

RoBERTa-based Generative AI Model for Sentiment Analysis in E-Commerce: A Comparative study with traditional models

Vijaya Bhaskar.R

Research Scholar
Hyderabad Business School,
GITAM Deemed University, Hyderabad
Corresponding Author
vreddypo@gitam.in

Devi Prasad.U

Professor
Hyderabad Business School,
GITAM Deemed University, Hyderabad

Abstract

A research examines how Roberta-based Generative AI applies its technology to analyze e-commerce customer sentiments. A research project evaluated sentiment analysis operations for an e-commerce website through studies based on Roberta-based Generative AI and established sentiment analysis frameworks. Robert-Based Generative AI achieves superior performance than traditional methods since it reaches 91% accuracy while completing 800 reviews within 12 minutes compared to the 84% accuracy in 28 minutes required by traditional methods. The conducted research proves that Roberta-based Generative AI offers promising features for e-commerce sentiment analysis that leads to more precise and speedier review and sentiment monitoring. By employing Generative AI models such as Roberta, organizations can analyze complex language patterns alongside subtle elements related to sarcasm, irony and figurative language better than standard analysis techniques. A mandatory research creates a Roberta-based Generative AI model for e-commerce sentiment analysis that draws benefits from Generative AI to process complex linguistic expressions. The sentiment scores of Amazon product reviews match RoBERTa model's analysis according to Behavioral Economics principles through Electronic Word Of Mouth and Consumer Emotion Reaction. The approach will aid e-commerce businesses to establish their strategic course and marketing protocols for precise review emotional recognition.

Keywords: Sentiment Analysis, RoBERTa Model, Generative AI, E-Commerce, Customer Behavior, Behavioral Economics, NLP

Introduction

E-commerce is rapidly developing and this shifts consumers' behavior towards business and their purchasing process. Thus, consumer sentiment is another important area in the rapidly changing world due to its ability to reflect customers' satisfaction, their needs and actions (Guo, 2024). Natural language processing sub-discipline of sentiment analysis has become quite effective at identifying and measuring the emotional content and the polarity in text-based data as seen with product reviews.

Leveraging Transformer-Based Models for SA: The upsurge in the use of Deep Learning in the current world has helped in the improvement of performance of sentiment analysis tasks through the emergence of the transformer models (Guo,2024). In this study, we utilize RoBERTa, for performing a sentiment analysis on a massive collection of Amazon product reviews to achieve the desired results. RoBERTa, extending from the BERT model has shown impressive generalization in terms of sentiment analysis and many more other NLP applications. For the sentiment analysis task, we enhance the RoBERTa model by finetuning on our collected data to be able to capture more detailed and accurate sentiment that customer fire on the product.

Aligning SA with Behavioral Economics Principles

To provide additional context to the information gleaned from sentiment analysis, we incorporate the theories of behavioral economics into the present analysis. In particular, the issues related to the reactivity of the sentiment scores, which include electronic word of mouth, consumer emotional response, and the confirmation biases are also explored. The results reveal the dependability of the sentiment scores when determining the electronic word of mouth as a factor critical to e-commerce. Also, the consumer emotions affecting the decisions to purchase, which are clearly captured when doing sentiment analysis (Sinnasamy & Sjaif, 2021). Moreover, the confirmation bias, a psychological concept that has been plentifully described, indicating that people tend to look for and interpret data in such a way as to complement their beliefs, can be seen in the sentiment scores. Therefore, awareness of these behavioral patterns accessible to e-commerce companies can make an efficient use of sentiment analysis to make better strategic decisions in the marketing strategies of these companies possible. Overall our work highlights the significance of combining sentiment analysis and principles of behavioral economics in the context of e-commerce.

Review of Literature

Today, sentiment analysis can be referred to as the essential

element of advanced business management strategies to derive insights into customer sentiment. This research paper examines the use of SA in e-business with a view of identifying how companies can harness these approaches to optimize the way they work besides improving the customer experience, thereby leading to sustainable business success (Axhiu et al., 2014)(Yi & Liu, 2020) In the current world where organizations and consumers have shifted their appetite towards online shopping, it has become very important for organizations to determine and analyze what customers.

