Review

Not peer-reviewed version

# Natural Language Processing for Analyzing Online Customer Reviews: A Survey, Taxonomy, and Open Research Challenges

Nadia Malik and Muhammad Bilal *

Posted Date: 15 January 2024

doi: 10.20944/preprints202312.2210.v2

Keywords: Natural Language Processing; Online Customer Reviews; E-Commerce; Sentiment Analysis; Opinion Mining

doi:10.20944/preprints202312.2210.v2



*Review*

# Natural Language Processing for Analyzing Online Customer Reviews: A Survey, Taxonomy, and Open Research Challenges

**Nadia Malik [1] and Muhammad Bilal [2]**

[1] Department of Management Sciences, COMSATS University Islamabad, Islamabad 4000, Pakistan
[2] Department of Computing and Information Systems, School of Engineering and Technology, Sunway University, Petaling Jaya 47500, Selangor, Malaysia
* Correspondence: Author: Muhammad Bilal[2]; Petaling Jaya 47500, Selangor, Malaysia; Email address: muhammadb@sunway.edu,my

**Abstract:** In recent years, e-commerce platforms have replaced conventional marketplaces. People are rapidly adopting internet shopping due to the convenience of making purchases from the comfort of their homes. Online review sites allow customers to share their thoughts on products and services. Customers and businesses are increasingly relying on online reviews to assess and improve the quality of products. Existing literature uses Natural language processing (NLP) to analyze customer reviews for different applications. Due to the growing importance of NLP for online customer reviews, this study attempts to provide a taxonomy of NLP applications based on existing literature. This study also examined emerging methods, data sources, and research challenges by reviewing 154 publications from 2013 to 2023 that explore state-of-the-art approaches for diverse applications. Based on existing research, the taxonomy of applications divides literature into five categories: sentiment analysis, review analysis, customer feedback and satisfaction, user profiling, and marketing and reputation management. It is interesting to note that the majority of existing research relies on Amazon user reviews. Additionally, recent research has encouraged the use of advanced techniques like Bidirectional Encoder Representations from Transformers (BERT), Long Short-Term Memory (LSTM), and ensemble classifiers. The rising number of articles published each year indicates increasing interest of researchers and continued growth. This survey additionally addresses open issues, providing future directions in online customer review analysis.

**Keywords:** natural language processing; online customer reviews; E-Commerce; sentiment analysis; opinion mining

---

## Introduction

In the digital era, online customer reviews influence consumer choices. In a global market full of alternatives, customers seek peer reviews for assurance and guidance. Reviews offer real-world insights on product and service pros and cons, going beyond marketing narratives [1]. Online reviews affect e-commerce sites like Amazon and Flipkart and hospitality platforms like TripAdvisor and Yelp. The influence of online customer reviews goes beyond consumers. Businesses that actively engage with internet reviews can learn about the strengths and weaknesses of products. Customer review openness and genuineness generate a sense of community among customers. Online customer review analysis is essential for businesses that want to be competitive, responsive, and customer-centric in the digital marketplace [2].

Businesses looking to improve their goods, services, and customer satisfaction must understand the sentiments and opinions of these reviews. Obtaining useful insights from online reviews is difficult due to their number and textual nature [3]. Traditional manual review analysis is impossible due to the large number of daily reviews across platforms. NLP uses methods and techniques to process human-readable text. In online customer review analysis, NLP extracts attitudes, views, and topics from textual data. Sentiment classification groups review as positive, negative, or neutral, giving a broad snapshot of customer opinion. Opinion mining goes further by finding and

 doi:10.20944/preprints202312.2210.v2



categorizing particular thoughts in reviews, providing a more detailed insight into feedback from customers [4].

NLP's capabilities go beyond sentiment analysis and opinion mining. Aspect-based sentiment analysis extracts sentiments relating to specific product or service attributes, giving businesses actionable recommendations for improvement [5]. Emotion analysis adds to customer sentiment by capturing review emotions. As NLP advances, academics and businesses may utilize topic modeling, summarization, and deep learning to gain insights from customer reviews' diverse and complicated language. Using NLP to analyze online customer reviews is a technological innovation and a strategic need for organizations seeking digital competitiveness [6,7].

This survey has three main objectives. It begins with a detailed study of NLP literature on online customer review analysis. This paper attempts to provide readers with an in-depth understanding of the methodologies, tools, and datasets utilized in this field by integrating the results obtained from several studies. Second, the survey attempts to classify and organize NLP-based review analysis applications. This taxonomy helps academics, practitioners, and policymakers navigate the complex world of online customer review analysis. Finally, the study identifies and explains the open research issues in this growing field to enable future research and developments. By achieving these objectives, the survey aims to advance the field and apply NLP to online consumer reviews to get valuable insights. This survey also influences how businesses understand respond to, and utilize online customer reviews.

The remainder of the paper is divided into sections. Section 2 discusses survey methods. Section 3 defines the NLP taxonomy in online customer reviews and categorizes the literature by application. Section 4 discusses the survey studies, while Section 5 discusses challenges and future research directions. Finally, Section 6 concludes the findings of this study.

## Survey methodology

A carefully formulated search query was used to include research on online reviews, consumer feedback, and NLP applications in electronic commerce. The search query included several relevant keywords and phrases to cover the topic thoroughly. Online Reviews, Customer Reviews, App Reviews, Google Apps Reviews, NLP for Online Reviews, Sentiment Analysis, Review Summarization, Opinion Mining, Aspect-Based Sentiment Analysis, Named Entity Recognition, Emotion Analysis, User Profiling, Anomaly Detection, Design, Defects, Quality, Comparison, Sale, Reputation, Helpfulness Prediction, Ranking, Rating Prediction, and Fake Reviews were searched The search query also included Yelp, Amazon, Google Reviews, TripAdvisor, Trustpilot, App Store Reviews, and Play Store Reviews. This strategy aims to ensure the retrieval of papers directly related to the use of data from popular online review platforms. The inclusion and exclusion criteria are provided in Table 1. The literature was filtered for relevance and quality using inclusion and exclusion criteria. Studies are needed to address NLP and online reviews in e-commerce. Conference papers, articles, and chapters published from 2013 and 2023 in English were included. Studies that did not fulfill the publication type and language criteria or the study area were excluded.

**Table 1.** Inclusion and Exclusion Criteria.

| Inclusion Criteria | Exclusion Criteria |
|---|---|
| Papers directly address topics related to online reviews, customer feedback, or NLP within the context of electronic commerce | Papers not focusing on the intersection of NLP and online reviews within the context of electronic commerce are excluded |
| Papers published from 2013 to 2023 | Papers published before 2013 are excluded |
| Papers written in English are included | Papers not written in English are excluded |
| Papers eligible for inclusion include conference papers, articles, and chapters | Papers that are not conference papers, articles, or chapters are excluded |

Figure 1 depicts a thorough screening and relevance assessment of each retrieved article. A comprehensive title and abstract screening was performed on 1256 identified papers. In this phase, papers that did not meet research objectives were removed. After thorough screening, 473 records

 doi:10.20944/preprints202312.2210.v2



were retained for further analysis. The papers that were kept went through a more thorough full-text examination afterward. In order to understand the contribution of each paper, this crucial step included a thorough analysis. This comprehensive screening allowed articles that made substantial contributions to NLP and online reviews in electronic commerce to be included. Following the full-text examination, 154 papers were chosen as matching the inclusion criteria. Relevant data points from chosen publications were methodically retrieved and categorized. Key results, methodology, and overall themes from each study were identified. Extracted data served as the foundation for later analysis and synthesis of the literature, so contributing to the overall goals of this survey.

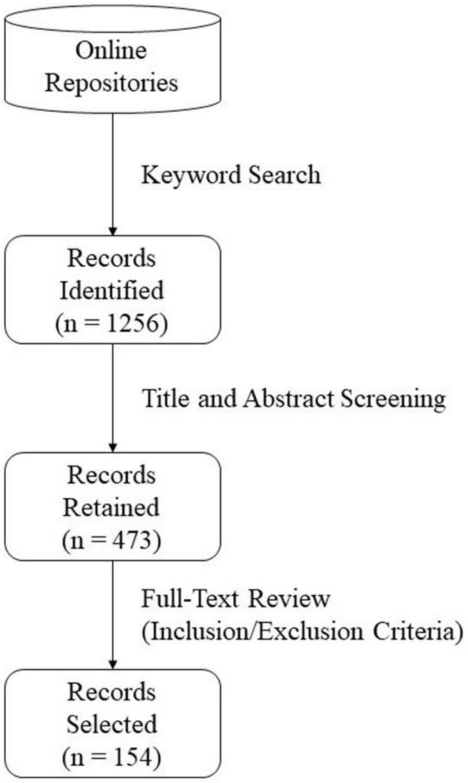


**Figure 1.** The flow of steps involved in the selection process of papers included.

## Taxonomy of NLP Applications in Online Customer Reviews

This section covers the taxonomy of NLP applications in this domain, focusing on sentiment analysis, review analysis and management, customer feedback and satisfaction, user profiling, recommendation systems, marketing, and brand management. Figure 2 depicts the taxonomy of NLP applications in online customer reviews. The taxonomy presented in this survey is critical for understanding the applications of NLP in the analysis of online customer reviews.

