Societal Transformation: AI and Big Data Journal
Vol. 2(2): 19-34, 2024
DOI: 10.20547/aibd.242202

Exploring User Perception and Attitudes towards ChatGPT Using Sentiment Analysis

Kulsoom Nasir Hussain*
Maria Andleeb†

Abstract

Sentiment analysis has been a topic of intense research. This study performed sentiment analysis to explore user perception and attitude towards ChatGPT using sentiment analysis. In this study, sentiment analysis techniques are applied to two datasets. These datasets include i.e. tweets and Reddit posts. It contains references to ChatGPT within the context of education. The study aims to reveal the user's sentiment, opinions, and comments expressed on these social media platforms regarding the use of ChatGPT in educational environments. The study adopts the following methodology. A diverse quantity of textual data is collected in real time from Twitter and Reddit. The collected data is preprocessed to improve its quality and to extract valuable information. A whole natural language processing pipeline is applied such as tokenization, stemming, lemmatization, the removal of stop words, normalization, and replacement of similar words with a common term. The textual data is classified into a sentiment category by two different sentiment analysis models i.e. Logistic regression and support vector machines. In addition, this study performed the sentiment analysis using lexicon-based approaches employing different NLP tools.

Keywords*: sentiment analysis, machine learning, social media, twitter, reddit*

1. Introduction

Artificial intelligence (AI) has immensely transformed many aspects of human life including education in recent times. AI-powered tools are increasingly being integrated into educational settings as they have a wide range of applications. This includes administrative tasks for direct assistance in education. One such AI use case which has received a lot of attention is the ChatGPT chatbot which serves various needs for its users. Despite its growing popularity, there is little or no comprehensive research about how this tool is perceived and used in educational settings. As AI technologies become increasingly integrated into educational systems, it is imperative to assess the public sentiment about these innovations. This can help in responsible implementations of future chatbots and address potential concerns. This research focuses on the sentiment

* NED University of Engineering & Technology, Karachi, Pakistan.*kulsoomnasir23@gmail.com*
† NED University of Engineering & Technology, Karachi, Pakistan

analysis performed on tweets and Reddit data on the topic of ChatGPT. The paper primarily explores the impact of ChatGPT on the educational system This study builds upon previous research by, [28] which explored user perception and attitudes towards chatbot systems using a user-centered approach to sentiment analysis.

The integration of AI in education has the potential to revolutionize traditional teaching methods, offering personalized learning experiences and enhancing educational outcomes. However, the reception of these technologies among educators, students, and the general public remains a complex and dynamic aspect that requires careful examination. Social media platforms, such as Twitter and Reddit, serve as valuable repositories of public opinions, providing a rich source of data to gauge the sentiments, perceptions, and discussions surrounding AI in education.

This research aims to delve into the collective sentiments expressed on social media platforms regarding ChatGPT's impact on education, uncovering not only positive sentiments but also potential challenges and concerns. Similar studies have been reported in literature. For instance, [16] discussed the attitude towards ChatGPT based on sentiment analysis on twitter and reddit. User perception and biases towards ChatGPT based on machine learning has been explored in [17]. Knowledge mapping has been has been used in [18]. By employing sentiment analysis, this study seeks to identify prevailing attitudes towards the integration of ChatGPT in educational settings. Rest of the sections of the paper has been organized as follows. The next section presents the literature review on the topic. This is followed by methodology and then results. Finally, conclusion and limitations/future work sections are presented.

2. Literature Review

The integration of artificial intelligence (AI) in the educational domain has been a subject of growing interest and scrutiny in recent years. As ChatGPT has emerged, researchers are increasingly focused on understanding its impact on the educational system. [1,2] analyzed the views and perspectives of early ChatGPT practitioner in educational settings. By analyzing social media data gathered from platforms like Twitter, Reddit, YouTube, and LinkedIn, the study specifically explores how ChatGPT is currently utilized and what its potential might be in the future. [3] explored how to integrate ChatGPT 3.5 into the design thinking method for creating new products. The goal of the study is to close the understanding gap regarding how AI could help the well-established design thinking methods, especially in the context of rapidly evolving related to technology environments.
[4] explores the opportunities, threats and possible implications of integrating ChatGPT into higher education. Thy specifically focused on software engineering (SE) and scientific writing in particular. (Iqbal, Ahmed, and Azhar 2022) evaluates the opinions and attitudes of 20 professors at a private institution in Pakistan towards the utilization of ChatGPT in learning environments. Semi-structured interviews were used in the study's qualitative methodology, which was based on the technology acceptance

model, to collect data. The results show that professors are tend to be apprehensive and have a bad attitude towards implementing ChatGPT into their teaching methods.

[5] used machine learning models to assess tweets in order to understand what everyone feels about ChatGPT. 217,622 distinct ChatGPT-related tweets from Kaggle were collected for the study as a dataset. Positive, negative, or neutral labels were manually assigned to these tweets. For sentiment analysis, a number of machine learning models were employed, including logistic regression, random forest classifier, k- nearest neighbors, decision tree classifier, and gradient boosting. The feature extraction methods used were Bag of Words and Term Frequency-Inverse Document Frequency (TF-IDF).

The use of AI in education has witnessed a surge in research and development, driven by the promise of personalized learning experiences, adaptive tutoring, and improved student outcomes. [6, 7] have explored the potential benefits of AI applications, emphasizing the ability to cater to individual learning styles and provide timely feedback. Social media platforms, such as Twitter and Reddit, have become invaluable sources for understanding public opinions and sentiments. [8,9] have employed sentiment analysis techniques to extract insights from user-generated content, enabling an understanding of attitudes towards emerging technologies.

