[1]Computer Science and Engineering (Artificial Intelligence)
[1]Shri Shankaracharya Institute of Professional Management and Technology , Raipur , India
[1]asharma@ssipmt.com, [2]priyanshsahu@ssipmt.com, [3]raj.choubey@ssipmt.com, [4]ryadu@ssipmt.com,
[5]rashmi.banchhor@ssipmt.com

Abstract* — The way we communicate has changed a lot because of communication systems and social media platforms like Twitter, Facebook. These platforms generate an amount of user-generated content every day. This content shows what people think and feel about things. To get information from this content is a big challenge. We need to analyze this data to understand what people think about a product or a service. This is important for things like customer feedback analysis, brand monitoring and decision support systems. But the way people write on media is informal they use slang, abbreviations and their grammar is not always correct. This makes it hard to analyze the data. Sentiment analysis is a part of Natural Language Processing (NLP). It is about finding and categorizing opinions, emotions and attitudes in text. We have used machine learning methods like Naïve Bayes, Support Vector Machines (SVM) and Logistic Regression to do sentiment analysis. These methods use feature extraction techniques like Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) to convert text into numbers. These methods are not very good at understanding the context of the text and the relationships between words. Recent advancements in learning, particularly transformer-based architectures have improved the performance of NLP systems. Transformers can process sequences of text at the same time and understand the context. Models like BERT, RoBERTa and DistilBERT have shown results in sentiment analysis tasks. These models create embeddings by looking at the text on both sides of a word. This helps to interpret the text accurately. In this research we propose a sentiment analysis system that uses transformer-based models. The system is a web-based application that can predict sentiment in time. The system has stages, such as data preprocessing, tokenization, embedding generation and model inference. We used a dataset of airline-related tweets to test the system. This dataset is a benchmark because it is informal and noisy. The results show that RoBERTa is more accurate than BERT and DistilBERT.. Distilbert is faster and uses less computational power making it good for real-time applications. BERT is a balance between accuracy and computational complexity. The proposed system is a solution for sentiment analysis in real-world applications. It uses the strengths of transformer-based models to improve the accuracy and reliability of sentiment classification. It can be used in domains, like social media monitoring, customer feedback analysis and decision support systems, where understanding user sentiment is important.. The system can be used to monitor media and understand what people think about a product or a service, It can be used to analyze customer feedback and improve services. It can be used to make decisions based on user sentiment. Overall the system is scalable, efficient and practical. It can be applied in domains where understanding user sentiment is important.

***Keywords*: Sentiment Analysis, Transformer-Based Models, RoBERTa, BERT, DistilBERT, Natural Language Processing (NLP), Deep Learning, Text Classification, Self-Attention Mechanism, Real-Time Prediction**

I. Introduction

Artificial Intelligence (AI) has become a disruptive technology that has had a profound impact on many sectors, such as healthcare, finance, education, and communication. Some of its numerous uses include the Natural Language Processing (NLP) which is significant in making machines understand, interpret and create human language. Sentiment analysis, a subset of NLP, is one of the most notable applications of NLP, and it aims at discerning the emotional tone or sentiment presented in written information. As social media networks, including Twitter, Facebook, Instagram, etc., rapidly expand, every day, users produce enormous quantities of unstructured textual information. This data is rich in information on the opinion of the population, customer satisfaction and trends in the market. Nevertheless, it is inefficient and impractical to analyze such great amount of data manually. Sentiment analysis systems that are automated are thus needed in order to extract meaningful information in time and in a cost-effective way. The classical methods of sentiment analysis used machine learning methods including Naive Bayes, Support Vector Machines (SVM) and Logistic Regression. These models are based on the feature engineering methods like Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF), which encode textual information in terms of numbers. These methods are reasonable in performance, but do not give us the contextual relationships and the word order of text, so their usefulness in the real world is limited. Recurrent Neural Networks (RNN) and Long Short-Term Memory (LSTM) models were introduced, which represented a breakthrough in the NLP. These models have been created to operate with sequential data and to obtain associations between words. Nevertheless, they do it at a word-by-word basis making them more computationally complex and less able to effectively capture longer-range dependencies. The latest developments in transformer-based architectures have radically transformed NLP. Transformer model is a model that uses self-attention to process sequences of data simultaneously and allows them to compute efficiently and understand the context better. Based on this architecture, other models, including BERT and RoBERTa, have demonstrated state-of-the-art performance on a variety of NLP tasks, including sentiment analysis. BERT provides a bidirectional encoding that helps to interpret the context of words in a sentence, whereas RoBERTa improves the technique and trains more data. DistilBERT is also more efficient in that it minimizes size without compromising performance of the model. These models have shown to be more accurate and robust than traditional and deep learning models. This work presents a transformer-based sentiment analysis system, where BERT, RoBERTa and DistilBERT models are incorporated into a web-based program. The system will be engineered to handle real-world textual data, and will be used to deliver correct sentiment classification in real time. Using the state-of-the-art transformer architectures, the proposed system is expected to break the constraints of the conventional solutions and enhance the overall performance. Figure 1: Retrieval-Augmented Generation

