Int. J. Advanced Networking and Applications
Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

Performance evaluation of GPT for sentiment analysis of movie reviews in Indian Languages

Nikita Desai
Department of Information Technology, Dharmsinh Desai University, Nadiad, Gujarat, India
Email: nikitadesai.it@ddu.ac.in
Ayush Thakor
Department of Information Technology, Dharmsinh Desai University, Nadiad, Gujarat, India
Email : ayushthakor1313@gmail.com
Dev Desai
Department of Information Technology, Dharmsinh Desai University, Nadiad, Gujarat, India
Email : devdesai702@gmail.com
Rajan Patel
Department of Information Technology, Dharmsinh Desai University, Nadiad, Gujarat, India
Email : rajan78053@gmail.com

---**ABSTRACT**--

This study evaluates GPT's performance in sentiment analysis of Indian-language movie reviews, focusing on Hindi and Gujarati. Trained primarily on English data, the model's accuracy in predicting sentiments beyond English is investigated using a 0-shot approach. The observed accuracies present a nuanced landscape of GPT: 61% for Hindi reviews, 59% for Gujarati reviews, 40% for sarcastic Hindi reviews, and 17% for sarcastic Gujarati reviews. The findings reveal varying levels of success, highlighting GPT's strengths and areas for improvement, especially in neutral sentiments and nuanced cases like positive sarcastic reviews. The research establishes a foundation for future NLP advancements in sentiment analysis across diverse languages. By explicitly discussing GPT's results, the study provides valuable insights into its potential as a tool for sentiment analysis in multilingual contexts. Emphasizing the need for continued refinement, the research outlines a roadmap to address unique challenges, marking a crucial step toward optimizing GPT's performance in languages beyond English.

Keywords – **GPT, Indian text, movie review, NLP, sentiment analysis, sentiment classification.**

Date of Submission: 28 November, 2023 Date of Acceptance:

I. INTRODUCTION

Sentiment analysis, a fundamental process in natural language processing (NLP), revolves around the extraction of subjective data and comprehension of expressed emotions within written content. It is widely utilized across domains like social media tracking, market research, analyzing customer feedback, and evaluating movie reviews. Precise sentiment analysis offers valuable perspectives into public sentiment, consumer choices, and user interactions.

Over the past few years, the progress made in the field of sentiment analy- sis has been greatly influenced by the emergence of extensive language models such as GPT (Generative Pre-trained Transformer)[1]. The GPT model, which is built upon the Language Model (LM) architecture, has demonstrated impressive capabilities in producing logical and contextually appropriate text. Nevertheless, its training predominantly centers around English data, leading to uncertainties about its effectiveness when applied to sentiment analysis tasks involving languages other than English. GPT has limited knowledge beyond 2021 and relies on internet resources of varying credibility[2].

Analyzing sentiment refers to locating meaningful information in text and interpreting its meaning using natural processing techniques, identifying whether the writer is positive, negative, or neutral[3] [4]. This study aims to examine how well GPT performs in analyzing sentiments in Indian language texts, particularly Hindi[5] and Gujarati. Given India's linguistic diversity, this research paper explores whether GPT can accurately capture the nuances of sentiment in non-English languages. By assessing GPT's ability to detect and comprehend specific sentiment nuances in Indian languages, we can evaluate its effectiveness for sentiment analysis in these contexts.

To gain a thorough grasp of sentiment analysis in Indian languages, we will begin by examining the current body of literature on sentiment analysis, natural language processing (NLP), and the processing of multilingual texts. Through a thorough analysis of past research on sentiment analysis in Indian languages, employing rule-based techniques, machine learning algorithms, and deep learning models[6], our objective is to understand the advantages and limitations associated with these methods.

To ensure consistent evaluation and comparison of GPT's sentiment analysis performance across different languages, we utilize a meticulously crafted prompt specification for research purposes. The specifics of this prompt specification will be thoroughly addressed in the Results and Analysis section.

Int. J. Advanced Networking and Applications
Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

The results of this research add to the growing understanding of sentiment analysis in multilingual natural language processing. Through the assessment of GPT's ability to analyze sentiments in Indian languages, we offer valuable insights into the strengths and weaknesses of this model. Our objective is to improve the creation of precise and reliable sentiment analysis models for a wide range of languages.

