2024 6th International Conference on Sustainable Technologies for Industry 5.0 (STI), 14-15 December, Dhaka

Detecting Discrepancy to Enhance the Accuracy of Sentiment and Emotion Analysis: A Study Based on Bangla E-commerce Dataset

Md Tanbeer Jubaer*, Barisha Chowdhury †, Md Minhazul Islam‡, Mohammad Harun Or Rashid§
*†‡*Department of Computer Science & Engineering*
§*Department of Humanities*
Rajshahi University of Engineering and Technology, Rajshahi, Bangladesh
E-mails: tanbeerjubaer@gmail.com, barishachowdhury24@gmail.com,
minannu.2001@gmail.com, harunmr9@gmail.com

***Abstract*—The growth of e-commerce platforms has led to a huge amount of data, providing useful insights into consumer behavior and preferences. However, the quality of this data is often affected by discrepancies, inaccuracies, and noise, making reliable sentiment and emotion analysis difficult. This paper proposes a novel approach to improve sentiment and emotion analysis in Bangla E-commerce dataset by first identifying and mitigating discrepancies. Using a robust approach these inconsistencies are thoroughly identified and resolved. Next, we apply state of the art model BERT for sentiment analysis and emotion recognition on the cleaned dataset, extracting sentiment polarity and emotion categories from text data related to Bangla e-commerce transactions.**

The experimental results show significant improvements in sentiment and emotion analysis accuracy after removing discrepancies. Initially, sentiment analysis achieved an accuracy of 93.96%, but this increased to 96.31% after addressing the discrepancies. Similarly, emotion analysis accuracy improved from 90.94% to 93.59% after cleaning the dataset. By conducting comparative analyses with traditional methods, we demonstrate the effectiveness of our approach in improving insights from e-commerce data. Our study highlights the importance of discrepancy detection as a key preprocessing step in e-commerce data analysis. By correcting data inconsistencies before sentiment and emotion analysis, we enable a more accurate understanding of consumer sentiments and emotions.

***Index Terms*—E-commerce, Discrepancy, Sentiment, Emotion, Review, BERT.**

I. Introduction

The post-pandemic era has overserved the booming of online-based business and e-commerce [1] where user-generated online reviews play a pivotal role in shaping consumer decisions across a myriad of domains. These reviews are often based on both numerical ratings and textual feedback, providing valuable insights into product quality, service satisfaction, and overall user experience. However, an intriguing phenomenon observed in such reviews is the occasional misalignment between the numerical rating assigned by users and the opinion expressed in their accompanying textual feedback. This difference between numerical and textual reviews has been conceptualized as a discrepancy during the analysis of such reviews. The discrepancy can stem from diverse underlying factors such as semantic misalignment, subjectivity, interpretation, contextual factors, reviewer bias, incentives, etc [2] [3] [4] [5]. This issue has already been addressed in English-centric NLP applications such as sentiment analysis, topic modeling, opinion mining, etc [6]. However, it remains unaddressed in Bangla-related applications. Several studies have tried to analyze the sentiment of Bangla reviews using different Machine-learning [7] [8] and deep learning [9] approaches. This paper aims to address the intricacies of discrepancy in the context of the Bangla language, an area that has not received much attention up to this point. This study further focuses on how detecting and scrapping the anomalies contributes to an enhanced accuracy of reviewers' sentiments, which will thus confirm a more reliable e-commerce platform as well as better decision-making of online business platforms. The state-of-the-art model Bangla BERT base [10] was employed to analyze the sentiment and emotion of the reviews with discrepancies and without discrepancies from a well-annotated dataset [11]. The model accuracy notably increased when the inconsistencies were addressed accordingly. It is expected that this process can enhance the interpretability and reliability of user-generated reviews, sentiments, and content analysis.

II. Literature Review

E-commerce has gained significant popularity, leading to extensive research in areas such as sentiment analysis, emotion analysis, and customer behavior analysis. To identify the underlying sentiment and emotion of the user evaluations, this study examines the impact of discrepancy and proposes a method that combines discrepancy detection with emotion and sentiment analysis.