[10,11] have highlighted the ability to generate coherent and contextually relevant responses from chatbots, positioning it as a powerful tool for diverse applications, including education. The literature underscores the need to assess its impact on educational practices and the learning experience. [12, 13] have highlighted issues related to data privacy, ethical considerations, and the potential for exacerbating educational inequalities. Understanding public sentiment is crucial for navigating these challenges. Teachers, students, parents, and policymakers may have varying expectations and concerns regarding the integration of AI technologies [14, 15].

Sentiments towards ChatGPT have been briefly studied in literature. In this direction, [23] presented analysis of attitudes toward ChatGPT. Sentiments of AI through social media analysis has been done in [24]. Perception of higher education faculty towards ChatGPT has been analyzed in [25]. Knowledge, perception and attitude of researchers towards ChatGPT has been analyzed in [27]. However, the research gap is that they have used survey and questionnaire to determine perception and attitude.

This section now discusses briefly research on sentiment analysis. [19] has discussed sentiment analysis towards a political party in Pakistan. A systematic literature review on sentiment analysis has been provided in [20]. With the arrival of Generative AI, the domain of sentiment analysis has changed extensively. This has been discussed in great deal in [21]. Besides text, multimodal sentiment analysis has been discussed in [22] that involves not only text but visuals and audio. Sentiment analysis of tweets using machine and deep learning has been done in [26]. The selected algorithms were decision trees, Naïve Bayes, LSTM, CNN etc.

Based on the study, it can be concluded that it is very essential to explore how ChatGPT is perceived by different stakeholders in the educational community. Most of the studies analyzed sentiment analysis towards ChatGPT based on questionnaire. Very limited studies are available for using machine learning/ deep learning for sentiment analysis. In summary, the literature suggests a dynamic landscape where the integration of AI in education, particularly through models like ChatGPT, holds immense promise but is accompanied by a need for careful examination of public sentiment and potential challenges.

3. Methodology

Evaluating the impact of chatbots in the educational system requires understanding user perceptions and experiences. Sentiment analysis of user interactions with chatbots offers valuable insights into their perceived benefits, limitations, and overall impact on learning. Here's a detailed methodology for conducting sentiment analysis for chatbots in education.

3.1. Processing Pipeline

This section briefly discussed the processing pipeline used for implementation.

- **Data acquisition**: utilized web scraping tools and APIs to collect data from social media platforms. The approach also employed web crawling tools for forum and news article extraction. In addition, conducted surveys among students and teachers to gather their direct opinions.

- **Data Cleaning & Preprocessing**: removed irrelevant information like URLs, hashtags, punctuation, and stop words. Then, normalized text by converting them into lower case and further processing such as stemming/lemmatization is also performed.

- **Labeling**: labelled data with sentiment categories like positive, negative, neutral, or mixed based on the context of the text. This can be done manually or using sentiment analysis tools with pre-trained lexicons. The approach utilized both the techniques i.e. first pre trained lexicons are used and then manual verification is performed.

- **Data Split:** Finally, divide the cleaned and labeled data into training, validation, and test sets for machine learning models.

- **Sentiment Analysis:** As the last step, the sentiment analysis was performed. For this purpose, choose appropriate machine learning algorithms for sentiment analysis. The approach trained the selected model on the labeled training data, optimizing hyperparameters for best performance on the validation set.

- **Model Evaluation:** Evaluate the trained model's performance on the test set using metrics like accuracy, precision, recall, and F1-score. Analyze error patterns and refine the model if necessary.

3.2 Data Acquisition

For data acquisition of reddit, the approach conducted thorough research to identify subreddits directly relevant to your research topic. The approach considered factors like subscriber count, activity level, and community engagement. Utilize Reddit's search functionality and explore related subreddits to uncover niche communities. The methodology created a Reddit developer account to obtain API credentials (client ID and client secret). Install the PRAW (Python Reddit API Wrapper) library to interact with the API. Authenticate the application using the provided credentials.

The approach utilize PRAW's functions to retrieve specific data such as posts: titles, text content, upvotes, downvotes, author, timestamp, and comments. text content, upvotes, downvotes, author, timestamp.

For the acquisition of Tweets data, the following steps were performed:

- Identify Twitter accounts with significant influence on your research topic.
- Consider official accounts, industry leaders, relevant organizations, or individuals.
- Utilize Octaparse's visual interface to define scraping rules:
- Target elements: tweets, usernames, dates, likes, retweets.
- Pagination: handle multiple pages of results.
- Extraction logic: extract relevant information accurately.
- Run the scraper to automatically collect tweets based on your settings.
- Downloaded extracted data in CSV or Excel format for further analysis.

3.3 Preprocessing

The following steps were performed after the data has been extracted:

- Clean text data.
- Remove irrelevant characters (punctuation, symbols, emojis).
- Convert text to lowercase for consistency.
- Remove stop words (common words like "the", "a", "and").

- Handle URLs, hashtags, and mentions appropriately (remove or extract relevant information).
- Potentially address slang, abbreviations, and typos using normalization techniques.
- Save preprocessed data in structured formats: JSON for nested data structures (Reddit), and CSV for tabular data (Twitter).

3.4.Sentiment Analysis: Lexicon-based Approach

As discussed earlier, two types of sentiment analysis techniques have been employed. The paper used lexicon-based approaches as well as machine learning based approaches. For the lexicon-based approach, the following steps were performed:

The approach choose a pre-built sentiment dictionary like VADER (includes sentiment intensity) or TextBlob (incorporates subjectivity).

- Consider domain-specific lexicons if your research requires specialized vocabulary (e.g., finance, politics).
- Assign positive, neutral, or negative scores to each word based on the chosen dictionary.
- The approach calculate the overall sentiment score of a text unit:
- Average score of individual word scores.
- Weighted score considering word importance and sentiment intensity.