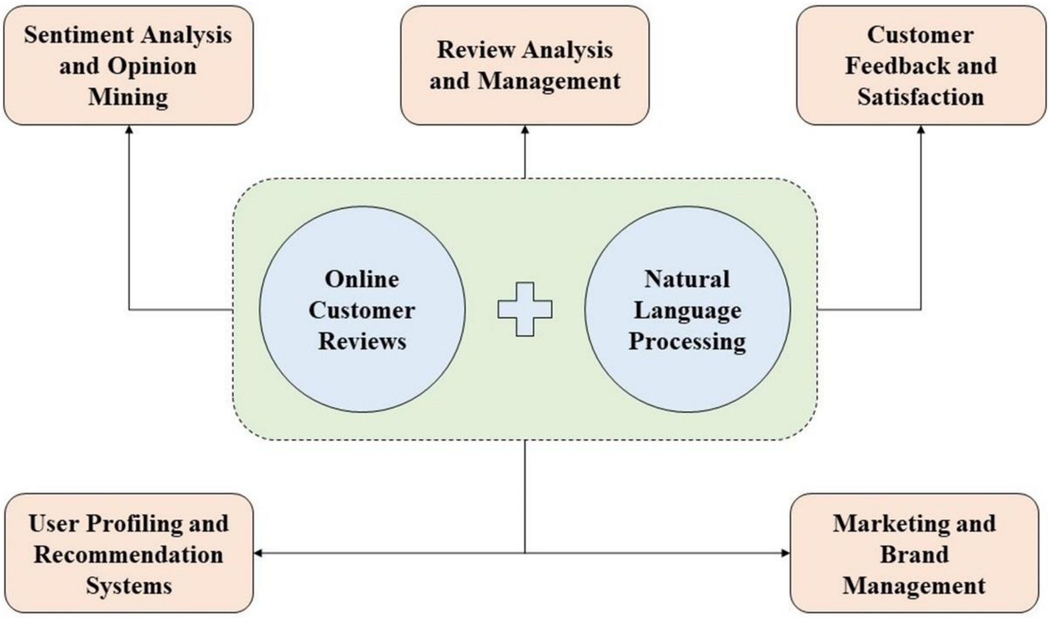


**Figure 2.** Taxonomy of Applications of NLP in online customer reviews.

*Sentiment Analysis and Opinion Mining*

Many studies examined the use of deep neural network strategies and traditional algorithms for sentiment extraction in various domains. A study used Bing Liu's aspect-based opinion-mining approach to tourism, addressing specific elements absent from the physical product model. The authors detected elements in online travel reviews and suggested advanced NLP-based sentiment categorization methods. They used a general tool architecture for Lake District TripAdvisor reviews and increased tourist performance, reaching a 92% F-measure for sentiment classification [8]. The study categorized sentiments and nuanced emotions from a large dataset of internet mobile phone reviews using Sentiment Analysis. This comprehensive strategy helped consumers and producers make informed decisions by emphasizing the relevance of online reviews in determining client needs as well as providing timely feedback [9].

A study also used cognitive computing-based Artificial Intelligence (AI) technologies to examine textual content and numerical ratings in online reviews. The study examined hotel reviews using sentiment analysis to uncover discrepancies between review content and scores [10]. The study also developed a Chinese sentiment mining method that outperformed other models on TripAdvisor reviews and included features to improve sentiment analysis [11]. A study also used Amazon and Twitter data to develop an NLP and machine learning book review sentiment categorization system. Users might analyze book public opinion and create user-friendly word clouds based on top attributes [12]. Genetic Algorithms for automated text sentiment analysis performed well on huge Amazon datasets and were highlighted for business and scientific applications [13].

Deep learning has been utilized in sentiment extraction using Convolutional Neural Network (CNN) and LSTM architectures to extract features from customer reviews [14]. Another study examined the role of opinion mining in e-commerce, using algorithms such as Naïve Bayes, SVM, Random Forest, and hybrid Support Vector Machine (SVM) to classify reviews as positive or negative, improving understanding and the use of opinion mining in online reviews [15]. NLP was also used for Amazon reviews to enhance service by comparing K-Nearest Neighbour, SVM, and Decision Tree classification algorithms to analyze customer feedback [16]. Deep learning techniques like word2vec for word embedding and CNN were used to evaluate social marketing tactics and help consumers make informed purchase decisions [17]. Another study used classical machine learning and deep learning to classify multiple affective attributes with over 90% accuracy using customer emotional needs from online product reviews [18]. A study modified Opinion Mining system

 doi:10.20944/preprints202312.2210.v2



evaluation by using a user profiling system to parameterize the system based on user preferences and improve results [19].

A study examined opinion mining and sentiment analysis of Amazon product reviews to increase accuracy. The Senti algorithm outperformed sentiment analysis APIs, enabling commercial, political, and financial decision-making [20]. Another research used Hybrid Attribute Based Sentiment Classification (HABSC) to successfully identify sentiment orientation in online consumer reviews. HABSC outperformed state-of-the-art approaches by integrating syntactic characteristics, implicit word relations, and domain-specific information to reveal differences in review content and ratings between local and international consumers on multinational social commerce platforms [21]. Additionally, Japanese restaurant reviews were examined to see how ethnic culture affects customer ratings. Bilingual text mining software highlighted cultural implications in social commerce by showing different emotion distribution patterns among Japanese and Western customers [22].

Researchers also examined sentiment analysis in NLP to evaluate emerging technologies with many aspects. Mobile reviews were identified using web scraping and machine learning techniques, specifically Decision Tree and SVM, with over 90% accuracy using 1-2 grams. User experience and marketing tactics improved by real-time product review sentiment analysis [23]. A study examined the importance of sentiment analysis in business analytics for product and market competitiveness. The research examined machine learning classification methods, including a Hybrid Algorithm, emphasizing the rising relevance of sentiment analysis in corporate strategy and product quality [24,25]. A study addressed the constraints of using customer ratings or review summaries for extracting useful data from online product reviews [26]. Two new corpora with full Word Clouds were produced using a General Approach and a Specific Approach to improve product analysis accuracy and efficiency. The approach sought nuanced consumer sentiment and product characteristics. Additionally, a new sentiment analysis algorithm improved the Dempster-Shafer algorithm [27]. This novel method treated reviews as sentences with sentiment orientations and ratings. The method outperformed the original algorithm on TripAdvisor and CitySearch datasets.

Lexicon-based analysis was used to evaluate Amazon books and writers [28]. The study used a bag-of-words technique to evaluate review positivity and negativity, emphasizing the importance of sentiment in market analysis and its ability to anticipate business trends. Another study compared LSTM, random forest, SVM, and eXtreme Gradient Boosting (XGBoost) for sentiment analysis in AI [29]. The findings highlighted uses in customer management systems and Twitter and e-commerce platforms. Two Aspect-Based Sentiment Analysis aspect extraction algorithms were presented to analyze unstructured social media reviews [30]. Using SemEval, Yelp, and Kaggle datasets, the hybrid technique predicted aspect categories accurately. An end-to-end sentiment analysis technique for negotiations was published [31]. The method reduced biases and enhanced sentiment categorization across datasets. Sentiment analysis improved user experience on online buying platforms [32]. Compared to Logistic Regression, Multinomial Naive Bayes, and SVM, Stochastic Gradient Descent has the greatest accuracy. A lexicon-based technique and logistic regression were used to analyze Web sentiment [33]. These methods successfully extracted sentiments from various web sources.

A unique Statistics-Based Outlier Detection and Correction Method study [34] highlighted the need for proper sentiment analysis in Amazon customer reviews. This technology improved sentiment analysis without data loss over previous systems. Various machine learning systems analyzed Amazon electronics product reviews for sentiment [35]. Logistic Regression had the best accuracy, demonstrating the relevance of sentiment analysis in customer recommendations. Another work used part-of-speech-based feature extraction and game-theoretic rough sets to reduce dimensionality in sentiment analysis [36]. The model outperformed other models and classifiers. A study on Amazon Electronics product reviews using machine learning due to the rising relevance of e-commerce [37]. Preprocessing methods were tested, and the Multi-Layer Perceptron classifier performed well. A publication introduced the BERT Base Uncased model to improve e-commerce platform review sentiment analysis, outperforming standard machine learning approaches [38]. A work using NLP for sentiment analysis addressed Amazon's growing customer review volume [39].

The Term Frequency - Inverse Document Frequency (TF-IDF) approach using unigram and SVM was the most accurate for Amazon product reviews.

An Ensemble Classifier study [40] stressed the importance of online reviews in understanding customer perspectives and needs. The Ensemble Classifier outperformed machine learning techniques in consumer feedback analysis. A paper used Naïve Bayes, random forest, and SVM algorithms to improve Amazon product sentiment analysis accuracy [41]. Aspect-based BERT models were used for tourist sentiment analysis [42]. The findings helped merchants improve their products and services and provided users with personalized recommendations. A mixed generative-discriminative strategy combining Fisher kernels and hidden Markov models improved textual sentiment analysis [43]. Amazon and IMDb user reviews showed that the method improved sentiment identification compared to established methods. The influence of digitalization on e-commerce and information overload was examined using machine learning algorithms on Amazon Fine Food reviews [44]. The project attempted to simplify review analysis so customers could quickly and accurately assess product opinions. Another research examined BERT-based sentiment analysis across domains [45]. The study showed that sentiment analysis should incorporate class label variations from various sources.