II. LITERATURE REVIEW

1) Conventional Sentiment Analysis Methods

Sentiment analysis first depended on rule-based and lexicon-driven methodologies whereby predefined dictionaries were employed to establish the polarity of text. These approaches were easy to use and understand, but they did not have the power to represent the contextual dependencies and semantic subtleties in a natural language. The human language is naturally complicated as explained by Pang and Lee [1], and sometimes rule-based systems cannot interpret sarcasm, ambiguity, and informal expressions encountered in real-life text in a way that is correct. In order to address these shortcomings, machine learning algorithms like Naive Bayes, Support Vector Machines (SVM), and Logistic Regression were developed. These models are based on such feature extraction methods like Bag-of-Words (BoW) and Term Frequency-Inverse Document Frequency (TF-IDF) which transforms the text into the form of numbers. Liu [2] highlighted that these methods enhanced classification performance over and above rule-based systems, but still had serious limitations, such as loss of word order and the failure to formulate contextual dependencies. Consequently, they do not work well with complicated data including social media text, where meaning in most cases is contextually dependent.

2) Deep Learning Model and Word Embeddings

Distributed word representations were a big breakthrough in Natural Language Processing. The use of techniques like Word2Vec [3] and GloVe [4] facilitated the representation of words as dense vectors in a continuous vector space, which captured semantic relationships between words. These embeddings enabled the models to learn similarities and relationships among words which enhanced them to perform better in sentiment analysis tasks. Nevertheless, these methods still give each word a predetermined representation, irrespective of how it is used in various sentences. Deep learning models also enhanced sentiment analysis by allowing automatic feature extraction and improved sequential data processing. Recurrent Neural Networks (RNN) were developed to work with sequences by keeping a hidden state which stores information about past inputs. RNNs are however affected by the vanishing gradient problem, which restricts their learning of long term dependencies [5]. To overcome this, Long Short-Term Memory (LSTM) networks have been proposed, which have memory cells and gating to store significant information in a longer sequence [6]. On the same note, Gated Recurrent Units (GRU) offer a computationally efficient variant, with a similar performance [7]. Although these advances have been made, sequential models like RNN, LSTM, and GRU have their limitations. They work stepwise through the text, making their computational processes more complex and limiting the ability to process information in parallel. Moreover, their capabilities of capturing long range dependencies are finite especially in large and complex data sets.

