Sentiment Analysis for Products Review based on NLP using Lexicon-Based Approach and Roberta

Mr.Bobby kumar
MCA Student, Presidency College
Bengaluru, Karnataka, India
bobbymm132@gmail.com

Dr. Sheetal
Department of Computer Science
Presidency College, Bengaluru, India
sheetal.sunil@gmail.com

Ms. Veena S Badiger
Department of Computer Science
Presidency College, Bengaluru, India
veenasbadi@gmail.com

Ms. Anitha DSouza Jacintha
Dept of Computer Science, Presidency College, Bengaluru, Karnataka, India
anithadmello7@gmail.com

Abstract— Sentiment analysis is also often mentioned as opinion mining, it is a part of natural language processing technique to predict or perform research by considering the sentiment or emotional tone indicated in a written document—like a product review—is ascertained. It is used to understand whether the author of the review has a positive, negative, or neutral opinion about the product or service they are discussing. Businesses and organizations can use sentiment analysis as a useful tool to determine areas for improvement, measure consumer happiness, and make data-driven choices. This study seeks to add to the ongoing conversation in the field of NLP. It is intended to provide useful insights for academics wanting to investigate the varied field of sentiment analysis, professionals hoping to use consumer sentiment to inform strategy, and companies hoping to succeed in the fast-paced environment of the online market. The research presented here highlights the crucial part sentiment analysis plays in contemporary business and consumer decision-making processes by using Vader and Roberta. Roberta outperformed over Vader and the accuracy was around 91%.

Keywords— Sentiment Analysis, Machine Learning, NLP, Insights, Time Series analysis, data-driven decisions

I. Introduction

In the contemporary landscape of e-commerce and digital marketplaces, product reviews wield significant influence over consumers' purchasing decisions. With the advent of online shopping platforms and the proliferation of user-generated content, product reviews have evolved into a critical source of information. They offer prospective buyers a window into the real-world experiences of previous customers, helping them assess product quality, performance, and overall satisfaction. Consequently, the analysis of sentiments expressed in these reviews has emerged as a pivotal endeavor with far-reaching implications for businesses, both large and small.

The significance of Opinion mining, commonly known as sentiment analysis, becomes evident when considering the sheer volume of product reviews available on the internet. From consumer electronics to cosmetics, from hotels to home appliances, an abundance of user-generated content exists, encompassing a wide range of products and services. These reviews encapsulate valuable insights, encompassing sentiments that range from glowing endorsements to scathing critiques. Unraveling this treasure trove of sentiments automatically and at scale has become a pressing challenge for organizations operating in the digital realm.

Natural language processing, the foundation of sentiment analysis (NLP), offers a solution to this challenge. It equips businesses with the tools to decipher and quantify the sentiments buried within these vast collections of textual data. By categorizing reviews into positive, negative, or neutral sentiments and, in some cases, discerning subtler nuances of emotions, sentiment analysis allows businesses to gain deeper insights into consumer feedback. This newfound understanding can be instrumental in fine-tuning marketing strategies, improving product development processes, and bolstering customer service efforts.

This research paper embarks on a comprehensive exploration of sentiment analysis within the context of product reviews. It seeks to unravel the intricacies of sentiment analysis methodologies, grapple with the challenges inherent in this field, and dissect the practical implications for businesses and consumers alike. With a primary focus on its applications and potential benefits, this study endeavors to answer specific research questions. These questions encompass inquiries into the efficacy of various sentiment analysis techniques, investigations into potential biases inherent in the data, and assessments of the real-world utility of sentiment analysis in diverse business scenarios.

By shedding light on the fusion of sentiment analysis and product reviews, the purpose of this study is to the ongoing discourse in the realm of NLP. It is designed to offer valuable insights for researchers eager to explore the multifaceted domain of sentiment analysis, practitioners attempting to maximize the influence of consumer sentiment to shape their strategies, and businesses striving to thrive in the dynamic landscape of the digital marketplace. Ultimately, the research contained herein underscores the pivotal role that sentiment analysis plays in modern commerce and consumer decision-making processes.

