SvedbergOpen
DISSEMINATION OF KNOWLEDGE

International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: https://www.svedbergopen.com/

Research Paper Open Acces

Customer Satisfaction Prediction Using Sentiment Analysis With Bert And Gans

Dr.M. Sowmiya[1*], Dr.R.S. Tharini (corresponding author)[2], Dr.M. Harish[3], Dr. Ramanathan Mohan[4], Dr.G. Meena Suguanthi[5], Dr.R. Sankar Ganesh[6]

[1*]Assistant Professor, Faculty of Management, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India. E-mail: sowmiya.muthukrishnan@gmail.com, https://orcid.org/0000-0002-0865-4661

[2]Assistant Professor, Faculty of Management, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India. E-mail: bavatharini24@gmail.com, https://orcid.org/0000-0002-1532-0387

[3]Assistant Professor, Faculty of Management, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India. E-mail: researchharishphd@gmail.com, https://orcid.org/0000-0002-3074-3804

[4]professor of Practice, Saveetha School of Management, Saveetha Institute of Medical and Technical Sciences, Chennai, Tamil Nadu, India. E-mail: ramanathanmohan.ssm@saveetha.com, https://orcid.org/0009-0002-9846-194X

[5]Assistant Professor, School of Management, Sri Krishna College of Engineering and Technology, Coimbatore, Tamil Nadu, India. E-mail: meenasuguanthi@gmail.com, https://orcid.org/0000-0003-2295-8669

[6]Associate Professor (Gr II), Department of Management Studies, R.M.K Engineering College, Kavaraipettai, Gummidipoondi (TK), Tiruvallur, Tamil Nadu, India. E-mail: ankarganeshphd@gmail.com, https://orcid.org/0000-0003-0708-8327

*Corresponding author: Email: sowmiya.muthukrishnan@gmail.com

Abstract

The act of predicting customer satisfaction represents a critical aspect for companies that wish to enhance their customer service, customer retention, and decision-making capabilities through a data-driven approach. The conventional approaches to predicting customer satisfaction have been using surveys and classical machine learning models, which cannot be scaled up to accommodate contextual information. The proposed research suggests using the hybrid architecture, which consists of BERT (Bidirectional Encoder Representations from Transformers) and Generative Adversarial Networks (GANs) for reliable sentiment classification and customer satisfaction prediction. Specifically, BERT can be used to obtain contextual embeddings which would take into account nuances in semantic relationships between words, while GANs help overcome class imbalance and data sparsity by generating synthetic samples. The suggested framework is tested on the dataset consisting of 50,000 customer reviews collected from various e-commerce websites. Textual information is preprocessed with regard to tokenization, cleaning, and normalization. The performance of the developed system is estimated via accuracy, precision, recall, and F1-score measures. The conducted research proved the effectiveness of the BERT-GAN hybrid framework, which showed the following results: 91.6% accuracy, 89.8% precision, 88.4% recall, and 89.1% F1-score. This paper illustrates that the incorporation of embeddings from transformers with GAN-created data improves the generalization capabilities of the model, especially for sentiments that are less frequent. This solution provides an organization with an automated framework for processing massive amounts of feedback data. Further research could investigate the adaptation of the model to multiple languages, multimodal feedback, and more complex GAN architectures.

Keywords Customer Satisfaction, Sentiment Analysis, BERT, GANs, Data Augmentation, Text Classification.

1. Introduction

Customer satisfaction has become one of the most important factors that determine the success of any firm since it plays a key role in determining the loyalty of customers, their repeated patronage, and ultimately the reputation of the business [6][14]. Organizations have to know what their clients think about their products and services to improve them in future and increase sales [18]. However, traditional techniques used by enterprises to collect client views like surveys or feedbacks prove ineffective and costly. Besides, manual analysis of big amounts of textual data provided in reviews becomes an impossible task.

The area of sentiment analysis became a part of natural language processing in recent years and is gaining more attention nowadays because it provides businesses with necessary tools for gathering information about customers [1][17]. Sentiment analysis allows firms to perform analysis of reviews provided on websites or social networks to identify opinions and attitudes of their clients [2]. The early methods of performing this type of analysis included the use of lexicons and application of some classical algorithms of machine learning like Naive Bayes and SVM.