3.5 Sentiment analysis using a machine learning approach

- Following steps were performed for using machine learning for sentiment analysis:
- Label a portion of your scraped data with sentiment categories (positive, neutral, negative).
- Split the labeled data into training, validation, and testing sets for model training and evaluation.
- Choose a suitable machine learning algorithm:
- Naive Bayes: efficient for large datasets, performs well with text data.
- Logistic Regression: interpretable model, good for binary classification (positive/negative).
- Support Vector Machines: robust to noise, effective for high-dimensional

data.

- Train the chosen model on the labeled training data.
- Evaluate the model's performance on the validation set using metrics like accuracy, precision, recall, and F1 score.
- Use the trained model to predict sentiment scores for new, unlabeled text data.

4. Results

This section presents the results of sentiment analysis. Figure 1 shows the ROC curves for support vector machine. The overall accuracy of the SVM model is 0.7848. This means that the model correctly classified 78.48% of the sentiment samples. The precision score reflects the proportion of samples predicted as a certain sentiment class that are actually of that class. In this case, the model is most accurate for positive sentiment, with 80% of samples predicted as positive being truly positive. The model's performance is close for neutral sentiment, with 79% of samples predicted as neutral being truly neutral. The model's precision for negative sentiment is not directly visible but appears to be lower than the other two classes. Figure 2 shows the precision and Figure 3 shows the F1-score for support vector machine.

