International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(Review Article)

Check for updates

Analyzing sentiments in e-commerce: Techniques, applications and challenges

Amrithkala M Shetty * and Manjaiah D.H

Department of Computer Science,Mangalore University, Karnataka, India.

International Journal of Science and Research Archive, 2024, 12(02), 2307–2320

Publication history: Received on 01 April 2024; revised on 11 August 2024; accepted on 14 August 2024

Article DOI: https://doi.org/10.30574/ijsra.2024.12.2.0843

Abstract

Sentiment Analysis (SA), a crucial component of Natural Language Processing (NLP),involves discerning sentiment within textual data. While widely applicable across diverse domains, this paper specifically focuses on SA within E-Commerce datasets. Serving as a concise review for beginners,the paper encompasses an overview of approximately twenty early works in SA. It elucidates the techniques employed, SA levels, and delineates the methodologies underpinning these approaches. In addition to elucidating the foundational methodologies, this paperprovides insights into the applications of SA. It sheds light on how SA techniques find practical utility in discerning and interpreting sentiment nuances within the realm of E-Commerce. Furthermore, the paper delves into the challenges intrinsic to SA, offering a comprehensive understanding of the intricacies associated with this NLP task.

Keywords: NLP-Natural Language Processing; Machine learning; Deep learning; CNN; LSTM

1. Introduction

Individuals are consistently influenced by the thoughts, ideas, and perspectives of others when making decisions. The surge of user-generated content, including comments, reviews, and opinions, through the proliferation of social media has become a significant factor in this dynamic. Both consumers and manufacturers stand to gain valuable insights from this wealth of information. Consumers, in particular, have adopted the practice of perusing reviews and opinions shared by others before making online purchases. This trend provides them with a comprehensive understanding of a product's strengths and weaknesses, aiding in informed decision-making. Simultaneously, manufacturers can leverage this user-generated content to gain valuable insights into the reception of their products in the market. However, the sheer volume of opinionated text data has become overwhelming for users. Despite the valuable perspectives offered, the abundance of information can be challenging to navigate. Scholars are increasingly intrigued by the exploration of methods to assess and distill the thoughts expressed in these vast datasets.

Sentiment Analysis, also known as Opinion Mining [1], is a captivating field that focuses on developing techniques to evaluate and condense sentiments from extensive opinionated text data. By understanding user sentiments, both consumers and manufacturers can make more informed decisions, contributing to a dynamic marketplace. The goal of SA is to examine attitudes or opinions on various subjects [2]. The rapid growth of internet-based applications, such as websites, social networks, and blogs, has led to a massive influx of opinions and reviews about goods, services, and everyday activities. Governments, research scholars, and businesses use SA to extract and evaluate public mood, obtain business insights, and enhance decision-making [3].

In this paper, we have conducted a review of existing studies focused on SA within eCommerce datasets. Additionally, we have examined papers based on the methodologies employed and the valuation matrices utilized. The discussion in this paper will encompass various levels of SA techniques, including machine learning models, methods, and deep learning models. It is a concise review paper comprising essential works tailored for beginners in SA. The selected

* Corresponding author: Amrithkala M Shetty

papers offer insights into the fundamentals of SA, its applications, and the associated challenges. These chosen papers, sourced from top journals, provide a clear understanding of this technique.

Section 2 of this paper explains the different levels of SA. Section 3 describes the application domain of SA. In Section 4, an overview of SA techniques is provided, encompassing both machine learning approaches and deep learning approaches. Section 5 includes information about the datasets used, while Section 6 discusses the challenges associated with SA. Section 7 contains the literature review, providing descriptions of each paper. Finally, the paper concludes in Section 8.

2. Sentiment analysis levels

The analysis levels can be done namely Document level, Sentence level, Phrase level and Aspect/Feature level analysis

2.1. Document level SA

In the realm of document-level sentiment analysis, the objective is to ascertain whether an entire opinion document conveys a positive or negative sentiment [4]. Although less commonly employed, this approach finds relevance in scenarios such as categorizing chapters or pages of a book into positive, negative, or neutral sentiments. Supervised and unsupervised learning techniques can both be utilized for the classification of documents at this level [5].

Within document-level sentiment analysis, two prominent challenges emerge: cross-domain and cross-language sentiment analysis [6]. For instance, a system might determine the overall sentiment, positive or negative, expressed in a product review. This analysis operates under the implicit assumption that each document encapsulates opinions related to a singular entity, like a specific product or service. However, for documents assessing or comparing multiple elements, a more nuanced analysis is required, making the approach less suitable for such document types [4].

2.2. Sentence level SA

At this analytical tier, our focus narrows down to individual sentences, where we conduct a detailed analysis to ascertain whether each sentence holds a positive, negative, or neutral polarity. This approach proves valuable when a document comprises multiple sentences, each carrying diverse sentiments [4]. However, achieving this objective necessitates the preliminary classification of sentences into objective (conveying factual information) or subjective (conveying opinions and views) categories [3].

To accomplish this, a similar methodology to document-level sentiment analysis is employed, with the classification of each sentence’s polarity based on available training data and processing resources. The polarities assigned to individual sentences can then be aggregated to deduce the overall sentiment of the document or utilized independently.

2.3. Phrase level SA

Sentiment analysis can also be executed at the phrase level, involving the mining of opinion words within phrases for subsequent classification. Phrases can encompass varying numbers of aspects, ranging from one to several. This approach, known as phrase-level sentiment analysis, is particularly beneficial in scenarios involving multi-line product reviews [5], [7].

Recent research has witnessed substantial exploration in this domain. While sentence-level analysis proves advantageous due to its ability to capture both positive and negative statements within a text, document-level analysis predominantly focuses on categorizing the entire document as subjective, either positively or negatively. Within this context, the word, as the simplest unit of language, plays a crucial role, and its polarity is inherently influenced by the subjectivity of the sentence or document in which it is embedded. Notably, the presence of adjectives significantly enhances the likelihood of a sentence being subjective [5].

