iJRASET
International Journal For Research in Applied Science and Engineering Technology

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

**Volume:** 11 **Issue:** X **Month of publication:** October 2023

**DOI:** https://doi.org/10.22214/ijraset.2023.56290

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue X Oct 2023- Available at www.ijraset.com*

# Social Media Sentiment Analysis

Dhruv Kudalkar[1], Tanya Mistry[2], Karan Thakkar[3], Prachi Satam[4]
*Department of Information Technology, Dwarkadas J. Sanghvi College of Engineering, Maharashtra, India*

***Abstract:*** ***In today's digital landscape, understanding and harnessing customer sentiment expressed across social media platforms is essential for the success of businesses, especially in the aviation and retail industries. This article delves into the profound significance of sentiment analysis in shaping strategic decisions and enhancing customer engagement in this digital age. Sentiment analysis emerges as a potent instrument for extracting invaluable insights from the vast reservoir of unstructured data that pervades social media platforms. We begin a thorough investigation of several social media sentiment analysis approaches in the context of this study. Both supervised and unsupervised procedures are included in our trip. In order to ensure a thorough knowledge of the applicability and subtleties of supervised approaches like Long Short-Term Memory (LSTM), which require labeled data for classification, we closely examine them. With SentiWordNet, we simultaneously explore the unsupervised domain, focusing on lexical sentiment analysis to sharpen the understanding of sentiment represented in the text. We provide a novel hybrid strategy that combines linear support vector machines (SVM) with latent semantic analysis (LSA). This combination delivers sentiment analysis that is more comprehensive and nuanced and is especially suited to the changing requirements of companies working in the digital era.***

***Keywords: Sentiment Analysis, Social Media, Long Short-Term Memory, SentiWordNet, Latent Semantic Analysis.***

## I. INTRODUCTION

Businesses are negotiating a complicated network of customer contacts, thoughts, and attitudes expressed on social media platforms in today's quickly expanding digital world. Understanding and exploiting client sentiment have emerged as critical success drivers in industries as diverse as aviation and retail, among the various difficulties and possibilities given by the digital age. This article will take you on a trip to discover the fundamental importance of sentiment analysis in defining strategic choices and raising consumer engagement to unparalleled heights.

We shed light on sentiment analysis' revolutionary potential as a strong tool for gleaning meaningful insights from the immense sea of unstructured data that pervades the arena of social media [2]. In the midst of this digital storm, sentiment analysis stands out as a bright spot and a strategic necessity for businesses. Sentiment analysis (or opinion mining) is the process of extracting sentiment, emotion, and opinion from text. It helps businesses navigate the ever-changing landscape of social media by providing valuable insights into consumer behavior, customer satisfaction, and new trends. One of the main attractions of sentiment analysis is its ability to handle the massive amount of unstructured social media data that is out there. From tweets to posts, reviews to comments, it’s a sea of information that’s way too big for the human brain to handle. Sentiment analysis algorithms can sort through that data, break it down into good, bad, or neutral sentiments, and measure the power of what’s being expressed [1]. With this analytical power, businesses can turn raw data into useful intelligence.

Sentiment analysis plays a multi-faceted role in the aviation and retail industries. For airlines, sentiment analysis provides a way to measure passenger satisfaction, identify crises, and improve customer service. In the retail industry, sentiment analysis drives product innovation, informs marketing strategies, and shapes brand perception. Sentiment analysis is not limited to the aviation or retail industries, however. It is used across a wide range of industries, from hospitality and retail to finance and healthcare [2].

This paper begins a thorough examination of social media sentiment analysis, considering a wide range of techniques and methodology. It examines both supervised approaches, such as Long Short-Term Memory (LSTM), which use labeled data for classification, and unsupervised methods, such as SentiWordNet, which focuses on lexical sentiment analysis. It also presents a hybrid technique that blends Latent Semantic Analysis (LSA) with Linear Support Vector Machines (SVM) to provide a more comprehensive view of sentiment analysis for organizations navigating the digital era [1][3]. The digital revolution has heralded a new era in which client emotion, as expressed and amplified on social media, holds the key to success for businesses of all sizes. This article invites readers to go on a voyage of discovery, navigating the dynamic environment of digital consumer engagement and utilizing the power of sentiment to design a brighter future for businesses in the digital age, via the lens of sentiment analysis.

©IJRASET: All Rights are Reserved 

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue X Oct 2023- Available at www.ijraset.com*

## II. LITERATURE SURVEY

We discovered several crucial elements that need to be taken into consideration after reading the evaluation papers of these suggested systems and browsing the material that is available on numerous platforms and research organizations.

### *A. Sentiment Analysis of Tweets Using Unsupervised Learning Techniques and the K-Means Algorithm*

The study's emphasis on classifying and analyzing content related to AFP affiliates on Twitter is in line with a growing trend in sentiment analysis, as well as social media mining. More and more literature emphasizes the need to understand user sentiments in various areas, including finance [1]. Studies in sentiment analysis have demonstrated the importance of using machine learning techniques to accurately measure public sentiment. As mentioned in the study above, classification and sentiment analysis are well-known methods in this field.

This allows for the detection of positive and negative sentiment expressed by users. Association rule learning provides insights into the relationship between words, which improves sentiment analysis accuracy [1]. When it comes to financial services, knowing user sentiments is essential for making informed decisions and managing risk. The study's integration of supervised learning to classify and regression, and unsupervised learning to cluster using methods like k-means and Gaussian mixture models reflects the variety of machine learning approaches used to analyze user sentiments. In addition, content analysis tools like word frequency calculation (WFM), lemmatization (LMT), and word cloud generation (WCG) are essential for finding trends and patterns within textual data [1].

### *B. Classification of Abusive Comments Using Various Machine Learning Algorithms*

The use of machine learning techniques to create prediction models to identify potentially offensive remarks is a huge step forward in online content management. Because of the exponential rise of user-generated material on websites and social media platforms, human moderation has become a challenging effort.

This has encouraged researchers to investigate automated techniques for efficiently filtering out rude or dangerous remarks. In this setting, machine learning algorithms have emerged as potent tools. To categorize and highlight potentially harmful content, text-based algorithms such as TFIDF (Term Frequency-Inverse Document Frequency), Logistic Regression, and Multinomial Naive Bayesian have been widely used. To detect abusive language or behavior, these algorithms use linguistic patterns and contextual information.