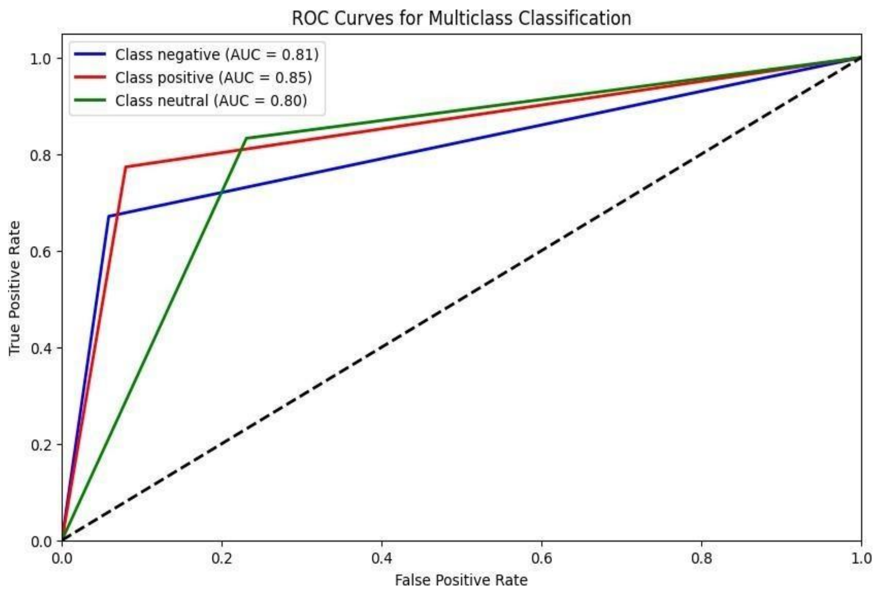

Figure 1: Support Vector Machine Model ROC Curves

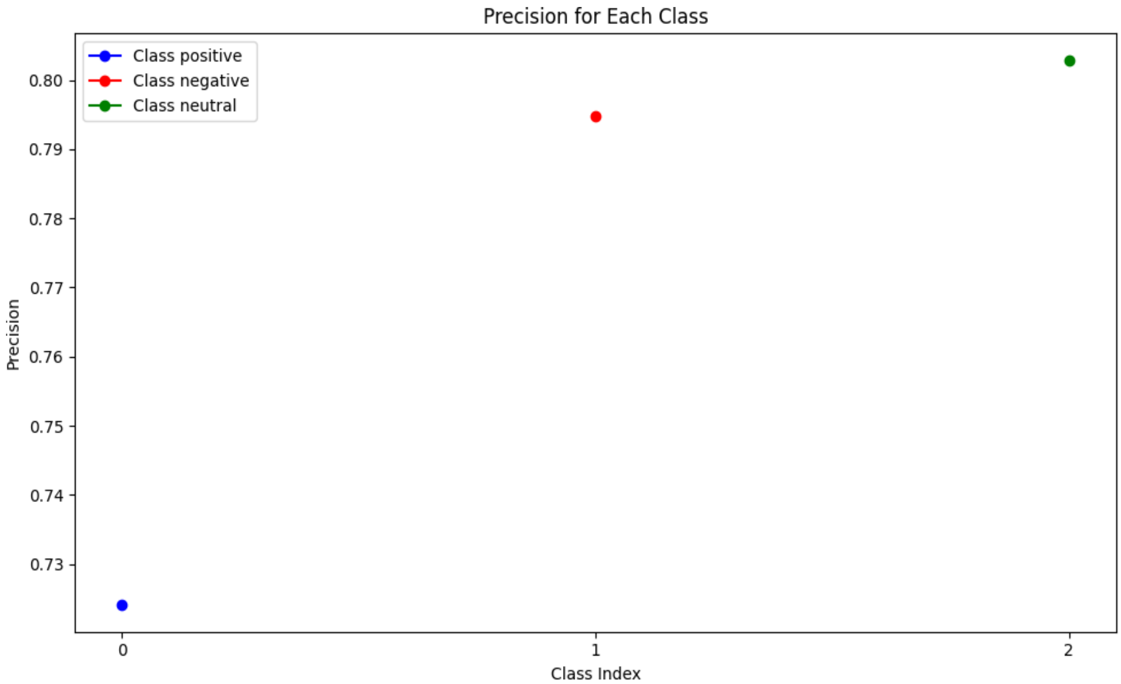

Figure 2: Precision for Support Vector Machine

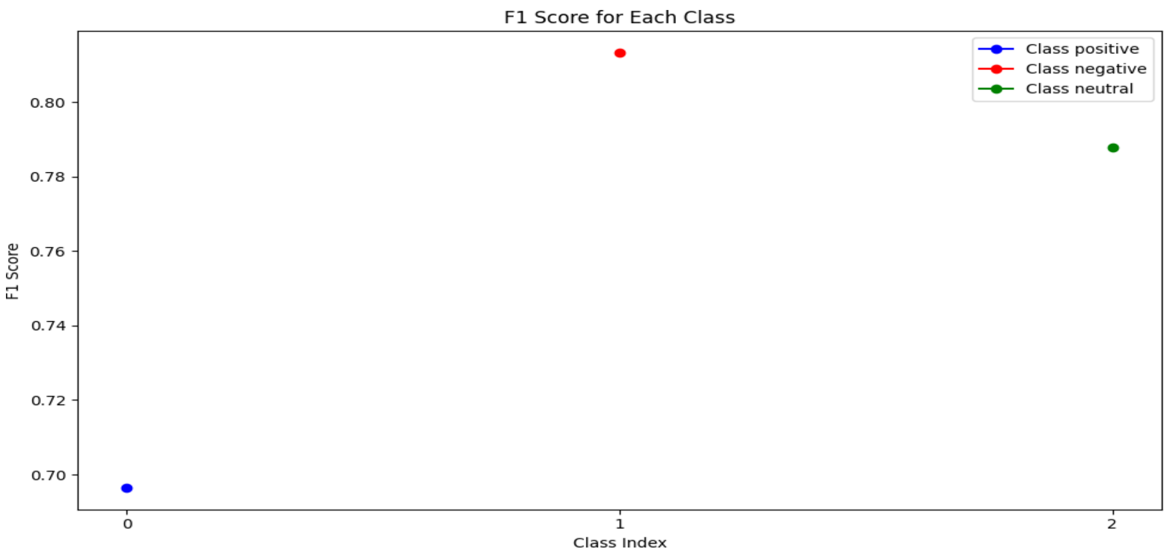

Figure 3: F1-score for support vector machine

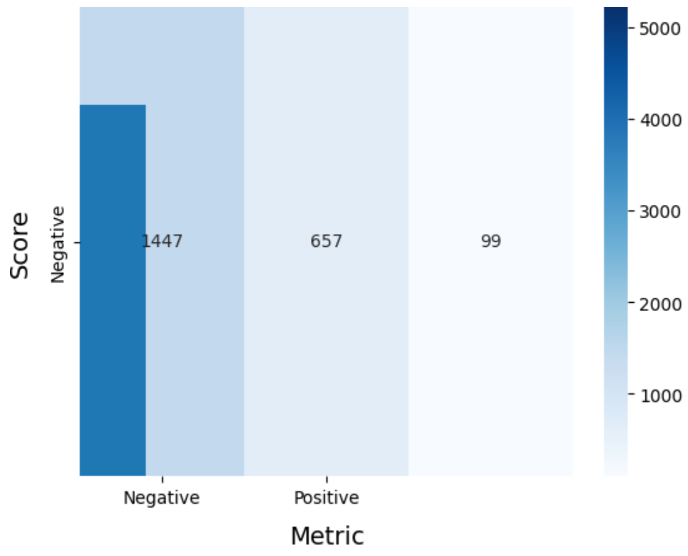

Figure 4: Accuracy of Logistic Regression model

Figure 4 shows the accuracy of logistic regression model. The figure indicates the overall accuracy of the logistic regression model, meaning it correctly classified 80.35% of the samples. Figure 5 shows the F1-score for logistic regression. Based on the F1-scores, the logistic regression model seems to perform reasonably well for all sentiment classes, with neutral sentiment slightly outperforming the others.

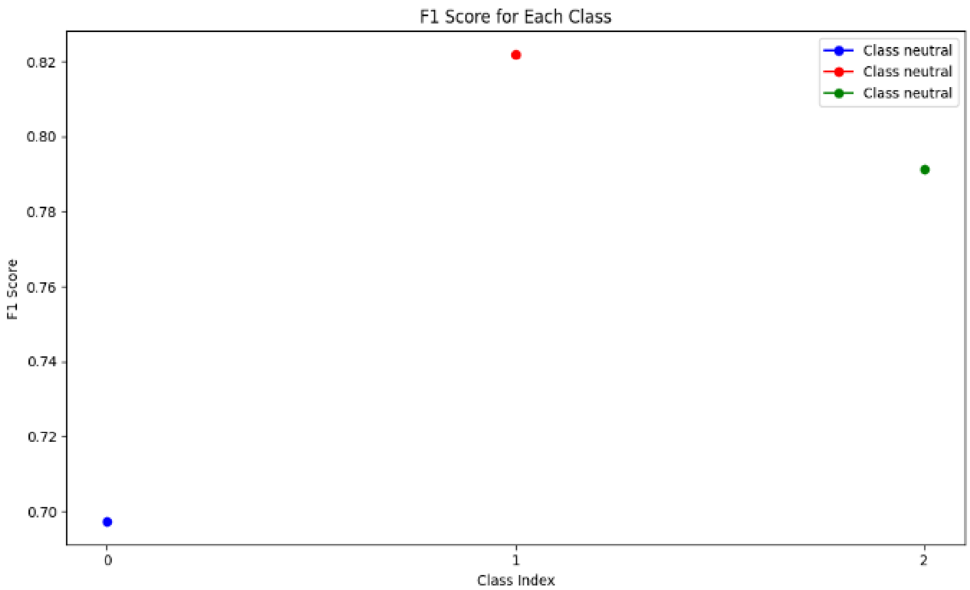

Figure 5: Logistic Regression Model F1-Score

Figure 6 shows the ROC curve for logistic regression. The ideal ROC curve is located in the upper left corner of the graph, where the TPR is high and the FPR is low. All three curves in the image appear to be smooth and relatively close to the upper left corner, suggesting good overall performance for the logistic regression model. The green curve (neutral) seems to be closest to the upper left corner, indicating the model might be slightly better at distinguishing neutral sentiment compared to the other classes. The red curve (negative) and blue curve (positive) are slightly further away from the upper left corner, suggesting the model might have slightly more difficulty distinguishing these classes compared to neutral sentiment. However, all three curves still achieve relatively high TPR and low FPR, indicating good performance for all sentiment classes.