2.4. Aspect/feature level SA

At the aspect level, Sentiment Analysis (SA) is executed to delve into the fine-grained details of opinions, considering that each sentence may encompass multiple aspects. Formerly referred to as feature-level sentiment analysis or future-based opinion mining and summarization, aspect-level analysis delves directly into opinions. Particularly prominent in business contexts, Aspect-Based Sentiment Analysis (ABSA) plays a pivotal role in enhancing business operations by pinpointing specific product qualities requiring improvement based on consumer feedback, thereby contributing to the product’s market success.

ABSA precisely identifies the aspects within a product review and assigns sentiment classes to each aspect mentioned in the review. For instance, in the phrase ”The camera of iPhone 11 is fantastic,” the sentiment is expressed toward the aspect ”camera” of the entity ”iPhone 11,” and it is classified as positive [3].

While classifying opinionated texts at the document or sentence level proves beneficial in various scenarios, these approaches often lack the nuanced details required for certain applications. Simply having a positive opinionated document about a specific object does not necessarily imply unanimous approval of all the object’s aspects or features. Similarly, a negative opinionated document doesn’t necessarily mean disapproval of every aspect. Authors commonly include both positive and negative elements of an object in opinionated texts, even if the overall sentiment is positive or negative. Such nuanced information is often absent in classifications conducted at the document or sentence level [8].

3. Application domain of SA

Sentiment Analysis (SA) holds diverse applications, ranging from assessing patient mental health conditions through social media posts to evaluating consumer opinions. Technological advancements like Blockchain, IoT, Cloud Computing, and Big Data have significantly expanded SA’s scope, making it applicable across various industries. Here are some major sectors where SA is extensively utilized:

3.1. Business Intelligence

SA plays a pivotal role in business intelligence applications such as credit checks and assessing corporate reputation. It offers valuable insights into customer perceptions, aiding in brand revitalization and strategic decision-making. Deep learning, with its automatic feature extraction and elimination of redundant features, is particularly effective in this domain. Businesses leverage SA data to enhance products, analyze user feedback, and formulate new marketing strategies. This application is not limited to product manufacturers; consumers can also use SA to make informed decisions based on reviews [3], [5].

3.2. Recommendation system

The objective of a recommendation system (RS) is to provide valuable suggestions to users regarding products or services that align with their interests. Customer reviews play a pivotal role in aiding individuals in their decision-making process when contemplating a purchase or service utilization online. Prior to making a buying decision, consumers often rely on perusing numerous user evaluations and feedback shared on social media platforms. It’s evident that users are more likely to opt for products that boast high ratings or receive positive customer reviews. Hence, employing sentiment analysis systems can be advantageous in recommending products [9], [10].

3.3. Government Intelligence

SA is a powerful tool for analyzing public opinions on various topics, including politics, religion, and social issues. Governments can leverage SA and machine learning techniques to understand public sentiments on policies, products, and services. During the 2012 US presidential elections, public sentiments expressed on social media were studied to inform policy decisions [3], [11]

3.4. Healthcare and medical domain

The healthcare industry has recently embraced SA to gather data on diseases, adverse drug reactions, epidemics, and patient sentiments. SA can contribute to understanding public perceptions of medical professionals, medications, illnesses, and treatments. This information can enhance patient treatment, public health monitoring, and epidemic management [3], [12].

3.5. Stock market

SA is employed for stock price prediction by analyzing news and social media content related to the stock market. Real-time information from platforms like Twitter and blogs is analyzed to predict stock price trends. Positive news tends to drive up stock prices, while negative news can lead to a downturn. Although not extensively explored, there is potential for using SA in stock market predictions [3], [5].

4. SA techniques

Sentiment Analysis (SA) utilizes a range of techniques, including lexicon-based, machine learning, and deep learning approaches, which are widely recognized. There is a constant quest among researchers for more effective methods that can achieve higher accuracy while minimizing computing costs. An emerging trend in this pursuit involves exploring nature-inspired algorithms, which are classified into physics-chemistry-based algorithms, bio-inspired algorithms, and other categories. This area of research is currently experiencing significant growth [13].

4.1. Lexicon Based approach

Lexicons consist of tokens, each representing the neutral, positive, or negative aspects of the text, with assigned scores. These scores, typically ranging from +1 for positive sentiment to -1 for negative sentiment, determine the polarity of each token. In the lexicon-based approach, scores for individual tokens are aggregated, with separate summations for positive, negative, and neutral scores. The overall polarity of the text is then determined based on the highest aggregated score in the final stage of analysis [5].

There are mainly two approaches used in Lexicon-based methods:

- Dictionary based approach
- Corpus based approach

4.1.1. Dictionary based approach

This approach involves the creation of a dictionary containing a limited number of words initially. To broaden the vocabulary, synonyms and antonyms for these words are included, often sourced from a thesaurus, online dictionaries, or resources like WordNet. The dictionary expansion process continues until further additions are not feasible, at which point manual review is employed to refine and enhance the lexicon [14].

4.1.2. Corpus based approach

This approach determines context-specific words by their sentiment orientation through two methods:

- Statistical approach: Words that frequently occur in positive contexts are assigned positive polarity, while those occurring frequently in negative contexts are assigned negative polarity. Words with similar frequencies in both positive and negative contexts are considered neutral in polarity [3], [14].
- Semantic approach: This method enhances sentiment analysis by associating sentiment values not only with individual words but also with phrases, taking into account synonyms, antonyms, and semantically related terms. By incorporating a broader semantic understanding, it aims to capture nuanced sentiment nuances more accurately [3], [5], [14].

4.2. Machine learning approach

Machine learning approaches in sentiment analysis encompass various methodologies such as supervised learning, unsupervised learning, semi-supervised learning, and reinforcement learning. A multitude of machine learning algorithms are applied in sentiment analysis tasks, including Support Vector Machine (SVM), Logistic Regression (LR), Decision Tree (DT), Maximum Entropy (ME), K-Nearest Neighbors (KNN), among others [3], [5], and more.