A study used a hybrid method for sentiment analysis of Amazon customer reviews using NLP, machine learning, and Deep Learning [46]. The results showed that sentiment research may improve brand value, advertising, and customer service. E-commerce platform sentiment analysis using SVM and CNN Models [47]. The methods were more accurate than others. A study investigated using machine learning for review sentiment analysis [48].

A study also attempted to utilize NLP to automate analysis of product reviews from various platforms such as Amazon [49]. The approach used machine learning to train a neural network to classify product reviews as positive, neutral, or negative. A study examined how varied NLP algorithms affected Yelp and Zappos data [50]. For consumer review data analysis, BERT and Neural Network were helpful, providing algorithm selection insights. The research used NLP methods like TF-IDF Vectorizer and Count Vectorizer to create a food industry model. Logistic Regression, Dummy Classifier, and Random Forest Classifier were used to efficiently analyze online review consumer sentiments, giving manufacturers significant product perception insights. The investigation found that the proposed sentiment analysis model worked [51]. Another article proposed a Bayesian network architecture for sentence-level sentiment analysis of e-commerce product reviews with automated rule creation and progressive retainability. The study met model requirements instantaneously, demonstrating its scalability and durability in opinion mining across themes [52].

A new method using agglomerative clustering for outlier detection and a stacked autoencoder with ensemble classification algorithms was developed to detect sarcastic tweets and reviews. This technique outperformed other algorithms in sarcasm prediction and sentiment identification with 99.3% accuracy [53]. With 1.5 million Amazon and Yelp reviews, a study introduced the 'Amazon and Yelp Reviews' dataset for sentiment analysis. The sentiment analysis method included daily data collecting, user comment preparation, and a Bidirectional LSTM (BiLSTM) model to achieve 87.3% accuracy. The dataset and methodology might be used for consumer feedback analysis and online reputation management [54]. Amazon values customer opinions and stresses the importance of customer satisfaction in organizational success. The article used NLP to turn text into numerical arrays for machine learning techniques. Five scores were assigned to Amazon reviews using supervised machine learning algorithms, such as SVM, Naïve Bayes, and Decision Tree [55]. Another article used machine learning to assess Amazon product review sentiment across categories. Text Blob, Logistic Regression, SVM, and Multinomial Naive Bayes improved sentiment classification accuracy, proving that various review sentiment ratings may be predicted [56]. A study [57] focused on sentiment polarity analysis for e-commerce customer reviews, while [58] presented an EESNN-SA-OPR method utilizing Collaborative Filtering (CF) and product-to-product similarity.

A study investigated business strategies for customer retention and attraction, employing NLP-based sentiment analysis [59]. The impact of internet reviews on consumer decisions was examined using a CNN model for text review sentiment classification. Comparative investigation showed the CNN model's 90% Amazon review accuracy. Stop words are crucial to sentiment analysis, and the CNN model outperforms other algorithms on big datasets [60]. A publication also introduced the Adaptive Particle Grey Wolf Optimizer with Deep Learning Based Sentiment Analysis (APGWO-DLSA) approach for sentiment analysis on online product reviews using NLP and machine learning algorithms. On the Cell Phones And Accessories (CPAA) dataset, the suggested technique was better, obtaining 94.77% accuracy [61]. NLP and LSTM were used to create a customer review summary model to handle the increase in textual material. The hybrid sentiment analysis method provided organizations with important insights due to its excellent accuracy (94.46%), recall (91.63%), and F1-score (92.81%) [62].

A study on mobile phone reviews utilized consumer reviews to improve post-purchase products. After testing SVM, Naïve Bayes, and Logistic Regression algorithms, the Random Forest (Unigram) classifier performed best on a balanced dataset, highlighting the importance of sentiment analysis in consumer feedback for product development [63]. LSTM and Naive Bayes were compared for sentiment analysis of online product reviews. Comprehensive assessments of varied internet items were conducted to better understand user attitudes [64]. A recent article uses data mining to analyze sentiment on Facebook, Instagram, Twitter, and Amazon. The research used consumer input to improve corporate strategy and predict customer requirements. Twitter data obtained via the API key was analyzed using NLP techniques, demonstrating their ability to provide organizations with important insights for personalized marketing and organizational benefit [65]. The summary of existing literature on sentiment analysis and opinion mining is given in Table 2.

**Table 2.** Summary of literature on sentiment analysis and opinion mining.

| Ref. | Year | Dataset | Description |
|---|---|---|---|
| [8] | 2014 | TripAdvisor | Introduced enhanced opinion mining for tourism, outperformed existing models in sentiment classification in Chile. |
| [9] | 2017 | Amazon, Flipkart | Utilized Sentiment Analysis on mobile reviews for consumer decision-making. |
| [10] | 2017 | TripAdvisor | Employed AI tools revealing mismatches between sentiment and scores in hotel reviews. |
| [11] | 2017 | TripAdvisor | Presented Chinese sentiment mining outperforming existing models in reviews. |
| [12] | 2017 | Amazon | Detailed a methodology using NLP and machine learning for sentiment classification in book reviews. |
| [13] | 2017 | Amazon | Presented two Genetic Algorithms for automated text sentiment analysis, outperforming existing methods. |
| [14] | 2018 | IMDb, Amazon | Investigated the application of deep neural networks, showcasing superior performance compared to traditional methods. |
| [15] | 2018 | Amazon | Highlighted the importance of opinion mining in e-commerce, aiming to develop a machine for sentiment analysis. |
| [16] | 2018 | Amazon | Utilized NLP and a UCI machine learning dataset to assess Amazon customer reviews. |
| [17] | 2018 | Amazon | Introduced a sentiment analysis approach for product reviews, utilizing deep learning with word2vec. |
| [18] | 2018 | Amazon | Introduced a cost-effective approach to product design through effective engineering. |
| [19] | 2018 | Yelp | Presented a new method to evaluate Opinion Mining system performance by incorporating user preferences. |
| [20] | 2018 | Amazon | Addressed the challenges of opinion mining in handling diverse online user data. |
| [21] | 2019 | TripAdvisor, Amazon | Proposed HABSC, a novel method leveraging syntactic features, implicit word relations, and domain-specific knowledge. |
| [22] | 2019 | Yelp | Explored the influence of ethnic culture on customer reviews in social commerce. |

| | | | |
|---|---|---|---|
| [23] | 2019 | Large Mobile Review Dataset | Investigated sentiment analysis in NLP. |
| [24] | 2019 | Amazon | Utilized in machine learning to perform sentiment analysis on E-commerce product reviews. |
| [25] | 2019 | Amazon | Explored sentiment analysis in NLP's context, emphasizing its pivotal role in Business Analytics. |
| [26] | 2020 | Amazon | Proposed a model to overcome the limitations of traditional online product analysis. |
| [27] | 2020 | TripAdvisor, CitySearch | Presented a novel sentiment analysis method, utilizing a two-point structure to enhance the Dempster–Shafer algorithm. |
| [28] | 2020 | Amazon | Explored sentiment analysis in e-commerce, specifically on Amazon, to assess book and author quality. |
| [29] | 2021 | Check | Compared the effectiveness of LSTM, random forest, SVM, and XGBoost in both binary and multiclass scenarios. |
| [30] | 2021 | SemEval, Yelp, Kaggle datasets | Proposed two methods for aspect extraction in Aspect-Based Sentiment Analysis from unstructured social media reviews. |
| [31] | 2021 | Amazon | Introduced a novel sentiment analysis approach with a customized negation marking algorithm. |
| [32] | 2021 | Amazon | Introduced a sentiment analysis mechanism employing machine learning algorithms. |
| [33] | 2021 | - | Investigated sentiment analysis on WWW content, utilizing a lexicon-based method and logistic regression in machine learning. |
| [34] | 2021 | Amazon | Investigated sentiment analysis and outlier detection in Amazon customer reviews. |
| [35] | 2021 | Amazon | Focused on sentiment analysis of Amazon electronics product reviews. |
| [36] | 2021 | Yelp, IMDb, Amazon | Introduced a sentiment analysis model addressing challenges in data pre-processing and classification uncertainty. |
| [37] | 2021 | Amazon | Investigated the role of machine learning, employing diverse classifiers and preprocessing techniques. |
| [38] | 2021 | Amazon | Improved sentiment analysis of E-commerce reviews by introducing the BERT Base Uncased model. |
| [39] | 2022 | Amazon | Employed NLP, utilizing term-based methods and N-grams. |
| [40] | 2022 | Amazon | Explored the importance of online product reviews, employing an Ensemble Classifier. |
| [41] | 2022 | Amazon | Employed SVM, random forest, and naive bayes algorithms to enhance sentiment analysis for Amazon products. |
| [42] | 2022 | TripAdvisor | Outlined a novel aspect-based sentiment analysis model, leveraging BERT, to extract sentiment and aspect-category information. |
| [43] | 2022 | Amazon, IMDb | Introduced a hybrid generative-discriminative approach using Fisher kernels and hidden Markov models. |
| [44] | 2022 | Amazon | Investigated the role of online reviews in the digitized e-commerce landscape, utilizing machine learning algorithms. |
| [45] | 2022 | Twitter, YouTube, Facebook, Amazon, TripAdvisor | Investigated sentiment analysis on Italian corpora using BERT-based models. |
| [46] | 2022 | Amazon | Proposed a hybrid approach, leveraging NLP, machine learning, and Deep Learning. |
| [47] | 2022 | Amazon | Employed SVM and CNN Models for Customer Review Sentiment Analysis. |
| [48] | 2022 | Amazon | Outlined a machine learning-based sentiment evaluation model for e-commerce shopper reviews. |
| [49] | 2022 | Amazon | Utilized NLP to automate the analysis of product reviews on platforms like Amazon. |
| [50] | 2022 | Yelp, Zappos | Investigated the impact of NLP models on consumer reviews from Yelp and Zappos. |