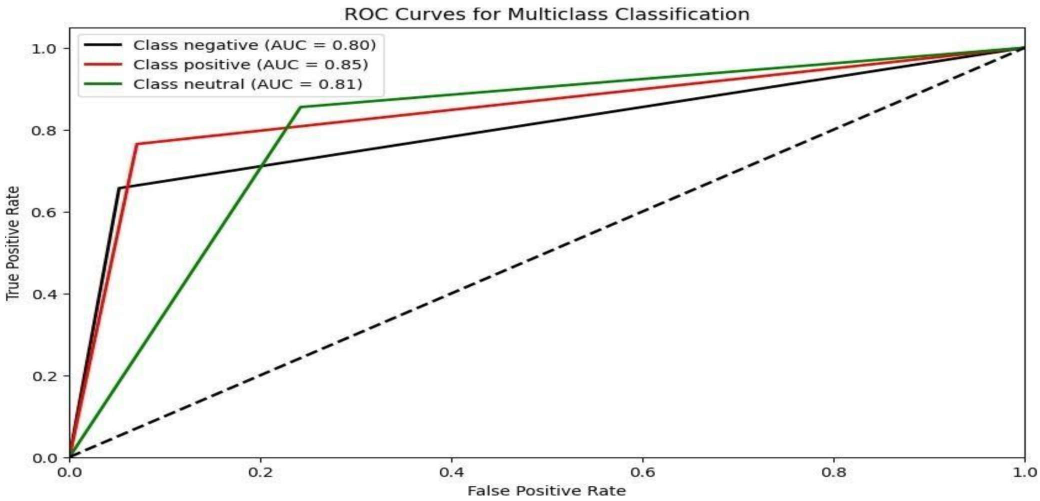

Figure 6: Logistic Regression Model ROC Curve

This section now reports the sentiment analysis score lexicon based approach. Figure 7 shows the sentiment analysis score for the tools SLA, NLTK_SIA, TextBlob and Flair.

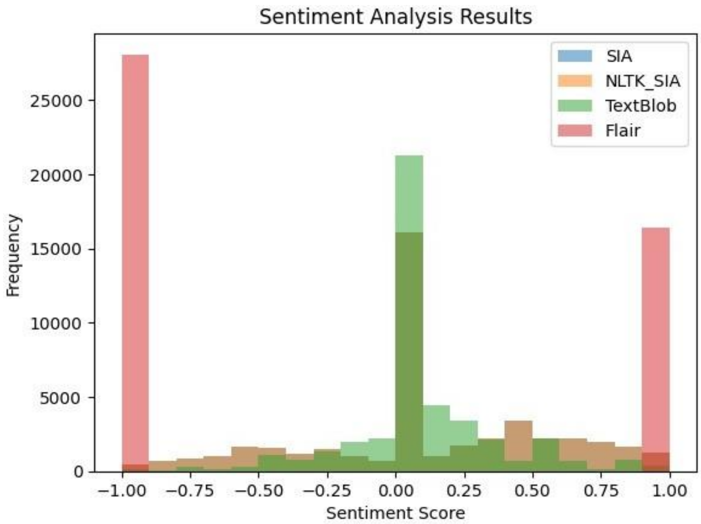

Figure 7: Sentiment Analysis results based on lexicon-based approach

Figure 8 shows the sentiment analysis intensity for SIA. Figure 9 shows the distribution for TextBlob. Figure 10 shows the distribution for NLTK and Flair.