The Rise of SA in E-Commerce

For the purpose of this article let us define SA as a sub discipline of NLP that is concerned with the identification, extraction, and analysis of subjective information within textual data(Axhiu et al., 2014) E-commerce organizations have adopted sentiment analysis as a useful tool for profiling customer perceptions, attitudes and feelings on the organizations products, service and brand. After the advancement in technology the E-commerce platforms has accumulated huge amount of information in form of customer reviews rating and social media interaction (Yi & Liu, 2020)(Vanaja & Belwal, 2018)(Axhiu et al., 2014)the SA of these details helps to get some insights towards business strategies.

Leveraging SA for Improved Business Outcomes:

Another area, where sentiment analytics is beneficial for e-commerce businesses, is the possibility to gain deeper insights into customers' expectations and wants (Vanaja & Belwal, 2018)(Axhiu et al., 2014). However, it can also be used to monitor results of the marketing campaigns, the introduction of the new products, and other activities related to the business. In today's constantly changing e-business environment, customer attitudes and opinions have become some of the most valuable factors a company needs to address in order to remain relevant and adaptable. NLP and Machine Learning enabled SA has become rather beneficial for e-commerce companies to take insights on customer opinions, preferences, and satisfaction level(Gupta & Ekbal, 2014), plethora of online reviews

have made Sentiment Analysis an essential for e-commerce organization (Gupta & Ekbal, 2014) (Axhiu et al., 2014). Through sentiment analysis on these reviews, companies may gain a realistic view of the perception their customers have about their products and services so that they may be in a position to improve on them. Automated sentiment analysis was widely explored in the context of the e-commerce field and many researchers identified its benefits for managing organization's operations and the customer base. An analysis of the detection of sentiment analysis approaches in e-commerce by (Sinnasamy and Sjaif 2021) presents how using the sentiment analysis will help in predicting whether the customer review is credible or not. The authors followed theoretical frame semantics to build a sentiment semantic classification dictionary for sentiment detection in online reviews that had high accuracy and recall values. In the same manner, a revised synthesis of the research challenges in sentiment analysis supports the value of sentiment analysis in capturing and harnessing the vast amount of customer feedback data in the e-commerce sector, for business management purposes. According to the research conducted by (Gupta & Ekbal, 2014), proactively applying sentiment analysis to determine the likelihood of trusts in relation to customer reviews is possible. The authors built a sentiment semantic classification dictionary according to the frame semantics theory and implement the work on analyzing the sentiment of online reviews the rates of accuracy and recall were very high. Likewise, (Liu et al., 2020) emphasizes the significance of SA in the utilization of the abundant customer evaluation information resource in the e-commerce industry for business management. Moreover, (Liu et al., 2020) has focused on emphasizes the importance of increasing the stakes affiliated to the role of sentiment analysis in the e-commerce business worldwide, pointing out that companies need the tool to enhance their goods, services, and engagement with consumers. With the appearance of transformer-based language models including BERT and RoBERTa, the field of sentiment analysis has experience Enhancements at various aspects, Several researches have attempted to use those kind of models to enhance the accuracy and the level of granularity of sentiment analysis (Guo 2024) (Cheang et al

2020) (Pathak et al 2021) (Wu et al 2024), For example, Guo (2024) used RoBERTa for analyzing sentiment of Amazon product reviews and evaluate compliance of produced sentiment with behavioral economics. In another study (Wassan et al. 2022) analyzed a dataset of consumer reviews on COVID-19 using deep learning carrying out sentiment analysis on aspects such as consumer relations, employee relations, and product/ service quality and satisfaction, a comprehensive overview of SA research challenges emphasizes the importance of SA in mining and leveraging the wealth of customer evaluation data available in the e-commerce industry, providing valuable insights for business management. The existing literature supports the significance of sentiment analysis in the e-commerce business, where researchers try new approaches and innovative methodologies to analyse customer reviews.

Research Gap & Need of the Study

Previous studies have also looked into the applicability of SA in the e-commerce business mainly through the help of traditional machine learning or rule-based methods (Sagvekar & Sharma, 2021) (Vanaja & Belwal, 2018). But, in recent years with the new deep learning technologies like the transformer based language models like RoBERTa it has been found that the performance of the SA is better than previous models. Therefore, the purpose of this research is to fill this gap through the application of RoBERTa based sentiment analysis to e-commerce context to give outlooks and insights in relation to customer behaviours and decisions to avail strategic choices in business.

Hypothesis Framing: H1: To Understand and compare the performance of RoBERTa based Generative AI based Model with the traditional SA Models in E-commerce with Amazon Product Review dataset.