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202312.2210.v2



| | | | |
|---|---|---|---|
| [51] | 2022 | Amazon | Introduced a model designed for the food industry, utilizing NLP techniques and machine learning classification algorithms. |
| [52] | 2023 | - | Introduced a Bayesian-network framework for automated sentence-level sentiment analysis on e-commerce websites. |
| [53] | 2023 | Amazon | Pioneered a method for effectively identifying sarcastic opinions in online content. |
| [54] | 2023 | Amazon, Yelp | Presented the 'Amazon and Yelp Reviews' dataset as a valuable resource for sentiment analysis. |
| [55] | 2023 | Amazon | Explored the use of NLP to analyze Amazon reviews. |
| [56] | 2023 | Amazon | Employed machine learning techniques, including NLP and deep learning algorithms, to analyze Amazon product reviews. |
| [57] | 2023 | Amazon, Yelp | Focused on sentiment polarity analysis for e-commerce customer reviews. |
| [58] | 2023 | Amazon | Presented an EESNN-SA-OPR method utilizing CF and product-to-product similarity. |
| [59] | 2023 | Amazon | Investigated business strategies for customer retention and attraction, employing NLP-based sentiment analysis. |
| [60] | 2023 | Amazon | Introduced a CNN model for sentiment analysis in internet reviews. |
| [61] | 2023 | Cell Phones and Accessories dataset, Amazon | Introduced the APGWO-DLSA method to enhance sentiment analysis in online product reviews. |
| [62] | 2023 | SemEval-2014, Sentiment140, STS-Gold | Outlined a consumer review summarization model using NLP and LSTM. |
| [63] | 2023 | Amazon | Investigated leveraging post-purchase customer reviews, particularly focusing on mobile phone reviews. |
| [64] | 2023 | Amazon | Utilized sentiment analysis to comprehensively assess online product reviews. |
| [65] | 2023 | Twitter | Conducted sentiment analysis through data mining on various platforms, including Twitter. |

*Review Analysis and Management*

There have been initiatives taken to address fake reviews and counterfeit goods in the context of online marketplaces. AI methods like NLP and topic analysis were used to detect counterfeit items on Amazon and eBay. Topic analysis of product and seller reviews identified deception-related keywords and concepts. The findings showed automated counterfeit detection might boost online marketplace trust and efficiency [66]. The fake Product Review Monitoring and Removal System (FaRMS) analyzed reviews from numerous platforms with 87% accuracy in English and Unique Urdu support to combat fake reviews. By providing honest product ratings, FaRMS sought to improve customer satisfaction [67].

A study examined how review length affects online purchasing decisions and questioned the idea that lengthier reviews are always better. Using Amazon reviews and powerful NLP, the study discovered that argumentation frequency altered the association between review length and helpfulness, showing that longer reviews were not always more helpful [68]. A novel approach combined business data and user reviews to improve relevance and diversity in machine-generated fake reviews. In response to traders' deception, the proposed model generated high-quality and diverse reviews [69]. The significance of vigilance in the face of manipulation on large online platforms was highlighted by stylometry-based algorithms that detected misleading online reviews [70].

A predictive model used BERT and deep learning to improve online product review usefulness evaluation and overcome previous model limitations [71]. Introducing Social Network Strength (SNS) elements to analyze the influence of friends and followers on review helpfulness helped overcome information overload in online customer reviews. Validated on Yelp, the methodology gave researchers, businesses, reviewers, and review platforms insights [72]. In order to combat the

 doi:10.20944/preprints202312.2210.v2



ubiquity of fake reviews, supervised machine learning was utilized to identify opinion spammers, which improved the accuracy of spotting fraudulent reviews on well-known platforms [73].

Addressing the critical issue of fake review detection, NLP techniques and machine learning models, including Naïve Bayes and random forest, were applied to combat the increasing prevalence of fake reviews in the E-commerce industry. The models demonstrated scalability, offering a solution for platforms to promptly identify and address spam reviews [74]. Another study aimed to identify the most effective feature combination for detecting fake reviews, highlighting the significance of behavior-related features in combination with text-related features, with verified purchase emerging as a crucial factor [75]. A hybrid CNN-LSTM deep learning model with sentiment analysis techniques was employed to assess the authenticity of customer reviews, proposing a solution to combat fraudulent reviews in the e-commerce sector [76].

Supervised machine learning and NLP techniques were utilized to identify and remove fake reviews from a dataset, focusing on major E-commerce companies to combat the prevalence of counterfeit product reviews impacting customer decisions and profits [77]. A Python-based system was introduced to detect fake product reviews on Amazon, using SVM techniques to distinguish between genuine and fake reviews and enhance the reliability of product reviews [78]. Lastly, an innovative method employing a CNN and adaptive particle swarm optimization with NLP techniques achieved a remarkable 99.4% accuracy in identifying fake online reviews, offering practical implications for consumers, manufacturers, and sellers in maintaining the trustworthiness of online reviews [79]. Another study proposed a generalized solution by fine-tuning the BERT model to predict review helpfulness, demonstrating superior performance compared to traditional bag-of-words methods [80]. The summary of existing literature on review analysis and management is given in Table 3.

**Table 3.** Summary of literature on review analysis and management.

| Ref. | Year | Dataset | Description |
|---|---|---|---|
| [66] | 2015 | eBay, Amazon | Proposed AI framework for automated detection of counterfeit products using NLP and topic modeling. |
| [67] | 2019 | Amazon, Flipkart, Daraz | Introduced an Intelligent Interface for detecting and eliminating fake product reviews on major e-commerce platforms with 87% accuracy. |
| [68] | 2019 | Amazon | Challenged the belief that longer online reviews are universally more helpful, suggesting impact depends on argumentation within the text. |
| [69] | 2019 | Yelp | Combated machine-generated fake reviews by integrating business information and user reviews using an encoder-decoder model. |
| [70] | 2020 | Amazon, Yelp, Google, Facebook | Addressed the challenge of detecting deceptive online reviews and ratings using a historical stylometry-based methodology. |
| [71] | 2020 | Amazon | Introduced a model combining BERT features with deep learning techniques to predict the helpfulness of online customer product reviews. |
| [72] | 2021 | Yelp | Predicted review helpfulness using regression and classification, incorporating SNS features. |
| [73] | 2021 | Yelp | Explored the influence of online reviews on consumer decisions, proposing a supervised approach to detect opinion spammers in reviews. |
| [74] | 2021 | Amazon, Yelp | Focused on using NLP techniques and machine learning models to detect and eliminate fake reviews. |
| [75] | 2022 | Amazon | Investigated optimal feature combinations for fake review detection, emphasizing the importance of behavior-related features. |
| [76] | 2023 | Amazon, Yelp | Introduced a hybrid CNN-LSTM deep learning model with sentiment analysis to detect fraudulent reviews. |
| [77] | 2023 | Reviews of 20 Hotels | Utilized supervised machine learning and NLP to detect and eliminate fake reviews, focusing on counterfeit product evaluations influencing customer decisions. |

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202312.2210.v2



| | | | |
|---|---|---|---|
| [78] | 2023 | Amazon | Outlined a Python-based SVM system to identify and differentiate fake product reviews. |
| [79] | 2023 | Ott, Amazon, Yelp, TripAdvisor, IMDb | Proposed an effective method using CNNs and adaptive particle swarm optimization with NLP to detect fake online reviews with 99.4% accuracy. |
| [80] | 2023 | Yelp | Tackled information overload in online reviews by proposing a solution using fine-tuned BERT models. |

### *Customer Feedback and Satisfaction*

A study adopted Bing Liu's aspect-based technique to identify customer preferences in TripAdvisor hotel and restaurant reviews to examine opinion mining in tourism. The approach demonstrated 90% precision and recall in extracting sentiment orientations, though struggled with explicit aspect expressions. Emphasizing the value of tourism product reviews, the research underscored the importance of aspect-based opinion mining in revealing customer preferences [81]. Another study focused on English online reviews of hotels, employing natural language preprocessing and sentiment analysis. Organizations emphasizing these techniques outperformed peers in growth, earnings, and performance metrics, offering practical implications for hotel managers to leverage social media reviews for strategic decision-making [82].

Introducing a novel method for hotel summaries from travel forums, a study incorporated author credibility and conflicting opinions. Using a new sentence importance metric and k-medoids clustering algorithm, the approach outperformed conventional methods, affirmed by subjects for providing more comprehensive hotel information [83]. In the realm of retail, an article proposed an online platform using NLP to analyze customer sentiments and streamline input through Speech-to-Text technology. The focus was on enhancing the shopping experience by understanding emotions expressed in reviews, suggesting smart shop solutions to improve overall customer satisfaction [84]. Another research delved into creating artificial personal shoppers for e-commerce platforms, emphasizing user engagement and trust. The study adapted existing information retrieval and NLP technologies, aiming to establish the groundwork for effective artificial personal shoppers in the online shopping domain [85].