Understanding discrepancies between ratings and reviews is crucial for accurate sentiment analysis. A noteworthy study by [12] analyzed 8600 reviews from the app store and discovered a 20% discrepancy between ratings and reviews. This highlights that customer reviews may not always be

979-8-3315-3197-3/24/$31.00 ©2024 IEEE

entirely fair. In this study [13], the authors examined various studies and emphasized the importance of considering discrepancies in sentiment analysis. They claimed that assuming a direct correlation between reviews and ratings is inaccurate, suggesting the need to assess the correlation between them for more accurate analysis.

Studies have demonstrated the value and way of detecting discrepancies between user reviews and ratings for sentiment and emotion analysis. Further research [6] identified discrepancies between user reviews and ratings. Using CNN, BiLSTM, GRU, and RNN models, they calculated sentiment scores and review ratings to detect discrepancies. Based on these scores, they designed an algorithm to distinguish between biased and unbiased ratings. Additionally, in [14], a custom CNN model was proposed to detect inconsistencies. Comparing the model output to an empirically separated dataset, their model achieved an impressive accuracy of 95%. These findings underscore the importance of detecting discrepancies for more precise sentiment and emotion analysis.

Advancements in sentiment and emotion analysis have increasingly leveraged machine learning (ML) techniques for improved performance. [15] conducted sentiment analysis on E-commerce feedback using CountVectorizer, TF-IDF, and classifiers like LR, MNB, and SVM, achieving 94% accuracy. Another study [16] compared ML approaches, finding SVM and Naive Bayes to be most effective. A review in [17] found deep learning (DL) models, such as RNN with attention mechanisms, outperforming traditional ML models, which sometimes struggle with implicit data features.

Recent research underscores DL's effectiveness in sentiment analysis. For example, [18] achieved 66.06% accuracy with a DL model on an imbalanced dataset, surpassing A-LSTM. A hybrid LSTM-CNN model in [19] reached 90.49% accuracy in emotion detection, while [20] used fast-text and multichannel CNN, attaining 79.83% validation accuracy. These studies highlight DL's superior capability for sentiment and emotion analysis.

In a study [21], sentiment analysis was applied to e-commerce data in Arabic, English, and Turkish, with RNN achieving the highest accuracies: 85.3%, 89.5%, and 87.1%, respectively, outperforming other models. Similarly, [22] developed a hybrid RoBERTa-LSTM framework for multilingual sentiment analysis across French, Arabic, and English, using oversampling techniques like RUS, SMOTE, and GPT to handle data imbalance. Both studies demonstrate the use of advanced machine learning approaches in e-commerce sentiment analysis, showcasing the adaptability of these models across different languages for more accurate sentiment prediction.

In the era of transformers, these models have surpassed both deep learning (DL) and machine learning (ML) models for sentiment classification. In [23], SkipBangla-BERT was employed for feature extraction, demonstrating the highest accuracy among various models. Another related study [24] yielded promising results by analyzing modern techniques such as Transfer Learning (LSTM, GRU) and transformers (Bangla-BERT). They found that Bangla-BERT outperformed other models for both binary and multiclass classification. In a curated e-commerce dataset, they achieved impressive accuracies of 94.5% for binary classification and 88.78% for multiclass classification. These findings underscore the remarkable performance of transformers in sentiment analysis tasks.

After reviewing these papers, it's evident that there is an opportunity to enhance user sentiment analysis by first detecting discrepancies, then removing these discrepancies, and applying an efficient model with transformers architecture to predict user sentiment and emotion. Combining these two steps discrepancy detection and sentiment detection creates a research gap and opens up a new horizon to assist e-commerce websites in showing more relevant products to customers. This not only boosts their reliability but also enhances sales by catering to the preferences and sentiments of users more accurately.

III. Materials and Methods

A. Dataset Description

Our research utilizes a comprehensive, well-annotated dataset for sentiment and emotion classification from Bangladeshi e-commerce reviews [11]. This dataset, containing 78,130 product reviews from Daraz and Pickaboo, captures consumer sentiment and emotions in Bangladesh's growing online market. The dataset's uniqueness lies in its bilingual content, with reviews in both Bengali and English, offering valuable insights into diverse customer feedback.

The data is annotated with five emotion labels—Happiness, Sadness, Fear, Love, and Anger—and divided into Positive and Negative sentiment classes. Following Shaver's model [25], emotions of Love and Happiness are grouped as Positive, while others are categorized as Negative. To simplify, the study merged Love and Happiness into a single class, 'Happy,' due to their similarity. The dataset was split into 80% training and 20% testing data for analysis.