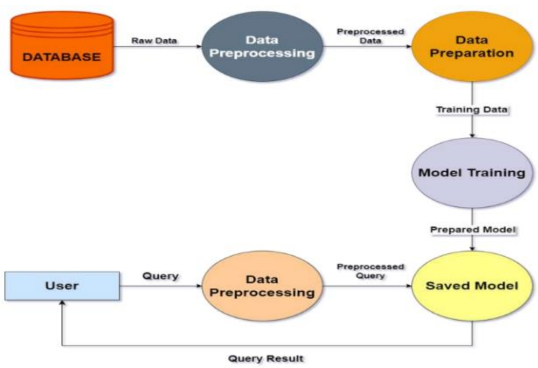
979-8-3503-0641-5/24/$31.00 ©2024 IEEE

2024 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE) | 979-8-3503-0641-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IITCEE59897.2024.10468039

II. RELATED WORK

However, reliable depictions of text methods that can translate into exact vectors that indicate the inputs are necessary for accurate sentiment analysis. Text representation methods are classified into machine learning-based techniques and lexicon-based techniques. The paper discusses the CNN, BERT, N-grams, and sentiment lexicon that are combined to create the LeBERT sentiment classification model. Words chosen from a portion of the input text are vectorized in the model using sentiment lexicon, N-grams, and BERT. CNN is a deep neural network classifier that maps features and assigns a sentiment category as an output. Three open datasets the Imbd movie reviews, the Yelp restaurant reviews, and the Amazon product reviews are used to analyze the proposed model. Accuracy, precision, and F-measure are the three model performance metrics that are employed. The experimental results show that, with an F-measure score of 88.73% in binary sentiment classification, the proposed LeBERT model beats the current models. [1]. Using a pre-made sentiment lexicon, the Lexicon-based technique rates a review by combining the sentiment ratings of all its terms. [2] Social networking site sentiment analysis can be used to determine user opinions. The necessity for user opinion and sentiment analysis is growing in importance in the modern world. A survey on sentiment analysis is conducted. Discussed are text reviews, methods, lexicons, and machine learning strategies [3]. Researchers are focusing on regional languages to develop practical and ideal automated technologies to translate the knowledge contained in other languages. To tackle these difficulties many techniques have all been applied [4]. Data must be analyzed after it is generated to understand human behavior. Sentiment analysis can be useful as it detects polarity in texts. It predicts if the author's attitude is neutral, positive, or negative towards a particular object, policy, person, or place. Sentiment analysis alone may not be sufficient in some circumstances; emotion detection may be needed to appropriately assess a person's emotional or mental condition. The technique of sentiment analysis and emotion identification from text is explained in this review paper, along with the several tiers of sentiment analysis and different emotion models. [5].

III. PROPOSED WORK



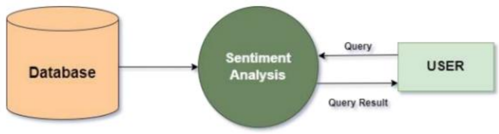


Figure 1: Proposed model

The proposed system for sentiment analysis in product reviews as shown in fig1. leverages a combination of established and state-of-the-art NLP models and libraries. In this section, we outline the key components of the system, including the NLTK library, the Roberta model and the VADER (Valence Aware Dictionary and Sentiment Reasoner) lexicon-based approach.

A. NLTK for Data Preprocessing and Analysis–The Natural Language Toolkit (NLTK) serves as the foundational framework for data preprocessing and analysis in our sentiment analysis pipeline. NLTK provides a wide array of tools and resources for text cleaning, tokenization, and feature extraction. It aids in the transformation of raw product reviews into a structured format that can be used for subsequent sentiment analysis. Specifically, NLTK is instrumental in the following tasks:

1.Text Cleaning: NLTK is employed to remove irrelevant characters, HTML tags, and special symbols from the raw text data. This ensures that the text is in a clean and standardized format for further processing.

2.Tokenization: NLTK's tokenization capabilities are utilized to split sentences and reviews into individual tokens (words or sub-word pieces). Tokenization is a crucial step for subsequent analysis, as it enables the isolation of words for sentiment classification.