In the following sections of this document, we will outline the dataset characteristics and features used in the experimentation. Additionally, we will elucidate the experimental configuration utilized to evaluate the sentiment predictions made by GPT, present the outcomes and analysis, and engage in a discussion regarding the implications of our discoveries. As per our research, no previous similar work is reported so far.

II. LITERATURE REVIEW

Sentiment analysis, a fundamental task in natural language processing (NLP), has garnered significant attention in recent years. Numerous studies have explored the application of sentiment analysis in various domains, including social media, product reviews, and movie reviews. While much of the research in sentiment analysis has been focused on English text, there is a growing interest in understanding sentiments expressed in languages other than English.

It has also been seen in [7] that ChatGPT's responses varied when evaluating the same task multiple times. Although their research suggested that the differences in successive evaluations are relatively small. Due to this there is a need to address the ethical implications, biases, privacy issues and copyright considerations associated with it which emphasize the importance of AI literacy in understanding the capabilities, limitations, and ethical considerations of AI models like ChatGPT[8] [9].

In the context of Indian languages, sentiment analysis has emerged as a burgeoning field of research. Studies have attempted to address the unique challenges posed by the rich linguistic diversity and cultural nuances of Indian languages. For instance, Das and Bandyopadhyay investigated sentiment analy- sis in Bengali movie reviews, employing machine learning algorithms to achieve promising results[10]. Similarly, Rao et al. explored sentiment analysis for Tamil movie review using deep learning models, highlighting the importance of considering linguistic variations in achieving accurate sentiment predictions[11].

However, it is crucial to acknowledge the limitations of existing sentiment analysis approaches when applied to Indian languages. It is noted by Desai et al, for Gujarati NLP systems, resources like parser, word sense disambiguation, anaphora resolver, and named entity recognizer are still lagging as compared to the English language counterparts [12]. Sharma et al. [13] explored sentiment analysis challenges in Hindi movie reviews and highlighted the need for comprehensive data-sets and model customization for accurate sentiment predictions.

Furthermore, the aspect of sarcasm detection in sentiment analysis has gained traction. Joshi and Sharma addressed sarcasm detection in English text[14], highlighting the importance of understanding contextual cues for accurate sentiment identification. However, the applicability of these techniques to Indian languages, with their unique linguistic characteristics, remains an open question.

III. EVALUATION SETUP

The current research paper focuses on assessing the effectiveness of GPT in analyzing sentiments expressed in movie reviews written in Hindi and Gujarati. The dataset comprises reviews in these languages, encompassing not only conventional positive and negative reviews but also sarcastic ones. Our evaluation primarily relies on the 0-shot results of GPT, wherein it generates outputs without specific training for the target language. The aim is to gauge the performance of GPT in sentiment analysis within this context.

GPT version 3 is used in experimentation. The prompt used for invoking the environment was as follows: *"Classify the overall sentiment expressed in the following movie review (answer as to whether overall sentiment is positive, negative or neutral)"*.

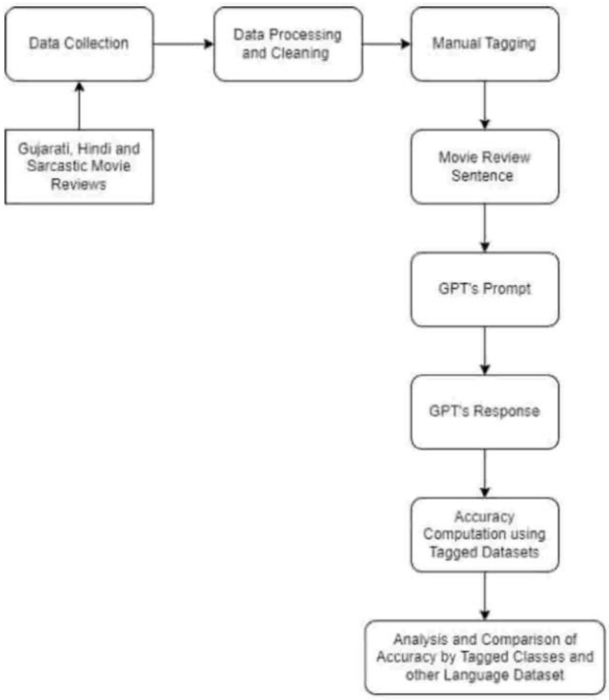

Fig. 1 Architecture flow diagram of Proposed Approach

GPT version 3 is used in experimentation. Of-course the GPT version 4(input capacity of 32k tokens) can give a more accurate rate result then version 3(input capacity of 8k tokens) [15]. The prompt used for invoking the environment was as follows:

- Long Reviews for movies: These reviews have a size ranging from 1000 to 1800 words. These reviews are likely to provide detailed insights and opinions about the movies, covering various aspects such as plot, acting, direction, and more. The data was collected from various sources like \cite{GujaratiReviews}, \cite{HindiReviews}. The reviews

Int. J. Advanced Networking and Applications
Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

were manually tagged with one of the 3 possible categories of sentiments.

- Sarcastic Reviews for movies: This dataset consists of very simple and short sarcastic one-liner movie reviews, typically of 15 to 20 words. Sarcastic reviews are often characterized by expressing the opposite of what the reviewer means, intending humor or irony. The sarcastic reviews were handcrafted with help from a language expert and also tagged as positive or negative as per the case.

Table 1. Dataset Characteristics

Data-Set Type	Language	Total Reviews	Positive	Negative	Neutral
Long Reviews	Hindi	895	350	273	272
Long Reviews	Gujarati	453	206	134	113
Sarcastic Reviews	Hindi	35	9	26	NA
Sarcastic Reviews	Gujarati	35	9	26	NA

For the long reviews, the goal of the evaluation is to test the GPT tool's ability to correctly classify the overall sentiment expressed in the movie reviews into three categories: positive, negative, or neutral. As sarcasm is only for flipping the polarity from positive to negative and vice-versa, we consider the sarcastic reviews to be classified into only two groups namely positive and negative. Further, one additional class of CND (Cannot Determine) is also used. This tag is assigned to indicate the model's inability to make a prediction..

IV. Results and Analysis

The GPT model's overall accuracy in sentiment classification for movie reviews is assessed. It represents the collective performance of the model in correctly classifying reviews into positive, negative, and neutral sentiments.

IV.I. Long(Non-Sarcastic) Movie Reviews.

IV.I.I. Accuracy (Hindi Reviews):

The GPT model achieved an accuracy of 75.00% in classifying Hindi movie reviews correctly for the positive reviews. The GPT model's lowest accuracy of 38% is in predicting the neutral reviews. In 67% of negative reviews, the model was successful in predicting the negative class.

Table 2. Confusion Matrix of Hindi movie reviews

			Predicted Tag				
Expected Tag		Total	Positive	Negative	Neutral	CND
	Positive	348	261	9	39	39
	Negative	271	33	181	38	19
	Neutral	269	44	106	101	18

IV.I.II. Failure Analysis (Hindi Reviews):

The GPT model's failure rate for opposite predictions, where it erroneously classifies reviews with sentiments opposite to their true labels (e.g., positive reviews misclassified as negative and negative reviews misclassified as positive), is reported as 3% and 12% respectively (as illustrated in figure 2). The reasons behind these opposite predictions, such as linguistic complexities and contextual ambiguities, are discussed in the next section.

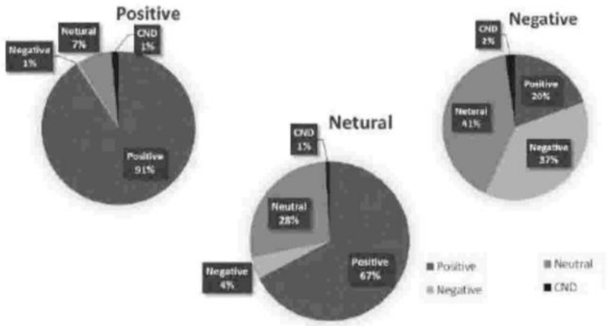

Fig. 2 GPT's predictions for positive, negative and neutral Hindi reviews

GPT exhibits significant failures in predicting positive and negative reviews as neutral. Specifically, it incorrectly classifies 11% of positive reviews as neutral and 14% of negative reviews as neutral as shown in figure 2. These errors indicate the model's challenges in accurately discerning the true sentiment of reviews, leading to misinterpretations where positive and negative content is mistakenly identified as neutral, which will be further explored in the subsequent section. The GPT model exhibits the highest level of uncertainty in classifying neutral reviews, where it predicts 16% of them as positive and 39% as negative as illustrated in figure 2. This indicates the challenges faced by the model in accurately determining the sentiment of neutral content, leading to ambiguous predictions with varying degrees of positivity and negativity.