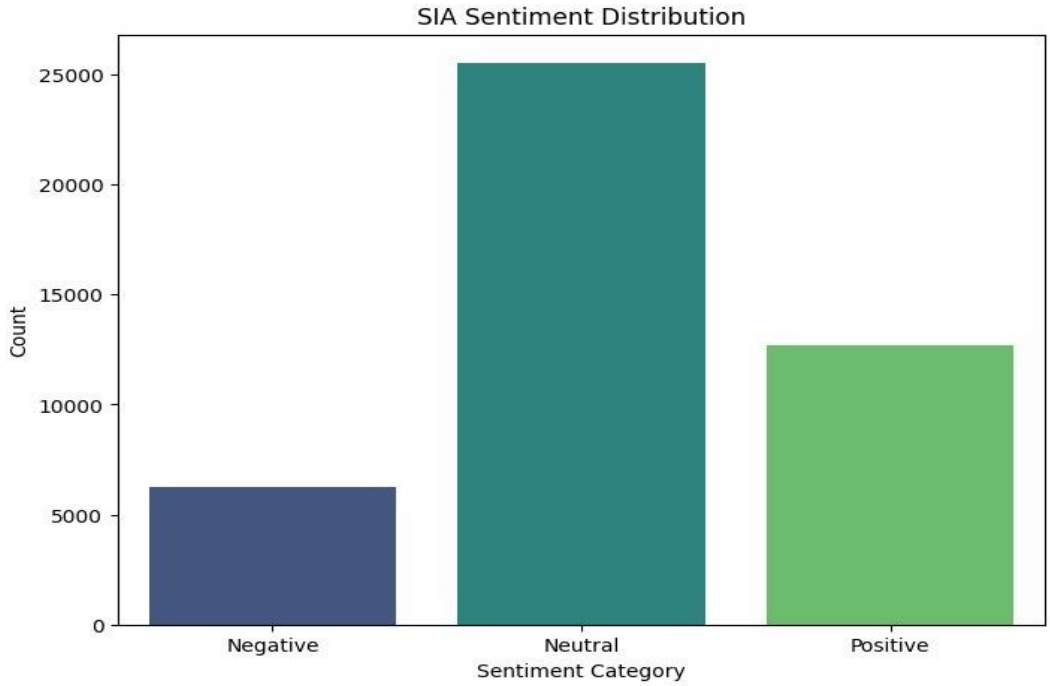

Figure 8: Sentiment intensity analyzer for SIA

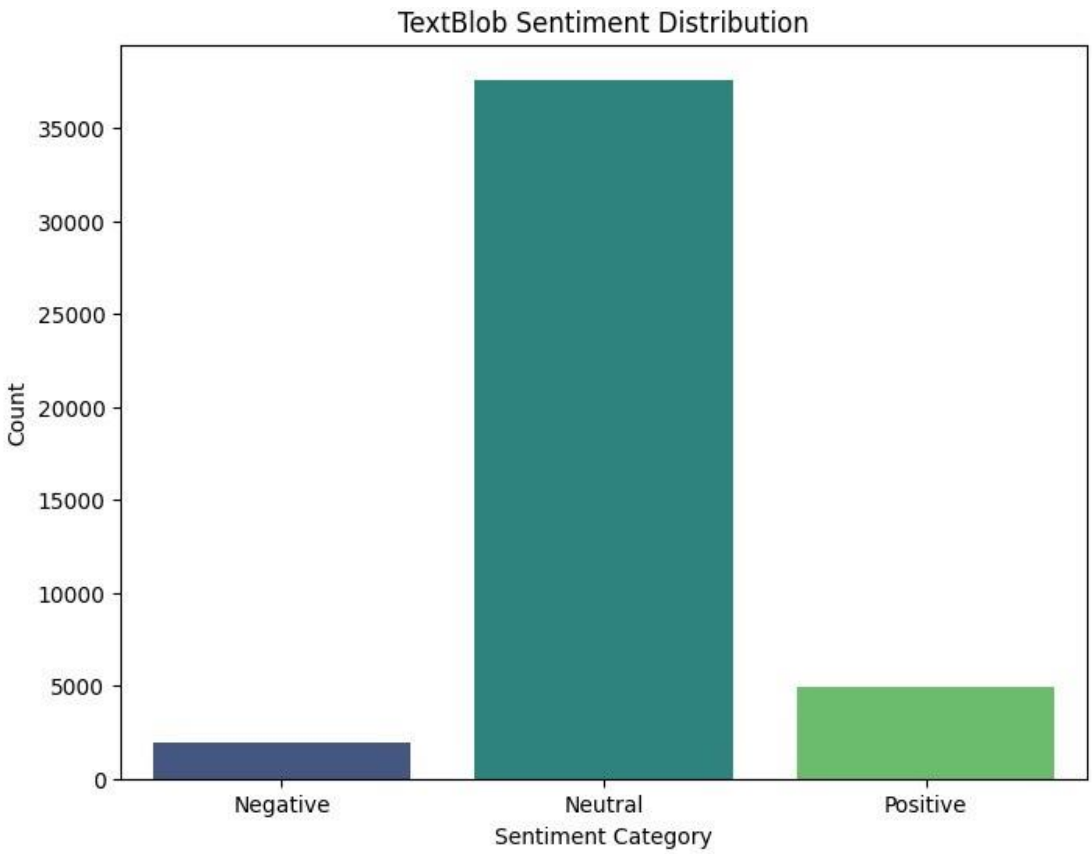

Figure 9: TextBlob Sentiment Distribution

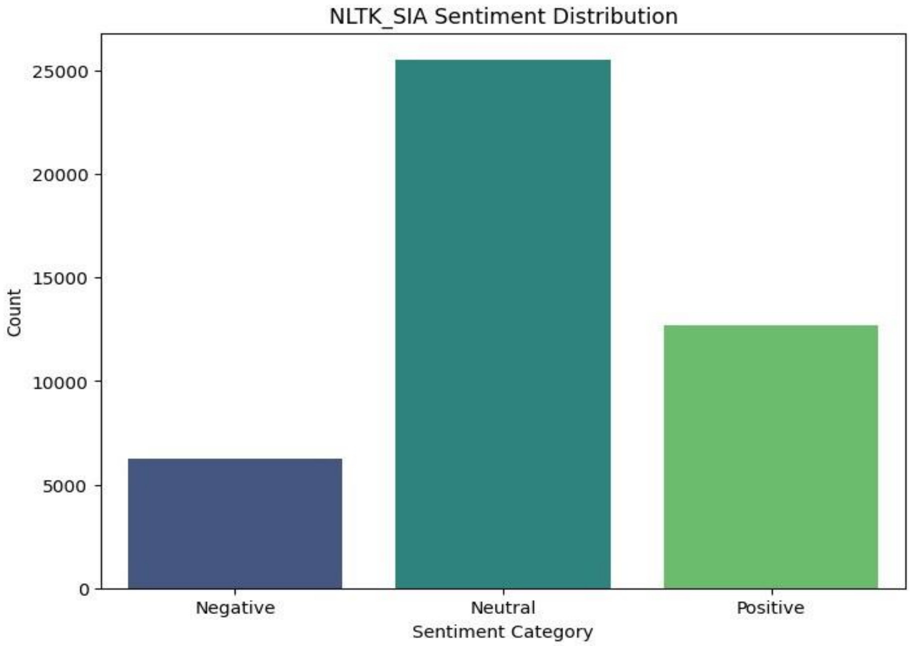

Figure 10: NLTK SIA Sentiment Distribution

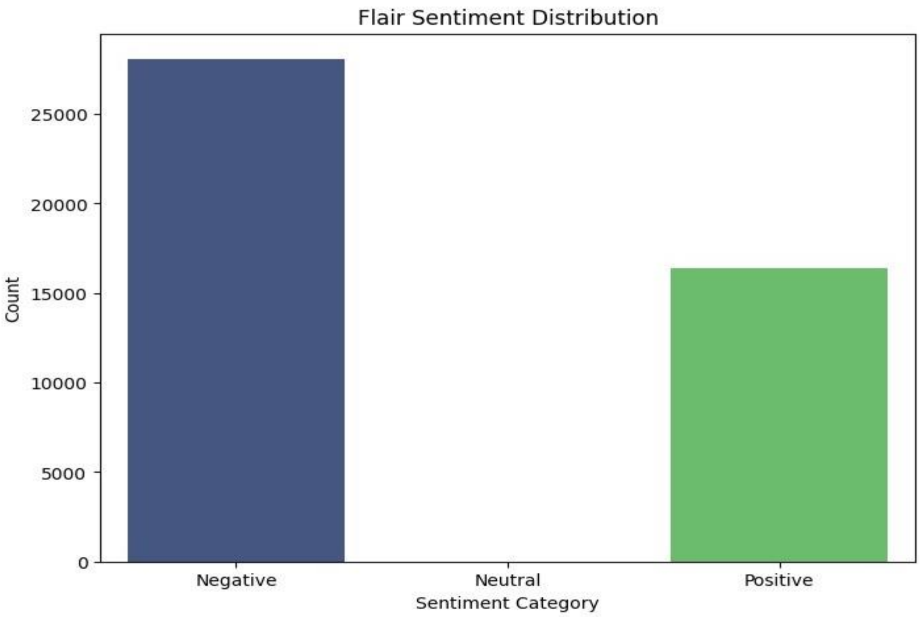

Figure 11: Flair Sentiment Distribution

5.Conclusion

In this research paper, sentiment analysis was conducted on Twitter and reddit data about ChatGPT using machine learning and lexicon-based approaches used. Table 1 shows the results summary. For logistic regression, we have found 81% accuracy and for SVM 78% accuracy has been achieved. The results for precision, recall and F1 score have been discussed in results section. Similar results have been reported in [16].

According to [28], the accuracy of Vader sentiment analysis compared to human classification was 60% for Twitter and 70% for Reddit for machine learning, the paper uses logistic regression and support vector machines. In [16], it is discussed that user attitude is 27% neutral, 20% negative, and 55% positive towards social media. In this study, for NLTK we found that 11% were negative, 59% neutral, and 30% were positive towards ChatGPT. For the analysis of twitter and Reddit data, the Reddit API and PRAW were used along with web scrapping for twitter dat. The analysis provided valuable insights into the sentiment expressed in social media content. Besides machine learning, multiple sentiment analysis algorithms such as those available in NLTK, TextBlob, SIA, Flair, and VADER, were employed to comprehensively evaluate the emotional tone/ sentiments of the text. The sentiment analysis conducted in this study serves as a foundation for further exploration and refinement. Future work may involve fine- tuning models, incorporating domain-specific lexicons, and adapting algorithms to evolving language trends. Below section discusses each of these points in detail.

Table 1: Summary of accuracy for the models for sentiment analysis

Model	Accuracy
Logistic regression	81%
SVM	78%

6. Limitations and Future Work

There are several limitations of proposed study. The findings are based on social media and reddit data. They may not truly represent all populations such as educators, students, and stakeholders in education. The fact of the matter is that social media users are more vocal, inclined towards technology, and engaged more in online discussions. This may lead to biasness in the data. Additionally, the posts on these platforms may not reflect the holistic picture due to several reasons such as online persona, social influence. In addition, the data may not represent minority and underrepresented groups. Some barriers such as language, culture, accessibility may limit their participation.