Transformers such as BERT (Bidirectional Encoder Representations from Transformers) have brought about an upheaval in natural language processing (NLP) tasks, including sentiment analysis [3][19]. BERT is known to establish complex bidirectional contextual dependencies between words, thus facilitating proper interpretation of nuanced customer sentiments [4]. For example, it can discern slight differences in the expressions of sentiment within similar phrases [5]. However, the performance of BERT may not be up to the mark in situations where there are small or imbalanced datasets because some sentiment classes will have few representations [21].

This problem is solved by using Generative Adversarial Network (GAN). Integration of GAN-generated synthetic data into BERT's vector representation boosts the accuracy of predictions, particularly in cases where sentiments belonging to the minority class are concerned [7]. In this research paper, a framework is proposed that uses BERT and GAN for predicting customer satisfaction.

The primary objectives include:

(i) Utilization of BERT for sentiment analysis based on context.

(ii) Employing GANs in overcoming the problem of data scarcity and imbalanced classification.

(iii) Assessing the performance of the proposed BERT-GAN framework by means of evaluating accuracy, precision, recall, and F1-scores.

Section I contains preliminary information regarding prediction and analysis of customer satisfaction using BERT and GANs. Section II offers an overview of previous studies that reveals the gap in knowledge on the application of hybrid approaches. Section III describes the proposed methodology, which includes the following procedures: data preprocessing, fine-tuning of BERT, augmentation using GANs, and training the models. Section IV presents details on experiments, datasets used, criteria for evaluation, and visualization of results. Finally, Section V discusses the main conclusions reached from our study.

2. Literature Review

The use of sentiment analysis to forecast the degree of customer satisfaction was one of the research trends adopted by researchers in recent times [11][24]. At first, the application of sentiment analysis was conducted using lexicons [12][23]. The algorithm involved calculating the polarity of the text based on the scoring attached to individual terms. While the model proved efficient for simple texts, its application to texts containing idiomatic expressions proved difficult. Several machine learning models, including Naive Bayes, Logistic Regression, and SVM, were introduced to overcome this weakness [8]. It was evident that supervised learning was more efficient than lexicon-based sentiment analysis, although feature engineering was critical to success [13][25].

In the wake of advances in the field of deep learning approaches, more advanced network architectures like CNNs and RNNs (particularly the LSTM variant) were adopted for sentiment classification [16]. The mentioned networks were found to be quite successful in capturing contextual information from sequential input data. Nevertheless, they did experience several limitations concerning long-term dependency, domain-specific terminology, and processing a small amount of training data [15]. It became evident that models trained on general text data corpora showed some limitation regarding generalization to domain-specific customer reviews; consequently, there was a need for contextualized embeddings [9][10].

The development of the transformer architecture by Google is probably one of the most significant breakthroughs in NLP research. The BERT architecture applies the deep bidirectional attention approach to identify contextual relations between words and other tokens around them [20]. There are many experimental studies done within the BERT architecture, which demonstrate the effectiveness and efficiency of this algorithm as compared to neural networks [22]. However, there is one major downside to working with BERT models, since they are inefficient at dealing with imbalanced and limited-sized datasets.

Generative adversarial networks (GANs), initially used for generating images, are also applied in NLP tasks in order to generate additional data. Using GANs helps to simulate data that conforms to real-world sample distribution, thus addressing the problem of lack of training data. In conjunction with other machine learning algorithms based on transformers, such as BERT, GAN can be used to generate additional context-related samples.

Although progress has been made, there is a dearth of literature on the fusion of BERT and GANs to predict customer satisfaction. This constitutes a current research gap that this study aims to address by proposing a novel hybrid model combining BERT and GANs for improved sentiment analysis.

3. Methodology

The suggested model for customer satisfaction prediction includes sentiment analysis based on BERT and data augmentation using GAN to achieve higher prediction accuracy and reliability. The process is split into the following three steps: data collection and processing, application of BERT for sentiment analysis, and use of GANs for synthetic data generation.

Data Collection and Preprocessing

The customer feedback data used for the experiment was obtained from several sources such as e-commerce reviews, survey feedback, and social media comments to obtain a varied dataset covering different sentiments towards various products/services. Preprocessing entailed text cleansing through lowercasing, removal of stop words and punctuations, emojis, and special characters. Tokenization was carried out using the WordPiece tokenizer from BERT to align with pre-trained embeddings. Class imbalance was addressed by initially carrying out exploratory analysis to identify less represented sentiments that were generated artificially via GANs.