Furthermore, by pooling the results of numerous classifiers, ensemble approaches such as Random Forests and Gradient Boosting improve prediction accuracy [2]. The use of deep learning techniques, notably Long Short-Term Memory (LSTM) neural networks, has increased the capacity of content moderation systems to capture sophisticated linguistic patterns and contextual information. The research recognizes the importance of machine learning algorithms in speeding content moderation, decreasing human participation, and guaranteeing a safer online environment with high accuracy rates. The growth of these models highlights the rising relevance of AI and machine learning in solving the issues of online content management in the digital age [2].

### *C. Sentiment Analysis on User Reaction for Online Food Delivery Services using BERT Model*

The use of machine learning models, namely the BERT (Bidirectional Encoder Representations from Transformers) model, to analyze sentiment in Facebook comments on online food delivery services is a significant achievement in the field of customer feedback analysis. The gathering and categorization of comments into positive, negative, and neutral categories addresses the requirement for effective feedback processing [3].

The study's novel technique has numerous significant advantages. It helps food delivery service operators gain useful insights from consumer feedback, identifying particular words and areas that need to be addressed. Businesses may improve their operations and customer happiness by applying the results. The application of machine learning models, namely the BERT (Bidirectional Encoder Representations from Transformers) model, to evaluate sentiment in Facebook comments on online food delivery services is a significant accomplishment in the field of consumer feedback analysis [3]. The necessity for successful feedback processing is met by accumulating and categorizing comments into positive, negative, and neutral categories [3]. The unique approach used in the study provides a number of major advantages. It assists meal delivery service providers in gaining important insights from client feedback, pinpointing specific words and regions that require attention. Businesses that use the data can enhance their operations and customer satisfaction.

©IJRASET: All Rights are Reserved 

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue X Oct 2023- Available at www.ijraset.com*

## III.SENTIMENT ANALYSIS ALGORITHMS

### *A. Supervised Learning*

#### *1) LSTM(Long Short-Term Memory)*

The LSTM (Long Short Term Memory) is a type of recurrent neural network (RNN) that is well-suited for learning long-term dependencies in time series data. Recurrent neural networks (RNNs) traditionally struggled to model long-term dependencies in sequential data due to the vanishing gradient problem, where gradients exponentially shrink as they are back-propagated over many time steps. To address this limitation, Hochreiter and Schmidhuber (1997) invented the Long Short Term Memory (LSTM) architecture. LSTM networks contain specialized memory cells with explicit gating mechanisms that can learn long-term dependencies in time series data. The LSTM cell state acts as a memory to store relevant information from potentially very long input sequences. Input, output, and forget gates then regulate the information flow into and out of the cell state appropriately. By maintaining a cell state that can preserve information over long durations and applying gating functions, LSTMs can overcome the vanishing gradient problem of plain RNNs. This makes LSTM networks extremely well-suited for learning long-term temporal dynamics, hence their widespread adoption for sentiment analysis, speech recognition, language translation, time series forecasting, and other sequence modeling tasks. Specifically for sentiment analysis, LSTM's ability to understand context, remember long-range dependencies, and aggregate information over sequences allows it to effectively analyze sentiment in textual data.

### *B. Unsupervised Learning*

#### *1) SentiWordNet*

The SentiWordNet is a well-known unsupervised method for sentiment analysis. In order to determine the overall sentiment of a text, it works by giving sentiment scores to specific phrases inside the text. SentiWordNet links sentiment scores to words using the WordNet lexical database. It makes the assumption that words with similar meanings frequently express opposite but related emotions. The program uses WordNet integration to map words to the appropriate synsets (groups of synonyms with a common meaning). Each synset is associated with sentiment scores, which normally range from 0 (neutral) to 1 (strongly positive) or -1 (strongly negative) and include positivity, negativity, and neutrality.SentiWordNet assesses specific words in a text using practical sentiment analysis by pulling the related sentiment ratings from the lexical resource. These word-level ratings are used to determine a sentence's or a document's overall emotion. This normally entails calculating a net sentiment score by adding the positivity and negativity numbers. Based on the net sentiment score, SentiWordNet enables researchers to specify precise thresholds for classifying sentiment as either positive, negative, or neutral. For instance, the sentiment might be classified as positive if the net score exceeds a predetermined threshold; on the other hand, it might be classified as negative if the net score is below the level.

### *C. Hybrid Learning*

#### *1) Latent Semantic Analysis and Linear SVM*

In sentiment analysis, hybrid learning combines the advantages of various techniques to increase the precision and resilience of sentiment categorization. Latent Semantic Analysis (LSA), a lexicon-based technique, and Linear Support Vector Machines(SVM), a machine learning algorithm, are two examples of hybrid approaches. For more accurate sentiment analysis, this hybrid model is created to collect both semantic and contextual information. LSA is used in this hybrid framework as a starting point. LSA works by identifying latent semantic links between words by examining the word co-occurrence patterns in a huge corpus of text. Words with related meanings are grouped together in a semantic space that is created as a result. By giving words numerical vectors depending on their semantic contexts, LSA successfully encodes the semantic meaning of those words. With the use of this lexicon-based methodology, it is possible to detect sentimental nuances that would otherwise go undetected. After LSA has converted the text into a numerical format that has been semantically enhanced, Linear SVM is used to classify the sentiment. An effective machine learning approach for classifying data is called linear SVM. In this instance, it uses the numerical text representations produced by LSA to train itself to recognize positive, negative, and neutral moods. The model is capable of processing complex, high-dimensional data because it learns the ideal decision boundary that maximizes the margin between various sentiment classifications. This hybrid approach's main benefit is that it combines the context-awareness and interpretability of LSA with the prediction strength of Linear SVM. SVM offers reliable sentiment categorization, while LSA catches the subtleties of language and context.
The model can handle a variety of texts thanks to this combination, including ones with delicate sentiment expressions and varied contextual clues. As a result, the hybrid model can produce findings for sentiment analysis that are more accurate and trustworthy, making it a useful tool in applications like social media monitoring, product evaluations, and analysis of consumer feedback.

©IJRASET: All Rights are Reserved 

*ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538*
*Volume 11 Issue X Oct 2023- Available at www.ijraset.com*

## IV. DATASET DESCRIPTION

The sentiment analysis implementation utilizes two publicly available datasets that contain customer feedback towards companies. The first dataset consists of tweets regarding consumer experiences with Virgin America Airlines flights. The second dataset contains Amazon customer reviews of various phone brands such as Samsung, Apple, Nokia, and Alcatel. These datasets enable training of supervised, unsupervised, and hybrid machine learning models approaches to perform sentiment analysis. The data is split into training and validation sets in a 5:1 ratio to evaluate model performance. Using real-world customer feedback data allows the models to learn nuanced sentiment and generalize to practical applications.

