An Analysis of Sentiment in E-Commerce Products Review: A Detailed Survey on Methods

Sathya.P[1], Anuratha.V[2], Elamparithi.M[3]

[1]Research Scholar, Department Of Computer Science, Kamalam College of Arts and Science, Anthiyur, Affiliated to Bharathiar University, Coimbatore, Tamil Nadu, India,
Email: sathyakamalraj1234@gmail.com
[2]Associate Professor ,Department Of Computer Science, Kamalam College of Arts and Science, Anthiyur, Affiliated to Bharathiar University, Coimbatore, Tamil Nadu, India, Email: profanuratha@gmail.com
[3]Associate Professor ,Department Of Computer Science, Kamalam College of Arts and Science, Anthiyur,Affiliated to Bharathiar University, Coimbatore, Tamil Nadu, India,
Email: profelamparithi@gmail.com

Received: 17.04.2024 Revised : 19.05.2024 Accepted: 25.05.2024

ABSTRACT
Review sites for different products and services have increased due to the development of several e-commerce websites. Nowadays, reviews are a simple way for people to learn more about the goods and services they intend to utilize. Here, categorizing the polarity of product reviews is a crucial function of sentiment analysis (SA). Today, sentiment research allows businesses to comprehend consumer attitudes towards goods and services in a worldwide economy. Negative statements significantly impact sentiment identification.A sentiment analysis that merely provides overall polarity when there are many reviews must be more comprehensive. Finding reviews of the product's particular parts (features) will be challenging. This study aimed to examine machine learning (ML) and deep learning (DL) models for predicting customer sentiments in product evaluations on e-commerce sites. The perspectives on the procedure, necessary activities, and tactics of SA that have been highlighted in many studies are presented in this study. Additionally, different difficulties with the sentiment classification process are discussed.

Keywords: Sentiment Analysis, E-Commerce, Product Review, Machine Learning, and Deep Learning.

1. INTRODUCTION

E-commerce is the term used to describe the buying and selling goods and services over the Internet. It has a lot of information, procedures, and tools for buyers and sellers, like smart device shopping, cash on delivery, and encryption for online payments. A research study claims that the COVID-19 epidemic has increased online sales. After receiving delivery, customers are accustomed to leaving reviews or remarks regarding the quality of the goods and services [1]. Users can interact with user opinions to establish a debate and receive recommendations and advice on goods or services. Additionally, by examining these attitudes, online retailers will better comprehend client expectations, offer better customer service, and improve sales. SA identifies whether a text block is promising, harmful, or neutral. SA is extracting words from their context to assess the social sentiment associated with a brand and help companies decide whether or not to sell their products [2]. The goal of SA is to take a look at public opinion in a way that will help businesses grow. It emphasizes polarity (positive, negative, and neutral) and emotions (happy, sad, furious, etc.).

The ML/DL-based method and the lexicon-based method are the two methodologies that are most frequently used to perform SA. To determine the overall sentiment of a sentence, the lexicon-based methodology uses a dictionary of terms annotated by sentiment. However, continuous maintenance and fine-tuning are needed, which raises the standard for implementation efforts [3]. ML is mainly used to comprehend, examine, and derive intelligent meaning from human language. Some commonly used ML algorithms are Support vector machine (SVM), random forest (RF), decision tree (DT), logistic regression (LR), and k-nearest neighbors (KNN). The DL has also recently made enormous progress, according to a SA [4]. This motivates us to survey the recent works regarding SA in E-commerce product reviews using ML and DL methods.

The remainder of the paper is divided into the following sections: The theoretical framework of SA is covered in Section 2. The studies on SA that combine ML and DL models are briefly described in Section 3. Sections 4 and 5 provide the discussion and conclusion with recommendations for further research.

2. BACKGROUND OF THE STUDY

The current topic in e-commerce product quality management is SA of online reviews, which manufacturers can use to determine how the general public feels about goods offered on e-commerce platforms. Customers can learn about other people's opinions about identical things in the meantime. The overall architecture of the sentiment classification system is depicted in Figure 1, which primarily consists of "4" phases, including dataset collection, text preprocessing, word embedding, and classification.