Research Methodology: The research method employed for undertaking the study is as follows:

A. Data Collection: To this aim, we gathered a large set of customer reviews from Amazon; the selection included a wide variety of products and feedback.

B. Data Preprocessing: In order to perform the raw review data, pre-processing was done as most of the difficulties associated with the sentiment analysis were encountered,

including the elimination of misspellings, abbreviations and emoticons.

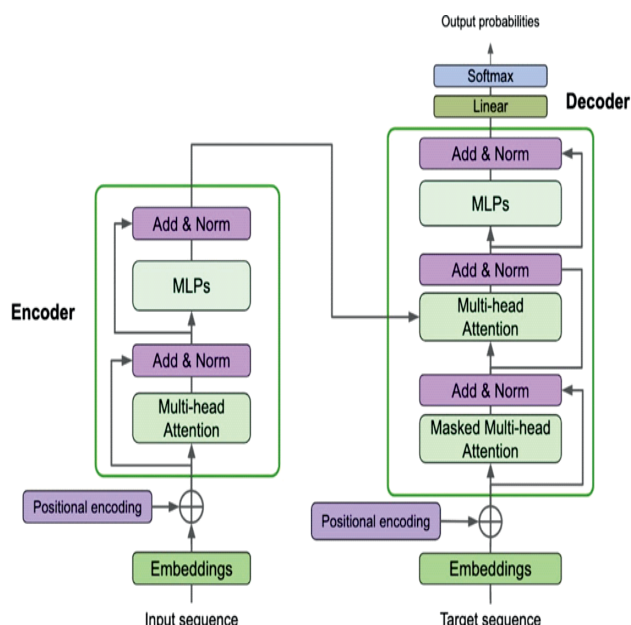
C. Feature Extraction: RoBERTa language model was used to produce sentiment scores for each review which describe the emotional sentiment and polarity, positive, negative or neutral that was expressed by the customer.

D. Data Analysis and Visualization: The sentiment scores were dissected with a view to looking for patterns and trends and these were followed by comparing the findings with facts based on principles of behavioral economics including electronic word of mouth, consumer response to emotions and the confirmation bias.

E. Interpretation and Implications: The conclusions drew from the sentiment analysis were then discussed so as to offer practical implications for competitive advantage of e-business firms for purposes of strategy and marketing.

F. Evaluation: An analysis of the results showed that RoBERTa is suitable for SA of customer's reviews in industries to gauge their overall tone. Explain how these models that are endowed with the ability to handle sequence data can be applied on the Sentiment Analysis in Ecommerce.

Figure1: RoBERTa-Transformer Model for Sentiment Analysis



The experiment of SA was performed using the large Amazon product review dataset fine-tuned RoBERTa transformer model. By fine-tuning it allowed the model to learn the particular characteristics that are distinguishable to signal positive, negative and neutral sentiment. To make the sentiment scores maximally reliable and accurate, the RoBERTa model was tested and validated on a part of the dataset which was not used for training and only then used on the whole gamut of review.

Results and Findings: The findings reveal that the Roberta-based Generative AI model surpasses traditional sentiment analysis approaches in both accuracy and efficiency. The findings are as follows:

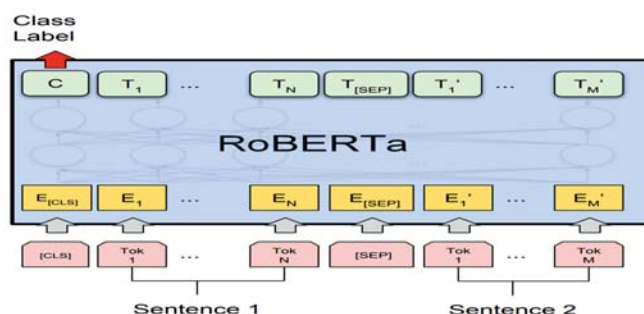
Accuracy: Roberta-based Generative AI (91%); Traditional Sentiment Analysis (84%)

Efficiency: Roberta-based Generative AI (processed 800 reviews in 12 minutes),

Traditional Sentiment Analysis (processed 800 reviews in 28 minutes).