## V. IMPLEMENTATION AND METHODOLOGY USED

### *A. Supervised Learning(LSTM)*

In this portion of the research article, we demonstrate a deep learning implementation of sentiment analysis that makes use of a Long Short-Term Memory (LSTM) neural network. With tweets about airlines as the sample dataset, the goal is to predict sentiment labels for text data. Data preparation is the first step of the procedure. We open a CSV file containing a dataset and choose the necessary columns, text and airline sentiment, which include the tweet text and associated sentiment labels. Upon closer inspection, we find that certain tweets are classified as "neutral," which lacks sentiment polarity. In order to get a dataset containing non-neutral attitudes, we filter away these neutral tweets. After factorization, the sentiment labels are transformed into numerical values that can be used as input for machine learning models.

The captured tweet text is used for additional processing. Tokenization is then used to turn the text input into number sequences. We employ the Keras Tokenizer with a 5,000 word vocabulary limit. This tokenizer creates a mapping between words and number indices after being fitted to the twitter data. The tweet text is encoded into integer sequences according to the vocabulary's length. We use the Keras pad sequences function to pad the input sequences to make sure they are all the same length. This process makes sure that every input sequence has a fixed length of 200, which is necessary for deep learning model training. The sequential neural network architecture is used to build the sentiment analysis model. We start with an embedding layer, which trains continuous vector space word representations. Understanding the semantic meaning of terms in the tweet text requires this layer. The use of spatial dropout prevents overfitting. The embedding layer is followed by the 50-unit LSTM layer. LSTM is a specific recurrent neural network design that excels at effectively capturing sequential data. It is suitable for sentiment analysis and other natural language processing tasks.

To further regularize the model and lessen overfitting, a dropout layer is used. Finally, we incorporate a dense layer with a sigmoid activation function and a single output unit. With a likelihood score for a good sentiment, this output unit acts as the sentiment prediction. The Adam optimizer, a popular optimization tool, and binary cross-entropy loss are employed in the model's construction. The measurement metric for evaluation is accuracy. Fitting the model to the padded sequences and associated sentiment labels is the process of training the model. The model's performance is evaluated along with the training process. The training history is saved, together with accuracy and loss data, and is displayed using matplotlib. In conclusion, this section describes how LSTM was used to create a deep learning-based sentiment analysis model. It demonstrates how deep learning techniques may be used to analyze sentiment in text data successfully and covers data pretreatment, model building, training, evaluation, and prediction.

### *B. Unsupervised Learning(SentiWordNet)*

In this section of the research article, we provide a sentiment analysis solution that makes use of the lexicon-based SentiWordNet dictionary. The goal is to forecast sentiment labels for text data and assess the performance of the model using a variety of measures, such as accuracy, precision, recall, F1-score, confusion matrix, and classification report. To make model evaluation easier, the dataset is split into a 30 percent validation/test set and a 70 percent training set. First, we import the necessary libraries for supporting data processing, model evaluation, and visualization, including NumPy, Pandas, scikit-learn (sklearn), and Matplotlib.

We also import special utility functions from "utils," which most likely contain crucial functions for lexicon-based sentiment analysis. The dataset is then loaded from a CSV file including Amazon mobile review data. The dataset has columns for "review.text" and "sentiment," among others. After preprocessing, which entails deleting stop words and lemmatizing the text, the'review.tokened' column is used for sentiment analysis. Using scikit-learn's train-test split function, the data is then divided into a training set(70 percent) and a validation/test set (30 percent).

©IJRASET: All Rights are Reserved 

This stage makes sure that we have labeled data for our sentiment analysis model's training as well as data for evaluating its effectiveness. We repeatedly go through the validation/test set while applying the "sentiment" function to each review in order to predict sentiment using the SentiWordNet vocabulary. SentiWordNet is probably used by the sentiment function to determine the sentiment score for each review and to apply a sentiment label based on a predetermined threshold. A NumPy array is created by compiling the predictions. We use label encoding to translate sentiment labels into numerical values to speed up evaluation. Given that the majority of machine learning algorithms require numerical input, this stage is essential. For this, the Scikit-Learn LabelEncoder is employed. Metrics for evaluating performance are computed and shown. Accuracy, precision, recall, and F1-score are some of these measurements.

To facilitate comparison between various models or techniques and easy comprehension of the results, they are given in a tabular manner. We also use bar charts to illustrate the performance indicators, giving a visual snapshot of how well the model categorizes sentiment. Additionally, a confusion matrix is created and shown to evaluate the model's aptitude for sentiment classification. For each sentiment class, the confusion matrix offers insights into true positive, true negative, false positive, and false negative predictions.

We also provide a categorization report that provides comprehensive details on the precision, recall, and F1-score for each sentiment class. This report makes it easier to comprehend the model's performance on a class-by-class basis. This section concludes by demonstrating how to use the SentiWordNet vocabulary to construct a sentiment analysis approach. Data preprocessing, sentiment analysis, model evaluation, visualizations, and thorough reporting are all covered. This extensive framework enables a full evaluation of the lexicon-based sentiment analysis method.

### *C. Hybrid Approach(Latent Semantic Analysis and Linear SVM)*

In this section of the research paper, we describe a hybrid approach to sentiment analysis, which performs sentiment classification using both lexicon-based techniques and machine learning techniques. The main goal of this strategy is to maximize the benefits of both approaches for more accurate sentiment analysis. First, we load the necessary modules for supporting data processing, model evaluation, and visualization, including NumPy, Pandas, scikit-learn (sklearn), and Matplotlib. Similar to earlier parts, we now load the dataset from a CSV file and get it ready for analysis.

To make model evaluation easier, the data is then divided into a training set (70 percent) and a testing set (30 percent). This division makes sure we have labeled data for the hybrid sentiment analysis approach's training and testing phases. A critical phase in the hybrid technique is feature extraction. Starting with the training and test data, the TF-IDF vectorization technique is used. Inverse document frequency (IDF), n-gram range, and stop words removal are some of the configuration options for the TF-IDF vectorizer. The TF-IDF features that have been converted are then prepared for processing. Latent Semantic Analysis (LSA) used on the TF-IDF characteristics is also a component of the hybrid technique.

LSA is a machine learning method that lowers the data's dimensionality while retaining semantic data. This stage entails building a pipeline with normalization and truncated singular value decomposition (TruncatedSVD). Following LSA, sentiment analysis is performed using the obtained LSA features.