3) Transformer-Based Models

The shortcomings of the sequential models gave rise to the invention of transformer-based architectures, which have revolutionized the NLP field immensely. Transformer model, proposed by Vaswani et al. [8], is a model that makes use of self-attention mechanisms to analyze the relationships between words in a sentence, irrespective of the position of the words. This allows the model to take into consideration both local dependencies and global dependencies and also to enable parallel computation. On the basis of this architecture, Bidirectional Encoder Representations of Transformers (BERT) was proposed by Devlin and others [9]. BERT works in both directions, i.e. it takes into consideration the left and the right context at the same time. This results in context-aware embeddings, which significantly improve performance in sentiment analysis tasks. Studies have demonstrated that BERT-based models are more effective compared to the traditional machine learning and deep learning models because they can comprehend the more complex linguistic structures. RoBERTa, which is a suggestion by Liu et al. [10], is an extension of BERT where the training process is optimized. It eliminates the following sentence prediction task, employs dynamic masking and is trained on larger data sets. These enhancements result in increased generalization and accuracy. Comparative studies [11], [12] show that RoBERTa is always better than BERT, especially when it is used on unstructured and noisy data like social media text. DistilBERT is a small size version of BERT that employs knowledge distillation methods to achieve small model size without compromising on its performance [13]. This is especially applicable in real-time applications due to the limitation of computational resources. Altogether, transformer-based models are the latest state-of-art in sentiment analysis because they can better represent the contextual relationships and complex language patterns.

4) Research Gap and Motivation

Judging by the available literature, it is clear that the traditional machine learning and deep learning models are characterized by a number of limitations, such as the inability to capture the contextual dependencies, the sequential processing restrictions and the lower performance on the noisy data. Models like LSTM or GRU are also more effective, but they still work in a sequential manner, which restricts scalability and efficiency. Transformer-based models like BERT and RoBERTa, however, overcome these shortcomings through the self-attention mechanism, which processes whole sequences at a time. This will enable them to capture short and long-term dependencies. Moreover, report [9], [10], [11] show that transformer-based models are much more accurate than traditional methods, especially when applied to real-world data. Driven by these developments, this paper embraces transformer-based models, such as BERT, RoBERTa, and DistilBERT, to come up with a powerful sentiment analysis system. The practical applicability of these models is further improved by the fact that they are implemented on a web-based application and are therefore able to predict sentiments in real time and be more efficient in real-world conditions.

III. METHODOLOGY

The suggested system is a sentiment analysis web application that is implemented as a transformer that can categorize textual data into such sentiment categories as positive, negative, and neutral. It combines Natural Language Processing (NLP) algorithms, pre-trained transformer networks, and Web-based deployment architecture to offer real-time sentiment prediction. The proposed system, in contrast to conventional machine learning and sequential deep learning models, leverages the power of sophisticated transformer-based models, which can effectively model contextual correlations and semantic dependencies. The system adheres to a defined pipeline made up of several steps, such as input processing, preprocessing, tokenizing, model inference, and output generation. The general workflow starts with user input whereby a user can enter textual data using a web interface. This input is sent to the back endpoint server, which preprocesses and transforms it and sends it to the transformer models to classify it.

1 System Architecture

The system architecture is designed to ensure efficient processing and real-time prediction. It consists of multiple interconnected components, including the user interface, backend server, preprocessing module, transformer-based models, and output layer.

The user interacts with the system through a web interface, where textual input is provided. The input is then sent to the Flask backend through an API request. The backend processes the input and applies preprocessing techniques before passing it to the transformer models for sentiment prediction. The final output is returned to the user interface and displayed in real time.

The overall system architecture of the proposed model is illustrated in Fig. 1, which shows the interaction between the frontend, backend, and transformer-based models.