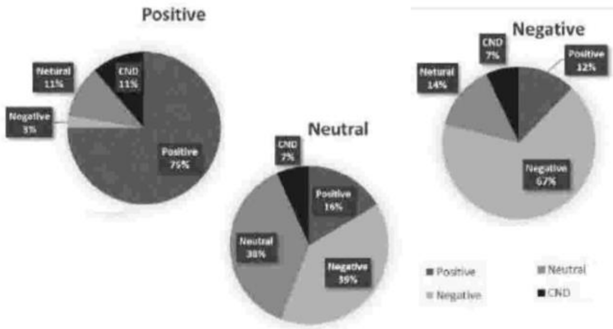

Fig. 3 GPT's predictions for positive, negative and neutral Gujarati reviews

IV.I.III. Accuracy (Gujarati Reviews):

The GPT model can give accuracy of 59.16%. Specifically, the GPT model achieved an accuracy of 91% for positive reviews. This indicates the proportion of positive reviews that the model accurately identified, showcasing its proficiency in recognizing positive sentiments. The GPT model's accuracy in classifying neutral

Int. J. Advanced Networking and Applications
Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

reviews is only 28%. While the model achieved an accuracy of 37% in classifying negative Gujarati reviews.

Table 3. Confusion Matrix of Gujarati movie reviews

Expected Tag		Predicted Tag				
	Total	Positive	Negative	Neutral	CND
Positive	206	187	1	15	3
Negative	134	26	50	55	3
Neutral	113	76	5	31	1

IV.I.IV. Failure Analysis (Gujarati Reviews):

The GPT model demonstrates a 1% failure rate for opposite predictions, where positive reviews are erroneously classified as negative, and a 20% failure rate for negative reviews misclassified as positive (as depicted in figure 3). These opposite predictions can be attributed to various reasons, including linguistic complexities and contextual ambiguities, which will be further explored in the subsequent section. GPT's performance shows notable shortcomings when predicting positive and negative reviews as neutral. It inaccurately classifies 7% of positive reviews as neutral and 41% of negative reviews as neutral, as illustrated in figure 3. These errors highlight the model's difficulties in accurately discerning the genuine sentiment of reviews, resulting in misinterpretations where positive and negative content is mistakenly identified as neutral, are discussed in the next section. The GPT model exhibits the highest level of uncertainty in classifying neutral reviews, where it predicts 67% of them as positive as shown in figure 3. This indicates the challenges faced by the model in accurately determining the sentiment of neutral content, leading to ambiguous predictions with varying degrees of positivity and negativity.

IV.II. Sarcastic Movie Reviews

Table 4. Confusion Matrix of sarcastic Hindi movie reviews

Expected Tag	Predicted Tag			
	Total	Positive	Negative	Neutral
Positive	9	0	9	0
Negative	26	7	14	5

Table 5. Confusion Matrix of sarcastic Gujarati movie reviews

Expected Tag	Predicted Tag			
	Total	Positive	Negative	Neutral
Positive	9	0	8	1
Negative	26	10	16	10

IV.II.I. Accuracy (Sarcastic reviews):

GPT's performance in classifying sarcastic reviews varies across sentiments and languages. It struggles with positive sarcastic reviews, showing 0% accuracy due to challenges in understanding context and tone. However, it achieves moderate success with 53.85% accuracy in negative sarcastic Hindi reviews and relatively low performance with 23.08% accuracy in negative sarcastic Gujarati reviews. Limited exposure to Gujarati text during training and the complexity of Gujarati sarcasm may have contributed to this.