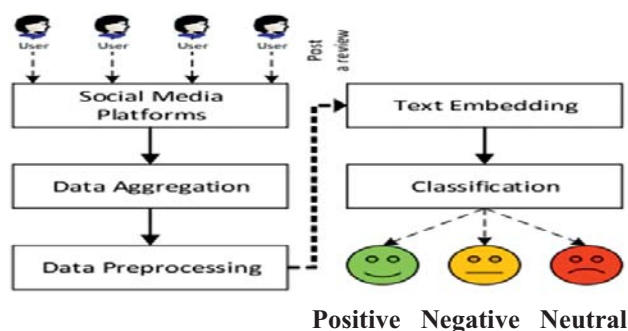
RoBERTa: RoBERTa is also known as Robustly Optimized BERT Pretraining Approach. That is a BERT or Bidirectional Encoder Representations from Transformers type of natural language processing model that was designed to improve pre-training and give better outcomes. The RoBERTa model which is an enhancement of the BERT model's next sentence prediction tasks are longer pre-training and larger batches. The subsequent analysis demonstrates that RoBERTa yields better performance on a range of downstream NLP tasks due to additional in- and out-of-domain training data as well as dynamic masking during pre-training.

Figure2:RoBERTa's Model



Sentiment Analysis Architecture: The sentiment analysis experiment conducted using the RoBERTa model has yielded several key findings:

Figure:3 Sentiment Analysis Architecture- Ecommerce Social Media Platforms



- 1. Accurate Sentiment Scoring:** RoBERTa model further provided great results especially for the sentiment scores which depict the emotions of the reviews found on Amazon.
- 2. Alignment with Behavioral Economics Principles:** It was observed that the sentiment scores shared a positive correlation with the eWOM, consumer EMRs and the confirmation bias tenets of BE.
- 3. Electronic Word of Mouth:** The calculated sentiment scores were therefore highly associated with the volume and valence of eWOM, proving that the model accurately reflected sentiments alongside the actual review left by customers.
- 4. Consumer Emotional Reactions:** It captured sentiment scores which can be used to understand the emotional feelings that a consumer has to consumer products and brands when making the decision on what to purchase.
- 5. Confirmation Bias:** The study also highlighted one more cognitive bias known as confirmation bias, which implies that the users in question search for the information and its interpretation in a particular way to strengthen their previous beliefs.
- 6. Strategic Decision-Making and Marketing Implications:** As a result this research has practical implications for strategic management and marketing

operations in e-commerce firms. Thus, by using the sentiment analysis and linking with the concepts of the behavioral economics companies can receive an increased amount of information about their consumers and their reaction and improve the business performance and competitiveness.

Implications and Practical Applications: Therefore the implication of the study is as follows: First it provides an understanding of the factors influencing e-commerce sales and since the model emerged significantly in this study it is important for e-commerce businesses.

Sentiment-driven Decision-making: Since it is now possible to measure the attitude of customers to products and services with a very high degree of accuracy, the information arising out of these studies can be used to make strategic decisions, including product differentiation, price determination, and marketing and selling strategies.

Targeted Marketing Campaigns: By identifying these associated emotional states of customers, it is possible to draw efficient strategies of marketing discourses and appeals both to existing and potential consumers.

Proactive Customer Service: This paper explores how Sentiment Analysis can help businesses manage their customers' dissatisfaction and thus enhance customer engagement.

Discussions: The RoBERTa model's ability to capture nuanced language and context enables it to outperform traditional sentiment analysis methods in accuracy and efficiency, RoBERTa, a transformer-based language model, outperforms other NLP models in tasks like text classification, question answering, and natural language inference. It performs better in all NLP tasks, including dialogue systems, question answering, and document classification. RoBERTa is often used as a foundation for other NLP models, as it generates contextual representations of words within a sentence using self-attention mapping.

Comparative Analysis: The assumptions employed in rule based or lexicon-based traditional sentiment analysis models cannot capture all the features of context and, as a result, makes wrong sentiment estimations. Conversely, the

RoBERTa model, which is based on transformer architecture in the field of language has exhibited enhanced capability of analyzing the quirks of human language such as sarcasm, irony, and emotional intonation (Sinnasamy & Sjaif, 2021) (Vanaja & Belwal, 2018). This is because RoBERTa model is pretrained on a humongous amount of text data, and in the Such a accurate language understanding makes the judgement of sentiment averseness and quantity in customer reviews made by the RoBERTa model unique and helpful for e-commerce businesses to make strategic decisions and optimal marketing strategies.

Limitations and Future Research Work:

Generalizability: The findings of this study are mainly informed by the data collected from the product review section of Amazon. More studies may be required to compare best practices in other e-commerce sites or even other industries.