Future work can focus on fine-tuning the parameters of the sentiment analysis models, including those used in logistic regression, support vector machines, and other algorithms. Hyperparameter optimization techniques, such as grid search or Bayesian optimization, can be applied to enhance model performance. Tailoring sentiment analysis models to specific domains can improve accuracy. Incorporating domain-specific lexicons, terminology, and context-aware features will make the models more adept at capturing nuances within specialized topics or industries.

Ensemble methods, such as combining predictions from multiple models, can be explored. Ensemble learning often leads to more robust and accurate results by leveraging the strengths of individual models. Techniques like stacking or bagging could be applied to create a more powerful sentiment analysis system.

Given the dynamic nature of language and the emergence of new words and phrases, incorporating dynamic language models, such as contextual embeddings (e.g., BERT, GPT), can enhance the ability of sentiment analysis models to grasp evolving language trends and better capture context. Addressing the challenges posed by sarcasm, irony, and other forms of nuanced language is crucial. Developing models that can recognize and interpret these linguistic nuances will contribute to more accurate sentiment predictions.

References

[1] R. H. Mogavi, C. Deng, J. J. Kim, P. Zhou, Y. D. Kwon, A. H. S. Metwally, A. Tlili, S. Bassanelli, A. Bucchiarone, S. Gujar, and L. E. Nacke, "Exploring user perspectives on ChatGPT: Applications, perceptions, and implications for AI-integrated education," *arXiv preprint arXiv:2305.13114*, 2023.

[2] P. Perera and M. Lankathilake, "AI in higher education: A literature review of ChatGPT and guidelines for responsible implementation," 2023. [Online]. Available: [Add URL if applicable]

[3] D. Miller, "Exploring the impact of artificial intelligence language model ChatGPT on the user experience," *Int. J. Technol. Innov. Manage. (IJTIM)*, vol. 3, no. 1, pp. 1–8, 2023.

[4] M. Neumann, M. Rauschenberger, and E. M. Schön, "We need to talk about ChatGPT: The future of AI and higher education," in *2023 IEEE/ACM 5th Int. Workshop Softw. Eng. Educ. Next Gener. (SEENG)*, May 2023, pp. 29–32.