- Support vector machine(SVM): Support Vector Machines tackle classification problems in big data scenarios by identifying an optimal decision boundary, or hyperplane, to separate different classes in n-dimensional space. This approach determines the boundary with the maximum margin, aided by support vectors, which are critical data points defining the hyperplane.
- Naive Bayes (NB): NB is a straightforward and efficient classification method commonly applied in sentiment analysis, text classification, spam filtering, and recommendation systems. It utilizes Bayes' theorem to predict unknown class labels, particularly effective in Natural Language Processing tasks.
- Logistic regression(LR): LR predicts the probability of binary events occurring and is widely used in binary classification tasks. It is frequently employed in scenarios where the outcome variable has two possible states, making it a popular choice in various machine learning applications.
- Decision tree(DT): Decision tree classifiers employ a tree-like structure where each internal node represents a test on a specific attribute. While effective for large datasets, decision trees may suffer from instability and overfitting, particularly with smaller datasets [3].

- Maximum Entropy(ME): Maximum Entropy classifiers, also known as probabilistic classifiers, are a version of the exponential model. Unlike naive Bayes, they do not assume conditional independence among features. Maximum Entropy classifiers find applications in various text classification tasks, including sentiment analysis [15].
- K-nearest neighbours (KNN): K-nearest Neighbors is a simple supervised machine learning algorithm utilized for classification and regression tasks. While not as commonly employed in sentiment analysis, KNN can yield favorable results when appropriately trained. It operates on the premise that the classification of a test sample will resemble that of its nearest neighbors, with the choice of K value critical for its performance [5].

4.3. Deep learning (DL) approaches

DL provides a layered structure for the hidden layers of neural networks. In contrast to traditional ML approaches that often require manual feature selection or definition, deep learning models autonomously learn and extract features from data, leading to improved accuracy and performance. Additionally, deep learning models typically optimize their hyperparameters automatically, alleviating the need for manual tuning [3].

DL and artificial neural networks have emerged as highly effective solutions for a multitude of tasks, including image and speech recognition, as well as natural language processing [16]. They excel in learning intricate patterns and representations from complex data, making them particularly well-suited for tasks where traditional methods may struggle to capture nuanced relationships.

- **Recurrent neural network(RNN**): RNN, a type of neural network, utilizes the output from the previous step as the input for the current step. Unlike traditional neural networks that treat inputs and outputs as independent entities, RNNs are well-suited for tasks that require consideration of sequential information, such as predicting the next word in a sentence. This is achieved through the incorporation of a hidden layer. The key component of RNN is the hidden state, which retains specific information about a sequence. RNNs possess a ”memory” that stores previously calculated information. By using the same parameters for each input, RNNs perform the same task on all inputs or hidden layers, thereby minimizing parameter complexity compared to other neural networks.
- **Convolutional Neural Networks (CNN)**: Deep learning, owing to its ability to handle large datasets, has emerged as a powerful tool. Among deep neural networks, Convolutional Neural Networks (CNNs) have gained prominence. Initially developed in the 1980s, CNNs found significant zip codes, and pin codes in postal services.
- CNN architecture typically comprises convolutional and pooling (or subsampling) layers, followed by a fully connected classification layer. Convolution layers employ filters to extract features from inputs, and the outputs of multiple filters can be combined. Pooling or subsampling layers reduce feature resolutions, enhancing CNNs' robustness to noise and distortion. Fully connected layers are responsible for classification tasks [16].
- **Long Short Term Memory network (LSTM):** Long Short Term Memory Networks, or ”LSTMs,” are a specialized variant of RNNs designed to capture long-term dependencies effectively. Initially introduced by Hochreiter and Schmidhuber in 1997, LSTMs have since been refined and popularized, proving highly effective across various domains. Specifically engineered to address the challenge of maintaining information over extended sequences, LSTMs excel at capturing and retaining new information for prolonged periods. The distinguishing feature of LSTM architectures lies in their gated units or cells within the hidden layer. Comprising four layers, these units collectively manage the cell state and output layer, which are then passed to the next hidden layer. Unlike traditional RNNs with a single tanh layer, LSTMs incorporate three logistic sigmoid gates along with a tanh layer, facilitating better long-term memory retention .

5. Datasets

E-commerce datasets can be obtained from Kaggle, the UCI Machine Learning Repository, Amazon Customer Review, and social media APIs. In this section, we have listed some of the e-commerce datasets.

- Amazon Product Data: This dataset, derived from a comprehensive Amazon review dataset provided by Stanford professor Julian McAuley, spans from May 1996 to July 2014. It comprises reviews accompanied by various attributes such as ratings, text, helpful votes, product description, category information, price, brand, and image features.
- Amazon and Best Buy Electronics: This dataset encompasses over 7000 reviews covering 50 electronic products sourced from both Amazon and Best Buy. It includes details like date, source, rating, title, reviewer metadata, and more.

- Women’s E-Commerce Clothing Reviews: Derived from Amazon, this retail dataset comprises approximately 23,000 authentic customer reviews and ratings specifically focused on women’s clothing.
- Multi-Domain Sentiment Dataset This dataset contains reviews, both positive and negative, for numerous Amazon products. Despite focusing on older products, it remains highly valuable. Sourced from the John Hopkins University Department of Computer Science, it includes reviews accompanied by binary-convertible star ratings ranging from 1 to 5.
- Flipkart Reviews The dataset comprises reviews provided by customers who have purchased products from Flipkart [17]. These reviews reflect the customers’ experiences with the products bought through Flipkart, accompanied by a corresponding rating. The dataset includes three columns: Product name, review and ratings.