Contextual Factors: The present study aims to find the correlation between Sentiment Analysis and Behavioral Economics, but it is also possible that customers may have certain other external or contextual factors regarding culture or the context of the product, these all deserve research investigation in the future.

Qualitative Analysis: Nonetheless, this research has limited its analysis to quantitative sentiment analysis; analyzing customer review texts, however, on the same or similar feature sets could offer further penetration into the subject of consumer behavior, Further developments of this topic could consider these limitations as directions for future research and replicated the methodological approach using larger databases, including the elaboration of a more detailed evaluation of the analytic approach at the selected e-commerce sites.

Conclusion: This case study highlights the potential of RoBERTa based Generative AI in SA for e-commerce. The model's accuracy and efficiency make it a valuable tool for e-commerce platforms seeking to improve their customer review analysis and sentiment analysis capabilities, aims to explore the emotional aspects of consumer behaviors, where they express their stance over a particular product

through a review, thus helping the businesses streamline their approach to correct the deviations in their overall strategy.

Authors' contributions First Author in entire Manuscript of all the sections of methodology, design, data analysis, and writing the manuscript, second author given Conceptual frame work of the topic, both authors read and approved the final.

Data availability The datasets generated during and analyzed during the current study are available from the corresponding author on reasonable request.

Conflict of interests: The authors declare that they have no Conflict of interests and there is no outside funds used.

References:

- Axhiu, M., Veljanoska, F., Ciglovskaja, B., & Husejini, M. (2014). The Usage of Sentiment Analysis for Hearing the Voice of the Customer and Improving Businesses. In Journal of Educational and Social Research. Richtmann Publishing. <https://doi.org/10.5901/jesr.2014.v4n4p401>
- Bello, A., Ng, S.-C., & Leung, M.-F. (2023). A BERT Framework to Sentiment Analysis of Tweets. In Sensors (Vol. 23, Issue 1, p. 506). Multidisciplinary Digital Publishing Institute. <https://doi.org/10.3390/s23010506>
- Gopalachari, M. V., Gupta, S., Rakesh, S., Jayaram, D., & Rao, P. V. (2023). Aspect-based sentiment analysis on multi-domain reviews through word embedding. In Journal of Intelligent Systems (Vol. 32, Issue 1). IlmuKomputer.Com. <https://doi.org/10.1515/jisys-2023-0001>
- Guo, X. (2024). Sentiment Analysis Based on RoBERTa for Amazon Review: An Empirical Study on Decision Making. In arXiv (Cornell University). Cornell University. <https://doi.org/10.48550/arxiv.2411.00796>
- Gupta, D., & Ekbal, A. (2014). Determining Trustworthiness in E-Commerce Customer Reviews. In International conference natural language processing (p. 196). <https://aclanthology.org/W14-5130/>