B. Preprocessing

Messy data can lead to incorrect conclusions, underscoring the need for data preprocessing in data mining. This process removes irrelevant or redundant information that could confuse the classification phase. The dataset used in this study [11] includes Bengali, Romanized Bengali (Banglish), and English text, with some reviews containing only emojis or single words like "good," "thanks," or "joss."

Since our review texts comprise Bengali, English, and Romanized Bengali, we implemented normalization [26] for Bengali texts before fitting them into our model. For non-Bengali texts, we initially explored language detection and translation models. To begin with, we employed the 'Langdetect'[27] library to detect the language of each data point. Subsequently, we utilized the "BanglaT5"[28] to translate English text to Bengali. Interestingly, BanglaT5 also performs transliteration, effectively resolving any issues with data points containing English characters. Table I depicts some samples of our data before preprocessing and after preprocessing.

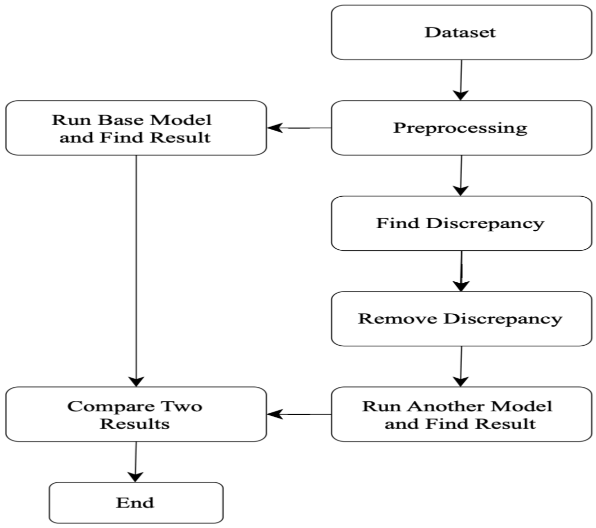


Fig. 1. Pipeline

Table I: Some Examples of preprocessed data

Before Preprocessing	After Preprocessing
Delay Delivery... Good Product.	বিলম্বে সরবরাহ। ভাল পণ্য।
Good phone according to my uses, Upgraded from Realme C11 2020, Except camera it's a good bump. Got a faster delivery within 12 hrs.	আমার ব্যবহার অনুযায়ী ফোন ভাল, রিয়েলম সি১১ ২০২০ থেকে আপগ্রেড করা হয়েছে, ক্যামেরা ছাড়া এটি একটি ভাল বাম্প। ১২ ঘন্টার মধ্যে দ্রুত ডেলিভারি পাওয়া যায়।

Table II: Example of Discrepancy in Text

Rating	Review	Emotion	Sentiment
1	Unofficial phone dhorai dise. please don't buy from daraz.	Love	Positive
5	অর্ডার করলাম Rainy Night পেলাম। Sunshower এইটা কেমন ব্যবহারের ভাই	Anger	Negative
5	আমি মোবাইলের সাথে ওয়ারেন্টি কার্ড পাইনি, ব্যবস্থা নিন প্লিজ	Anger	Negative

The original database reveals discrepancies among reviews, ratings, emotions, and sentiments. For example, in the first entry in Table II, the reviews and ratings align, but the sentiments and emotions appear misclassified, reflecting inconsistencies in emotional tone. In the 2nd and 3rd entries, the sentiments and emotions align with the content but conflict with the ratings. These mismatches introduce ambiguity and compromise label accuracy, which could impact model performance. Our research aims to detect and address these discrepancies, enhancing the reliability of sentiment and emotion classifications in the dataset.