6. Challenges of SA

Analyzing and comprehending emotions in natural language poses significant challenges for sentiment or emotion analysis, as machines require training to approach human-like understanding. Moreover, understanding the subtleties of different languages adds another layer of complexity. However, sentiment analysis (SA) software is progressively improving to tackle these obstacles as data science evolves. The following outlines the primary challenges that need to be addressed in SA:

6.1. Sarcasm detection

Sarcasm, defined as ”a sharp, bitter, or cutting expression or remark,” is a fundamental aspect of human social interaction, capable of both eliciting

However, recognizing sarcasm can be challenging as it often involves indirect expression of emotions. One approach to detecting sarcasm involves using irony to mock or convey disdain. The use of irony to mock or convey contempt.

Numerous methods have been proposed for sarcasm detection. Yu du et al. [18] introduced a dual-channel convolutional neural network that analyzes both the semantic content and sentimental context of the target text. Additionally, they incorporated SenticNet to enrich a long short-term memory (LSTM) model with common sense. Furthermore, they implemented an attention mechanism to account for the user’s expression patterns.

By combining these approaches, researchers aim to improve the accuracy of sarcasm detection algorithms, thereby enhancing our understanding of sarcastic communication in digital interactions.

6.2. Negation handling

Negation words, like ”not,” ”neither,” and others, play a crucial role in sentiment analysis (SA) as they can invert the polarity of a text, significantly impacting its sentiment. For example, the statement ”This movie is good” expresses a positive sentiment, whereas ”The movie is not good” conveys a negative sentiment due to the negation term ”not.”

Unfortunately, some SA systems overlook negation words either because they are included in stop-word lists or because they are considered to have a neutral sentiment value in lexicons and are thus ignored. However, neglecting negation words can lead to inaccurate sentiment analysis results. It’s essential to handle negation terms carefully because not all occurrences of these words imply negation. For instance, in phrases like ”not only but also,” the term ”not” does not negate the sentiment. Therefore, distinguishing between instances where negation words affect sentiment and where they do not is crucial for accurate sentiment analysis [2]. Partha Mukherjeea, Youakim Badra et al. [19] provided a novel end-to-end SA method for dealing with negations, which also included scope marking and negation identification. They proposed a customized negation marking algorithm for explicit negation detection and conducted experiments on SA of Amazon reviews, specifically reviews of cell phones, using a variety of machine learning algorithms, including Naive Bayes, Support Vector Machines (SVM), Artificial Neural Network (ANN), and Recurrent Neural Network (RNN). RNN had the best accuracy of 95.67 percent .

6.3. Spam detection

Spam detection is crucial in sentiment analysis (SA) as online opinions significantly impact customer purchasing decisions. Fraudulent reviews, known as spam reviews, can harm a brand’s reputation and manipulate perceptions of goods or services. Spam reviews are created to promote or demote specific products, often for profit or fame, a practice referred to as review spamming [20]. Distinguishing fake reviews from genuine ones is challenging due to the lack of obvious differences between them [21]. As customer purchasing decisions are influenced by online comments, spam and fake reviews can harm a brand’s reputation and artificially influence users perceptions about goods, services, businesses or other entities [21]. The problem of creating a spam detection system that can detect fake reviews among a large number of reviews is extremely difficult because there are no obvious differences between reviews.

6.4. Audio-visual data

Analyzing sentiment in videos poses unique challenges compared to text-based data. In addition to transcribing spoken content, video sentiment analysis must also consider visual elements such as brand logos. Social media videos further complicate the task by including user comments alongside the video content. To address these challenges, a sophisticated sentiment analyzer is required, capable of extracting insights from both text and video data.

6.5. Low-resource languages

Most SA research is done in languages like english and others with rich linguistic resources. Low-resource languages are those that experience a shortage of linguistic resources. Algorithms for supervised learning are used for languages having linguistic resources. This issue can be solved using a variety of techniques, including building linguistic resources from scratch and applying unsupervised, semi-supervised, and transfer learning techniques.

6.6. SA of code-mixed data

Natural language processing has a significant challenge when dealing with code-mixed data because of how drastically different its properties are from the traditional structures of standard languages. The need for SA of social media text, most of it is code-mixed, is growing. Due to the complexity of mixing at various levels of the text, systems trained on monolingual data fail when applied to code-mixed data. To construct models specifically for code-mixed data, there are, however, very few resources accessible. Even though semi-supervised or unsupervised methods have been utilized in a lot of research on multilingual and cross-lingual SA, supervised methods still outperform them. There are very few datasets accessible for popular languages like English-Spanish, English-Hindi, and English-Chinese [22].

7. Literature review

In this section, we present recent research works related to sentiment analysis.

Chang YC, Ku CH, and Le Nguyen DD proposed a study to address a research gap by identifying perceived aspect-based sentiments and predicting unrated sentiments across different categories [23]. Their research extends the application of data-driven and visual analytics approaches to enhance understanding of customer satisfaction within the airline industry, particularly in the context of the COVID-19 pandemic. Moreover, the study aims to improve existing sentiment analysis methodologies.Neha Nandal, Rohit Tanwar, et al. implemented aspect-level sentiment detection, which targets specific features of the reviewed item [11]. They applied this approach to Amazon customer reviews by first selecting aspect phrases for each review from crawled data. Subsequently, the system preprocesses the data and assigns a rank based on whether the sentiment expressed is positive or negative. Long Mai and Bac Le proposed a framework combining Aspect-based Sentiment Analysis (ASA) and Sentiment Strength Analysis (SSA) for automatically gathering, organizing, and analyzing YouTube comments related to a specific product [24]. Their joint approach leveraged the complementary benefits of SSA and ASA. They experimented with various architectures, including BGRU, ELMO, and BERT, to determine the optimal architecture for Joint Sentiment Analysis (JSA). Through multiple trials, they found that the BERT model yielded the best results for JSA.