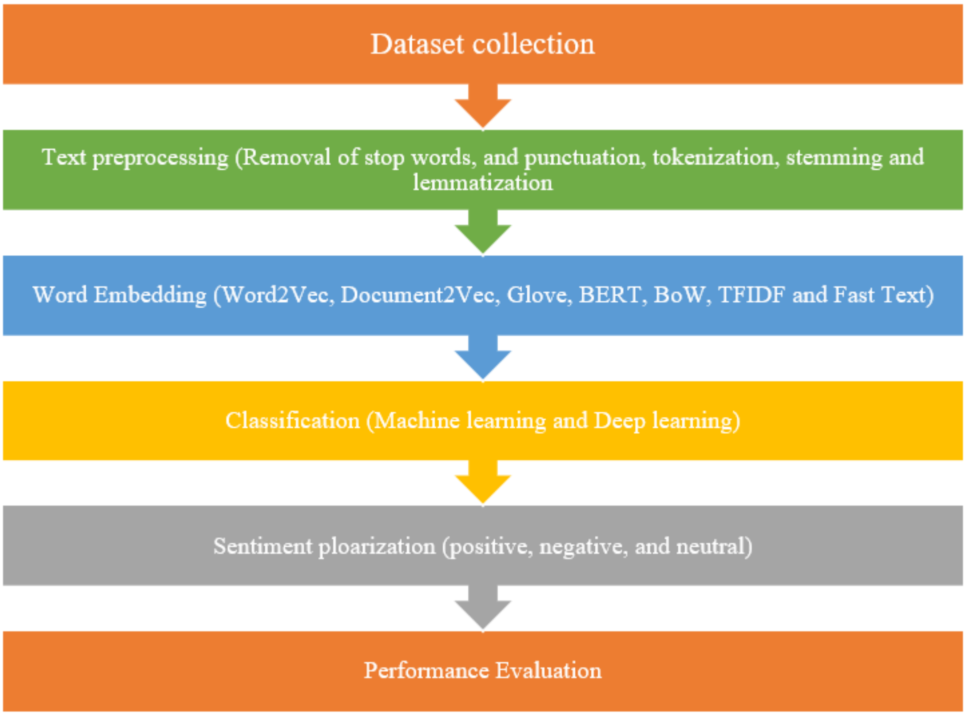

Figure 1. General structure of sentiment classification system

2.1 Dataset collection

The current topic in e-commerce product quality management is SA of online reviews, which manufacturers can use to determine how the general public feels about goods offered on e-commerce platforms. Customers can learn about other people's opinions about identical things. The overall architecture of the sentiment classification system is depicted in Figure 1, which primarily consists of "4" phases, including dataset collection, text preprocessing, word embedding, and classification.

2.2 Text Preprocessing

Preprocessing techniques are used to get the best results by preparing the data for the model. Null values, lowercase letters, spelling corrections, tokenization, stop words, eliminating punctuation, and lemmatization were among the preprocessing operations. Based on the review scores, each review in the dataset is labelled [6]. A rating is deemed suitable if it is more significant than three stars, neutral if equal to three stars, and negative if it is less than three stars.

2.3 Word Embedding

Word embedding uses numerical and vector representations of each word. Word embedding describes texts containing precise replicas of words with the same meaning. Word embedding, in particular, is unsupervised word representation learning that resembles semantic similarity [7]. Word embedding represents Individual words as real-valued vectors in a predetermined vector space. The method is frequently called DL since each word is assigned to a single vector, and the vector values are learned to resemble a neural network. Word embedding can be done using different techniques, including N-gram, Bag of Words (BoW), term frequency-inverse document frequency (TF-IDF), Word2vec, Fast Text, etc.

2.4 Classification

One of the data mining approaches used to identify a class of existing data is classification. The classification is carried out because several data rows are known to be categorized and labelled. The extracted dataset's product reviews are categorized here as good, negative, or neutral. The sentiment, which is typically regarded as a key influencing factor for future and prospective customers to make wise purchasing choices, is predicted using a variety of powerful learning algorithms, including ML and DL [8]. This can improve user experience and help organizations construct models or make decisions to improve customer connections. ML is a subfield of artificial intelligence that focuses on computer-implemented statistical and algorithmic models. It is unnecessary to manually create the features because DL can automatically learn them from the data.

3. Review On Methods Used For Sentiment Analysis

SA in e-commerce product reviews is a procedure to look at online product reviews to find overall opinions or feelings against the products. Several scholars have proposed techniques to address issues with SA as well as the analysis and mining of customer reviews. In this part, a thorough analysis of earlier work is offered. The ML techniques for analyzing the sentiment in product reviews on online stores are covered in Table 1. ML is the most used method for SA, particularly when labelled data is already available.