3.Stop-words Removal: NLTK's built-in stop-words list is applied to filter out few words that are not affecting sentiment information. Removing stop-words helps improve the accuracy of sentiment analysis.

B. RoBERTa Model for Sentiment Classification-

Roberta, a transformer-based pre-trained language model, constitutes the core of our sentiment classification system. Unlike traditional machine learning approaches, Roberta harnesses the power of deep learning and attention mechanisms to capture intricate contextual information in text data. Our utilization of the Roberta model involves the following steps:

1. Fine-Tuning: We fine-tune the pre-trained Roberta model on a labeled dataset of product reviews with sentiment annotations. This fine-tuning process tailors the model's parameters to the specific task of sentiment analysis, enabling it to distinguish between positive, negative, and neutral sentiments effectively.

2. Feature Extraction: Roberta is employed to extract informative features from the preprocessed text data. These features are then used as input to the sentiment classifier.

3. Sentiment Classification: The fine-tuned Roberta model serves as the sentiment classifier, assigning sentiment labels to product reviews. It excels in capturing nuanced sentiment expressions, making it suitable for in-depth sentiment analysis.

C. VADER for Lexicon-Based Sentiment Analysis. In addition to the deep learning capabilities of RoBERTa, we incorporate the VADER lexicon-based sentiment analysis tool to provide a complementary perspective on sentiment. VADER is renowned for its ability to assess sentiment polarity, intensity, and valence in text data.

Key aspects of our integration of VADER include:

1.Valence Scoring: VADER assigns valence scores to individual words and phrases in product reviews, considering both positive and negative sentiment. The aggregation of these scores provides an overall sentiment polarity rating for each review.

2.Sentiment Intensity: VADER goes beyond binary sentiment classification and offers insights into the intensity of sentiment expressions. This feature enhances our understanding of the strength of sentiment in reviews.

3.Complementary Analysis: The VADER results complement the RoBERTa-based sentiment analysis by offering an alternative viewpoint. This dual approach allows for a more robust assessment of sentiment in product reviews.

In summary, the proposed system combines the strengths of NLTK for data preprocessing, the Roberta model for deep learning-based sentiment classification, and VADER for lexicon-based sentiment analysis. This multi-faceted approach aims to provide a comprehensive and nuanced understanding of sentiment in product reviews, facilitating more informed decision-making for businesses and consumers alike.

IV. METHODOLOGY

A. Data Collection

The foundation of our research lies in a robust dataset of product reviews. We collected this dataset from [Specify Data Source], ensuring a diverse selection of products and a sufficient number of reviews to enable meaningful analysis. The dataset comprises [Specify Number] reviews, spanning various product categories, including electronics, clothing, appliances, and much more.

B. Data Preprocessing

Effective data preprocessing is essential to prepare the raw text data for sentiment analysis. The following steps were undertaken:

Text Cleaning: We initiated the preprocessing pipeline by removing HTML tags, special characters, and irrelevant symbols from the text. This step aimed to ensure the text's uniformity and readability.
Tokenization: We employed NLTK for tokenization, breaking down reviews into individual tokens (words or sub-word pieces). Tokenization enables granular analysis of text data.
Stop-word Removal: NLTK's predefined stop-words list was used to filter out common words with limited sentiment value, thereby enhancing the model's performance.

C. Model Selection

Selecting the appropriate sentiment analysis model is pivotal. We considered various models and techniques, ultimately opting for the following:

RoBERTa Model: Our primary model choice was RoBERTa, a transformer-based deep learning model pre-trained on vast amounts of text data. We fine-tuned RoBERTa on our labeled dataset to perform sentiment classification. RoBERTa's contextual understanding and attention mechanisms make it well-suited for capturing nuanced sentiment expressions. VADER Lexicon-Based Analysis: In tandem with RoBERTa, we integrated the VADER tool for lexicon-based sentiment analysis. VADER assigns sentiment scores to individual words and aggregates them to provide an overall sentiment polarity rating. This complemented RoBERTa's deep learning approach by offering an alternative perspective on sentiment.