C. Discrepancy Detection

In this section, we embark on the critical task of detecting discrepancies, which is paramount for our research, as previously outlined. To accomplish this, we propose the following workflow for detecting discrepancies:

Input:

- Dataset D with attributes: rating, sentiment, emotion
- Discrepancy conditions:
  - (rating $\leq 3 \wedge$ sentiment = Positive)
  - (rating $\geq 4 \wedge$ sentiment = Negative)
  - (emotion $\in$ {happiness, love} $\wedge$ sentiment = Negative)
  - (emotion $\in$ {Sadness, Anger, Fear} $\wedge$ sentiment = Positive)

Output:

- Cleaned dataset D' with reduced discrepancies
- Model M for automatic discrepancy detection

Steps:

1) **Identify Discrepancies:**
   - $D_d \leftarrow \{r \in D \mid (r.\text{rating} \leq 3 \wedge r.\text{sentiment} = \text{Positive}) \vee (r.\text{rating} \geq 4 \wedge r.\text{sentiment} = \text{Negative}) \vee (r.\text{emotion} \in \{\text{happiness, love}\} \wedge r.\text{sentiment} = \text{Negative}) \vee (r.\text{emotion} \in \{\text{Sadness, Anger, Fear}\} \wedge r.\text{sentiment} = \text{Positive})\}$
2) **Manual Verification:**
   - Confirm true discrepancies in D_d.
3) **Remove Verified Discrepancies:**
   - $D' \leftarrow D - D_d$
4) **Assume D' is Cleaned:**
   - Proceed with D' assuming minimal discrepancies.
5) **Build Discrepancy Detection Model(any statistical model):**
   - Define features $F = \{\text{rating}, \text{sentiment}, \text{emotion}\}$.
   - Label D' with discrepancy flags no.
   - Label D_d with discrepancy flags yes.
   - merge D' and D_d to make training datasetDt.
   - Split Dt into training and test sets $Dt_{\text{train}}, Dt_{\text{test}}$.
   - Train model M on Dt_{train}.
   - Evaluate M on Dt_{test}.
   - Deploy M for future discrepancy detection.

This algorithm depicts how we detect discrepancies from the dataset as the dataset is labeled with such features. We assume that the labeling is correct. Based on that we will implement this algorithm.

D. Vectorization Approach

In this section, preprocessed text data is converted into embeddings using BERT's vectorization. BERT is highly effective in Natural Language Understanding (NLU) tasks, including text classification and question answering, due to its rapid development and strong performance with minimal data.

Tokenization: Reviews are tokenized using the 'csebuetnlp/banglabert' [10] auto tokenizer. Special tokens [CLS] and [SEP] are added at sequence boundaries, and tokens are mapped to vocabulary indexes. Sequences are set to a fixed length of 512, with padding or truncation as needed, and attention masks differentiate real from padded tokens.

Table III: Hyperparameter values for training model

Model	Hyperparameter Values	Accuracy for Sentiment Analysis	Accuracy for Emotion Analysis
BERT	checkpoint = "csebuetnlp/banglabert", max_position_embeddings = 512, num_epochs = 4, learning_rate = 5e-6, weight_decay = 5e-7	93.96% (With discrepancy) 96.31% (Without discrepancy)	90.94% (With discrepancy) 93.59% (Without discrepancy)

Embedding Generation: Each token is mapped to a high-dimensional vector, with BERT retrieving contextual embeddings based on the token's surrounding words.

This is how we create embeddings from a sample review in our dataset. These vectors are condensed representations of the review's meaning. We use them to make different decisions.

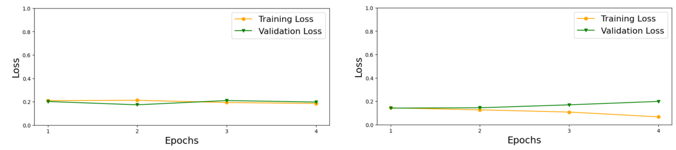
IV. Experimental Outcomes

A. Experimental Settings

The dataset is examined from different angles with the help of the Bangla BERT base [10], a Bengali language-specific version of the BERT architecture, to obtain a noticeably improved outcome. Since datasets with brief texts yield poor results [29], texts containing fewer words are removed from the entire dataset to distinguish the difference between the model's performances. It is seen that excluding texts of less than two words was beneficial in this instance and assisted the model in producing better results. All these are done while keeping the learning rate and epoch constant. The dataset is tokenized using the Bangla BERT base tokenizer. The analysis is performed by adjusting epochs and learning rates to observe accuracy variations in sentiment and emotion analysis. The model training, conducted on Kaggle's P100 GPU, showed that a learning rate of 5e-6 with 4 epochs yielded the highest accuracy. These hyperparameter values were also effective for emotion analysis. Table **??** outlines the specific hyperparameter settings for BERT fine-tuning.