A unique approach assessed customer loyalty through sentiment analysis of online reviews, achieving a 94% accuracy in determining loyalty types. Leveraging tokenization, lemmatization, and SentiWordNet, the study utilized a fuzzy logic model with rule-based systems, surpassing previous methods [86]. Addressing the inadequacy of general e-commerce platforms for vitamins and nutraceuticals, another study employed NLP to extract insights from user-generated product reviews. The system provided a five-point rating system, summarized commonly discussed topics, and offered representative reviews, empowering consumers with tailored information for informed decisions [87]. A proposed a rapid customer loyalty model for e-commerce with a 72% loyalty rate from Amazon.com reviews [88]. Similarly, sentiment analysis and opinion mining in Yelp datasets using ABSA provided business strategies based on one-year forecasted data, emphasizing the importance of leveraging online reviews for improving customer satisfaction [89].

Analyzing user-generated hotel review data comprehensively, a study employed various techniques, achieving high precision (0.95) and recall (0.96). Visual analytics revealed patterns in user ratings, emphasizing the potential for improving business services and product quality [90]. Investigating the impact of latent content factors on online review helpfulness, the study found that argument quality and valence significantly influenced review helpfulness. This approach surpassed previous manifest content and reviewer-related factors, enhancing understanding and addressing sentiments for improved customer satisfaction [91]. Scrutinizing online complaints related to hotel guest experiences, a study distinguished patterns between Asian and non-Asian guests, revealing service failures in different domains and stages of the hotel guest cycle [92].

Preprints.org (www.preprints.org) | NOT PEER-REVIEWED | Posted: 15 January 2024 doi:10.20944/preprints202312.2210.v2



Leveraging logistic regression and NLP, another study discerned sentiment and topics among tourists in Cyprus, offering insights into the nuanced relationship between tourist culture, purchasing power, and reviews [93]. Addressing the challenge of efficiently processing user feedback, a study introduced a crowdsourcing method for classifying app store reviews, indicating the potential of crowd workers as an affordable and reliable resource for classifying user reviews [94]. Investigating parental preferences for childcare using Yelp reviews, the study revealed variations in satisfaction based on income levels, emphasizing safety, learning environment quality, and child-teacher interactions as pivotal factors [95]. Exploring opinion summarization in Web 3.0 platforms, a study proposed a novel graph-based abstractive technique, comparing it with extractive methods for coherence and completeness in generating summaries [96].

Enhancing review-based question answering systems using NLP models, a study addressed the challenge of manual handling of product-related queries on online platforms. The proposed enhancements, including BERT, significantly improved response effectiveness, achieving a BLEU score of 0.58 [97]. Surveying visitor reviews of Croatia's Plitvice Lakes National Park, a study utilized multidimensional scaling, sentiment analysis, and NLP to identify key topics and discern strengths and weaknesses, providing valuable insights for protected natural areas [98]. Investigating wine packaged tours in Tuscany, a study identified critical elements influencing success through text mining and sentiment analysis on TripAdvisor reviews, highlighting the significance of tour guides in consumer satisfaction [99]. Furthermore, a study introduced a hierarchical attention network-based framework for analyzing Amazon Smartphone reviews [100].

A study used sentiment analysis to classify smartphone reviews and predict product ratings based on user feedback [101]. Addressing information overload in Community-based Question Answering (CQA) platforms, a study introduced a CQA summarization task. Evaluating various summarization methods, the research provided a robust baseline for CQA summarization, contributing to the user experience in navigating overwhelming information [102]. Assessing pre-trained transformers for sentiment extraction, a study applied five models to an Amazon database of automotive products, suggesting their potential for practical applications like product monitoring and market research [103]. Lastly, employing machine learning and NLP, a study demonstrated the effectiveness of text summarization in efficiently handling and comprehending extensive online product review data [104]. The summary of existing literature on customer feedback and satisfaction is given in Table 4.

**Table 4.** Summary of literature on customer feedback and satisfaction.

| Ref. | Year | Dataset | Description |
|---|---|---|---|
| [81] | 2013 | TripAdvisor | Extracted consumer preferences from tourism reviews using aspect-based opinion mining with a 35% average extraction. |
| [82] | 2016 | TripAdvisor | Analyzed English hotel reviews in Chinese cities for managerial insights using NLP, text mining, and sentiment analysis. |
| [83] | 2017 | TripAdvisor | Generated hotel summaries from travel forums, considering author credibility and conflicting opinions for improved performance. |
| [84] | 2018 | - | Integrated eCommerce and offline retail using NLP and Speech-to-Text for efficient checkout and smart shop solutions. |
| [85] | 2018 | Alibaba, Amazon, eBay | Leveraged automatic speech recognition to create artificial personal shoppers for eCommerce, enhancing user trust through human-like conversations. |
| [86] | 2018 | Amazon | Measured customer loyalty using aggregated sentiment scores and fuzzy logic with 94% accuracy on Amazon.com data. |
| [87] | 2019 | Amazon | Used NLP to extract insights from user-generated reviews in the nutraceutical retail vertical for better decision-making. |
| [88] | 2019 | Amazon | Proposed a rapid customer loyalty model for e-commerce with a 72% loyalty rate from Amazon.com reviews. |

| | | | |
|---|---|---|---|
| [89] | 2019 | Yelp | Predicted customer concerns for restaurants using sentiment analysis and opinion mining on Yelp datasets. |
| [90] | 2019 | TripAdvisor | Integrated sentiment analysis, aspect extraction, and visual analytics for improved hotel reviews analysis. |
| [91] | 2019 | Yelp | Examined online review helpfulness using the Elaboration Likelihood Model, revealing the impact of latent content factors. |
| [92] | 2020 | TripAdvisor | Applied Aspect-Based Sentiment Analysis to categorize hotel-related service failures, highlighting cultural differences. |
| [93] | 2020 | TripAdvisor | Analyzed sentiment and topics among Cyprus tourists using logistic regression and NLP. |
| [94] | 2020 | App Store Reviews | Investigated a crowdsourcing approach for efficiently classifying user feedback on app stores and social media. |
| [95] | 2020 | Yelp | Examined parental preferences for child care using Yelp.com reviews, revealing income-dependent satisfaction variations. |
| [96] | 2021 | Standard Opinosis Dataset | Explored opinion summarization in Web 3.0 e-commerce platforms using abstractive and extractive techniques. |
| [97] | 2022 | Amazon | Enhanced review-based question answering systems using advanced NLP models like BERT and BART. |
| [98] | 2022 | TripAdvisor | Surveyed reviews of Croatia's Plitvice Lakes National Park to identify management topics, strengths, and weaknesses. |
| [99] | 2022 | TripAdvisor | Investigated success factors of wine tours in Tuscany using text mining and sentiment analysis. |
| [100] | 2022 | Amazon | Introduced a hierarchical attention network-based framework for analyzing Amazon Smartphone reviews. |
| [101] | 2022 | Amazon | Used sentiment analysis to classify smartphone reviews and predict product ratings based on user feedback. |
| [102] | 2022 | Amazon QA Corpus (COQASUM) | Introduced a novel CQA summarization task to address information overload in Community-based Question Answering platforms. |
| [103] | 2022 | Amazon | Investigated the utility of pre-trained transformers in extracting customer sentiment from online reviews. |
| [104] | 2023 | Amazon | Applied machine learning, NLP, and deep learning for text summarization of product reviews, reducing reading time and enhancing understanding. |

### *User Profiling and Recommendation Systems*

In addressing challenges related to service recommendation accuracy and incomplete modalities in recommender systems, two innovative algorithms were introduced. Value Features and Distributions for Accurate Service Recommendation (VFDSR) leverages fine-grained value features extracted from customer reviews to enhance personalized service recommendations, demonstrating superior performance on a Yelp dataset [105]. Learning to recommend with missing modalities (LRMM), on the other hand, tackles incomplete modalities through modality dropout and a multimodal sequential autoencoder, outperforming existing methods in real-world Amazon data experiments and proving robust in mitigating data-sparsity and the cold-start problem [106]. The integration of data mining, human psychology, and NLP aimed to enhance recommender-based mobile applications. The strategy generated "wh" questions from recommended items, utilized a web scraper for relevant information, and strategically employed human-computer interaction psychology to increase user engagement. Survey results confirmed an improved hit rate, supporting the method's effectiveness on platforms like Amazon [107].

Another study focused on the rising use of intelligent personal assistants in business workflows, introducing an explanation mode feature for speech interaction in Enterprise Resource Planning software. Task attraction was identified as pivotal for usefulness, emphasizing the supplementary role of intelligent personal assistant alongside traditional input methods [108]. Advancements in personalized advertising and recommender systems were explored with Double Attention for Click-Through Rate Prediction (DAMIN), an enhanced model incorporating a double attention mechanism into the Deep Interest Network. Experimental results on Amazon datasets demonstrated DAMIN's superiority, improving AUC by 4%–5% compared to classical models [109]. TripAdvisor data was leveraged to enhance hotel customer targeting through a fine-tuned BERT model and a multi-criteria recommender system. Outperforming a benchmark single-criteria system, the approach considered nuanced hotel aspects, demonstrating superior performance [110]. In the Pakistani fashion industry, user interests were extracted from social media using Latent Dirichlet Allocation (LDA), Latent Semantic Analysis (LSA), and BERT for topic modeling, sentiment analysis tools, and K-Means for clustering. Empirical validation demonstrated moderate agreement between human and machine evaluations [111].