Table 1. Comparison among existing methods using ML methods

Author & Ref. No	Method used	Dataset used	Outcomes	Benefits	Limitations
Huiliang Zhao et al. [9]	Local Search Improvised Bat Algorithm based Elman Neural Network (LSIBA-ENN)	Amazon	Accuracy= 93.91%, Precision=90.9 %, Recall=90.01%	The optimized ENN reduced the overall loss and improved accuracy	After determining the gradient on the entire dataset, ENN weights were modified. Therefore, it might take longer to converge to the minimum if the dataset is too huge.
M.P. Geetha and D. Karthika Renuka [10]	Bidirectional Encoder Representations from Transformers (BERT)	Amazon	Accuracy=88.48 %, Precision=88.09 %, Recall=86.22%, F1-Measure=89.41 %	It offered high accuracy and the BERT algorithm required less training time	Due to the learning structure, and complex pattern, the model was unsuitable.
Murat Demircan et al. [11]	SVM, RF, DT, LR, and KNN	Hepsiburada product dataset	F1-score of SVM was 0.86%, RF was 0.78%, DT was 0.82%, LR was 0.75%, and KNN was 0.76%	Little effort is needed to build a model since even in complex domains, it was simple to understand.	Multiple ML algorithms often cause the model to be overfit.
Emre Deniz et al. [12]	Binary Relevance (BR) and Multi-label K-Nearest Neighbor (Ml-KNN)	Turkish e-commerce sites	MicroP=0.9157, MicroR=0.8837, and Micro-F1=0.8925	Efficiently uses computing resources, and is simple to use and understand	The association between labels is ignored by the BR technique, and the Ml-KNN performed poorly when classifying

					imbalanced datasets.
James Mutinda et al. [13]	Naïve Bayes (NB), KNN, DT, and SVM	Amazon dataset	F-measure of NB was 83.84%, KNN was 88.12%, DT was 86.72%, and SVM was 90.15%	The capability of handling complex relationships between variables and improved generalization performance even when faced with noisy data.	If there were many more traits than samples, the performance would be poor.
Pinar Savci and Bihter Das [14]	SVM, RF, and Multinomial Naive Bayes (MNB)	Sample data taken from e-commerce sites	Accuracy of SVM was 86.7%, RF was 87.5%, and MNB was 78.3%	By averaging or casting a vote among three models, it lowers the variation and bias of the forecasts.	Due to the requirement for training and storing many models, it was time-consuming and costly in terms of computation.

Recently, SA has made substantial use of DL. Based on its neural network, a well-behaved DL approach can determine whether or not its prediction is accurate.

Gagandeep Kaur and Amit Sharma [3] presented a DL model called long short-term memory with hybrid feature extraction strategies for SA of consumer reviews. The data from the STS-Gold dataset was utilized, and the hybrid features, such as aspect-related and review-related features, were extracted from the pre-processed data. The hybrid features were given to LSTM for sentiment polarity classification. The system attained 94.46% precision, 91.63% f-score, and 92.81% recall. Naveen Kumar Gondhi et al. [15] suggested a word2Vec-based LSTM model for SA on E-commerce product reviews. The system used the Word2Vec model for vectorizing the input data, and the vectorized scores were passed to LSTM for classification. The system attained the highest accuracy, recall, and f-score of 89%, 90% and 93% when tested on the Amazon Review dataset 2018. Mohammad Eid Alzahrani et al. [16] developed a combination of CNN and LSTM frameworks for SA of E-commerce data. The data was pre-processed initially by performing the following:" removal of stop words and punctuation, tokenization and lowercasing. The pre-processed data was given to CNN-LSTM for polarity prediction. CNN was used to extract the features, and LSTM was used to learn long-term dependencies in the extracted features. The system attained a higher accuracy of 94% on the Amazon product review public dataset.