A Linear Support Vector Machine (Linear SVC) classifier is trained using the LSA-transformed training data to accomplish sentiment categorization. Following that, sentiment labels for the LSA-transformed testing data are predicted using the trained model. Calculating important performance metrics, such as accuracy, precision, recall, and F1-score, is a part of evaluating the hybrid technique.

These measures shed light on how well the hybrid model performs sentiment classification. Bar charts are also used to visually depict the performance data for simple comparison and interpretation. Understanding how the hybrid strategy works in comparison to other models or approaches is made easier with the help of this visualization. A confusion matrix is created and displayed to provide further depth on the performance of the model. For each sentiment class, the confusion matrix lists the true positive, true negative, false positive, and false negative forecasts.

The evaluation also includes a thorough classification report with class-specific precision, recall, and F1-score values. The development of a hybrid approach for sentiment analysis, integrating lexicon-based and machine learning techniques, is summarized in this portion of the research report. Model training, feature extraction, reporting, assessment metrics, and data pretreatment are all covered. By combining the advantages of the two techniques, this hybrid methodology seeks to increase the accuracy of sentiment analysis.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

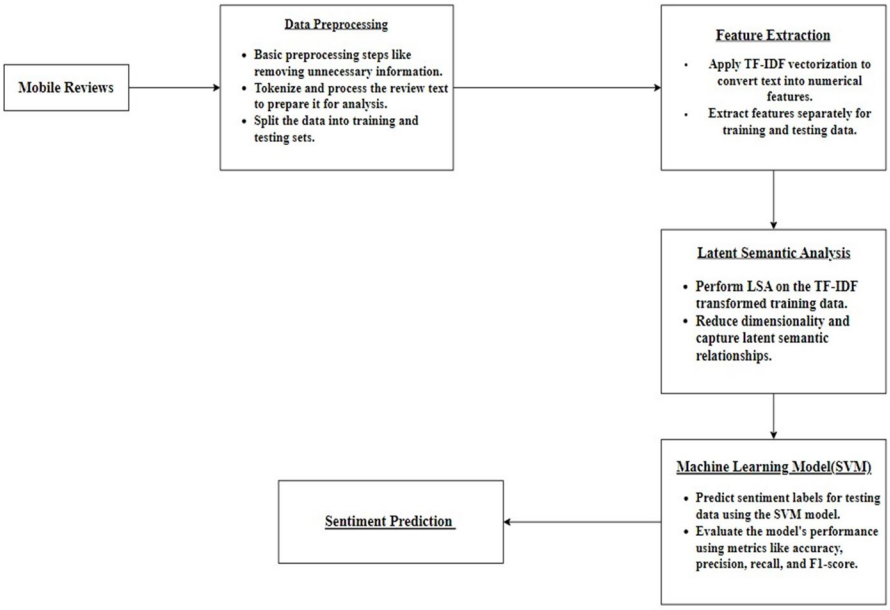



Figure 1 : System Architecture For The Hybrid Approach

## VI. EVALUATION PARAMETERS

### A. *Precision*

Precision measures the proportion of positive predictions made by a model that are actually correct. It is calculated as the number of true positives (Truepos) divided by the sum of true positives and false positives (Falsepos). A true positive refers to when the model accurately predicts the positive class. A false positive is when the model incorrectly predicts the positive class.

Precision = Truepos/ (Truepos + Falsepos)

### B. *Recall*

Recall is a metric that shows how good a model is at finding all the positive cases. It is calculated by taking the number of true positives (Truepos), which are cases correctly identified as positive, and dividing by the total number of actual positive cases. A high recall means the model is accurately labeling most of the positive examples. Recall is related to false negatives (Falseneg), which are cases the model incorrectly predicts as negative when they are actually positive.

Recall = Truepos/ (Truepos + Falseneg)

### C. *F1-Score*

F1 score is a metric that combines precision and recall to measure a model's accuracy. It takes the harmonic mean of precision and recall, giving equal weight to both metrics. The F1 score ranges from 0 to 1, with 1 being perfect precision and recall. F1 provides a single metric for model performance that balances precision and recall.

F1 score = 2 * (precision * recall) / (precision + recall)

## VII. PERFORMANCE ANALYSIS

In the performance analysis section, a comprehensive and in-depth study of the system's capabilities is performed. These results highlight the ability of the proposed system to successfully classify social media sentiments.

### A. *Supervised Analysis*

Fig 2 shows how the supervised learning model correctly classifies airline reviews into positive and negative sentiments. Fig 3 shows how the model correctly distinguishes between mobile review sentiments.

```
Review:  The plane was old, dirty, and cramped, making for an extremely uncomfortable 6-hour flight.
1/1 [==============================] - 0s 32ms/step
Predicted label:  negative
Review:  The flight attendants were very friendly and helpful!
1/1 [==============================] - 0s 30ms/step
Predicted label:  positive
```

Figure 2: Sentiments Correctly Recognised By The Supervised Learning Model

```
Review:  I couldn't hear the caller properly because the earpiece volume was too low.
1/1 [==============================] - 0s 29ms/step
Predicted label:  negative
Review:  The camera captures amazing landscape and portrait shots with excellent color accuracy
1/1 [==============================] - 0s 34ms/step
Predicted label:  positive
```

Figure 3: Sentiments Correctly Recognised By The Supervised Learning Model

Fig 4 and Fig 5 show the confusion matrices for the two datasets respectively. This shows how accurate the model is in identifying sentiments properly.