An innovative product recommender model for e-commerce platforms analyzed customer reviews using NLP, sentiment analysis, and clustering algorithms. Experiments on Amazon datasets showed notable enhancement in multi-node cluster setups over single-node configurations [112]. A graph-based movie recommender system incorporating user sentiments and emotions demonstrated superior performance. Utilizing BERT for sentiment analysis and a Kaggle dataset, the proposed IGMC-based models outperformed conventional and state-of-the-art graph-based systems [113]. The impact of cognitive absorption dimensions on continuous use intention in AI-driven Recommender Systems was investigated, revealing that curiosity and focused immersion significantly influenced continuous use intention [114].

The study proposed a novel approach to enhance trip suggestions for tourists by integrating neural networks and deep learning techniques. The hybrid framework combined Neural Network-LSTM for Point of Interest recommendations and BERT for sequential trip recommendations, demonstrating superior performance on TripAdvisor and Yelp datasets [115]. A weighted hybrid recommendation method combining user reviews, rating data, and sentiment analysis achieved improved precision scores on the Amazon Reviews dataset, integrating CF for enhanced recommendations [116]. FusionSCF addressed issues in Recommendation Systems by integrating CF with sentiment analysis of textual user reviews. Using e-commerce datasets, the model combined weighted ratings and sentiment scores to enhance recommendations, demonstrating the effectiveness of the sentiment-based model over traditional CF methods. The study also explored the impact of fake reviews on the filtering system [117]. The summary of existing literature on user profiling and recommendation systems is given in Table 5.

**Table 5.** Summary of literature on user profiling and recommendation systems.

| Ref. | Year | Dataset | Description |
|---|---|---|---|
| [105] | 2017 | Yelp | Introduced VFDSR, a service recommendation algorithm using fine-grained value features from customer reviews, demonstrating superior performance. |
| [106] | 2018 | Amazon | Proposed LRMM, a framework for multimodal learning in content-based recommendation, excelling in rating prediction and handling data-sparsity. |
| [107] | 2018 | Amazon | Utilized data mining, psychology, and NLP to enhance recommender-based mobile apps' profitability and usability. |
| [108] | 2020 | - | Investigated intelligent personal assistants in business workflows, introducing an explanation mode for speech interaction in ERP software. |
| [109] | 2021 | Amazon | Introduced DAMIN, a deep learning model showing improved click-through rate prediction. |
| [110] | 2021 | TripAdvisor | Refined a BERT model for a multi-criteria hotel recommender system, outperforming single-criteria benchmarks. |

| | | | |
|---|---|---|---|
| [111] | 2022 | Roman Urdu Tweets, Google Reviews | Explored user interests in the Pakistani fashion industry using LDA, LSA, BERT, sentiment analysis, and K-Means clustering. |
| [112] | 2022 | Amazon | Outlined a product recommender model using NLP on customer reviews, showing notable performance gains in multi-node clusters. |
| [113] | 2022 | Rotten Tomatoes, Amazon | Introduced a graph-based movie recommender system, outperforming conventional models on Kaggle datasets. |
| [114] | 2023 | Amazon | Explored the influence of curiosity and focused immersion on AI-driven Recommender Systems in e-commerce. |
| [115] | 2023 | TripAdvisor, Yelp | Proposed a tourist recommendation system using Neural Network-LSTM and Bidirectional Encoder Representations from Transformer. |
| [116] | 2023 | Amazon | Proposed a weighted hybrid recommendation system using sentiment analysis and CF, resulting in enhanced precision. |
| [117] | 2023 | Amazon, Flipkart | Introduced FusionSCF, a model addressing cold-start and long-tail issues in Recommendation Systems by combining CF with sentiment analysis. |

*Marketing and Brand Management*

Advanced methods for opinion mining in concise e-commerce feedback remarks were studied to create seller rating profiles. The novel approaches integrated opinion mining and dependency relation analysis to propose an algorithm for extracting dimension ratings. The computation of dimension weights from ratings was framed as a factor analytic problem and solved through matrix factorization. The algorithm demonstrated efficacy on eBay and Amazon datasets, achieving 93.1% and 89.64% accuracy in identifying dimensions and ratings, respectively [118]. CommTrust, a novel approach to the 'all good reputation' problem in e-commerce trust models, leveraged free-text feedback comments to create a multidimensional trust model. The algorithm, combining NLP, opinion mining, and topic modeling, effectively mitigated universally high seller reputation scores on eBay and Amazon, providing a more reliable ranking of sellers based on trust [119].

Another study combined Opinion Mining and CF algorithms to analyze Yelp data, highlighting inconsistencies between textual reviews and star ratings. The research explored the impact of restaurant popularity on user ratings, yielding noteworthy results in Root Mean Squared Error (RMSE) [120]. The significance of consumer reviews in e-commerce was emphasized in a distinct model that focused on fine-grained analysis of feedback comments. The methodology, validated on Amazon and Flipkart, revealed notable discrepancies in trust scores, enhancing the understanding of seller trust profiles [121]. The impact of Amazon's Verified Purchase badge on review helpfulness and product ratings was investigated, revealing significant increases in review helpfulness and product ratings for verified purchase reviews [122].

To enhance Amazon Search's relevance ranking, the study employed a diverse set of relevance algorithms, emphasizing the significant impact on customer satisfaction and financial outcomes [123]. Various methods for analyzing consumer opinions on platforms like Amazon.com were explored, introducing a hybrid approach that effectively ranked products based on text reviews, Question Answer (QA) data, and star ratings, enhancing sales predictions [124]. The study addressed the financial and reputational impact of product issues in over the counter (OTC) pain relief products, utilizing Amazon's product reviews to identify safety and efficacy concerns through "smoke word" dictionaries and sentiment analysis [125]. Another research investigated whether models trained on a dataset could accurately reflect human proficiency in online review writing, employing Knowledge Tracing to track the development of reviewers' skills over time [126].

The challenge of navigating through verbose customer reviews was addressed through a multi-criteria decision-making approach to recommend optimal products on platforms like Flipkart and Amazon [127]. The detection of ironic opinions in social networks and e-commerce was explored, comparing feature-based irony detection with a novel approach using character language model classifiers, showing competitive accuracy in experiments [128]. The evolving landscape of consumer behavior in e-commerce was examined, proposing an algorithmic solution to mitigate inaccuracies in user-generated reviews and enhance the decision-making process using NLP techniques [129]. The Ranking Hotels using Aspect Level Sentiment Analysis (RHALSA) algorithm was introduced, effectively evaluating and ranking hotels based on user reviews through aspect-level sentiment analysis on a Tripadvisor dataset [130].

Leveraging user-generated content for marketing was explored through sentiment analysis tools, proposing a framework to derive new scores reflecting consumer sentiments for distinct product features on Amazon [131]. The impact of technology on people's lifestyles and decision-making processes was investigated using Yelp as a social network example, emphasizing the importance of reviews analysis in monitoring changes in business public opinion over time [132]. A Feature-Based Product (FBP) Recommendation system using NLP and sentiment analysis on Amazon mobile product reviews was proposed, demonstrating the effectiveness of SVM in suggesting the best company products for user-requested features [133]. The Level of Success model (LOS) was introduced, employing NLP, review quantification, and image analysis to contribute valuable insights for effective product market evaluation in the Amazon online market review context [134].

Quantifying Online Brand Image (OBIM) by analyzing consumer reviews was explored, introducing a model that evaluated associations' favorability, strength, and uniqueness through sentiment and co-word network analysis [135]. Using Python for preprocessing NLP features, the study focused on Sunshine product recommendations on Amazon, revealing insights for quarterly sales forecasting and product development trends based on customer text reviews [136]. A novel method, Tagging Product Review (TPR), was introduced to summarize e-commerce product reviews, achieving high tag relevance scores for both popular and cold products on Amazon [137]. AmazonRep, a reputation system considering review sentiment, helpfulness votes, review timing, and user credibility, proved effective in generating and presenting reputations for diverse products on Amazon [138]. Reputation generation for diverse entities using customer reviews was addressed through a unified reputation value integrating helpfulness, time, rating, and sentiment. The method outperformed three existing systems, offering a comprehensive approach and visualizations for numerical reputation, opinion categories, and top reviews [139].

In a novel design approach, the paper used collage placement to validate sustainable features for French Press coffee carafes extracted from Amazon reviews. The study revealed a disparity between customer perceptions and engineered sustainability, emphasizing the importance of understanding diverse perspectives. Participants evaluated products based on social, environmental, and economic sustainability, highlighting the efficacy of the collage method in assessing sustainability perceptions. Demographic variations in sustainability perceptions further underscored the method's relevance [140]. Another research focused on online sales strategies for Amazon products, using sentiment analysis and opinion mining for microwave ovens, baby pacifiers, and hairdryers. Mathematical models evaluated product reputation trends, predicting potential success or failure and proposing design features for enhanced desirability [141].