[5] S. Sharma, R. Aggarwal, and M. Kumar, "Mining Twitter for insights into ChatGPT sentiment: A machine learning approach," in *2023 Int. Conf. Distrib. Comput. Electr. Circuits Electron. (ICDCECE)*, Apr. 2023, pp. 1–6.

[6] R. S. J. de Baker and P. S. Inventado, "Chapter X: Educational data mining and learning analytics," *Comput. Sci.*, vol. 7, pp. 1–16, 2014.

[7] W. Holmes, K. Porayska-Pomsta, K. Holstein, E. Sutherland, T. Baker, S. B. Shum, O. C. Santos, M. T. Rodrigo, M. Cukurova, I. I. Bittencourt, and K. R. Koedinger, "Ethics of AI in education: Towards a community-wide framework," *Int. J. Artif. Intell. Educ.*, pp. 1–23, 2022.

[8] B. Pang and L. Lee, "Opinion mining and sentiment analysis," *Found. Trends® Inf. Retr.*, vol. 2, no. 1–2, pp. 1–135, 2008.

[9] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment strength detection for the

social web," *J. Am. Soc. Inf. Sci. Technol.*, vol. 63, no. 1, pp. 163–173, 2012.

[10] Z. Lin, P. Xu, G. I. Winata, F. B. Siddique, Z. Liu, J. Shin, and P. Fung, "CAIRE: An empathetic neural chatbot," *arXiv preprint arXiv:1907.12108*, 2019.

[11] J. E. Brown and J. Halpern, "AI chatbots cannot replace human interactions in the pursuit of more inclusive mental healthcare," *SSM-Mental Health*, vol. 1, p. 100017, 2021.

[12] D. Buckingham and R. Willett, Eds., *Digital Generations: Children, Young People, and the New Media*. Routledge, 2013.

[13] J. J. Voogt, M. C. Kars, E. L. van Rensen, M. M. Schneider, M. Noordegraaf, and M. F. van der Schaaf, "Why medical residents do (and don't) speak up about organizational barriers and opportunities to improve the quality of care," *Acad. Med.*, vol. 95, no. 4, pp. 574–581, 2020.

[14] J. Su and Y. Zhong, "Artificial Intelligence (AI) in early childhood education: Curriculum design and future directions," *Comput. Educ.: Artif. Intell.*, vol. 3, p. 100072, 2022.

[15] N. Wang and M. Johnson, "AI education for K-12: Connecting AI concepts to high school math curriculum," in *Workshop on Educ. Artif. Intell. K-12, 28th Int. Joint Conf. Artif. Intell.*, Jan. 2019.

[16] J. Örnfelt, "Users' attitude towards ChatGPT: A sentiment analysis on Twitter & Reddit," 2023.

[17] M. R. Baker and A. Utku, "Unraveling user perceptions and biases: A comparative study of ML and DL models for exploring Twitter sentiments towards ChatGPT," *J. Eng. Res.*, 2023.

[18] J. Zhou, Z. Liang, Y. Fang, and Z. Zhou, "Exploring public response to ChatGPT with sentiment analysis and knowledge mapping," *IEEE Access*, 2024.

[19] M. Mehdi, N. Islam, A. Irfan, and S. Habib, "Sentiment analysis of public reaction towards Pakistan Democratic Movement," *Pak. J. Eng. Technol.*, vol. 5, no. 2, pp. 66–72, 2022.

[20] Y. Mao, Q. Liu, and Y. Zhang, "Sentiment analysis methods, applications, and challenges: A systematic literature review," *J. King Saud Univ.-Comput. Inf. Sci.*, p. 102048, 2024.

[21] J. O. Krugmann and J. Hartmann, "Sentiment analysis in the age of generative AI," *Customer Needs Solut.*, vol. 11, no. 1, p. 3, 2024.

[22] U. Singh, K. Abhishek, and H. K. Azad, "A survey of cutting-edge multimodal

sentiment analysis," *ACM Comput. Surv.*, vol. 56, no. 9, pp. 1–38, 2024.

[23] R. Koonchanok, Y. Pan, and H. Jang, "Public attitudes toward ChatGPT on Twitter: Sentiments, topics, and occupations," *Social Netw. Anal. Min.*, vol. 14, no. 1, p. 106, 2024.

[24] W. Qi, J. Pan, H. Lyu, and J. Luo, "Excitements and concerns in the post-ChatGPT era: Deciphering public perception of AI through social media analysis," *Telemat. Inform.*, vol. 92, p. 102158, 2024.

[25] Y. Mamo, H. Crompton, D. Burke, and C. Nickel, "Higher education faculty perceptions of ChatGPT and the influencing factors: A sentiment analysis of X," *TechTrends*, vol. 68, no. 3, pp. 520–534, 2024.

[26] I. Hussain, S. M. H. Zaidi, U. Khan, and A. Ahmed, "Sentiment analysis classification of ChatGPT tweets using machine learning and deep learning algorithms," *J. Comput. Biomed. Inform.*, vol. 8, no. 01, 2024.

[27] A. S. Abdelhafiz, A. Ali, A. M. Maaly, H. H. Ziady, E. A. Sultan, and M. A. Mahgoub, "Knowledge, perceptions and attitude of researchers towards using ChatGPT in research," *J. Med. Syst.*, vol. 48, no. 1, p. 26, 2024.

[28] B. Thapa, "Sentiment analysis of cybersecurity content on Twitter and Reddit," in *Proc. Comput. Sci. Inf. Technol. (CSIT)*, pp. 95–107, 2022.

[29] K. N. Hussain, "Exploring user perception and attitudes towards a chatbot system with NLP: A user-centered approach to evaluating chatbot user experience," M.S. thesis, NED Univ. Eng. Technol., 2024.