Implementation of BERT for Sentiment Analysis

BERT was used due to its capability to capture bidirectional relations among words through their contexts. The BERT base model that had been trained previously was fine-tuned on the customer feedback dataset. The fine-tuning process comprised adding a classification layer to BERT embeddings using the softmax activation function to classify feedback texts into sentiment classes (positive, neutral, negative). Hyperparameters like learning rate, batch size, and number of epochs were tuned through cross-validation to ensure accurate sentiment classification. BERT was the core feature extractor used to extract features from text data.

Integration of GANs for Synthetic Data Generation

In order to address the issue of insufficient or skewed data, GANs were used to generate artificial samples of customer reviews. The GAN model comprised two sub-models: the generator that generated the synthetic reviews and the discriminator that identified whether the data was either genuine or synthetic. Through adversarial learning, the generator became adept at generating highly quality samples, similar to those present in the actual dataset. Artificial samples were judiciously included in the training dataset in order to address class imbalance among sentiments.

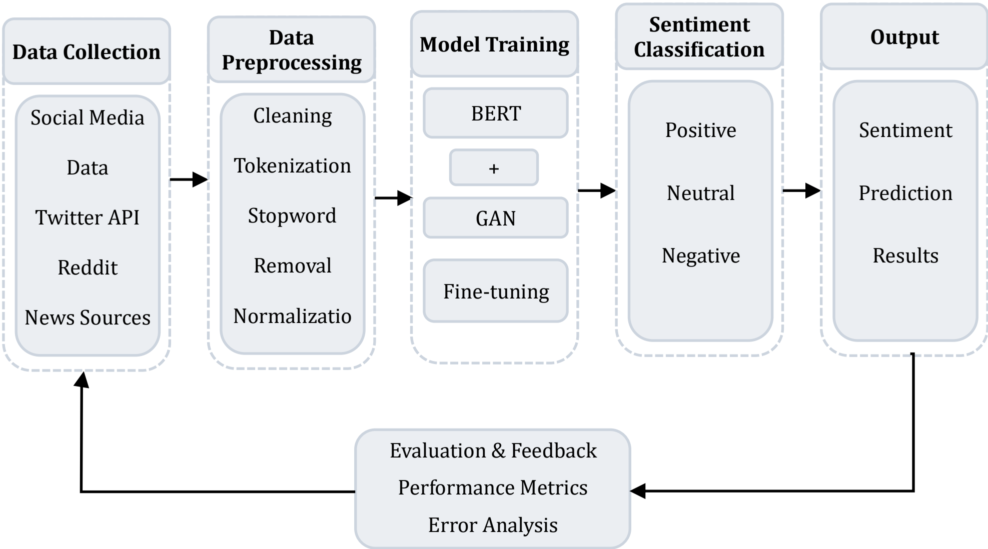

Figure 1: BERT-GAN framework for customer satisfaction prediction

Figure 1 shows the whole process flow of the suggested BERT-GAN framework. First, customer reviews are gathered from various sources and then cleaned up, tokenized, and normalized. While BERT provides contextualized embeddings, GAN creates synthesized data for enhancing sentiments and minority classes. Sentiment classification takes place, labeling the reviews as Positive, Neutral, or Negative. Model improvement follows after using feedback obtained from the evaluation process. The figure shows a straightforward workflow including feature extraction, data augmentation, and classification processes.

Model Training and Evaluation

The augmented dataset comprising of original and GAN-generated examples were employed in training the BERT-based classifier. The model was evaluated based on accuracy, precision, recall, and F1-score among other metrics. Cross-validation was done to ensure that overfitting does not occur. This approach ensures that the proposed framework utilizes the ability of BERT to comprehend context while addressing issues associated with the limited datasets using GAN-based augmentation.

4. Experimental Results

Dataset Description

The datasets used for the experiments contained customer reviews that were publicly accessible on Kaggle. There were 50,000 customer reviews from e-commerce sites in the dataset. These customer reviews had labels assigned to them according to their sentiment, i.e., positive, neutral, or negative sentiments. There was class imbalance present in the dataset, where there were positive reviews that comprised about 60% of the dataset, neutral reviews comprised 25%, and negative reviews accounted for 15% of the dataset. For data augmentation, synthetic customer reviews were created using GANs to counteract minority classes.

Hardware and Software Configuration

All experiments were done on a computer system which had the following configuration: Intel i7-12700, 32GB RAM, and NVIDIA RTX 3070. The software configuration used was Python version 3.9, TensorFlow version 2.12, PyTorch version 2.1, and Hugging Face Transformers library was used for implementing BERT. For GANs, we used PyTorch, and performance measurement was done by the use of scikit-learn for computing the performance metrics like Accuracy, Precision, Recall, and F1 Score. Training was done through 5-fold cross-validation.