The study on online product reviews from Flipkart and Amazon employed sentiment analysis and a bag of words model to assess the impact on third-party sellers. Categorizing reviews and conducting topic modeling, the findings emphasized the importance of considering both product and seller reviews for a seamless delivery and defect-free product, benefiting consumers and sellers alike [142]. Introducing a novel approach for computing reputation scores, the paper utilized a BiLSTM, Recurrent Neural Network (RNN) and NLP techniques to analyze textual opinions on online platforms like IMDB and Amazon. Experimental results demonstrated the method's effectiveness, aligning closely with ground truth and suggesting practical applicability for reputation generation [143].

A study on TripAdvisor reviews and online weather data used NLP to assess the impact of weather conditions on tourists' intention to revisit a destination. Enriching the dataset with weather information and hotel ratings, the findings identified factors like heat index and weather disparities influencing revisit intention, providing valuable insights for destination managers [144]. In a unique reputation generation system, Twitter content was evaluated to determine credible reputation scores for products. Integrating sentiment orientation, user credibility, and tweet credibility, the system's computed values closely aligned with ground truth scores from various platforms, suggesting practical applications for consumers and businesses in decision-making processes on e-commerce platforms [145].

Analyzing Amazon and iHerb reviews, the research on sweetness in food products identified opportunities for less sweet products catering to a healthier consumer base. The study employed manual curation, NLP, and machine learning to reveal the impact of sweetness on product liking, suggesting potential benefits for health-conscious customers and manufacturers [146]. Challenging the belief that longer product reviews are uniformly more helpful, the study utilized advanced machine learning methods to analyze Amazon reviews' sentence-level argumentation. Contrary to prevailing views, longer reviews with frequent shifts between positive and negative arguments were perceived as less helpful, with implications for optimizing customer feedback systems and improving reviewer guidelines [147].

The analysis of customer reviews for small domestic robots on Amazon addressed failure types and their impact on customer experience. Technical failures, particularly related to Task Completion and Robustness, significantly impacted customer experience more than Interaction or Service failures. An NLP model predicted failure content in reviews, providing insights for prioritizing crucial issues for robotic system improvement [148]. Using TripAdvisor reviews, the study explored the Memorable Tourist Experience (MTE) concept, employing NLP and machine learning to analyze terms and relationships. Comparative analysis of UNESCO sites revealed shared MTE elements and validated hypotheses, emphasizing the value of reviews as supplementary data in tourism studies [149].

Sentiment analysis on Amazon customer product reviews investigated digitization's impact on the e-commerce sector, utilizing SVM and deep learning techniques. The study provided valuable insights for businesses in the dynamic e-commerce market, indicating the effectiveness of both SVM and deep learning approaches in discerning sentiments [150]. Introducing the NLP-AHP method, the research assessed online shopping platform reviews through an empirical examination of microwave oven reviews on Amazon. The method swiftly identified crucial comments and temporal patterns, offering a valuable tool for data-driven decision-making to enhance product quality and refine sales strategies [151].

The analysis of Banglish text on social media in Bangladesh employed NLP techniques and machine learning models for product market demand assessment. Results indicated high accuracy in demand analysis, providing valuable insights into popular smartphone choices by gender in the Bangladeshi market [152]. The study on managerial responses to online customer complaints and negative reviews integrated justice theory and service recovery literature. Positive managerial responses influenced future review valence, with rational cues to procedural unfairness complaints enhancing future valence. The paper provided insights for both theory and practical applications [153].

An NLP analysis of Amazon reviews explored user satisfaction with physical activity trackers. Sentiment analysis and a Transformer-based language model classified technical aspects and user sentiments, revealing nuanced perspectives on product satisfaction [154]. The study on TripAdvisor used deep learning models based on the Myers-Briggs Type Indicator (MBTI) to discern consumers' personalities from electronic word-of-mouth (e-WOM). Findings linked specific discussion themes to personality traits, offering insights for personalized marketing messages and optimizing communication strategies [155].

The research addressed the challenge of assessing product quality in e-commerce, introducing the QLeBERT approach. Combining a quality-related lexicon, N-grams, BERT, and BiLSTM for classification, QLeBERT achieved superior performance, providing a deeper understanding of textual input for predicting product quality [156]. An algorithm utilizing language-transformer technologies automated product requirement generation from E-Shop reviews. The study showcased the transformative potential of transformer-enhanced opportunity mining in requirements engineering, efficiently extracting critical user needs from consumer reviews to enhance product improvement [157].

The impact of the "Amazon effect" on consumer perceptions of service attributes in offline/online retailers was explored. Analyzing social media comments using NLP, the study identified triggers for the Amazon effect, highlighting widespread dissatisfaction and reduced satisfaction with other retailers influenced by elevated consumer expectations shaped by Amazon [158]. The study on CF recommendation systems utilized sentiment analysis on user reviews to derive implicit ratings, introducing novel approaches that demonstrated effectiveness in enhancing CF performance [159].

To address issues in review-based recommender systems, the paper introduced the Time-Varying Attention with Dual-Optimizer (TADO) model, combining dual-optimizer network, BERT, and time-varying feature extraction. Tested on Amazon Product Reviews datasets, TADO outperformed state-of-the-art techniques by significant margins, offering improved classification and regression losses for enhanced performance [160]. Focusing on categorizing customer reviews on Amazon, the study employed machine learning techniques to enhance the e-commerce shopping experience. The model predicted sentiment, aiding users in making informed purchasing decisions by categorizing customer reviews based on inherent attributes [161]. The summary of existing literature on marketing and brand management is given in Table 6.

**Table 6.** Summary of literature on marketing and brand management.

| Ref. | Year | Dataset | Description |
|---|---|---|---|
| [118] | 2013 | eBay, Amazon | Introduced a novel algorithm combining opinion mining and dependency relation analysis for accurate e-commerce feedback comments. |
| [119] | 2014 | eBay, Amazon | Outlined CommTrust, a trust evaluation approach in e-commerce addressing the 'all good reputation' issue. |
| [120] | 2015 | Yelp | Experimentally integrated Opinion Mining and CF, revealing user inconsistencies in star ratings alignment. |
| [121] | 2016 | Amazon, Flipkart | Pioneered an approach to compute seller trust in e-commerce through fine-grained analysis of user feedback comments. |
| [122] | 2016 | Amazon | Investigated the influence of Amazon's Verified Purchase badge on review helpfulness and product ratings. |
| [123] | 2016 | Amazon | Detailed efforts to enhance Amazon Search's relevance ranking using diverse algorithms and NLP techniques. |
| [124] | 2016 | Amazon | Investigated methods for analyzing consumer opinions, proposing a hybrid approach to effectively rank products. |
| [125] | 2017 | Amazon | Focused on automating the discovery of safety and efficacy concerns in OTC joint and muscle pain relief products using "smoke word" dictionaries. |
| [126] | 2018 | Amazon | Investigated the ability of models to gauge skill acquisition in online review writing, focusing on the evolution of this skill over a sequence of reviews. |
| [127] | 2018 | Flipkart and Amazon | Introduced a method using multi-criteria decision-making to recommend the best product based on sentiment analysis. |
| [128] | 2018 | Twitter, Amazon, SemEval-2018 | Detected irony in social media and e-commerce texts using character language model classifiers. |
| [129] | 2018 | - | Addressed challenges of inaccurate user-generated reviews in E-commerce using NLP techniques. |
| [130] | 2018 | TripAdvisor | Presented Ranking Hotels using Aspect Level Sentiment Analysis (RHALSA) algorithm. |
| [131] | 2019 | Amazon | Outlined a methodological framework utilizing sentiment analysis and NLP techniques to analyze online reviews. |
| [132] | 2019 | Yelp | Examined the influence of technology, proposing a methodology to analyze reviews for assessing business attractiveness. |

| | | | |
|---|---|---|---|
| [133] | 2019 | Amazon | Introduced a Feature-Based Product Recommendation system using NLP and sentiment analysis. |
| [134] | 2020 | Amazon | Introduced the Level of Success model (LOS) for assessing product market impact in the evolving e-commerce landscape. |
| [135] | 2020 | Amazon | Proposed a model to quantify Online Brand Image (OBIM) by analyzing consumer reviews. |
| [136] | 2020 | Amazon | Employed Python, AHP, and ARIMA/OLS models to analyze Amazon "Sunshine" product sales. |
| [137] | 2020 | Amazon | Introduced an innovative Tagging Product Review (TPR) system that employed an unsupervised approach. |
| [138] | 2020 | Amazon | Detailed the development of AmazonRep, a reputation system for Amazon customers. |
| [139] | 2020 | IMDb, TripAdvisor, Amazon | Introduced a holistic approach to reputation generation from customer reviews. |
| [140] | 2021 | Amazon | Investigated differentiating e-commerce products based on customer perceptions of sustainability. |
| [141] | 2021 | Amazon | Examined online sales strategies for various products using sentiment analysis. |
| [142] | 2021 | Amazon | Utilized natural processing methods to assess the impact of online product reviews on third-party sellers. |
| [143] | 2021 | IMDb dataset. | Employed Bi-LSTM, RNN and NLP techniques to compute reputation scores for companies. |
| [144] | 2021 | TripAdvisor | Utilized NLP and classification techniques to investigate the impact of weather conditions. |
| [145] | 2021 | IMDb, Amazon, TripAdvisor, Yelp | Outlined a novel reputation generation system for Twitter. |
| [146] | 2022 | Amazon, iHerb | Examined sweetness in online food product reviews. |
| [147] | 2022 | Amazon | Challenged the assumption that longer product reviews are universally more helpful. |
| [148] | 2022 | Amazon | Investigated domestic robot failures using reviews. |
| [149] | 2022 | TripAdvisor | Investigated the Memorable Tourist Experience (MTE) concept through reviews. |
| [150] | 2022 | Amazon | Examined sentiment analysis of customer reviews using SVM, LSTM, and CNN techniques. |
| [151] | 2022 | Amazon | Employed NLP-AHP for analyzing online shopping reviews, providing actionable insights. |
| [152] | 2022 | - | Analyzed demand for smartphones in the market through social media data using NLP and machine learning models. |
| [153] | 2023 | TripAdvisor | Developed a model integrating justice theory and service recovery literature for online customer complaints. |
| [154] | 2023 | Amazon | Investigated user satisfaction with physical activity trackers using sentiment analysis. |
| [155] | 2023 | TripAdvisor | Employed text classification and topic modeling to discern consumer personalities. |
| [156] | 2023 | Amazon | Introduced QLeBERT, a model leveraging a quality-related lexicon to predict product quality. |
| [157] | 2023 | Amazon | Outlined an algorithm leveraging language-transformer technologies to automate product requirement generation. |
| [158] | 2023 | Amazon, Facebook | Explored the impact of the 'Amazon effect' on consumer perceptions of service attributes in Italian consumer electronics retailers. |
| [159] | 2023 | Amazon, Yelp | Introduced CF methods leveraging sentiment analysis on user reviews. |
| [160] | 2023 | Amazon | Introduced a novel TADO model for review-based recommender systems. |