Joni Salminen, Chandrashekhar Kandpal, et al. conducted experiments using ULMFiT and GPT-2 language models to generate fake product reviews based on an Amazon e-commerce dataset [25]. They created a dataset for a classification task to identify fake reviews, utilizing the superior performance of the GPT-2 model. Their findings indicate that while identifying fake reviews can be challenging for humans, machine learning models, particularly GPT-2, can effectively detect fake reviews generated by humans, demonstrating the potential of automated systems in combating fraudulent reviews.

Vandana Yadav et al. proposed a machine learning-based system for aspect-based sentiment analysis in the Hindi language [26]. Their system classified each sentence in Hindi reviews, specifically from the Amazon dataset, into categories such as positive, negative, neutral, or conflictual. Praphula Kumar Jain et al. introduced a machine learning model optimized with cuckoo optimization for sentiment analysis using data from the website https://www.airlinequality.com [27].

Xun Wang et al. proposed a framework for sentiment analysis focusing on harshness in product evaluations [28]. Their approach involved creating a probabilistic graphical model incorporating harshness and user feedback, followed by

Bayesian-based inference for sentiment mining. Experimental results demonstrated superior performance compared to existing approaches, particularly in aligning with professional assessments. M. Shivkumar and Srinivasaulu Reddy Uyyala developed an intelligent system for classifying customer review words into different aspects with four labels, using long short-term memory and fuzzy logic [29]. Their model, tested on benchmark datasets for Amazon reviews, outperformed state-of-the-art approaches, and also considered geographical location and current trends. Najla M. Alharbi et al. evaluated various deep learning methods for predicting customers' opinions based on mobile phone reviews from Amazon.com [30]. Their study analyzed deep learning algorithms and feature extraction strategies, showing the effectiveness of different approaches based on accuracy metrics. Abhishek Kumar et al. developed a sentiment detection model using decision tree and random forest classifiers, which demonstrated high efficiency among various methods tested [31].

Hai Wan et al. proposed a method for sentiment detection from both the target and the aspect, achieving excellent performance in detecting target-aspect-emotion triples [32]. Chao Wu et al. introduced a multiple element joint detection model (MEJD) for aspect-based sentiment analysis, which demonstrated effectiveness on restaurant review datasets from SemEval [33]. Their model integrated bidirectional long short-term memory, graph convolutional network with attention mechanisms, and BERT embeddings for aspect-sentence joint input representation. Feng Xu et al. introduced a continuous naive Bayes learning framework for sentiment classification of product reviews on large-scale, multi-domain e-commerce platforms [34]. They aimed to maintain the high computational efficiency of the standard naive Bayes model while addressing the challenges of handling large-scale and multi-domain data. Their approach extended the naive Bayes parameter estimation mechanism to a continuous learning paradigm. This continuous learning framework allowed for the incorporation of both generic classification knowledge from old domains and domain-specific knowledge from new domains, thereby improving distribution learning. By adapting to evolving data distributions, their method enhanced sentiment classification performance across diverse e-commerce domains.E. Suganya and S. Vijayarani utilized web scraping to gather online product reviews for their study [35]. Subsequently, they conducted opinion or sentiment analysis on these collected reviews using various classification models, including KNN (K-Nearest Neighbors), SVM (Support Vector Machine), RF (Random Forest), CNN (Convolutional Neural Network), and a hybrid SVM-CNN model proposed in their work.

Their experiments with these classification models yielded promising results, indicating the effectiveness of each approach in sentiment analysis of online product reviews. Md. Mahfuzur Rahman et al. conducted sentiment analysis of mobile app users' opinions in their study [36]. The main objective of their work is to assist developers in assessing whether users' opinions regarding their apps are positive or negative. To achieve this goal, they employed machine learning classifiers such as KNN (K-Nearest Neighbors), Random Forest (RF), SVM (Support Vector Machine), Decision Tree, and Naive Bayes in their NLP-based approaches. Anu J Nair et al. utilized Logistic Regression sentiment analysis, VADER sentiment analysis, and BERT sentiment analysis to determine the sentiment of tweets in their proposed study [37]. These analysis approaches were tailored to be more sensitive to sentiment expressions in social media environments. The study aimed to provide valuable insights into public opinion for government or health officials, enabling them to make informed decisions based on the acquired results.

Naveen Kumar Gondhi et al. investigated sentiment analysis in the context of online shopping reviews, focusing on reviews from Amazon's cell phone and accessories category [38]. They employed deep learning techniques, specifically Long Short-Term Memory Networks (LSTM), to classify sentiment. The model's performance was evaluated using accuracy, precision, recall, and F1 score metrics, with the F1 score reaching 93%, indicating strong performance.

In another study by Kubrusly et al. [39], three tree-based classification methods, including Classification Tree, Random Forest (RF), Gradient Boosting, and XGBoost, were utilized to classify reviews from apparel e-commerce sites. The Gradient Boosting technique demonstrated consistent results, while XGBoost showed signs of overfitting when the number of trees was too high. The Classification Tree excelled at detecting negative reviews but was less effective at identifying positive ratings. Furthermore, in a study aimed at establishing a relationship between review attributes and product recommendations in the e-commerce industry, five machine learning algorithms—Logistic Regression (LR), Support Vector Machine (SVM), Random Forest (RF), XGBoost, and Light GBM—were employed [40]. The study utilized a dataset of reviews of women's clothing on Amazon. Among the algorithms, Light GBM yielded the maximum Area Under the Curve (AUC) value and accuracy, indicating superior performance in this context. The RBF Kernel SVM algorithm, however, had the worst results, with an accuracy score of 81%. In [41] 2500 hotel reviews using Indonesian language were applied in this research. Both the Word2Vec model and the Long-Short Term Memory (LSTM) model are used in the study. LSTM and Word2Vec variables were combined. According to a number of observable parameters, both models are used. The Word2Vec architecture Skip-Gram classification system, with Hier- archical Softmax as a technique of assessment, would be the ideal scheme for sentiment classification utilising Word2Vec and LSTM models.

For the LSTM, an accuracy value of 85.96% was obtained by setting the dropout value to 0.2 and the learning rate value to 0.001.