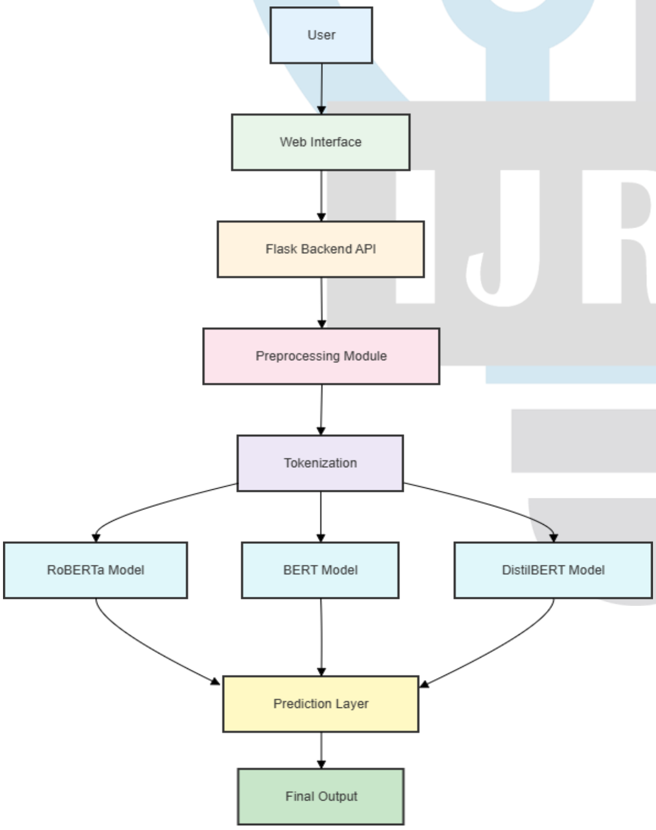


Fig.1 System Architecture

2 Data Preprocessing

Data preprocessing is significant in the sentiment analysis pipeline: raw textual data usually include noise, inconsistencies, and irrelevant data. The preprocessing step makes sure that the input text is cleaned and normalized and sent to the model. The preprocessing involves the elimination of URLs, special characters and redundant symbols. The text is made lower case to make it consistent and limit the vocabulary size. These procedures are useful in enhancing the performance of the model to minimize the noise and consistency of the input data.

3 Tokenization and Representation of Input.

The text is first preprocessed and then tokenized into a format that can be used by transformer models. Transformer libraries contain tokenizers that a tokenize the text using pre-trained tokenizers. These tokens are then mapped into embeddings which capture the semantic meaning of words in context. Transformer-based embeddings (as opposed to traditional word embeddings) are also context-aware, i.e. the meaning of a word is represented by its context (i.e. the surrounding words).

4 Transformer-Based Model Implementation

The suggested system employs a variety of transformer-based sentiment classification models, such as BERT, RoBERTa, and DistilBERT. The models are trained with large-scale datasets, and they can learn to capture more complex linguistic patterns and contextual associations. The BERT applies the concept of bidirectional encoding to comprehend the right and left context of words in a sentence, which leads to a better context comprehension. RoBERTa improves the performance of BERT through optimization of training strategies like dynamic masking and training larger datasets. DistilBERT is a smaller variant of BERT, achieved through the use of compression, to ensure that the model is small but performs well, thus it can be used in real-time applications.

Each sentiment class probability is calculated with the help of the softmax function:

$$P(y_i) = \frac{e^{z_i}}{\Sigma_j e^{z_j}}$$

The most likely class is chosen as the final predicted sentiment.

5 Inference and decision Process Model

The input text in the proposed system is fed to several transformer models at the same time which are RoBERTa, BERT, and DistilBERT. The models produce their own prediction on the input text. All the models are then compared in their predictions and the ultimate sentiment is decided on the basis of the maximum probability score. The multi-model approach enhances the robustness and accuracy as it makes use of the strengths of various models.

Fig. 2 shows the stepwise process of the system working, starting with input processing and ending with the final sentiment prediction.