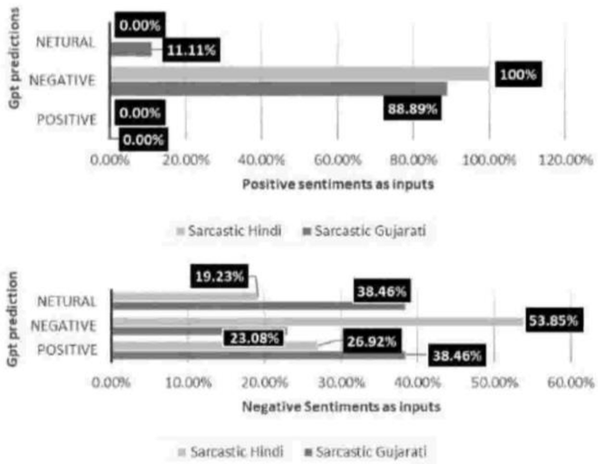

Fig. 4 GPT's predictions for positive and negative sarcastic reviews

IV.II.I. Failure Analysis (Sarcastic reviews):

GPT fails to predict positive sarcastic reviews in Hindi and Gujarati, misclassifying all positive Hindi sarcastic reviews as negative (100% failure) and 88.89% of positive Gujarati sarcastic reviews as negative, with 11.11% as neutral. The model struggles with negative sarcastic reviews, with a failure rate of 46.15% for Hindi and an even higher 76.92% for Gujarati. Improving sarcasm detection requires addressing these challenges and incorporating diverse and relevant training data for both languages.

V. Observations

V.I. Observation on Long Movie Review predictions by GPT

V.I.I. Word limit issue:

GPT struggles to generate responses for lengthy Hindi movie reviews lacking proper paragraph formatting. In such cases, it may produce inconclusive results like *"Hmm...something seems to have gone wrong."*. We found these in 8.49% Hindi reviews and 1.54% Gujarati reviews. Therefore, we have categorized these reviews as Can Not be Determined (CND). The optimal word limit for a Gujarati movie review that GPT can effectively respond to is approximately 350 words. However, it's worth noting that there are variations in the number of characters per word, with some words being longer and others shorter. As a result, a rough estimate based on character count suggests that GPT can respond to reviews of up to 1600 characters (excluding spaces). Nevertheless, it is important to consider that the appropriateness of the response may vary depending on the complexity and length of the text. On average, it can

Int. J. Advanced Networking and Applications
Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

be assumed that GPT can handle approximately 1500 characters.

V.I.II. ERRORS OF OPPOSITE PREDICTIONS:

In analyzing Gujarati and Hindi movie reviews with multiple sentences, GPT primarily prioritizes movie-specific content like actor/actress performance and director's character, while disregarding general review portions. This focused approach ensures that the model avoids predicting opposite sentiments, preventing abrupt shifts from positive to negative or vice versa.

It is exceptionally rare, occurring only once within our datasets, for GPT to predict a positive movie review as negative (1% in Hindi and 0.22% in Gujarati Movie Reviews). Conversely, we frequently observed instances where the model predicts a shift from negative to positive sentiment (3.69% in Hindi and 5.74% in Gujarati Movie Reviews). Interestingly, even when the positive sentences, typically related to the movie or actor's performance, are positioned at the beginning of the review, the model consistently maintains a positive prediction. Surprisingly, when we experiment by relocating those positive sentences to the end of the review instead, we find that all predictions remain unaltered, with the model persistently predicting a positive sentiment.

– Negative review given Positive classification:

Example 5.1.1: तारीफ करनी होगी इस फिल्म के निर्माताओं की, जिन्होंने ऐसे अहम सब्जेक्ट पर पूरी ईमानदारी के साथ फिल्म बनाई।*(The makers of this film have to be praised, who made a film on such an important subject with complete honesty.)*

Example 5.1.1 shows a sentence from a negative Hindi movie review, one sentence that praises the director's integrity GPT gives a positive class, even though the overall sentiment of the movie was negative.