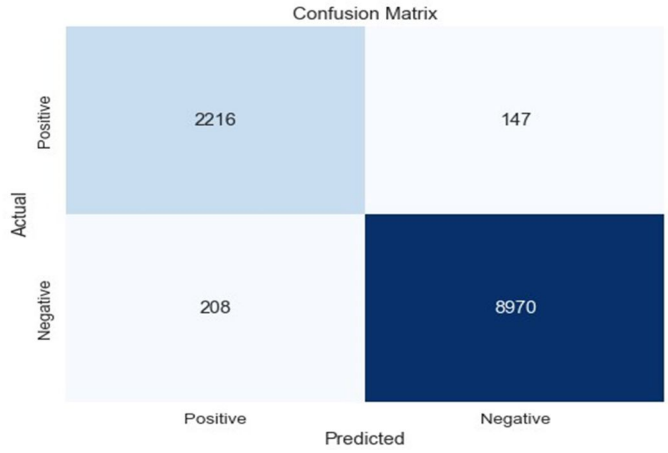



Figure 4: Confusion Matrix For Airline Reviews Dataset

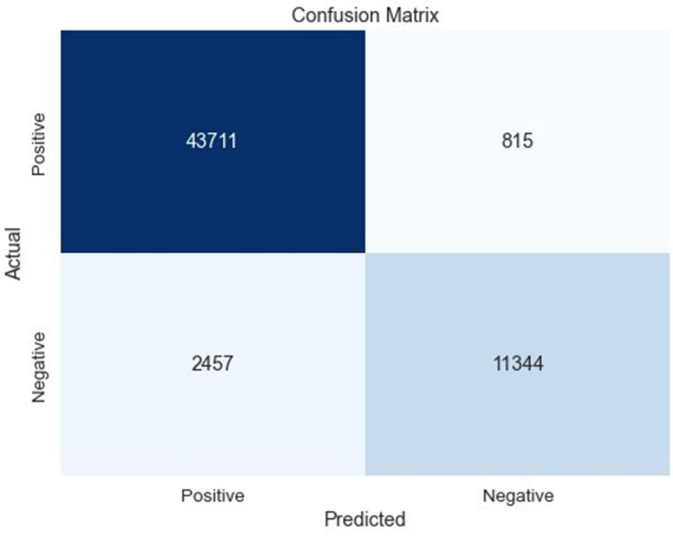



Figure 5: Confusion Matrix For Mobile Reviews Dataset

Table I : Precision, Recall and F1-Score for positive and negative sentiments(Airline Reviews Dataset)

| | Precision | Recall | F1-Score |
|---|---|---|---|
| Positive | 0.91 | 0.94 | 0.93 |
| Negative | 0.98 | 0.98 | 0.98 |
| Accuracy | | | 0.97 |
| Macro Avg | 0.95 | 0.96 | 0.95 |
| Weighted Avg | 0.97 | 0.97 | 0.97 |

Table II : Precision, Recall and F1-Score for positive and negative sentiments(Mobile Reviews Dataset)

| | Precision | Recall | F1-Score |
|---|---|---|---|
| Positive | 0.95 | 0.98 | 0.96 |
| Negative | 0.93 | 0.92 | 0.87 |
| Accuracy | | | 0.94 |
| Macro Avg | 0.94 | 0.90 | 0.92 |
| Weighted Avg | 0.94 | 0.94 | 0.94 |

*B. Unsupervised Analysis*

*1) Mobile Reviews Dataset*

Using the classification report and confusion matrix, we give a thorough analysis of the findings in order to assess the efficacy of our classification model. Our approach does remarkably well at differentiating between cases that are positive and those that are negative. The categorization report offers information on F1-score, precision, and recall for both positive and negative classes. 97 percent of the examples labeled as negative were, in fact, negative, according to our outstanding precision of 0.97 for negative class predictions. The model's capacity to precisely identify positive cases was demonstrated by the precision for the positive class, which was even greater at 0.98. Excellent recall rates of 97 percent for negative and 98 percent for positive occurrences were achieved for both classes. The F1-score, which combines precision and recall, was consistently high, at 0.97 for the negative class and 0.98 for the positive class. This demonstrates the model's capacity to maintain a strong balance between precision and recall.

Table III : Precision, Recall and F1-Score for positive and negative sentiments(Mobile Reviews Dataset)

| | Precision | Recall | F1-Score |
|---|---|---|---|
| Positive | 0.97 | 0.97 | 0.97 |
| Negative | 0.98 | 0.98 | 0.98 |
| Accuracy | | | 0.98 |
| Macro Avg | 0.98 | 0.98 | 0.98 |
| Weighted Avg | 0.98 | 0.98 | 0.98 |

Additionally, both the macro and weighted average F1-scores—which offer a comprehensive evaluation of the model's performance—were 0.98, reiterating the accuracy of the predictions produced for all classes. This demonstrates the model's capacity to effectively generalise to the entire dataset. We looked at the confusion matrix to have a better grasp of the predictions made by the model. It demonstrates that the algorithm correctly identified 12,038 of the actual positive events as positive, but that 205 incorrect negative predictions were made. In contrast, the model accurately predicted 6,934 of the actual negative instances to be negative, while making 236 incorrect positive predictions. These numbers show how sensitive to both classes the model is, as well as how it tends to minimise false positives and false negatives. The outcomes from our algorithm show that it performs well at correctly categorizing sentiment.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

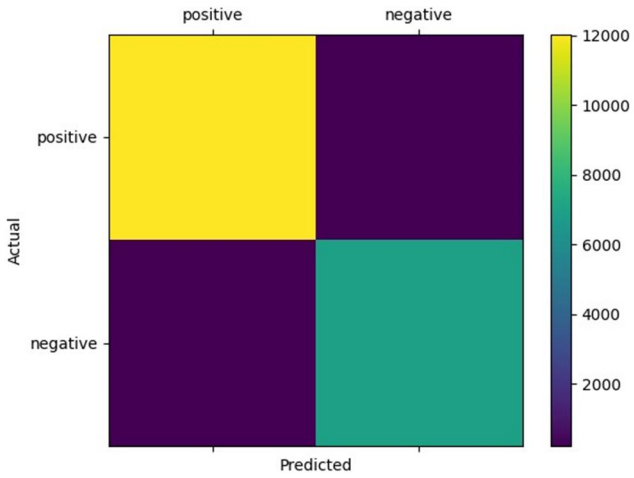



Figure 6: Confusion Matrix For Mobile Reviews Dataset

```
Review: The worst phone used ever
Predicted Sentiment: negative
Review: This is the phone with the best features
Predicted Sentiment: positive
```

Figure 7: Sentiments Correctly Predicted By The Unsupervised Learning Model

*2) Airline Reviews Dataset*

We outline the performance metrics of the predictions made by our algorithm in the categorization report. The precision, recall, and F1-score are important measures of how well the model categorizes sentiments. The precision for the 'negative' class is 0.86, which means that 86 percent of all cases expected to be 'negative' were properly classified. The precision for the 'positive' class is also 0.86, which indicates that 86 percent of the instances correctly classified as such were expected. Recall for the 'negative' class is 0.86, meaning that 86 percent of all real 'negative' instances were correctly identified by the model. Similar to the 'negative' class, the recall for the 'positive' class is 0.87, indicating that 87 percent of all real 'positive' instances were successfully captured by the model. Precision and recall are balanced by the standardized statistic known as the F1-score.