Recently, SA has made substantial use of DL. Based on its neural network, a well-behaved DL approach can determine whether or not its prediction is accurate. Jing Zhang et al.[17] presented customer preference extraction for air purifiers and was based on fine-grained SA of online reviews. From the review content, the product attributes were initially retrieved. Then, attribute-sentiment pairs were formed by coupling sentiment phrases with the respective product's attributes to recognize sentiment orientation via sentiment value computation. Finally, using demand analysis and extraction techniques based on the Kano model, it was possible to determine consumer preferences for the product's features from sentiment orientations. Here, the system's performance was analyzed using online reviews of air purifiers sold in the Chinese market that T-mall.com had indexed. Amjad Iqbal et al. [7] proffered recurrent neural network-based LSTM and Deep LSTM classifiers for the SA of consumer reviews on Amazon. The manual feature extraction achieved 0.97 maximum accuracy for both Amazon-Products and Amazon-Fine-Food Reviews datasets. M. Sivakumar and Srinivasulu Reddy Uyyala [18] presented LSTM and fuzzy logic to perform aspect-based SA on mobile phone reviews. The LSTM generated positive and negative scores for the pre-processed collected reviews and gave it to the fuzzy logic for classifying the sentiment polarity into highly positive, negative, negative, and positive. The system was tested on Amazon video game reviews, amazon product reviews, and Amazon cell phone reviews and attained an accuracy of 83.82%, 90.92% and 96.93%.

J. Shobana and M. Murali [4] suggested an optimized LSTM model for polarizing the reviews of Amazon, Demonetization, Trip Advisor and Book review datasets. Once essential preprocessing was done, the system utilized a skip-gram-based word2vec model for representing the input into vector data. Finally, the optimized LSTM was utilized to classify the review as positive or negative polarity, in which an adaptive particle swarm algorithm was employed to tune the weights of the LSTM. The system achieved a maximum accuracy and f-score of 96.8% and 80.45%. Jun Liu et al. [19] presented an optimal gradient decent network-based cross-domain sentiment-aware word embedding model for the sentiment polarization of Amazon reviews. The skip-gram and continuous bag-of-words models were utilized to polarize the reviews of the products, in which the model was optimized using a gradient descent approach. You Zhang et al. [20] presented an interactive attention model for sentiment classification online customer reviews. The input review texts were initially transformed into distributed representations using an embedding layer. Then, the interactive layer learned the bilinear relationship between the users and their purchased products. Finally, the learned information was classified into positive and negative sentiments. The system attained a maximum accuracy of 97.2% when experimenting with IMDB, Yelp, and Amazon datasets.

DISCUSSION

SA classifies a customer's opinions or emotions towards the products as positive, negative, and neutral. As a result, various SA methods are employed, reflecting the study of unstructured customer evaluations to lead and produce relevant and helpful data. Before now, most works had exclusively used local textual characteristics. More additional data, such as user and product information, might impact literature regarding context, meaning, and sentiment. As a result of the inclusion of this data, the sentiment classification model can estimate the sentiment polarity of a review that a user would write about a specific product. In other words, reviews made by different users regarding the same products or by the same users about different items may have reviews with different sentiment polarities. The works above have employed ML and DL to combine commonly used and high-performing external factors, including users, time, aspects, and themes, in order to address these shortcomings. In recent years, ML algorithms have attracted much interest and have been applied in various applications. The most popular ML techniques include SVM, RF, NB, DT, etc. These algorithms offer superior results with greater accuracy, but the ML approach necessitates a sizable labelled dataset with manual annotation. Recent tests have demonstrated that DL-based automated feature engineering and shallow classification beat state-of-the-art human feature engineering. To forecast the sentiment, the majority of the authors utilized an LSTM model. Due to their capacity for long-term memory, LSTMs belong to a class of DL approaches that are significantly better at managing long-term dependencies. Additionally, the LSTM are significantly less prone to the vanishing gradient problem.

CONCLUSION

The field of SA is expanding in today's digital world. This study concluded that SA was used in e-commerce in recent publications. It covers a variety of papers that support various methods used to detect polarity or categorize internet evaluations into various sentiments. This clearly shows that different types of characteristics have different distributions, and neither categorization model consistently performs better. Additionally, it has been discovered that various feature types and classification algorithms can be effectively integrated to overcome their shortcomings, capitalize on one another's advantages, and ultimately improve sentiment classification performance. Overall, DL-based SA of e-commerce product reviews is a promising method that can assist e-commerce platforms in remaining competitive and meeting the changing needs of their customers. It also offers better accuracy compared to ML methods. Future user evaluations should be summarized using a variety of sentiment summarizing methods.

REFERENCES

[1] Hossain, Md Jahed, Dabasish Das Joy, Sowmitra Das, and Rashed Mustafa. "Sentiment analysis on reviews of e-commerce sites using machine learning algorithms." In 2022 International Conference on Innovations in Science, Engineering and Technology (ICISET), pp. 522-527. IEEE, 2022.