– Positive review given Negative classification:

Example 5.1.2: गालियों में महिला के बारे में ही बुरा क्यों कहा जाता है, इस पर भी वे रोष जताती हैं और उनकी महिला पात्र ऐसी गालियां बकती हैं जिनमें पुरुषों को बुरा कहा जाता है। ये तीनों महिलाएं आपस में सहेलियां हैं और आपसी दु:ख को बांटती रहती हैं। दु:ख से भरी जिंदगी में कुछ पल खुशियों के भी वे चुरा लेती हैं।

The 5.1.2 is a sentence from a positive Hindi movie review, where due to usage of negative words like "दु:ख" (sadness) repeated often, GPT misclassified the review as negative.

V.I.III. ERRORS OF NEUTRAL CLASS:

When there are multiple flow changes in sentences from positive to negative, GPT may struggle to accurately predict neutrality and tends to mis-predict such reviews as positive. This limitation arises due to the model's complexity and its reliance on the context within individual sentences. GPT's predictive behavior may vary, and it might be sensitive to the ordering and nuances of the provided text, leading to biased predictions in favor of positive sentiments, as seen in below example 5.1.3

Example 5.1.3: ફિલ્મનું ટાઇટલ ફરાઝ છે પરંતુ તેનું શૌર્યઅંતમાં દેખાય છે. તેની બેકસ્ટોરી પર થોડું કામ કરવું જરૂરી હતું. ભાવનાત્મક દ્રષ્ટિએ ફિલ્મ હચમચાવી નાખે તેવી નથી. જોકે, એ વાત તો સ્વીકારવી રહી કે, ફિલ્મની નિયત સારી અને પ્રાસંગિક છે. લેખનની અમુક મર્યાદા દેખાઇ આવે છે. *(The title of the film is Faraaz but his heroism shows up at the end. His backstory needed some work. Emotionally, the film is not earth shattering. However, it has to be admitted that the film's plot is good and relevant. Some limitations of the writing are visible.)*

In some Hindi movie reviews at the end there is one section named "क्यों देखें" *(Why to watch the movie)*,as shown in example 5.1.4. The purpose of that section is to state some reasons why a person should watch the movie. However, due to the positioning of this section, GPT wrongly predicts the class of the review. Example 5.1.4: क्यों देखें: बस अगर आप शाहरुख खान के पक्के फै न हैंतो अपने चहेते स्टार का यह बदला रूप भी देख आए।" *(Why to watch:only if you are Sharukh khan's die hard fan then go to see his new form.)*

The below review in Example 5.1.5 is a snippet of the final lines of a negative movie review. But GPT struggles to detect the negativity accurately, leading to its misclassification of the overall sentiment as neutral.

Example 5.1.5 : 'बदमाश कंपनी' की सबसे बड़ी प्रॉब्लम इसकी स्क्रिप्ट है। परमीत सेठी ने इसे अपनी सहूलियत के हिसाब से लिखा है। मन मुताबिक कहानी को ट्विस्ट दिए हैं, भले ही वो विश्वसनीय और सही नहीं हो। इससे दर्शक स्क्रीन पर दिखाए जा रहे घटनाक्रमों से जुड़ नहीं पाते। *(The biggest problem of Badmaash Company is its script. Parmeet Sethi has written it according to his convenience. Twists have been given to the story according to the mind, even if it is not reliable and correct. Due to this, the audience does not get connected with the events shown on the screen.)*

In the positive reviews like example 5.1.6, some critics acknowledged certain flaws in the movie; however, they still recommended watching it, indicating a neutral stance overall. Such a case is seen in following example 5.1.6

Example 5.1.6: कुछ कमियों के बावजूद इस शुद्ध देशी फिल्म को एक बार देखा जा सकता है। *(Despite some shortcomings, this pure country film can be seen once.)*

V.I.II. ISSUES WITH THE NEUTRAL REVIEWS CLASSIFICATION IN GPT:

In neutral reviews, there are often an unbalanced number of positive and negative sentences. If there are more negative sentences mentioned, the prediction tends to be classified as a negative review. On the other hand, if there are more positive sentences, the prediction is labeled as a positive review. It appears that the presence of specific keywords like "એવરેજ" *(average)* and star ratings such as "2.5 out of 5" in Gujarati reviews is not consistently guiding GPT's sentiment prediction accurately. The model might be relying more on other sentences in the review, leading to mispredictions. For instance, in below example 5.1.7 even though this final sentence of review indicates a neutral stance, the overall sentiment predicted by GPT is positive. Example 5.1.7: આ ફિલ્મને 2.5 મિર્ચીઝ આઉટ ઓફ 5.*(2.5 mirchies(points) out of 5 for this movie)*

These issues are indicative of the limitations of language models like GPT when it comes to understanding context and context-specific keywords. While GPT has shown impressive language generation capabilities, it may still struggle with context-based sentiment analysis, especially when the sentiment is conveyed through non-explicit or nuanced expressions.