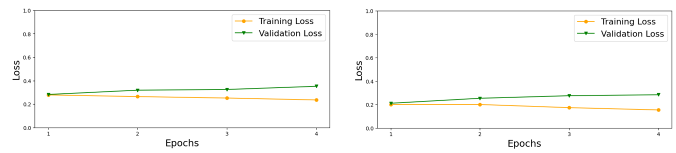
B. Result Analysis

This section is dedicated to discussing the results obtained and the proficiency of the BERT model to detect the sentiment and emotion of the reviews. With a learning rate of 5e-6 and epochs of 4, an accuracy of 93.96% was exhibited by the model in detecting the underlying sentiment on the dataset containing discrepancy. Removing the discrepancy, the accuracy elevated to 96.31%. With the same configuration, the model demonstrated an emotion detection accuracy of 90.94% on the dataset characterized by discrepancies, with a subsequent enhancement to 93.59% on the dataset devoid of such irregularities. Fig.2a and 2b summarize the overall training loss and validation loss changes on the dataset with and without discrepancy respectively over each epoch for the sentiment model. Where 3a and 3b represent the training and validation loss for the emotion model.
Fig.4 illustrates the accuracies achieved by our various models. It is evident from this figure that when we removed the discrepancies and trained our model, it exhibited better performance compared to the model trained with data containing discrepancies.



(a) With Discrepancy (b) Without Discrepancy

Fig. 2. Sentiment Model Training and Validation Loss



(a) With Discrepancy (b) Without Discrepancy

Fig. 3. Emotion Model Training and Validation Loss

In Fig. 5a and 5b we present the confusion matrix for sentiment analysis, comparing scenarios with and without discrepancies. It's worth noting that in the scenario with no discrepancy, we removed some records containing discrepancies, resulting in a smaller test data size of 20% in this case.

Fig.6a and 6b also illustrate the same confusion matrix analysis for emotion analysis in our dataset. It is evident that the model trained without discrepancies fails to make any correct predictions for the class of fear. Therefore, although its overall accuracy is higher, the model trained with discrepancies performs better in this particular case.

Fig. 7a and 7b showcase the ROC curves for sentiment analysis with and without discrepancies. Additionally Fig. 8a and 8b display the ROC curves for emotion analysis with and without discrepancies. It's noticeable that the result for class 3 (fear) is downgraded in the scenario without discrepancies. However, despite this, the overall accuracy of the model without discrepancies is better.

C. Discussion

Analyzing sentiment and emotion in e-commerce reviews provides valuable insights into customer perceptions and experiences. This study focuses on a method that combines discrepancy detection with sentiment and emotion analysis. We

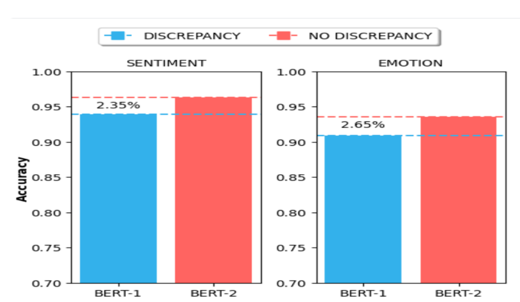
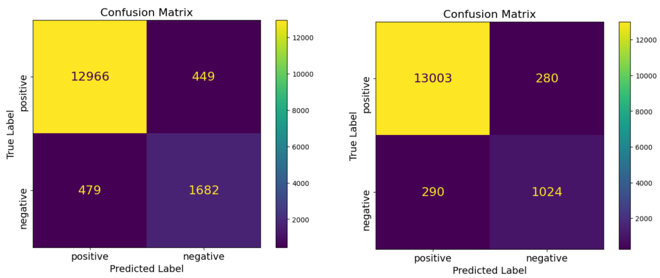
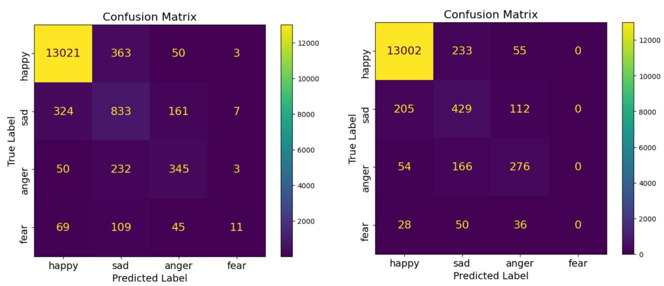