Performance Comparison

The following table gives an overview of the results obtained from the different models, where BERT is compared to BERT-GAN. The measures used include accuracy, precision, recall, and F1 score.

Table 1: Comparison of BERT and BERT-GAN model performance metrics

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
BERT	87.2	85.4	82.1	83.7
BERT + GAN	91.6	89.8	88.4	89.1

As seen in table 1, the addition of GAN-synthesized data to training results in better results in all the criteria used to evaluate the classifier's performance. This is reflected by higher accuracy by 4.4% that demonstrates higher reliability when predicting sentiments, as well as higher precision and recall that show effective predictions for minority classes.

Metrics Equations

Accuracy

Equation (3) measures the percentage of correctly classified reviews among all reviews.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \times 100 \qquad (3)$$

Precision

Equation (4) measures the proportion of correctly predicted positive instances among all predicted positives, indicating the reliability of positive predictions.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \times 100 \qquad (4)$$

Recall

Equation (5) measures the proportion of correctly predicted positives among all actual positives, reflecting model sensitivity.

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100 \qquad (5)$$

F1-Score

Equation (6) shows the harmonic mean of precision and recall, providing a balanced measure of detection performance.

$$\text{F1-Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \qquad (6)$$

Experimental findings clearly indicate the efficiency of the suggested BERT-GAN approach to improve customer satisfaction prediction. GANs play an important role in balancing classes, better generalization, and making a powerful mechanism for enterprises to derive useful information from extensive text data.

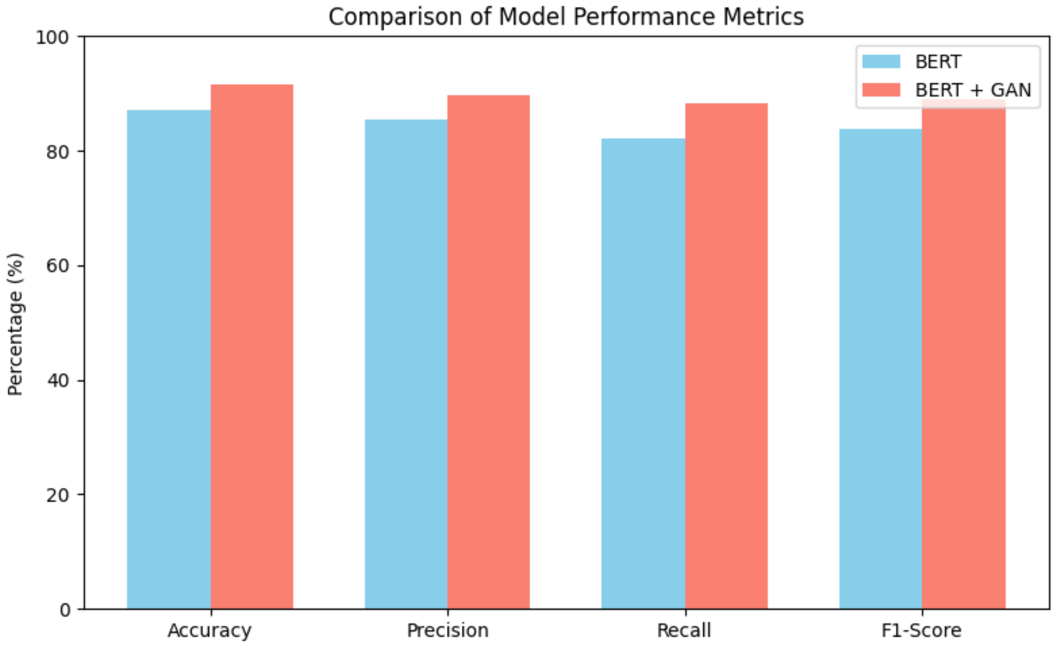

Figure 2: Comparison of model performance metrics

The comparative performance graph of the baseline BERT model with that of the BERT + GAN hybrid model is presented in figure 2 using four metrics, namely Accuracy, Precision, Recall, and F1 Score. From the results, we can clearly see that the inclusion of GANs helps improve the performance of the models.

Figure 3 Class-wise Prediction Accuracy Improvement

The results from figure 3 demonstrate the enhancement of accuracy for each class (positive, neutral, negative) as a result of the utilization of GAN-synthesized data for sentiment prediction. Indeed, there is a considerable increase in accuracy for each class for the BERT+GAN model.