Int. J. Advanced Networking and Applications
Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

V.II. Observation of Sarcastic reviews

The performance of GPT in classifying positive sarcastic reviews is inadequate, as it tends to misclassify them as negative reviews. The primary challenge lies in handling simple negation in Hindi movie reviews, where phrases like "अगर आपको अच्छी एक्टिंग पसंद नहीं हैतो आपको येफिल्म बिल्कुल नहीं देखनी चाहिए" જો તમને સારી એક્ટિંગ પસંદ ન હોય તો તમારેઆ ફિલ્મ બિલકુલ ન જોવી જોઈએ. *(If you don't like good acting then you should not watch this movie at all.)* use negative words but actually express a positive sentiment. The model struggles to discern such nuances, resulting in misclassifications. Improving GPT's understanding of negation and sarcasm in Hindi/Gujarati movie reviews is crucial for enhancing its sentiment classification accuracy.

VI. Conclusion and Future work

The research conducted on sentiment analysis of movie reviews using GPT-3.5 in Indian languages highlights notable observations regarding sentiment detection. While the model demonstrates commendable accuracy in identifying positive sentiments, challenges arise in understanding negative and neutral sentiments. Additionally, the inclusion of sarcastic reviews tends to skew the predictions toward positive sentiments, posing further complexities in the analysis. Moreover, the presence of sarcastic language adds complexity to the sentiment analysis process. The model's tendency to misinterpret sarcastic reviews as positive sentiment calls for specialized strategies to identify and handle sarcasm effectively. Developing techniques to distinguish sarcastic language from genuine positive expressions can significantly enhance the model's overall performance. To address the challenges associated with negative and neutral sentiment identification, a more extensive and diverse data-set is required. Expanding the data-set to include a broader range of negative and neutral sentiments will aid in fine-tuning the model and improving its accuracy.

Moving forward, future research should concentrate on enhancing GPT-3.5's performance in identifying negative and neutral sentiments in Indian languages. Fine-tuning the model on a diverse data-set, specifically curated to include a wide range of negative and neutral expressions, is paramount to improving accuracy beyond positivity. Additionally, specialized strategies to handle sarcastic language should be developed, focusing on sarcasm detection to differentiate between genuine positive sentiments and sarcastic expressions. Domain-specific training on movie-specific data-sets can also be explored to optimize the model for movie review sentiment detection. Addressing data availability concerns by creating larger and more diverse data-sets in Indian languages will further contribute to improved sentiment analysis. Lastly, researchers should delve into interpretability techniques for model decision-making transparency, and multi-model approaches combining GPT-3.5 with other sentiment analysis techniques, which may offer more robust sentiment analysis capabilities. Pursuing these research directions will advance sentiment analysis in Indian languages, yielding accurate and reliable solutions for diverse linguistic landscapes

References

[1] Wang, Z., Xie, Q., Ding, Z., Feng, Y., Xia, R.: *Is chatgpt a good sentiment analyzer? a preliminary study.* arXiv preprint arXiv:2304.04339 (2023)

[2] Doskaliuk, B., Zimba, O.: *Beyond the keyboard: Academic writing in the era of chatgpt.* Journal of Korean Medical Science **38**(26) (2023)

[3] Nandwani, P., Verma, R.: *A review on sentiment analysis and emotion detection from text. Social Network Analysis and Mining* **11**(1), 81 (2021)

[4] Rani, S., Kumar, P.: *A journey of indian languages over sentiment analysis: a systematic review. Artificial Intelligence Review* **52**, 1415–1462 (2019)

[5] Sharma, A., Ghose, U.: *Toward machine learning based binary sentiment classi- fication of movie reviews for resource restraint language* (rrl)–hindi. *IEEE Access* (2023)