Fig. 4. Comparison of Overall Performance with and without Discrepancy



(a) With Discrepancy (b) Without Discrepancy

Fig. 5. Sentiment Analysis Confusion Matrix

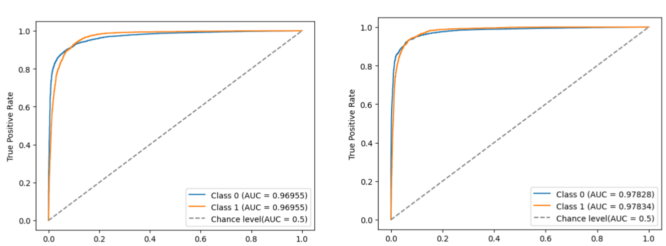


(a) With Discrepancy (b) Without Discrepancy

Fig. 6. Emotion Analysis Confusion Matrix

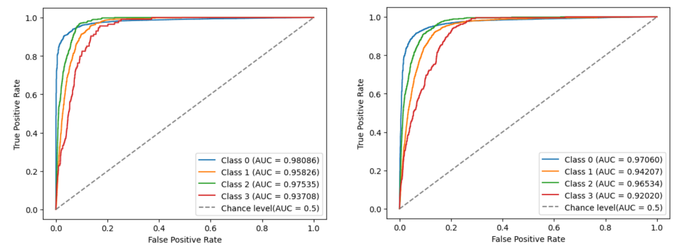
experimented with various approaches to identify discrepancies in the dataset [11] and compared results with and without their inclusion. Our approach introduces a comprehensive discrepancy detection methodology, including an algorithm and mitigation strategies specifically for Bangla reviews. While previous studies in other languages have explored discrepancy detection and mitigation [6], our method demonstrates superior performance and process efficiency. Unlike prior approaches that rely on sentiment and rating scores, our study uses a direct query-based system, yielding more reliable results. By analyzing sentiment and emotion both before and after discrepancy mitigation, we observed significant improvements, highlighting the importance of addressing discrepancies to enhance sentiment and emotion analysis in e-commerce.

We initially used the state-of-the-art BERT model, known for its exceptional text processing capabilities due to its



(a) With Discrepancy (b) Without Discrepancy

Fig. 7. Sentiment Analysis ROC



(a) With Discrepancy (b) Without Discrepancy

Fig. 8. Emotion Analysis ROC

attention mechanism and contextual coherence. This study has significant implications for both academia and industry. Academically, it addresses data discrepancies by developing a system that detects inconsistencies within datasets. From an industrial perspective, especially for e-commerce websites, and other industries that heavily rely on customer reviews and ratings. These platforms play a crucial role in influencing purchasing decisions. For instance, e-commerce websites could introduce a new feature displaying a summary of user reviews, indicating the number of positive and negative reviews, or the overall emotion of customers—whether they are happy or unhappy with their shopping experience. Furthermore, there's potential for implementing systems that prevent users from submitting inconsistent reviews. For example, if a user writes a positive review but gives a low rating, an automated system could detect the discrepancy and prompt the user to reconsider their rating or provide suggestions for a more appropriate rating based on the content of their review.

Industry 5.0 emphasizes a collaborative relationship between humans and advanced technologies, prioritizing resilience, human-centric AI, personalization, and ethical data use [30]. By improving sentiment and emotion analysis for the Bangla e-commerce space, our work addresses discrepancies that can hinder accurate AI insights. This encourages e-commerce businesses to develop robust supply chains that can withstand disruptions which in turn promotes resilience. Additionally, by enhancing the precision of sentiment analysis, the study supports ethical AI practices, ensuring that technology respects and accurately reflects human emotions and social contexts. This approach aligns with Industry 5.0's commitment to enhancing human well-being, trustworthiness, and ethical AI, making it a valuable contribution to the development of responsible, customer-focused AI systems.