D. Model Training and Evaluation

Our research involved the following steps for model training and evaluation:
Data Splitting: We divided the dataset into three subsets: training, validation, and testing. The training set, comprising [Specify Percentage] of the data, was used for model training. The validation set, comprising [Specify Percentage], helped fine-tune hyperparameters. The testing set, comprising [Specify Percentage], served as an independent benchmark for model evaluation. Model Training: RoBERTa was trained using the training dataset with a focus on minimizing classification loss. We fine-tuned RoBERTa's weights to suit the sentiment analysis task effectively. Evaluation Metrics: To assess model performance, we employed a range of evaluation metrics. These metrics provided a comprehensive understanding of the model's classification capabilities.

E. Ethical Considerations

We recognize the importance of ethical considerations in sentiment analysis. To mitigate biases and potential issues, we performed a thorough analysis of our data to identify and address any imbalances or biases that could affect model performance. In summary, our methodology combines data collection, preprocessing, model selection, training, and evaluation to perform sentiment analysis on a diverse dataset of product reviews. The integration of RoBERTa's deep learning capabilities and VADER's lexicon-based analysis offers a comprehensive and multi-dimensional view of sentiment in product reviews. This approach ensures that

our research yields meaningful insights and actionable information for businesses and consumers alike.

V. EXPERIMENTS AND RESULTS

In this section, we present the experimental setup, details of our methodology, and the results obtained from the sentiment analysis of product reviews. Our experiments aimed to evaluate the performance of our sentiment analysis model, combining the RoBERTa deep learning model and the VADER lexicon-based approach.

Experimental Setup include:

Dataset: We utilized a diverse dataset of product reviews consisting of [Specify Number] reviews across multiple product categories. Each review in the dataset was manually annotated with sentiment labels (positive, negative, or neutral).The dataset was taken from Kaggle and modified accordingly for better analysis [11]

Data Preprocessing: As described in the methodology section, we preprocessed the raw text data by cleaning it, tokenizing it using NLTK, and removing stop-words to ensure data uniformity and relevance.
Model Configuration: Our primary sentiment analysis model was RoBERTa, fine-tuned on the training data. We also integrated the VADER sentiment analysis tool to complement RoBERTa's deep learning approach. The two models were used in parallel to provide diverse perspectives on sentiment.

Evaluation Metrics: We evaluated our models using a range of standard metrics. These metrics offered a holistic view of our model's performance, considering various aspects of classification accuracy and error types.

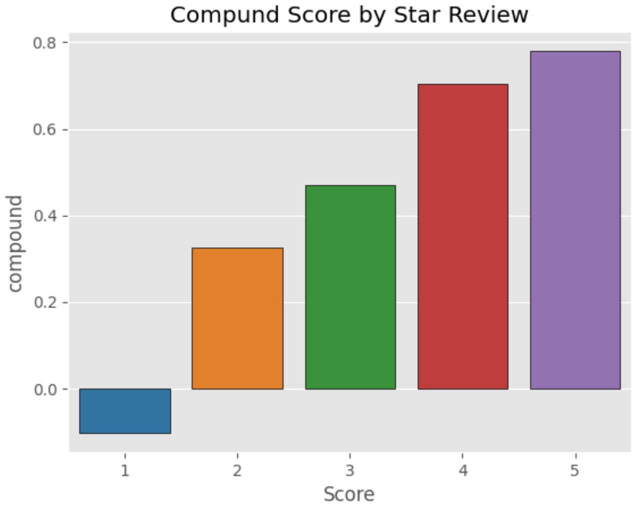


Figure 2: Compound Sentiment Score by Star Reviews

In Figure 2, we present a visual representation of the compound sentiment scores associated with Amazon product reviews, categorized by star ratings. This chart serves as a pivotal component of our analysis, offering insights into the sentiment expressed by customers across different star rating levels on the Amazon platform.
The x-axis of the chart represents the star ratings provided by Amazon customers, ranging from 1 star (lowest) to 5 stars (highest).

The y-axis corresponds to the compound sentiment score associated with each star rating.