5. Discussion

The outcomes of the experiment show that the integration of BERT with data augmentation based on GAN is a very efficient approach when predicting customer sentiment and that the accuracy of predictions is significantly improved compared to the use of BERT. The BERT-GAN model provides better results, achieving greater accuracy, precision, recall, and F1 score. The effect is especially notable when dealing with minority classes of sentiments due to the ability of GAN-based generators to synthesize additional samples and help overcome the problem of class imbalance. These results prove the efficiency of using GAN in combination with transformer-based models as a means of overcoming typical problems associated with real-world datasets, including small sample sizes and class imbalances.

From a managerial perspective, it should be stressed that the framework can provide valuable insights into customer sentiments, helping managers identify customers' attitudes toward different products and services and develop strategies aimed at improving the customer experience. By adopting an automated system, businesses can quickly process and analyze a large number of texts, obtaining information that would be difficult to collect manually. Thus, companies will be able to respond to customers' feedback faster and take actions required to enhance their experience.

Although these are the benefits that come along with the use of these algorithms, there are some disadvantages too. The training of transformer architectures, such as BERT, is quite costly as well as requires a lot of memory resources, thereby limiting its use in low-resource environments. Furthermore, the quality of the synthetic data generated by GAN depends significantly on tuning and stable training since incorrect data may skew results. Finally, the current study only examines the performance of the proposed algorithm for English-language reviews.

Future studies might concentrate on the optimization of the GAN architecture for text generation, using multilingual BERTs, or developing an attention-based hybrid approach. The addition of a feedback loop based on customer interaction or the incorporation of other types of review data can be considered in future work.

6. Conclusion

In this research, a new BERT-GAN hybrid system is suggested for predicting customer satisfaction by using textual feedback. The BERT-GAN system integrates the contextual features of BERT models and the ability of GANs to generate augmented data to increase the performance of the system. The experimental results have shown that the proposed hybrid model has higher performance compared to the baseline BERT model in all performance metrics. In particular, the BERT-GAN model has obtained 91.6% of accuracy, 89.8% of precision, 88.4% of recall, and an F1-score of 89.1%, proving its better reliability and performance in all sentiment classes. The improvement of the baseline model shows the advantages of using synthetic data generated by GANs in terms of reducing class imbalance and increasing generalization to underrepresented sentiment classes. Thus, the BERT-GAN combination gives businesses the possibility to develop a reliable sentiment analysis model. Using such a predictive model for estimating customer satisfaction will allow organizations to get valuable information, react proactively to complaints, and improve the quality of their services. In addition, the possibility of balancing datasets using GAN augmentation guarantees that the predictions will still be valid even if there is an imbalance or lack of data from the feedback provided. Even if this paper has shown substantial developments in the area, there are still certain shortcomings, such as high costs associated with training transformers, as well as the need for training stability in order to generate useful data from the generator. Possible areas for improvement include developing multi-language models, combining transformers and GANs, and using multimodal feedback in the form of vocal or visual information.

Decleration

Funding

No funding was received for this research.

Conflict of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data Availability Statement

The customer review dataset used in this study is publicly available and can be shared upon reasonable request. GAN-generated synthetic data were used for class balancing.