Table IV : Precision, Recall and F1-Score for positive and negative sentiments(Airline Reviews Dataset)

| | Precision | Recall | F1-Score |
|---|---|---|---|
| Positive | 0.86 | 0.87 | 0.86 |
| Negative | 0.86 | 0.86 | 0.86 |
| Accuracy | | | 0.86 |
| Macro Avg | 0.86 | 0.86 | 0.86 |
| Weighted Avg | 0.86 | 0.86 | 0.86 |

The F1-score for both classes is 0.86, showing a fair balance between recall and precision. With an overall accuracy of 0.86, the model correctly categorized 86 percent of all occurrences. Taking into account the intricacy of sentiment analysis, this is a reasonable outcome. According to the confusion matrix, the model accurately predicted 1932 actual "positive" instances and 299 actual "negative" instances. The model correctly identified 1856 of the real "negative" occurrences while incorrectly classifying 305 as "positive". Together, these findings show how reliable the model is in correctly categorizing sentiments, with balanced performance for both positive and negative sentiments.

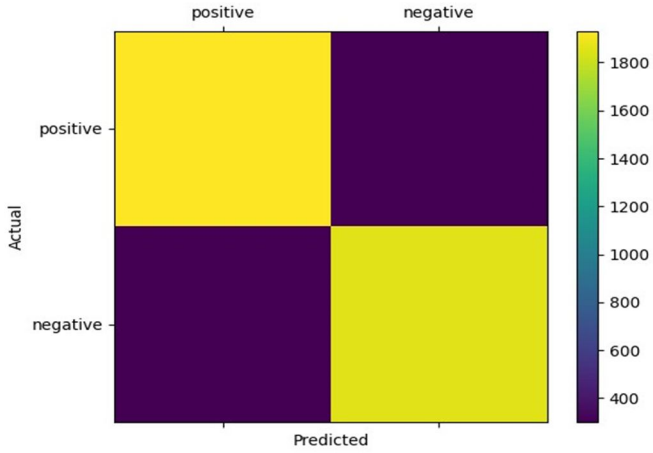



Figure 8: Confusion Matrix For Airline Reviews Dataset

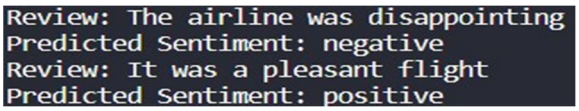



Figure 9: Sentiments Correctly Predicted By The Unsupervised Learning Model

*C. Hybrid Analysis*

*1) Mobile Reviews Dataset*

We assess the effectiveness of our sentiment classification model in the provided classification report. The precision, recall, and F1-score metrics in the report are essential for evaluating the model's performance. The precision for the 'negative' class is 0.73, meaning that 73 percent of all cases expected to be in the 'negative' category were properly identified. Recall for "negative" is 0.69, indicating that 69 percent of all real instances of "negative" were correctly detected by the model. The F1-score for 'negative' emotion is 0.71, which is a balanced indicator of precision and recall. The precision for the 'positive' class is 0.91, which indicates that 91 percent of the instances correctly classified as such were expected.

Table V : Precision, Recall and F1-Score for positive and negative sentiments(Mobile Reviews Dataset)

| | Precision | Recall | F1-Score |
|---|---|---|---|
| Positive | 0.91 | 0.92 | 0.91 |
| Negative | 0.73 | 0.69 | 0.71 |
| Accuracy | | | 0.87 |
| Macro Avg | 0.82 | 0.81 | 0.81 |
| Weighted Avg | 0.87 | 0.87 | 0.87 |

Additionally, the recall for "positive" sentiment is 0.92, meaning that 92 percent of all genuine "positive" instances were recorded by the model. 'Positive' sentiment has an F1-score of 0.91, which indicates high accuracy. The model's overall accuracy is 0.87, which indicates that 87 percent of all cases were properly identified. This high level of accuracy highlights the model's prowess in sentiment analysis. The precision, recall, and F1-score for the macro average are 0.82, 0.81, and 0.81, respectively, and give an overview of the model's performance for both classes. The model performed well when taking into account potential class imbalances; the weighted average precision, recall, and F1-score are all 0.87. Regarding the confusion matrix, the model accurately predicted 12,300 actual "positive" cases while predicting 1,049 actual "negative" ones. The model successfully identified 2,883 of the real "negative" occurrences while misclassifying 1,267 as "positive". Collectively, these findings demonstrate the model's effectiveness in correctly categorizing sentiments, with particularly excellent performance in recognizing 'positive' attitudes. The model also shows success in detecting 'negative' emotions, but there is still potential for development as evidenced by the class's somewhat lower recall value.

The outcomes from our algorithm show that it performs well at correctly categorizing sentiment.

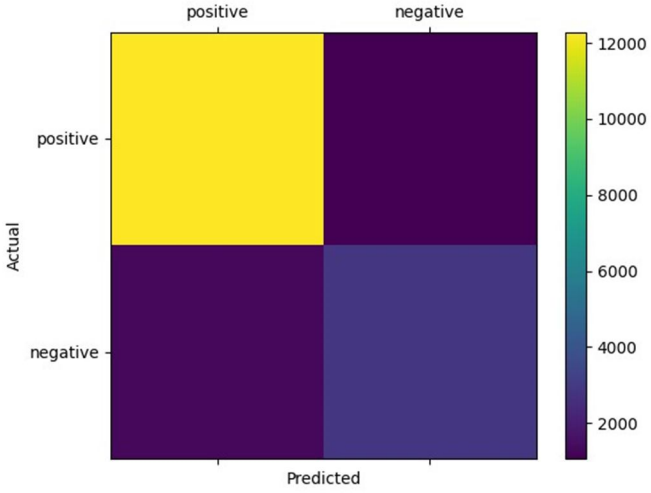



Figure 10: Confusion Matrix For Mobile Reviews Dataset

```
Review The screen was cracked and the touch was not working properly
Predicted sentiment for test sentence: negative
Review The phone had amazing features and its camera was awesome!
Predicted sentiment for test sentence: positive
```