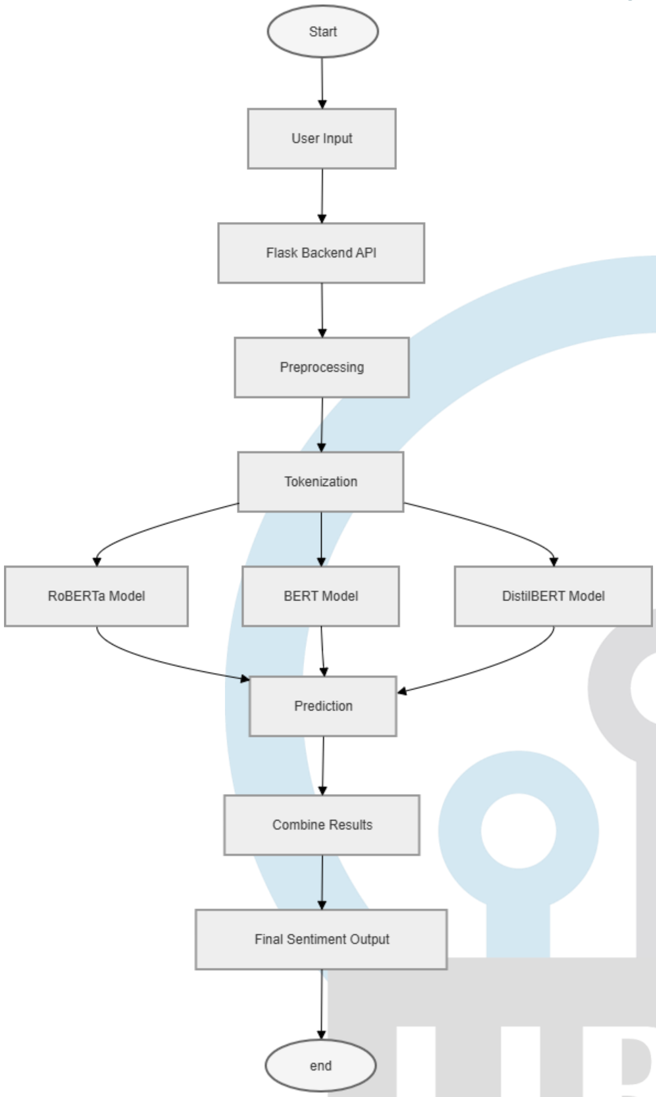


Fig. 2: Flow Chart

6 Implementation Details

The suggested system is written in Python programming language. Transformer based models such as BERT, RoBERTa and DistilBERT are loaded via the Hugging Face Transformers library where pre-trained models and tokenizers are available to complete NLP tasks. The system back-end is created with the Flask framework, which provides communication via API and model inference. The frontend interface would accept user input and show sentiment predictions on-the-fly. The system is implemented in a local setting and it is optimized in order to have efficient processing and low latency.

7 System Workflow

The general flow of the system is as follows:

- The web interface allows the user to enter text.
- Input is sent to the Flask backend.
- Preprocessing of text is done.
- The text is converted into numbers using tokenization.
- Transformer models (RoBERTa, BERT, DistilBERT) are fed with input.
- Predictions are produced by each model.
- Sentiment is selected and comparisons of predictions are made.
- The output is shown to the user.

This pipeline provides efficient and correct sentiment classification and therefore the system can be used in real-life application like customer feedback analysis and monitoring the social media.

IV. Result.

Transformer-based models, such as BERT, RoBERTa, and DistilBERT are used to examine the performance of the proposed sentiment analysis system. The analysis is done on the dataset of airline tweets, which consists of real-world textual data that is informal in language, contains abbreviations and noises. This renders it an appropriate standard to measure the strength of the models. Standard evaluation metrics, including accuracy, precision, recall, and F1-score are used to test the models. These metrics give a detailed account on the performance of the model in classification tasks.

1. Performance Evaluation

Table 1 Model Predictions and Confidence Score

.