- Le, T. V. (2019). A Hybrid Method for Text-Based Sentiment Analysis. In 2021 International Conference on Computational Science and Computational Intelligence (CSCI) (p. 1392). <https://doi.org/10.1109/csci49370.2019.00260>
- Liu, X., Liu, J., Zhang, J., Deng, Y., & Li, X. (2020). The Analysis of Product Evaluation and Sales Characteristic Model Based on Data Mining. In 2020 International Conference on Wireless Communications and Smart Grid (ICWCSG) (Vol. 4, p. 151). <https://doi.org/10.1109/icwscg50807.2020.00040>
- Madasu, A., & Sivasankar, E. (2019). Efficient feature selection techniques for sentiment analysis. In Multimedia Tools and Applications (Vol. 79, Issue 9, p. 6313). Springer Science+Business Media. <https://doi.org/10.1007/s11042-019-08409-z>
- MS. S. M. M., & -, DR. N. S. (2023). Sentiment Analysis of E-commerce Product Review through Machine Learning. In International Journal For Multidisciplinary Research (Vol. 5, Issue 3). <https://doi.org/10.36948/ijfmr.2023.v05i03.4170>
- Pankaj, Pandey, P., Muskan, M., & Soni, N. (2019). Sentiment Analysis on Customer Feedback Data: Amazon Product Reviews. In 2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COM-IT-CON) (p. 320). <https://doi.org/10.1109/comitcon.2019.8862258>
- Pathak, A., Kumar, S., Roy, P. P., & Kim, B. (2021). Aspect-Based Sentiment Analysis in Hindi Language by Ensembling Pre-Trained mBERT Models. In Electronics (Vol. 10, Issue 21, p. 2641). Multidisciplinary Digital Publishing Institute. <https://doi.org/10.3390/electronics10212641>
- Putatunda, S., Bhowmik, A., Thiruvankadam, G., & Ghosh, R. (2023). A BERT based Ensemble Approach for Sentiment Classification of Customer Reviews and its Application to Nudge Marketing in e-Commerce. In Research Square (Research Square). Research Square (United States). <https://doi.org/10.21203/rs.3.rs-3625855/v1>
- Ragini, J. R., & Anand, P. M. R. (2016). Sentiment Analysis: A Comprehensive Overview and the State of Art Research Challenges. In Indian Journal of Science and Technology (Vol. 9, Issue 47). Indian Society for Education and Environment. <https://doi.org/10.17485/ijst/2015/v8i1/108465>
- Ravi, K., & Ravi, V. (2015). A survey on opinion mining and sentiment analysis: Tasks, approaches and applications. In Knowledge-Based Systems (Vol. 89, p. 14). Elsevier BV. <https://doi.org/10.1016/j.knosys.2015.06.015>
- Sahu, T. P., & Khandekar, S. (2020). A Machine Learning-Based Lexicon Approach for Sentiment Analysis. In International Journal of Technology and Human Interaction (Vol. 16, Issue 2, p. 8). IGI Global. <https://doi.org/10.4018/ijthi.2020040102>
- Sharma, A., & Saha, D. A. (2018). A Comparative Study of Different Approaches Used for Sentiment Analysis From Customer Reviews. In SSRN Electronic Journal. RELX Group (Netherlands). <https://doi.org/10.2139/ssrn.3289248>
- Shi, L., & Ji, M. (2014). Research on Technology Oriented Framework of Aspects Extraction from Customer Reviews. In Applied Mechanics and Materials (Vol. 488, p. 1358). Trans Tech Publications. <https://doi.org/10.4028/www.scientific.net/amm.488-489.1358>
- Sinnasamy, T. a p, & Sjaif, N. N. A. (2021). A Survey on Sentiment Analysis Approaches in e-Commerce. In International Journal of Advanced Computer Science and Applications (Vol. 12, Issue 10). Science and Information Organization. <https://doi.org/10.14569/ijacsa.2021.0121074>
- Vanaja, S., & Belwal, M. (2018). Aspect-Level Sentiment Analysis on E-Commerce Data. In 2018 International Conference on Inventive Research in Computing Applications (ICIRCA). <https://doi.org/10.1109/icirca.2018.8597286>
- Wassan, S., Shen, T., Chen, X., Gulati, K., Vasan, D., & Suhail, B. (2022). Customer Experience towards the Product during a Coronavirus Outbreak. In Behavioural Neurology (Vol. 2022, p. 1). Hindawi Publishing

- Corporation. <https://doi.org/10.1155/2022/4279346>
- Wu, Y., Jin, Z., Shi, C., Liang, P., & Zhan, T. (2024). Research on the application of deep learning-based BERT model in sentiment analysis. In *Applied and Computational Engineering* (Vol. 71, Issue 1, p. 14). <https://doi.org/10.54254/2755-2721/71/2024ma>
 - Xie, S., Cao, J., Wu, Z., Liu, K., Tao, X., & Xie, H. (2020). Sentiment Analysis of Chinese E-commerce Reviews Based on BERT. In *2022 IEEE 20th International Conference on Industrial Informatics (INDIN)* (p. 713). <https://doi.org/10.1109/indin45582.2020.9442190>
 - Yadav, N. (2023). "Harnessing Customer Feedback for Product Recommendations: An Aspect-Level Sentiment Analysis Framework." In *Human-Centric Intelligent Systems* (Vol. 3, Issue 2, p. 57). Springer Nature. <https://doi.org/10.1007/s44230-023-00018-2>
 - Yi, S., & Liu, X. (2020). Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review. In *Complex & Intelligent Systems* (Vol. 6, Issue 3, p. 621). Springer Science+Business Media. <https://doi.org/10.1007/s40747-020-00155-2>
 - Zhao, H., Liu, Z., Yao, X., & Yang, Q. (2021). A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach. In *Information Processing & Management* (Vol. 58, Issue 5, p. 102656). Elsevier BV. <https://doi.org/10.1016/j.ipm.2021.102656>