V. Conclusion

In conclusion, E-commerce is spreading more nowadays. So, user interaction and engagement are becoming more important. To make those steps easier, it is necessary to recommend more personalized and reliable products to increase profit. Here, we preprocessed data on different aspects. We detected discrepancies, removed them, and finally applied the BERT model to find the result.

In this analysis, our primary goal was to accurately predict user sentiment and emotion. While we considered many factors, there are limitations to acknowledge. One such limitation is our handling of one or two-word reviews, which we chose to remove from the dataset. However, these shorter reviews could potentially enhance the data pool, an area for future exploration. Another limitation is our discrepancy detection technique. While effective in using dataset annotations for querying, real-time applications face the challenge of predicting discrepancies without explicit labeling, presenting a more complex task for future work. Looking ahead, we aim to overcome these limitations and strengthen our model. One direction is to develop a more advanced approach for detecting discrepancies, possibly utilizing machine learning techniques. We also aim to improve the model's ability to generate coherent sentences from short texts using large language models (LLM), broadening its applicability. Finally, we envision creating a real-time system for e-commerce websites to implement the findings of this research, helping improve decision-making and enhance user experiences.

References

[1] A. Srivastava and G. Kumari, *E-COMMERCE IN THE POST COVID-19 ERA: CHALLENGES, OPPORTUNITIES AND CONSUMERS' PREFERENCES*, pp. 19–26. 10 2021.

[2] M. Ott, Y. Choi, C. Cardie, and J. T. Hancock, "Finding deceptive opinion spam by any stretch of the imagination," *arXiv preprint arXiv:1107.4557*, 2011.

[3] M. Hu and B. Liu, "Mining and summarizing customer reviews," in *Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD04, ACM, Aug. 2004.

[4] Mudambi and Schuff, "Research note: What makes a helpful online review? a study of customer reviews on amazon.com," *MIS Quarterly*, vol. 34, no. 1, p. 185, 2010.

[5] D. Mayzlin, Y. Dover, and J. Chevalier, "Promotional reviews: An empirical investigation of online review manipulation," *American Economic Review*, vol. 104, pp. 2421–55, August 2014.

[6] S. Sadiq, M. Umer, S. Ullah, S. Mirjalili, V. Rupapara, and M. Nappi, "Discrepancy detection between actual user reviews and numeric ratings of google app store using deep learning," *Expert Systems with Applications*, vol. 181, p. 115111, Nov. 2021.

[7] M. T. Akter, M. Begum, and R. Mustafa, "Bengali sentiment analysis of e-commerce product reviews using k-nearest neighbors," in *2021 International Conference on Information and Communication Technology for Sustainable Development (ICICT4SD)*, IEEE, Feb. 2021.

[8] M. A. Shafin, M. M. Hasan, M. R. Alam, M. A. Mithu, A. U. Nur, and M. O. Faruk, "Product review sentiment analysis by using nlp and machine learning in bangla language," in *2020 23rd International Conference on Computer and Information Technology (ICCIT)*, IEEE, Dec. 2020.

[9] R. K. Das, M. Islam, M. M. Hasan, S. Razia, M. Hassan, and S. A. Khushbu, "Sentiment analysis in multilingual context: Comparative analysis of machine learning and hybrid deep learning models," *Heliyon*, vol. 9, p. e20281, Sept. 2023.

[10] A. Bhattacharjee, T. Hasan, W. Ahmad, K. S. Mubasshir, M. S. Islam, A. Iqbal, M. S. Rahman, and R. Shahriyar, "BanglaBERT: Language model pretraining and benchmarks for low-resource language understanding evaluation in Bangla," in *Findings of the Association for Computational Linguistics: NAACL 2022*, (Seattle, United States), pp. 1318–1327, Association for Computational Linguistics, July 2022.

[11] M. R. A. Rashid, K. F. Hasan, R. Hasan, A. Das, M. Sultana, and M. Hasan, "A comprehensive dataset for sentiment and emotion classification from bangladesh e-commerce reviews," *Data in Brief*, vol. 53, p. 110052, Apr. 2024.

[12] R. Aralikatte, G. Sridhara, N. Gantayat, and S. Mani, "Fault in your stars: an analysis of android app reviews," in *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data*, CoDS-COMAD '18, ACM, Jan. 2018.