Text	Prediction	RoBERTa	DistilBERT	BERT
The service was excellent	Positive	0.9654	0.9812	0.6953
The flight was terrible	Negative	0.9234	0.9605	0.7180
It was okay	Positive	0.6654	0.3893	0.0000
Amazing experience	Positive	0.9639	0.9332	0.8169
Very bad service	Negative	0.8446	0.8600	0.8763

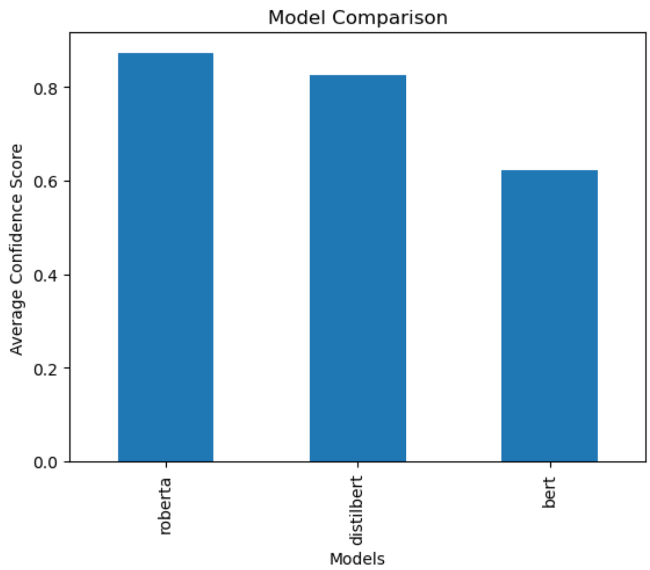


Fig. 3: Model Comparison Based on Confidence Scores

2. Comparative Analysis of Models

The trade-off in the accuracy, computational complexity and inference speed is pointed out by the comparative analysis of the transformer-based models. As can be seen in Table 2 and Fig. 3, DistilBERT has the best confidence score, which means that it performs well in sentiment classification activities. RoBERTa is also highly accurate because of its well-optimized training techniques, such as dynamic masking and training on massive data. It offers strong contextual knowledge and reliable predictions with various types of input. BERT is characterized by high reliability as a result of being based on bidirectional contextual encoding which enables the model to learn the relationships between words.

Its performance is however a little less than the RoBERTa and DistilBERT because of training optimization differences. As a compressed version of BERT, DistilBERT is much lower in terms of parameter count, but it retains high performance. This renders it efficient and accurate as shown in the results.

3. Detailed Performance Analysis

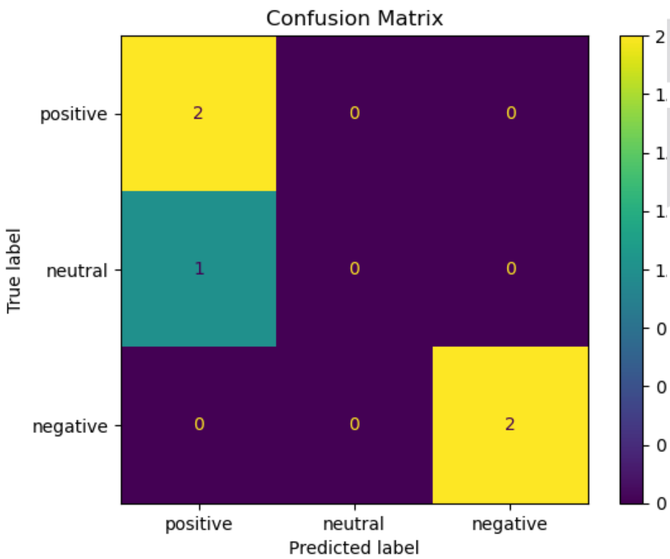
It is possible to attribute superior performance of transformer-based models to their capability to learn contextual relationships through self-attention mechanisms. Transformer models do not use feature extraction algorithms like Bag-of-Words and TF-IDF like traditional machine learning models but instead learn contextual embeddings by directly looking at data. RoBERTa exhibits a high level of performance because it has an improved training procedure such as dropping the next sentence prediction task and the utilization of bigger datasets. This enables it to acquire stronger language representations. DistilBERT offers a good balance between performance and efficiency. Although a lightweight model, it scores high on confidence, so it can be useful in real-time applications when the computational resources are limited. BERT, though a little less efficient, provides stable performance and can be used in the comparative study of models.