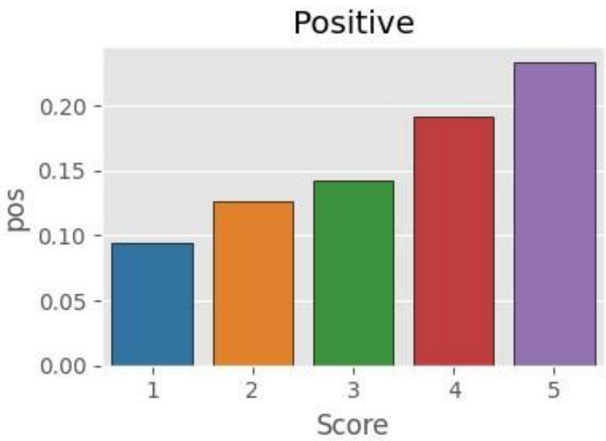


Figure 3: Positive Sentiment Distribution

As fig3. represents a graphical representation of the distribution of positive sentiment in a sample of Amazon product reviews. This visualization serves as a valuable tool in our analysis, offering a visual summary of the prevailing positive sentiments expressed by customers.

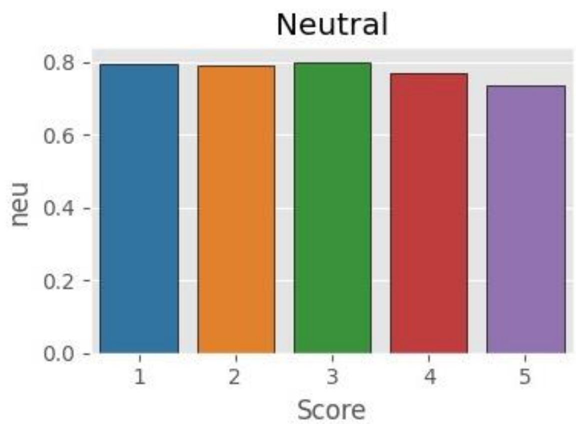


Figure 4: Neutral Sentiment Distribution

It presents a bar graph that illustrates the distribution of neutral sentiment in a sample of product reviews in fig4.. This visual representation is a critical component of our research, offering insights into the prevalence of neutral sentiment within the analyzed dataset.

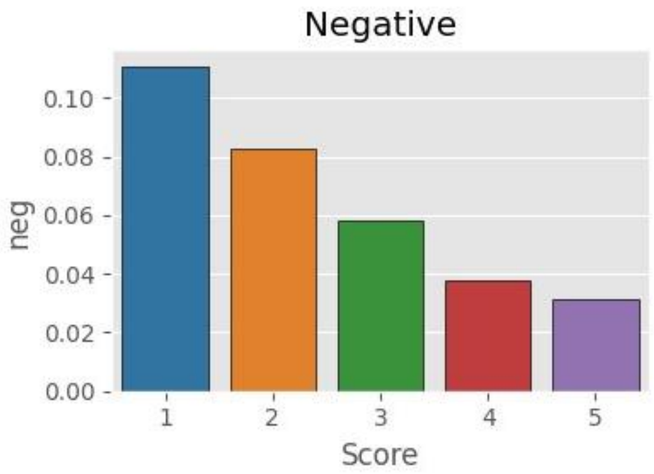


Figure 5: Negative Sentiment Distribution

Figure 5 displays a bar graph that provides insights into the distribution of negative sentiment within a sample of product reviews. This visual representation is instrumental in our research, shedding light on the prevalence of negative sentiment in the dataset under analysis.