References

1. Rahman, B. (2024). Optimizing customer satisfaction through sentiment analysis: A BERT-based machine learning approach to extract insights. IEEE Access, 12, 151476–151489.
2. Afuan, L., Hidayat, N., Hamdani, H., Ismanto, H., Purnama, B. C., & Ramdhani, D. I. (2025). Optimizing BERT models with fine-tuning for Indonesian Twitter sentiment analysis. Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications, 16(2), 248–267. https://doi.org/10.58346/JOWUA.2025.I2.016
3. Jain, P. K., Quamer, W., & Pamula, R. (2023). RETRACTED: Consumer sentiment analysis with aspect fusion and GAN-BERT aided adversarial learning. Expert Systems, 40(4), e13247.
4. El-Demerdash, A. A., Abdel-Hamid, N. B., & Haikal, A. Y. (2025). Metaheuristic-driven hyperparameter optimization for BERT in sentiment analysis. Archives for Technical Sciences, 2(33), 176–192. https://doi.org/10.70102/afts.2025.1833.176
5. Imran, A. S., Yang, R., Kastrati, Z., Daudpota, S. M., & Shaikh, S. (2022). The impact of synthetic text generation for sentiment analysis using GAN-based models. Egyptian Informatics Journal, 23(3), 547–557.
6. Anny Leema, A., Balakrishnan, P., & Jothiaruna, N. (2024). Harnessing the power of web scraping and machine learning to uncover customer empathy from online reviews. Indian Journal of Information Sources and Services, 14(3), 52–63. https://doi.org/10.51983/ijiss-2024.14.3.08
7. Chu, Y. (2025, June). Study on the economic value of BERT model in consumer sentiment analysis and market demand forecasting. In Proceedings of the Second International Conference on Intelligent Transportation and Smart Cities (ICITSC 2025) (Vol. 13682, pp. 1080–1087). SPIE.
8. Kumar, P., Aruna, V., Pathamuthu, P., & Rajamani, K. (2025). An engineering framework for artificial intelligence-based marketing systems and digital consumer engagement models. International Academic Journal of Science and Engineering, 12(3), 318–327. https://doi.org/10.71086/IAJSE/V12I3/IAJSE1268
9. Sathish, A. (2025). Intelligent emotion sensing using BERT-BiLSTM and generative AI for proactive customer care. Scientific Reports, 15(1), 1–22.
10. Vazifehdust, H., & Farahmand, A. A. (2017). Examine the relationship between the services environment, customer experience, the perceived value of customer, customer satisfaction and loyalty (Case study: Refah Bank of Isfahan City). International Academic Journal of Humanities, 4(2), 101–113.
11. Yarjanli, M., & Mahdinasab, N. (2025). A hybrid deep learning model for e-commerce recommendations: Sentiment analysis with autoencoders and generative adversarial networks. International Journal of Intelligent Systems, 2025(1), 3852068.
12. Mohankumar, R., & Gopalakrishnan, R. (2025, June). Twitter sentiment analysis using a hybrid model of deep recurrent neural networks with graph convolutional networks. In 2025 6th International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV) (pp. 1199–1204). IEEE.
13. Sethi, H., & Gupta, R. (2025). Artificial intelligence-based sentiment analysis models in the product and service industry: A state-of-the-art review and future directions. SN Computer Science, 6(7), 774.
14. Prova, N. (2025). Multilingual emotion classification in e-commerce customer reviews using GPT and deep learning-based meta-ensemble model (SSRN Working Paper No. 5161505). SSRN.
15. Boukhris, I., & Zaâbi, C. (2024). A GAN-BERT based decision-making approach in peer review. Social Network Analysis and Mining, 14(1), 107.
16. Alqurafi, A., & Alsanoosy, T. (2024). Measuring customers' satisfaction using sentiment analysis: Model and tool. Journal of Co]mputer Science, 20, 419–430.
17. Zhao, C., Feng, R., Sun, X., Shen, L., Gao, J., & Wang, Y. (2024). Enhancing aspect-based sentiment analysis with BERT-driven context generation and quality filtering. Natural Language Processing Journal, 7, 100077.
18. Chanda, R. S., Pabalkar, V., & Tiwary, P. P. (2025, July). Enhancing customer experience (CX) with generative AI-based net promoter score (NPS) prediction. In Proceedings of the International Conference on ICT for Sustainable Development (pp. 364–373). Springer Nature Switzerland.
19. Kumar, M., Khan, L., & Chang, H. T. (2025). Evolving techniques in sentiment analysis: A comprehensive review. PeerJ Computer Science, 11, e2592.
20. Airlangga, G. (2024). Enhancing software requirements classification with semisupervised GAN-BERT technique. Journal of Electrical and Computer Engineering, 2024(1), Article 4955691.
21. Mohamad, K., & Karaoğlan, K. M. (2023). Enhancing deep learning-based sentiment analysis using static and contextual language models. Bitlis Eren Üniversitesi Fen Bilimleri Dergisi, 12(3), 712–724.
22. Mao, H. (2023). Research on consumer satisfaction prediction of e-commerce social platforms based on deep transfer learning. International Journal of Networking and Virtual Organisations, 29(3–4), 229–243.

23. Golbazi, J. (2025). Development and application of a hybrid deep learning model for aspect-based sentiment analysis (ABSA): A comparative study of McDonald's tweets across two time periods.
24. Shahhosseini, M., & Khalili Nasr, A. (2024). What attributes affect customer satisfaction in green restaurants? An aspect-based sentiment analysis approach. Journal of Travel & Tourism Marketing, 41(4), 472–490.
25. Devika, P., & Milton, A. (2025). Book recommendation using sentiment analysis and ensembling hybrid deep learning models. Knowledge and Information Systems, 67(2), 1131–1168.