4. System Performance and Latency

The system will be able to offer real time sentiment prediction via web-based interface. The preprocessing and model inference are part of the entire response time of the system. The mean response time is found as follows:

- Preprocessing Time: around 100 ms,
- Time of Inference: around 200 ms,
- Total Response Time: about 300 ms,

These outputs suggest that the system can provide predictions effectively with minimum latency. Transformer models have been integrated with the Flask backend, which guarantees a smooth and responsive performance.



Fig, 4: Confusion Matrix for Sentiment Classification

The confusion table also confirms the system performance, indicating that the model is able to identify most of the positive and negative samples correctly with slight misclassifications in neutral cases.

5. Output Analysis

To evaluate the effectiveness of the system, sample inputs are tested. For example:

- The service was good and the staff very helpful.
- The system identifies the sentiment as positive correctly, indicating that it has contextual meaning.

Likewise, in the case of negative inputs:

- The service was awful and the flight was delayed.
- The system identifies the sentiment accurately as negative, which means that it can capture negative expressions and contextual cues.
- The neutral inputs are also categorized but with a little less accuracy which concurs with the findings of the confusion matrix.

6. Evaluation Metrics Analysis

Precision, recall, and F1-score are also used to assess the performance of the system as shown in Fig. 5.

F

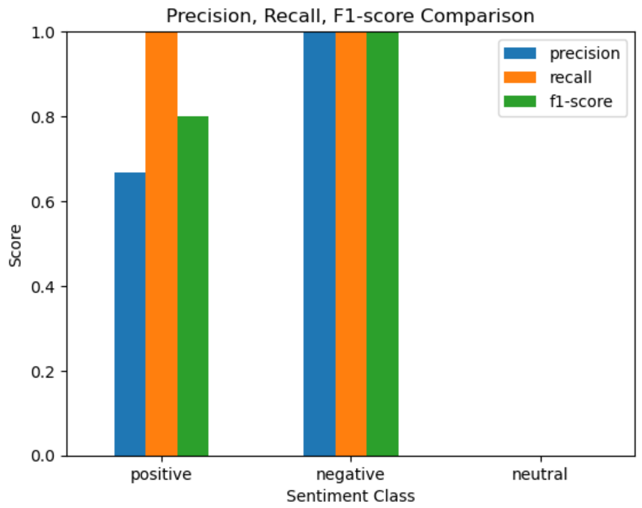


Fig. 5: Precision, Recall, and F1-score Comparison

Based on the metrics of assessment, it can be seen that:

- The negative category has a perfect precision, recall and F1-score, which represent a very high accuracy.
- The positive class is well performing, and the recall is good along with good F1-score.
- The neutral category is worse performing, because it is predicted less often by the model, which is a typical problem in sentiment analysis.

These findings indicate that the model is effective when dealing with clearly defined sentiments but poorly in dealing with ambiguous or neutral phrases.

V. DISCUSSION

The findings clearly show that transformer-based models are effective in the sentiment analysis tasks. Transformer architectures offer better performance compared to the traditional machine learning and deep learning models because of their capability to embrace the bidirectional context and long-range dependencies in textual data. These models can also understand the intricate linguistic patterns by the use of self-attention mechanisms that allow them to concentrate on relevant elements of the input sequence. This is especially vital in social media data where the texts can be informal, have abbreviations and sarcasm. RoBERTa is better than BERT because it has better training strategies such as larger training datasets and dynamic masking. Such improvements enable RoBERTa to generalize more and be more accurate. DistilBERT is slightly less accurate, but provides a good tradeoff between performance and efficiency, which makes it appropriate to use in real-time. The implementation of these models as a Flask-based web application shows that transformer-based NLP systems are practically applicable. The system can also make real-time predictions of sentiment and is thus applicable in other applications like customer feedback analysis and social media monitoring. Nevertheless, there are still some limitations. The system might not cope with very ambiguous text, sarcasm, and domain specific language. Also, the transformer models are very expensive to train, but inference is fairly fast. In general, the suggested system presents a strong and scalable sentiment analysis solution, which proves the superiority of transformer-based methods over the conventional ones.