Figure 11: Sentiments Correctly Predicted By The Hybrid Learning Model

*2) Airline Reviews Dataset*

We evaluate the effectiveness of our sentiment classification model in the provided classification report. In order to assess the model's efficacy, the report provides the precision, recall, and F1-score measures. The precision for the 'negative' class is 0.91, meaning that 91 percent of the cases expected to be in the 'negative' category were properly identified. Recall for "negative" is 0.98, indicating that 98 percent of all real instances of "negative" were correctly detected by the model. The F1-score for 'negative' sentiment is 0.94, indicating a balanced measure of precision and recall. The precision for the 'positive' class is 0.87, which indicates that 87 percent of the cases correctly classified as such were present. However, the recall for "positive" sentiment is 0.60, meaning that only 60 percent of all actual "positive" instances were picked up by the model. The 'positive' sentiment F1-score is 0.71. The model's overall accuracy is 0.90, which means that 90 percent of all occurrences were properly identified. This high level of accuracy demonstrates the model's efficiency in sentiment analysis. The precision, recall, and F1-score for the macro average are 0.89, 0.79, and 0.83, respectively, giving an overview of the model's performance for both classes. The precision for the 'positive' class is 0.87, which indicates that 87 percent of the cases correctly classified as such were present. However, the recall for "positive" sentiment is 0.60, meaning that only 60 percent of all actual "positive" instances were picked up by the model. The 'positive' sentiment F1-score is 0.71. The model's overall accuracy is 0.90, which means that 90 percent of all occurrences were properly identified. This high level of accuracy demonstrates the model's efficiency in sentiment analysis. The precision, recall, and F1-score for the macro average are 0.89, 0.79, and 0.83, respectively, giving an overview of the model's performance for both classes.
The outcomes from our algorithm show that it performs well at correctly categorizing sentiment.

Table VI : Precision, Recall and F1-Score for positive and negative sentiments(Airline Reviews Dataset)

| | Precision | Recall | F1-Score |
|---|---|---|---|
| Positive | 0.87 | 0.60 | 0.71 |
| Negative | 0.91 | 0.98 | 0.94 |
| Accuracy | | | 0.90 |
| Macro Avg | 0.89 | 0.79 | 0.83 |
| Weighted Avg | 0.90 | 0.90 | 0.90 |

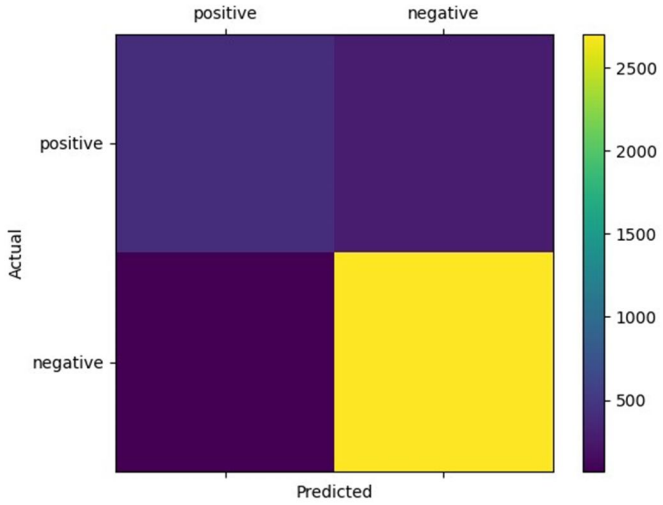



Figure 12: Confusion Matrix For Airline Reviews Dataset

```
Review It was an unpleasant experience, the cabin crew were not helpful
Predicted sentiment for test sentence: negative
Review Their business class is awesome! Totally worth it
Predicted sentiment for test sentence: positive
```

Figure 13: Sentiments Correctly Predicted By The Hybrid Learning Model

## VIII. CHALLENGES FACED

There are many difficult tasks involved in conducting sentiment analysis, whether using lexicon-based methods, machine learning strategies, or hybrid approaches. The lexical coverage of sentiment analysis systems is one of the main problems. These lexicons are made up of predefined terms with sentiment scores that may be missing expressions with specific subject knowledge or complex context. The correct use of negations and intensifiers in text is another problem because words like "not good" or "very bad" can change or intensify sentiment, making it difficult for right interpretation. Complexity is further increased by the use of ambiguous language and sarcasm in text data. For sentiment analysis techniques, detecting sarcasm, which transmits emotions in opposition to the literal text, can be particularly difficult. The performance of the model may be impacted by data imbalance, when one sentiment class predominates over another. Sentiment analysis methods frequently lack contextual awareness; as a result, they treat each text independently and may fail to recognize how sentiment is influenced by a larger context. Additional difficulties include ensuring model generality, particularly in domains with limited data, and improving interpretability. Last but not least, optimizing model hyperparameters for performance requires a lot of work. For sentiment analysis to improve in terms of accuracy and dependability, these issues must be resolved.

## IX. CONCLUSION

This study has investigated a variety of sentiment analysis methods, including lexicon-based, machine learning, and hybrid approaches. Each approach had particular advantages and disadvantages. Lexicon-based methods provide clarity and interpretability but struggle with complexities of context and language peculiar to a certain area. Machine learning methods, especially LSTM-based models, demonstrate a strong capacity to recognize intricate patterns in text data. However, they need a lot of labeled data for training, and they could have trouble efficiently processing negations and sarcasm. The hybrid technique aims to capitalize on the advantages of both paradigms, improving overall sentiment analysis performance. It combines lexicon-based and machine learning methodologies.