[2] Wankhade, Mayur, Annavarapu Chandra Sekhara Rao, and Chaitanya Kulkarni. "A survey on sentiment analysis methods, applications, and challenges." Artificial Intelligence Review 55, no. 7 (2022): 5731-5780.

[3] Kaur, Gagandeep, and Amit Sharma. "A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis." Journal of Big Data 10, no. 1 (2023): 5.

[4] Shobana, J., and M. Murali. "An efficient sentiment analysis methodology based on long short-term memory networks." Complex & Intelligent Systems 7, no. 5 (2021): 2485-2501.

[5] Chamekh, Ameni, Mariem Mahfoudh, and Germain Forestier. "Sentiment analysis based on deep learning in e-commerce." In International Conference on Knowledge Science, Engineering and Management, pp. 498-507. Cham: Springer International Publishing, 2022.
[6] Birjali, Marouane, Mohammed Kasri, and Abderrahim Beni-Hssane. "A comprehensive survey on sentiment analysis: Approaches, challenges and trends." Knowledge-Based Systems 226 (2021): 107134.
[7] Iqbal, Amjad, Rashid Amin, Javed Iqbal, RoobaeaAlroobaea, Ahmed Binmahfoudh, and Mudassar Hussain. "Sentiment analysis of consumer reviews using deep learning." Sustainability 14, no. 17 (2022): 10844.
[8] Huang, Huang, Adeleh Asemi, and Mumtaz Begum Mustafa. "Sentiment Analysis in E-commerce Platforms: A Review of Current Techniques and Future Directions." IEEE Access (2023).
[9] Zhao, Huiliang, Zhenghong Liu, Xuemei Yao, and Qin Yang. "A machine learning-based sentiment analysis of online product reviews with a novel term weighting and feature selection approach." Information Processing & Management 58, no. 5 (2021): 102656.
[10] Geetha, M. P., and D. Karthika Renuka. "Improving the performance of aspect based sentiment analysis using fine-tuned Bert Base Uncased model." International Journal of Intelligent Networks 2 (2021): 64-69.
[11] Demircan, Murat, Adem Seller, Fatih Abut, and Mehmet Fatih Akay. "Developing Turkish sentiment analysis models using machine learning and e-commerce data." International Journal of Cognitive Computing in Engineering 2 (2021): 202-207.
[12] Deniz, Emre, Hasan Erbay, and Mustafa Coşar. "Multi-label classification of e-commerce customer reviews via machine learning." Axioms 11, no. 9 (2022): 436.
[13] Mutinda, James, Waweru Mwangi, and George Okeyo. "Lexicon-pointed hybrid N-gram Features Extraction Model (LeNFEM) for sentence level sentiment analysis." Engineering Reports 3, no. 8 (2021): e12374.
[14] Savci, Pinar, and Bihter 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 35, no. 3 (2023): 227-237.
[15] Gondhi, Naveen Kumar, Eishita Sharma, Amal H. Alharbi, Rohit Verma, and Mohd Asif Shah. "Efficient long short-term memory-based sentiment analysis of e-commerce reviews." Computational intelligence and neuroscience 2022 (2022).
[16] Alzahrani, Mohammad Eid, Theyazn HH Aldhyani, Saleh Nagi Alsubari, Maha M. Althobaiti, and Adil Fahad. "Developing an intelligent system with deep learning algorithms for sentiment analysis of E-commerce product reviews." Computational Intelligence and Neuroscience 2022 (2022).
[17] Zhang, Jing, Aijia Zhang, Dian Liu, and Yiwen Bian. "Customer preferences extraction for air purifiers based on fine-grained sentiment analysis of online reviews." Knowledge-Based Systems 228 (2021): 107259.
[18] Sivakumar, M., and Srinivasulu Reddy Uyyala. "Aspect-based sentiment analysis of mobile phone reviews using LSTM and fuzzy logic." International Journal of Data Science and Analytics 12, no. 4 (2021): 355-367.
[19] Liu, Jun, Shuang Zheng, Guangxia Xu, and Mingwei Lin. "Cross-domain sentiment aware word embeddings for review sentiment analysis." International Journal of Machine Learning and Cybernetics 12 (2021): 343-354.
[20] Zhang, You, Jin Wang, and Xuejie Zhang. "Personalized sentiment classification of customer reviews via an interactive attributes attention model." Knowledge-Based Systems 226 (2021): 107135.