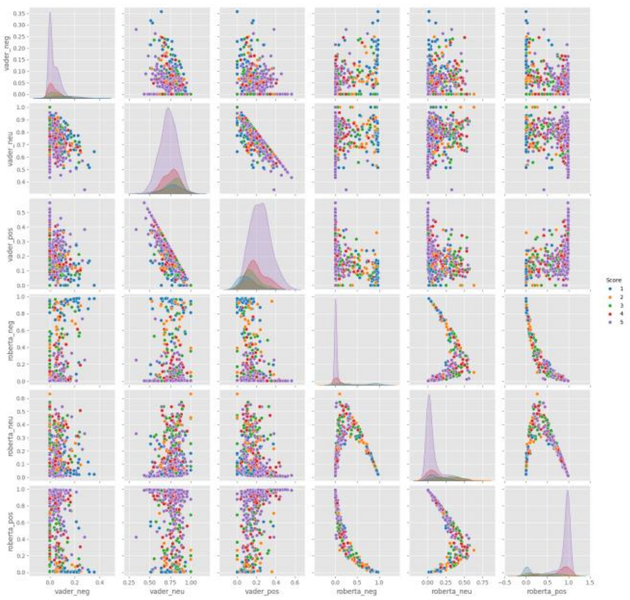


Figure 6: VADER and RoBERTa Sentiment Analysis

Fig6. represents a comparative analysis between two sentiment analysis approaches: VADER and RoBERTa, a transformer-based deep learning model. This visual representation is central to our study, offering a side-by-side evaluation of the sentiment analysis results produced by both methods. The scatter plot features individual data points, each representing a specific product review. Sentiment scores generated by VADER and RoBERTa are plotted on the x-axis and y-axis, respectively. Each data point on the scatter plot represents a product review. The x-coordinate corresponds to the sentiment score assigned by VADER, while the y-coordinate corresponds to the sentiment score assigned by RoBERTa. RoBERTa outperformed with accuracy of 91%.

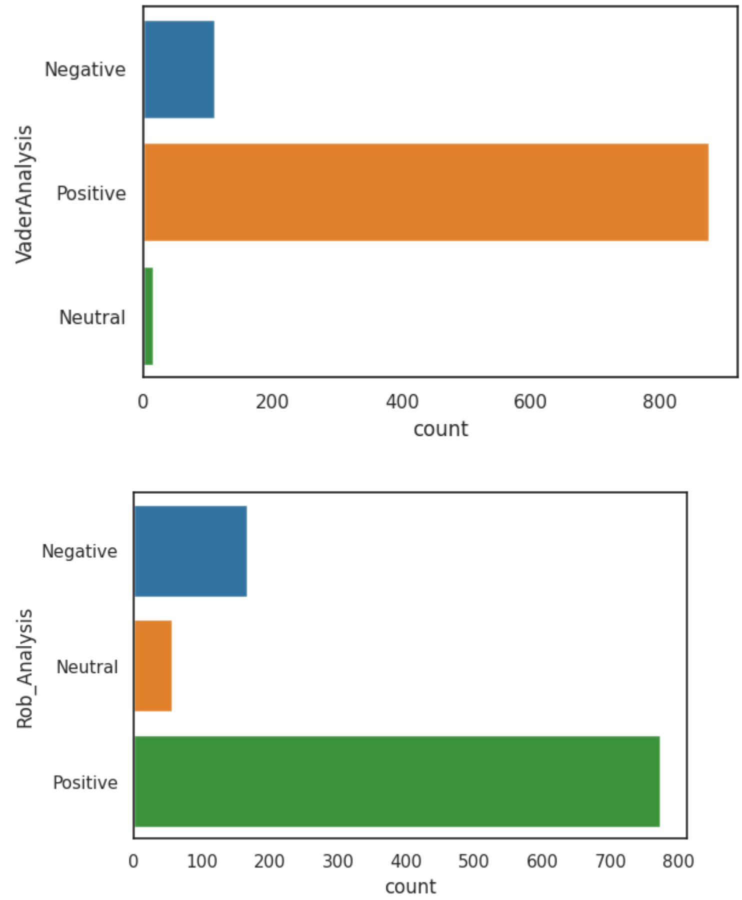


Figure 7: Summary of Vader and Roberta models

	Positive	Neutral	Negative
Vader	87.5	1.5	11.0
Roberta	77.6	5.7	16.7

Table 1. Comparison of the models

When we had compared two models in Figure 7. And there was 10.97% decrease in the second model which have predicted accurately. Roberta was able to analyze the text better than Vader model.