[13] A. Almansour, R. Alotaibi, and H. Alharbi, "Text-rating review discrepancy (trrd): an integrative review and implications for research," *Future Business Journal*, vol. 8, Feb. 2022.

[14] M. A. Kassem, A. A. Abohany, A. A. A. El-Mageed, and K. M. Hosny, "A novel deep learning model for detection of inconsistency in e-commerce websites," *Neural Computing and Applications*, Mar. 2024.

[15] H. Jahan, A. Kowshir Bitto, M. Shohel Arman, I. Mahmud, S. Fahad Hossain, R. Moni Saha, and M. M. H. Shohug, *Sentiment Analysis of E-commerce Consumer Based on Product Delivery Time Using Machine Learning*, p. 649–661. Springer Nature Singapore, 2022.

[16] A. Alslaity and R. Orji, "Machine learning techniques for emotion detection and sentiment analysis: current state, challenges, and future directions," *Behaviour amp; Information Technology*, vol. 43, p. 139–164, Dec. 2022.

[17] P. Nandwani and R. Verma, "A review on sentiment analysis and emotion detection from text," *Social Network Analysis and Mining*, vol. 11, Aug. 2021.

[18] S. Sharmin and D. Chakma, "Attention-based convolutional neural network for bangla sentiment analysis," *AI amp; SOCIETY*, vol. 36, p. 381–396, July 2020.

[19] M. Hoq, P. Haque, and M. N. Uddin, *Sentiment Analysis of Bangla Language Using Deep Learning Approaches*, p. 140–151. Springer International Publishing, 2021.

[20] U. Singh, A. Saraswat, H. K. Azad, K. Abhishek, and S. Shitharth, "Towards improving e-commerce customer review analysis for sentiment detection," *Scientific Reports*, vol. 12, Dec. 2022.

[21] P. Savci and B. Das, "Prediction of the customers' interests using sentiment analysis in e-commerce data for comparison of arabic, english, and turkish languages," *Journal of King Saud University - Computer and Information Sciences*, vol. 35, p. 227–237, Mar. 2023.

[22] N. Hicham and H. Nassera, "Improving emotion classification in e-commerce customer review analysis using gpt and meta□ensemble deep learning technique for multilingual system," *Multimedia Tools and Applications*, Aug. 2024.

[23] M. S. Islam and K. M. Alam, "Sentiment analysis of bangla language using a new comprehensive dataset bangdsa and the novel feature metric skipbangla-bert," *Natural Language Processing Journal*, vol. 7, p. 100069, June 2024.

[24] Z. Ahmed, S. S. Shanto, and A. I. Jony, "Advancement in bangla sentiment analysis: A comparative study of transformer-based and transfer learning models for e-commerce sentiment classification," *Journal of Information Systems Engineering and Business Intelligence*, vol. 9, p. 181–194, Nov. 2023.

[25] P. Shaver, J. Schwartz, D. Kirson, and C. O'Connor, "Emotion knowledge: Further exploration of a prototype approach.," *Journal of Personality and Social Psychology*, vol. 52, no. 6, p. 1061–1086, 1987.

[26] T. Hasan, A. Bhattacharjee, K. Samin, M. Hasan, M. Basak, M. S. Rahman, and R. Shahriyar, "Not low-resource anymore: Aligner ensembling, batch filtering, and new datasets for Bengali-English machine translation," in *Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, (Online), pp. 2612–2623, Association for Computational Linguistics, Nov. 2020.

[27] M. Danilák, "GitHub - Mimino666/langdetect: Port of Google's language-detection library to Python. — github.com." https://github.com/Mimino666/langdetect. [Accessed 27-04-2024].

[28] A. Bhattacharjee, T. Hasan, W. U. Ahmad, and R. Shahriyar, "Banglanlg: Benchmarks and resources for evaluating low-resource natural language generation in bangla," *CoRR*, vol. abs/2205.11081, 2022.

[29] S. Kiritchenko, X. Zhu, and S. M. Mohammad, "Sentiment analysis of short informal texts," *Journal of Artificial Intelligence Research*, vol. 50, pp. 723–762, 2014.

[30] X. Xu, Y. Lu, B. Vogel-Heuser, and L. Wang, "Industry 4.0 and industry 5.0—inception, conception and perception," *Journal of manufacturing systems*, vol. 61, pp. 530–535, 2021.