8. Conclusion

This paper provides a comprehensive overview of SA techniques applied to E-commerce datasets. Offering insights into major works in SA, this review paper serves as an invaluable resource for beginners seeking to grasp the intricacies of this topic. By summarizing key papers, it facilitates a foundational understanding for newcomers while also standing as a potential catalyst for further research. The research's potential is evident in its capacity to expand through the inclusion of additional review papers, thereby contributing to the continual advancement of SA in the context of E-commerce

Table I Literature Review

Paper	Methods	Algorithm	Metrics	Dataset	Observation
[23]	Aspect-level sentiment analysis and visual analytics are employed to concentrate on diverse facets of airline service and explore how ABSA can glean insights into service quality.	DL and word embedding techniques	Precision, recall, and F1 score	COVID-19	ABSA and visual analytics will be used to better understand customer satisfaction before and during the pandemic.
[11]	A novel approach based on aspect-level sentiment identification that focuses on the item's features has been proposed.	SVM with three kernels, linear, polynomial, and radial basis function (RBF), were used.	MSE, accuracy, precision, recall, confusion matrix and ROC curves.	Amazon reviews data) customer (crawled	Bipolar words sentiment analysis. in
[24]	ASA and SSA collaborated to develop a novel framework for automatically collecting, filtering, and analyzing YouTube comments for a certain product.	GRU, BGRU, BGRU-CRF, BGRU-CRF-ELMO BERT, BERT-CRF.	Accuracy, For ASA we use ,precision, recall and F1-score	YouTube comments for sentiment analysis (YCSA) was developed using Vietnamese comments.	BERT model provided the best performance for JSA.
[25]	Fake review creation and detection on an Amazon e-commerce dataset.	ULMFiT and GPT-2, to generate fake product reviews based on an Amazon e-commerce dataset.	Amazon Review Data (2018)	Precision, F1-score recall,	Fake review detection. Fortunately, machine learningclassifiers perform significantly better in this regard, recognizing reviews written by other

					machines with perfect accuracy.
[26]	The analysis of Hindi customer reviews from Amazon was presented.	Support vector machine	Accuracy	Data scraped from the website https://www.airlinequalit y.com	Amazon reviews and identification of ABSA category aspects in Hindi
[27]	They analyzed and developed a consumer recommendation prediction system using reviews and ratings obtained from airline review websites.	To predict airline recommendations, we proposed a cuckoo optimized machine learning model.	Accuracy, F1 measure, and Area Under the Curve (AUC)	Data scraped from the website https://www.airlinequalit y.com	To produce robust consumer recommendation prediction systems.
[28]	Product analysis method based on sentiment analysis based on user harshness analysis	Harshness-aware sentiment analysis framework for product review	Accuracy, Precision, Recall, F-measure	IMDB and Amazon	Removes both harsh and tolerant reader's comments, which is not the best strategy.
[29]	Amazon review sentiment analysis based on location	Fuzzy logic and long-term short memory	Precision, recall, F-score, and Accuracy	Amazon cell phones Review, Amazon video games Review, Consumer Review of Amazon products	Consumer review sentences are classified based on the customer's location and current trends.
[30]	Evaluating different deep learning approaches to accurately predict the opinion of customers on Amazon mobile reviews	RNN and its four variants and Word embedding	Accuracy, Recall, Precision, and F1-score	Amazon Phone reviews	Accurately predict Customer reviews on Amazon phone.
[31]	Proposed a model for detecting sentiments using sentiment analysis	Decision tree Classifier and random forest classifier	Amazon reviews	Accuracy, Precision, Recall	To find the sentiment associated with the review of the dataset and obtain higher accuracy.
[32]	Proposed a novel approach for joint target-aspect sentiment detection	TAS-BERT Model	Micro F1-score	SemEval-2015 Task12 and SemEval-2016 Task5	Target and aspect joint sentiment detection and also detects implicit targets.
[33]	Proposed an end-to-end multiple element joint detection	MEJD model uses the Bi-LST and the GACN module to extract	Micro F1-score	SemEval 2015 Task 12 and SemEval 2016 Task 5	Target and aspect joint sentiment detection and also detects implicit targets.

	model (MEJD), which effectively extracts all (target, aspect, sentiment) triples from a sentence	the interactive information between aspect and sentence, which can extract target-aspect-sentim triples effectively	ent		
[34]	On an e-commerce platform, a continuous naive Bayes learning system for sentiment categorization of large-scale and multi-domain product reviews was presented.	Naive Bayes model	F-score	Amazon multi-domain review sentiment data sets	To learn on a continuous basis from increasing reviews and multiple domains.
[35]	Online product reviews are analysed using classification models	KNN, SVM, Random Forest, CNN and proposed hybrid SVM-CNN	Precision, Recall, F-Score, Accuracy	Reviews are scraped from online shopping websites such as Amazon, Flip kart and Snap Deal	Sentiment Analysis product Reviews from multiple Web Pages is done using machine learning algorithms.
[36]	Sentiment analysis of mobile application users has been explained	KNN, Random Forest (RF), SVM, Decision Tree, Naive Bayes	Mobile App Review dataset	Precision, Recall, F1-score,Accuracy	Is to assist developers in identifying whether users' opinions on their apps are positive or negative.
[37]	To conduct sentiment analysis and examine the various emotions of a large public in relation to COVID-19.	Logistic regression, Valence Aware Dictionary and sentiment Reasoner (VADER), Bidirectional Encoder Representation from Transformers (BERT)	Accuracy, Precision, Recall, F1-score	Dataset related to tweets with different hashtags related to COVID -19	Is to develop a sentiment analysis algorithm to investigate user sentiment in COVID tweets.
[38]	Describes the sentiment analysis process and its requirements.	LSTM has been combined with word2vec representation	Accuracy, precision, recall, and F1 score	Amazon Review dataset 2018	The main goal of this is to verify the LSTM model's functionality with a large amount of data.
[39]	Is to conduct a Text Mining approach on a set	Random Forest, Gradient Boosting and XGBoost	Accuracy, precision, recall, F1	Women's Clothing E-Commerce Reviews	Classification Tree is strong at detecting bad reviews but poor at