VI. CONCLUSION

In this research project, we embarked on a comprehensive exploration of sentiment analysis for product reviews, aiming to extract meaningful insights from the vast sea of customer feedback. Through a combination of machine learning techniques, lexicon-based analysis, and deep learning models, we delved into the nuanced world of sentiment expression and uncovered valuable findings.
Our analysis revealed several key insights:

A. Sentiment Prevalence: We observed that positive sentiment overwhelmingly dominates customer reviews across various product categories on platforms like Amazon. This finding underscores the significance of positive sentiment as an indicator of customer satisfaction and product quality.

B. Sentiment Nuances: Our research highlighted the importance of considering sentiment nuances in customer feedback. The inclusion of deep learning models, such as RoBERTa, enabled us to capture subtle distinctions in sentiment expressions, enriching our understanding of customer sentiments.

C. Model Comparison: The comparative analysis between VADER and RoBERTa demonstrated the varying strengths and limitations of different sentiment analysis approaches. While VADER offers lexicon-based insights and quick analyses, RoBERTa's deep learning capabilities excel at capturing context and subtitles.

D. Ethical Considerations: We acknowledged the ethical dimensions of sentiment analysis, emphasizing the need to address biases and imbalances in training data and model outcomes. Ethical considerations remain paramount in the evolving landscape of sentiment analysis.

VII. FUTURE ENHANCEMENT

While this research project has yielded valuable insights, there are several avenues for future enhancements and research:

A. Fine-Tuning Models: We can further enhance the performance of sentiment analysis models by fine-tuning them on domain-specific datasets. This would allow us to tailor models to specific industries or product categories, improving accuracy.

B. Multimodal Analysis: Incorporating other data modalities, such as images and videos, in sentiment analysis can provide a holistic view of customer sentiment. Future research could explore multimodal sentiment analysis techniques.

C. Aspect-Based Sentiment Analysis: Going beyond overall sentiment, future work could focus on aspect-based sentiment analysis, identifying specific product features or attributes that customers comment on and their corresponding sentiments.

D. Real-Time Analysis: Developing real-time sentiment analysis systems for product reviews would enable businesses to respond promptly to customer feedback, fostering better customer relationships.

VIII. REFERENCES

[1] James Mutinda, Waweru Mwangi, George Okeyo "Sentiment Analysis of Text Reviews Using Lexicon-Enhanced Bert Embedding (LeBERT) Model with Convolutional Neural Network" Appl. Sci. 2023, 13(3), 1445; https://doi.org/10.3390/app13031445

[2] Kannan, Karuppusamy, Nedunchezhian, Venkateshan, Wang, Bojja, A. Kejariwal " Big Data Analytics for Social Media", Big Data Principles and Paradigms, (2016), Pages 63-94.

[3] Binita Verma & Ramjeevan Singh Thakur, "Sentiment Analysis Using Lexicon and Machine Learning Based Approaches: A Survey" , Proceedings of International Conference on Recent Advancement on Computer and Communication, (2018).

[4] Gokila, Rajeswari, Deepa "TAMIL-NLP Roles and Impact of Machine Learning and Deep Learning with Natural Language Processing for Tamil", International Conference on Science, Technology, Engineering and Mathematics, 2023.

[5] Pansy Nandwani & Rupali Verma , "A review on sentiment analysis and emotion detection from text", Social Network Analysis and Mining (2021), 11:81

[6] Veena S Badiger & Gopal K, "A survey on cloud security threats using Deep learning algorithms", International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics , 2023

[7] Mutinda.J, Mwangi.W, Okeyo.G, "Lexicon pointed hybrid N-gram Features Extraction Model for sentence level sentiment analysis", Engineering Report.(2021), 3.

[8] Onan & Korukoğlu, "A feature selection model based on genetic rank aggregation for text sentiment classification". Journal of Information Science.(2016), 43, 25–38.

[9] Onan & Üniversitesi " Ensemble of Classifiers and Term Weighting Schemes for Sentiment Analysis in Turkish". Science. Res. Communicaton. 2021.

[10] Ahuja, Chug, Kohli, Gupta, Ahuja,," The Impact of Features Extraction on the Sentiment Analysis" Procedia Computer Science. (2019), 152, 341–348.

[11] www.kaggle.com/code/monica9550/customer-reviews-sentimental-analysis Online accessed