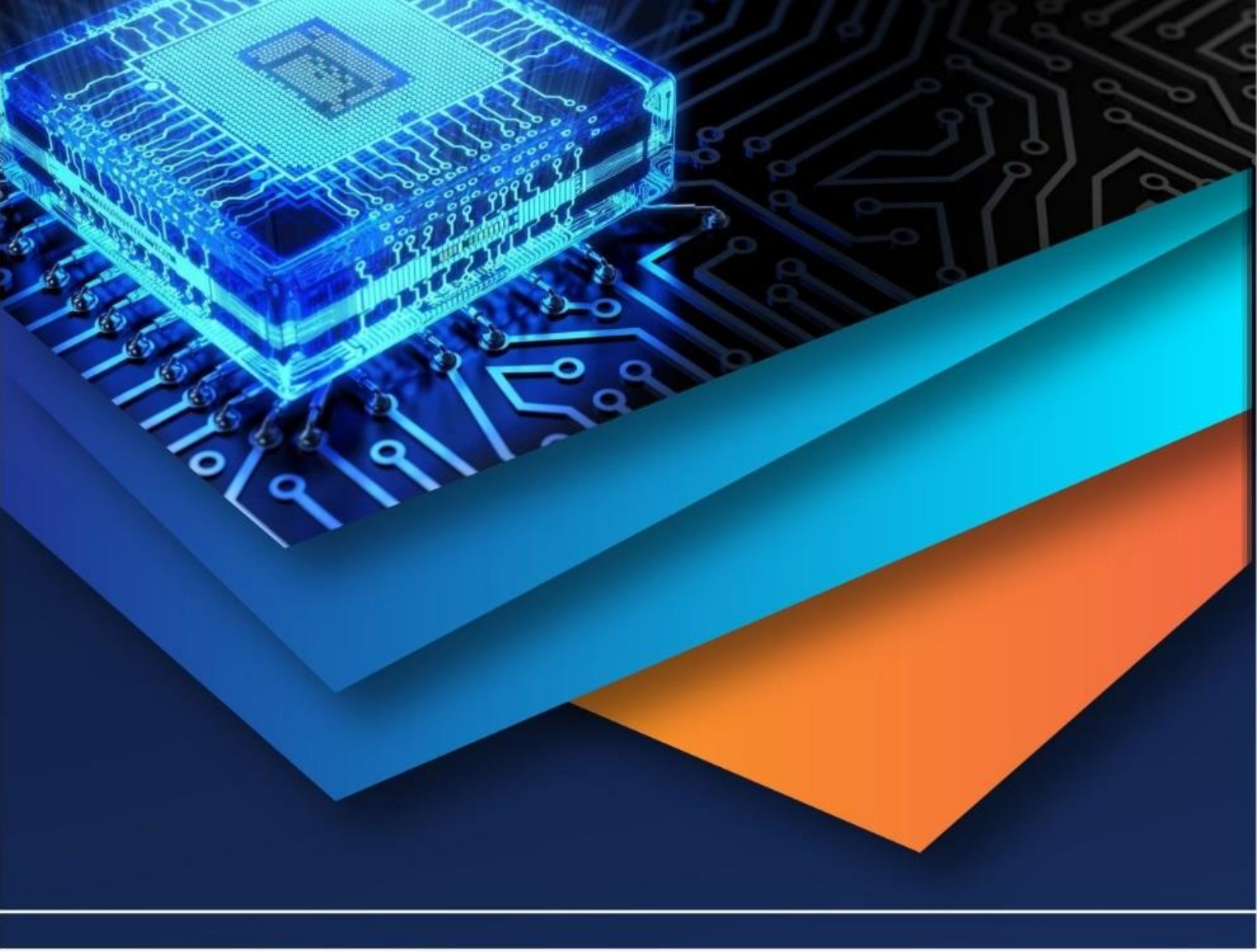
The difficulties with sentiment analysis have been clear throughout the project. The difficulties of performing sentiment analysis tasks are highlighted by lexical restrictions, resolving negations, sarcasm detection, addressing data imbalance, and assuring model generalization. Nevertheless, this study has shown how useful these approaches are for gleaning sentiment from textual data and offering insightful data on consumer feedback, public opinion, and social media sentiment tracking.

## X. FUTURE SCOPE

The current study has revealed a number of opportunities and difficulties in the field of sentiment analysis, opening up several intriguing directions for further investigation. First and foremost, one of the most important directions is the creation of domain-specific sentiment lexicons. Sentiment analysis's accuracy and applicability in specialized contexts can be considerably improved by customizing lexicons to certain industries or domains, such the financial or healthcare sectors. It also holds enormous promise to investigate cutting-edge machine learning methods, including deep learning systems like Transformers. Particularly in complex textual data sources like social media and news articles, these models show the potential to capture deep contextual nuances and semantics, enhancing sentiment classification accuracy. Furthermore, interpretable machine learning models deserve consideration. The adoption of sentiment analysis models in crucial applications like legal and healthcare, where model judgments need to be justified and explained, can be facilitated by improving the transparency and interpretability of the models. In our globalized and linguistically diverse society, it is also crucial to research sentiment analysis in code-mixed languages and multilingual languages. Last but not least, ongoing research into sentiment analysis in developing online platforms, including new social media networks and messaging apps, is still important. The continuous relevance and applicability of sentiment analysis in our quickly evolving digital environment will be ensured by adapting sentiment analysis approaches to new platforms and data kinds.

## REFERENCES

[1] Iparraguirre-Villanueva, O., Guevara-Ponce, V., Sierra-Liñan, F., Beltozar-Clemente, S., & Cabanillas-Carbonel, M. (2022). Sentiment analysis of tweets using unsupervised learning techniques and the K-Means algorithm.
[2] Chandrika, C. P., & Kallimani, J. S. (2020). Classification of abusive comments using various machine learning algorithms. In Cognitive Informatics and Soft Computing: Proceeding of CISC 2019 (pp. 255-262). Springer Singapore.
[3] Biswas, Jahanur & Rahman, Md & Biswas, Al Amin & Khan, Md Akib & Rajbongshi, Aditya & Niloy, Hasnaine. (2021). Sentiment Analysis on User Reaction for Online Food Delivery Services using BERT Model. 10.1109/ICACCS51430.2021.9441669.
[4] Onan, A. (2021). Sentiment analysis on product reviews based on weighted word embeddings and deep neural networks. Concurrency and Computation: Practice and Experience, 33(23), e5909.
[5] Sindhura, S., Praveen, S. P., Safali, M. A., & Rao, N. (2021, September). Sentiment analysis for product reviews based on weakly-supervised deep embedding. In 2021 Third International Conference on Inventive Research in Computing Applications (ICIRCA) (pp. 999-1004). IEEE.
[6] Muhammad, W., Mushtaq, M., Junejo, K. N., & Khan, M. Y. (2020). Sentiment analysis of product reviews in the absence of labelled data using supervised learning approaches. Malaysian Journal of Computer Science, 33(2), 118-132.
[7] Nguyen, H. D., Le, T., Tran, K., Luu, S. T., Hoang, S. N., & Phan, H. T. (2021). Multi-level sentiment analysis of product reviews based on grammar rules. In New Trends in Intelligent Software Methodologies, Tools and Techniques (pp. 444-456). IOS Press.
[8] Chauhan, P., Sharma, N., & Sikka, G. (2021). The emergence of social media data and sentiment analysis in election prediction. Journal of Ambient Intelligence and Humanized Computing, 12, 2601-2627.
[9] Xu, Q. A., Chang, V., & Jayne, C. (2022). A systematic review of social media-based sentiment analysis: Emerging trends and challenges. Decision Analytics Journal, 3, 100073.
[10] Singh, N. K., Tomar, D. S., & Sangaiah, A. K. (2020). Sentiment analysis: a review and comparative analysis over social media. Journal of Ambient Intelligence and Humanized Computing, 11, 97-117.
[11] Nemes, L., & Kiss, A. (2021). Social media sentiment analysis based on COVID-19. Journal of Information and Telecommunication, 5(1), 1-15.
[12] Chakraborty, K., Bhattacharyya, S., & Bag, R. (2020). A survey of sentiment analysis from social media data. IEEE Transactions on Computational Social Systems, 7(2), 450-464.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)