	of customer reviews		score and specificity		detecting positive ratings.
[40]	Figure out the insight Correlation between review features and product recommendation based on NLP	Logistic Regression, Support Vector Machine (SVM), Random Forest, XGBoost and LightGBM	Accuracy, precision, recall, F1 score	Women's E-Commerce ClothingReview dataset	generate a deeper comprehensionof customer sentiment and grasp customer psychology in e-commerce transaction industry.
[41]	Analyzing the performance of LSTM using word2vec	Long-Short Term Memory (LSTM) model and Word2Vec model	Accuracy	hotel reviews	Analysing the performance of Word2Vec and LSTM for classifying sentiment in hotel reviews with various parameters.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] F. Pozzi, E. Fersini, E. Messina, and B. Liu, Sentiment Analysis in Social Networks. Elsevier Science, 2016. [Online]. Available: https://books.google.co.in/books?id=aS2lCgAAQBAJ

[2] B. Liu, Sentiment analysis and opinion mining, Synthesis lectures on human language technologies, vol. 5, no. 1, pp. 1–167, 2012.

[3] M. Birjali, M. Kasri, and A. Beni-Hssane, A comprehensive survey on sentiment analysis: Approaches, challenges and trends, Knowledge-Based Systems, vol. 226, p. 107134, 2021.

[4] B. Liu, Opinions, Sentiment, and Emotion in Text, ser. Sentiment Analysis: Mining Opinions, Sentiments, and Emotions. Cambridge University Press, 2015. [Online]. Available: https://books.google.co.in/books?id=6IdsCQAAQBAJ

[5] M. Wankhade, A. C. S. Rao, and C. Kulkarni, A survey on sentiment analysis methods, applications, and challenges, Artificial Intelligence Review, pp. 1–50, 2022.

[6] P. Bhatia, Y. Ji, and J. Eisenstein, Better document-level sentiment analysis from rst discourse parsing, arXiv preprint arXiv:1509.01599, 2015.

[7] T. T. Thet, J.-C. Na, and C. S. G. Khoo, Aspect-based sentiment analysis of movie reviews on discussion boards, Journal of Information Science, vol. 36, pp. 823 – 848, 2010.

[8] B. Liu et al., Sentiment analysis and subjectivity. Handbook of natural language processing, vol. 2, no. 2010, pp. 627–666, 2010.

[9] S. B. Bhonde and J. R. Prasad, Sentiment analysis-methods, applications & challenges, International Journal of Electronics Communication and Computer Engineering, vol. 6, no. 6, p. 634, 2015.

[10] A. Ziani, N. Azizi, D. Schwab, M. Aldwairi, N. Chekkai, D. Zenakhra, and S. Cheriguene, Recommender System Through Sentiment Analysis, in 2nd International Conference on Automatic Control, Telecommunications and Signals, Annaba, Algeria, Dec. 2017. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01683511

[11] A. Rao and R. Vijayakumar, Empower good governance with public assessed schemes by improved sentiment analysis accuracy, Electronic Government, an International Journal, vol. 16, p. 118, 01 2020.

[12] R. Bose, R. Dey, S. Roy, and D. Sarddar, Sentiment Analysis on Online Product Reviews, 06 2019, pp. 559–569.

[13] A. Yadav and D. Vishwakarma, A comparative study on bio-inspired algorithms for sentiment analysis, Cluster Computing, vol. 23, 12 2020.

[14] N. Gupta and R. Agrawal, Chapter 1 - application and techniques of opinion mining, in Hybrid Computational Intelligence, ser. Hybrid Computational Intelligence for Pattern Analysis and Understanding, S. Bhattacharyya, V. Sna´sˇel, D. Gupta, and A. Khanna, Eds. Academic Press, 2020, pp. 1–23. [Online]. Available:https://www.sciencedirect.com/science/article/pii/B9780128186992000019

[15] S. Mishra, P. Mallick, H. Tripathy, G. Chae, and B. Mishra, Impact of AI and Data Science in Response to Coronavirus Pandemic, ser. Algorithms for Intelligent Systems. Springer Singapore, 2021. [Online]. Available: https://books.google.co.in/books?id=q8I5EAAAQBAJ

[16] N. C. Dang, M. N. Moreno-Garc´ıa, and F. De la Prieta, Sentiment analysis based on deep learning: A comparativestudy, Electronics, vol. 9, no. 3, p. 483, 2020.

[17] H. Edison. (2024) Flipkart reviews sentiment analysis. Last accessed: May 8, 2024. [Online]. Available: and J. P. McCrae, A sentiment analysis dataset for code-mixed Malayalam-English, in Proceedings of the 1st Joint Workshop on Spoken Language https://www.kaggle.com/datasets/harishedison/flipkart-reviews-sentimTeecnht-naonlaolgyiseiss for Under-resourced languages (SLTU)

[18] Y. Du, T. Li, M. S. Pathan, H. K. Teklehaimanot, and Z. Yang, An effective sarcasm detection approach based on sentimental context and individual expression habits, Cognitive Computation, vol. 14, no. 1, pp. 78–90, 2022.

[19] P. Mukherjee, Y. Badr, S. Doppalapudi, S. Srinivasan, R. Sangwan, and R. Sharma, Effect of negation in sentences on sentiment analysis and polarity detection, Procedia Computer Science, vol. 185, pp. 370–379, 2021, publisher Copyright: © 2021 Elsevier B.V.. All rights reserved.; 2021 Complex Adaptive Systems Conference ; Conference date: 16-06-2021 Through 18-06-2021.