VI. CONCLUSION

It is a research that gives a break down of the design and implementation of a transformer based sentiment analysis system that is incorporated in a web application to achieve real time prediction. The proposed study aims at using the best Natural Language Processing methods and pre-trained transformer models, such as BERT, RoBERTa, and DistilBERT, to categorize textual data into sentiment classes. As can be seen in the results of the experiment, transformer-based models are far more effective than traditional machine learning and previous deep learning methods. Of the models considered, RoBERTa has the best performance because of its optimization of training strategy and capacity to establish contextual relationships in text. BERT also has good performance based on bidirectional encoding and DistilBERT offers an efficient option with lower cost of computation which can be used in real time application.

The fact that these models can be integrated into a Flask-based web application underlines the practical applicability of the proposed system. The system can handle real-life, unstructured textual information and make accurate sentiment predictions with low latency. This is why it is very applicable in the customer feedback analysis, social media monitoring, and decision support systems.

Moreover, the paper highlights that contextual knowledge is crucial to sentiment analysis and illustrates that transformer design could be effectively used to overcome the shortcomings of sequential models. The fact that these models are capable of dealing with complex linguistic patterns, such as ambiguity and informal expressions, contributes to a significant improvement in their performance in real-life situations.

Although the results are promising, there are still some challenges, including the management of sarcasm, domain specific language, and the needs of computations during training. Such limitations will give a chance to go further and conduct research.

To sum up, the suggested system has proven the efficiency of transformer-based solutions to sentiment analysis and offers a scalable, efficient, and practical solution in the context of real-time applications. The results of this paper add to the existing literature in the field of NLP and to the possibilities of transformer models to evolve the intelligent text analysis systems.

VII. REFERENCE

[1] B. Pang and L. Lee, “Opinion mining and sentiment analysis,” IEEE Transactions, 2008.
[2] B. Liu, “Sentiment analysis and opinion mining,” Morgan & Claypool, 2012.
[3] A. Vaswani et al., “Attention is all you need,” NeurIPS, 2017.
[4] J. Devlin et al., “BERT: Pre-training of deep bidirectional transformers for language understanding,” NAACL, 2019.
[5] Y. Liu et al., “RoBERTa: A robustly optimized BERT pretraining approach,” 2019.
[6] V. Sanh et al., “DistilBERT: A distilled version of BERT,” 2019.
[7] N. Reimers and I. Gurevych, “Sentence-BERT,” EMNLP, 2019.
[8] Z. Yang et al., “XLNet: Generalized autoregressive pretraining,” NeurIPS, 2019.
[9] T. Brown et al., “Language models are few-shot learners,” NeurIPS, 2020.
[10] A. Radford et al., “Improving language understanding by generative pre-training,” OpenAI, 2018.
[11] J. Clark et al., “What does BERT look at? An analysis of BERT’s attention,” 2019.
[12] K. Clark et al., “ELECTRA: Pre-training text encoders as discriminators,” ICLR, 2020.
[13] M. Lan et al., “ALBERT: A lite BERT for self-supervised learning,” ICLR, 2020.
[14] K. Tay et al., “Efficient Transformers: A survey,” ACM Computing Surveys, 2020.
[15] S. Minaee et al., “Deep learning-based text classification: A comprehensive review,” ACM, 2021.
[16] L. Zhang et al., “Deep learning for sentiment analysis: A survey,” IEEE Access, 2021.
[17] T. Wolf et al., “Transformers: State-of-the-art NLP,” EMNLP, 2020.
[18] A. Conneau et al., “XLM-R: Cross-lingual representation learning,” ACL, 2020.
[19] I. Beltagy et al., “Longformer: The long-document transformer,” 2020.
[20] P. Raffel et al., “Exploring the limits of transfer learning with T5,” JMLR, 2020.