[6] Shah, P., Swaminarayan, P., Patel, M.: Sentiment analysis on film review in gujarati language using machine learning. *International Journal of Electrical and Computer Engineering* **12**(1), 1030 (2022)

[7] Jukiewicz, M.: *The future of grading programming assignments in education: The role of chatgpt in automating the assessment and feedback process*

[8] Tai, A.M.Y., Meyer, M., Varidel, M., Prodan, A., Vogel, M., Iorfino, F., Krausz, R.M.: Exploring the potential and limitations of chatgpt for academic peer- reviewed writing: Addressing linguistic injustice and ethical concerns. *Journal of Academic Language and Learning* **17**(1), T16–T30 (2023)

[9] A El-Seoud, S., Ayman, S.E., Nagaty, K., H Karam, O.: The impact of chatgpt on student learning/performing. Karam, Omar, *The Impact of ChatGPT on Student Learning/performing* (August 6, 2023) (2023)

[10] Hassan, A., Amin, M.R., Mohammed, N., Azad, A.: *Sentiment analysis on bangla and romanized bangla text (brbt) using deep recurrent models*. arXiv preprint arXiv:1610.00369 (2016)

[11] Shanmugavadivel, K., Sampath, S.H., Nandhakumar, P., Mahalingam, P., Subra- manian, M., Kumaresan, P.K., Priyadharshini, R.: An analysis of machine learning models for sentiment analysis of tamil code-mixed data. *Computer Speech & Language* **76**, 101407 (2022)

[12] Desai, N.P., Dabhi, V.K.: *Resources and components for gujarati nlp systems: a survey.* Artificial Intelligence Review **55**(7), 1–19 (2022)

[13] Pradhan, R., Sharma, D.K.: An ensemble deep learning classifier for sentiment analysis on code-mix hindi–english data. *Soft Computing p.* 1 (2022)

Int. J. Advanced Networking and Applications
Volume: 15 Issue: 04 Pages:100-200 (2023) ISSN: 0975-0290 (Special Issue – ASCIS 2023)

[14] Savini, E., Caragea, C.: A multi-task learning approach to sarcasm detection (stu- dent abstract). In: Proceedings of the AAAI *Conference on Artificial Intelligence.* pp. 13907–13908. No. 10 in 34 (2020)

[15] de Winter, J.C.: *Can chatgpt pass high school exams on english language compre-* hension. Researchgate. Preprint (2023)

[16] Bennett, C..C.L.: https://www.iamgujarat.com (2023)

[17] Webdunia.com: https://hindi.webdunia.com/ (2023)

[18] Sohail, S.S., Madsen, D.Ø., Farhat, F., Alam, M.A.: *Chatgpt and vaccines: Can ai chatbots boost awareness and uptake?* Annals of Biomedical Engineering pp. 1–5 (2023)

Author Biographies

Nikita Desai is an experienced Associate Professor with a demonstrated history of working in the education industry for more than 22 years. Skilled in C++, Artificial Intelligence (AI), Data Structures, Natural Language Processing (NLP), and Algorithms. She is a strong education professional with a Masters of Engineering focused in Computer Engineering from Dharmsinh Desai University, Nadiad, Gujarat, India.

Ayush Thakor, a student at Dharmsinh Desai University, is passionate about technology and eager to learn. With a strong focus on website development, artificial intelligence, NLP, and machine learning, Ayush possesses robust problem-solving skills. Dedicated to staying abreast of the latest industry trends, he brings a keen interest and strong technical skills to the field. Ayush's academic background in Information Technology underlines his commitment to contributing to the ever-evolving tech industry.

Dev Desai, an Information Technology student at Dharmsinh Desai University, is passionate about software development, NLP, artificial learning, blockchain, and web development. Known for his strong problem-solving skills, Dev is dedicated to achieving excellence in technology and making valuable contributions to the field.

Rajan Patel, student of Information Technology at Dharmsinh Desai University, passionate about Data Science, Machine Learning and Computer Vision analytics. Strong technical background with deep interest in entrepreneurship and strong soft skills. Rajan wants to contribute his technological learnings to integrate it with medical science in order to make it more compatible.