[20] N. Hussain, H. Turab Mirza, G. Rasool, Hussain, and M. Kaleem, Spam review detection techniques: A systematic literature review, Applied Sciences, vol. 9, no. 5, 2019. [Online]. Available: https://www.mdpi.com/2076-3417/9/5/987

[21] E. F. Cardoso, R. M. Silva, and T. A. Almeida, Towards automatic filtering of fake reviews, Neurocomputing, vol. 309, pp. 106–116, 2018. [Online]. Available: and Collaboration and Computing for Under-Resourced Languages (CCURL). Marseille, France: European Language Resources association, May 2020, pp. 177–184. [Online]. Available: https://aclanthology.org/2020.sltu-1.25 https://www.sciencedirect.com/science/article/pii/S092523121830531S9ervices, vol. 64, p. 102771, 2022. [Online]. Available:

[22] B. R. Chakravarthi, N. Jose, S. Suryawanshi, E. Sherly, https://www.sciencedirect.com/science/article/pii/S0969698921003374

[23] Y. Chang, C. Ku, and D. Nguyen, Predicting aspect-based sentiment using deep learning and information visualization: The impact of covid-19 on the airline industry, Information and Management, vol. 59, no. 2, Mar. 2022, funding Information: This research was supported by the Ministry of Science and Technology of Taiwan under grant MOST 107-2410-H-038 -017 -MY3, MOST 107-2634-F-001-005, and MOST 109-2410-H-038 -012 -MY2. © 2021 Elsevier B.V.

[24] L. Mai and H. B. Le, Joint sentence and aspect-level sentiment analysis of product comments, Annals of Operations Research, pp. 1–21, 2021.

[25] J. Salminen, C. Kandpal, A. M. Kamel, S. gyo Jung, and B. J. Jansen, Creating and detecting fake reviews of online products, Journal of Retailing and Consumer https://www.sciencedirect.com/science/article/pii/S092523121830531S9ervices, vol. 64, p. 102771, 2022. [Online]. Available:

[26] V. Yadav, P. Verma, and V. Katiyar, E-commerce product reviews using aspect based hindi sentiment analysis, in 2021 International Conference on Computer Communication and Informatics (ICCCI), 2021, pp. 1–8.

[27] P. K. Jain, E. A. Yekun, R. Pamula, and G. Srivastava, Consumer recommendation prediction in online reviews using cuckoo optimized machine learning models, Computers Electrical Engineering, vol. 95, p. 107397, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0045790621[04013]63P3. Muhammad, R. Kusumaningrum, and A. Wibowo,

[28] X. Wang, T. Zhou, X. Wang, and Y. Fang, Harshness-aware sentiment mining framework for product review, Expert Systems with Applications, vol. 187, p. 115887, 2022.

[29] M. Sivakumar and S. R. Uyyala, Aspect-based sentiment analysis of mobile phone reviews using lstm and fuzzy logic, International Journal of Data Science and Analytics, vol. 12, no. 4, pp. 355–367, 2021.

[30] N. M. Alharbi, N. S. Alghamdi, E. H. Alkhammash, and J. F. Al Amri, Evaluation of sentiment analysis via word embedding and rnn variants for amazon online reviews, Mathematical Problems in Engineering, vol. 2021, 2021.

[31] A. Kumar, A. P. Singh, and R. Jayaraman, Sentimental analysis for e-commerce site, in Data Intelligence and Cognitive Informatics, I. J. Jacob,S. Kolandapalayam Shanmugam, and R. Bestak, Eds. Singapore: Springer Singapore, 2022, pp. 73–83.

[32] H. Wan, Y. Yang, J. Du, Y. Liu, K. Qi, and J. Z. Pan, Target-aspect-sentiment joint detection for aspect-based sentiment analysis, in Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, no. 05, 2020, pp. 9122–9129.

[33] C. Wu, Q. Xiong, H. Yi, Y. Yu, Q. Zhu, M. Gao, and J. Chen, Multiple-element joint detection for aspect-based sentiment analysis, Knowledge-Based Systems, vol. 223, p. 107073, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0950705121003361

[34] F. Xu, Z. Pan, and R. Xia, E-commerce product review sentiment classification based on a na¨ıve bayes continuous learning framework, Information Processing Management, vol. 57, p. 102221, 02 2020.

[35] E. Suganya and S. Vijayarani, Sentiment Analysis for Scraping of Product Reviews from Multiple Web Pages Using Machine Learning Algorithms, 01 2020, pp. 677–685.

[36] M. M. Rahman, S. S. M. M. Rahman, S. M. Allayear,M. F. K. Patwary, and M. T. A. Munna, A sentiment analysis based approach for understanding the user satisfaction on android application, in Data Engineering and Communication Technology, K. S. Raju, R. Senkerik, S. P. Lanka, and V. Rajagopal, Eds. Singapore: Springer Singapore, 2020, pp. 397–407.

[37] A. J. Nair, V. G, and A. Vinayak, Comparative study of twitter sentiment on covid - 19 tweets, in 2021 5th International Conference on Computing Methodologies and Communication (ICCMC), 2021, pp. 1773–1778.

[38] N. K. Gondhi, Chaahat, E. Sharma, A. H. Alharbi, R. Verma, and M. A. Shah, Efficient long short-term memory-based sentiment analysis of e-commerce reviews, Computational Intelligence and Neuroscience, vol. 2022, p. 3464524, Jun 2022. [Online]. Available: https://doi.org/10.1155/2022/3464524

[39] J. Kubrusly, A. L. Neves, and T. L. Marques, A statistical analysis of textual e-commerce reviews using tree-based methods, Open Journal of Statistics, vol. 12, no. 3, pp. 357–372, 2022.

[40] X. Lin, Sentiment analysis of e-commerce customer reviews based on natural language processing, 04 2020, pp. 32–36.

[41] Muhammad, R. Kusumaningrum, and A. Wibowo, Sentiment analysis using word2vec and long short-term memory (lstm) for indonesian hotel reviews, Procedia Computer Science, vol. 179, pp. 728–735, 01 2021.