| | | | |
|---|---|---|---|
| [161] | 2023 | Amazon | Utilized machine learning to categorize product reviews, eliminating redundancy and preprocessing text with NLP tools to train a model capable of predicting sentiment. |

**Discussion**

The NLP applications in e-commerce research presented in Section 3 provide important insights into varied data sources and approaches. Figure 3 shows the distribution of selected and reviewed publications from 2013 to 2023. This survey includes 154 publications that represent the evolution of NLP research for online consumer reviews across time. The rising number of articles published from 2013 to 2023 shows a growing interest in the application of NLP in online customer review analysis. As businesses realize the strategic benefit of knowing the opinions of customers, researchers have explored creative sentiment analysis applications in several areas. The continual increase of articles shows a dedication to tackle emerging challenges including fake review identification, cross-domain sentiment transfer learning, and multi-modal analysis.

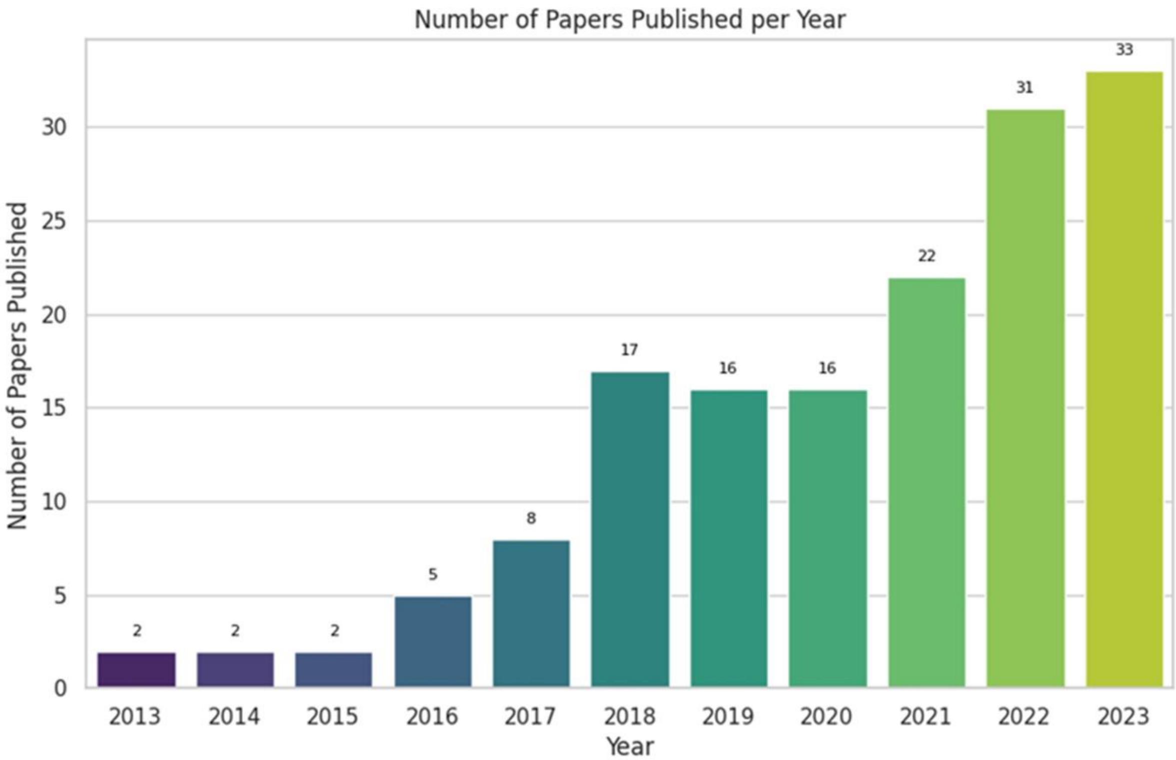


**Figure 3.** Distribution of numbers of papers published per year from 2013 to 2023.

Sentiment analysis is a major focus of existing NLP research on online consumer reviews. Early sentiment classification models [8,12,24] were gradually replaced by advanced techniques such deep neural networks [14,38,60] and the incorporation of cutting-edge models like BERT [38,49]. This trend highlights the need for more complex and context-aware sentiment analysis to better grasp online reviewer opinions and emotions. The study of sentiment polarity analysis for e-commerce customer reviews [57] and the suggested EESNN-SA-OPR approach [58] using CF and product-to-product similarity demonstrate the integration of sentiment analysis with user profile and recommendation systems. Review analysis and management are crucial to NLP applications, as shown in research on handling various online user data[20,30]. Novel aspect-based sentiment analysis methods [42,47] and HABSC, which uses grammatical characteristics and domain-specific information [21], demonstrate a rising focus on extracting granular insights from reviews. The algorithm QLeBERT [156] predicts product quality using a quality-related vocabulary, emphasizing sentiment analysis and product

 doi:10.20944/preprints202312.2210.v2



evaluation and management. These efforts improve review understanding and digital consumer opinion management.

Another important NLP application in this field is customer feedback and satisfaction analysis. The use of aggregated sentiment ratings and fuzzy logic to estimate customer loyalty [86] illustrates efforts to quantify consumer happiness using sentiment analysis. The quick customer loyalty model for e-commerce [88] and the Level of Success model (LOS) [134] also recognize the need to turn sentiment analytics into customer satisfaction and retention efforts. The use of NLP to extract insights from user-generated reviews in nutraceutical retail [87] and the study of sweetness in online food product reviews [146] demonstrate the multifaceted uses of NLP in customer satisfaction. Implementing NLP with recommendation systems shows its importance in user experiences and purchase decisions. FusionSCF [117], which combines CF with sentiment analysis, addresses cold-start and long-tail recommendation system problems. The DAMIN model [109] for click-through rate prediction and BERT models for a multi-criteria hotel recommender system [110] demonstrate how recommendation algorithms may be tailored to user preferences and review sentiments. The study of intelligent personal assistants in company processes [108] shows how NLP-driven recommendation systems affect user interactions and experiences beyond e-commerce.

NLP applications in online customer review analysis affect marketing and brand management. The 'Amazon effect' on consumer perceptions of service attributes in Italian consumer electronics retailers [158] and sentiment analysis of online sales strategies for various products [141] show how sentiment insights affect marketing strategies. The creation of an Amazon customer reputation system [138] and the study of a comprehensive approach to reputation generation from customer reviews [139] emphasize the importance of sentiment analysis in brand image management. NLP is used to analyze smartphone demand using social media data [152], demonstrating the importance of sentiment analysis in marketing and customer preferences. Online customer review NLP applications demonstrate language processing technology adaptability and influence. NLP is essential for extracting insights from online consumer-generated material, including sentiment analysis, review management, customer feedback analysis, recommendation systems, and marketing and brand management.

Existing literature uses data from several review platforms. Figure 4 shows that Amazon's enormous customer review dataset is regularly used in studies. The usage of Amazon datasets across a number of years suggests a persistent interest in sentiment analysis in the context of online shopping. This implies that online product reviews influence customer perceptions and decisions. TripAdvisor is famous in travel sentiment research studies. Hotel and visitor reviews on the platform are useful for understanding customer attitudes. The regularity with which TripAdvisor is mentioned emphasizes its significance as a source of sentiment-rich data in the travel industry. Studies also employ several datasets, indicating a move towards cross-domain sentiment analysis. Combining IMDb reviews with Amazon data or using varied datasets from Yelp, IMDb, and Kaggle shows a holistic approach to sentiment analysis that considers opinions from different areas. Including information from tourism, e-commerce, and social media shows that sentiment research may be applied across industries. Studies like [66] provide an AI framework for counterfeit goods identification, demonstrating sentiment analysis beyond review websites.

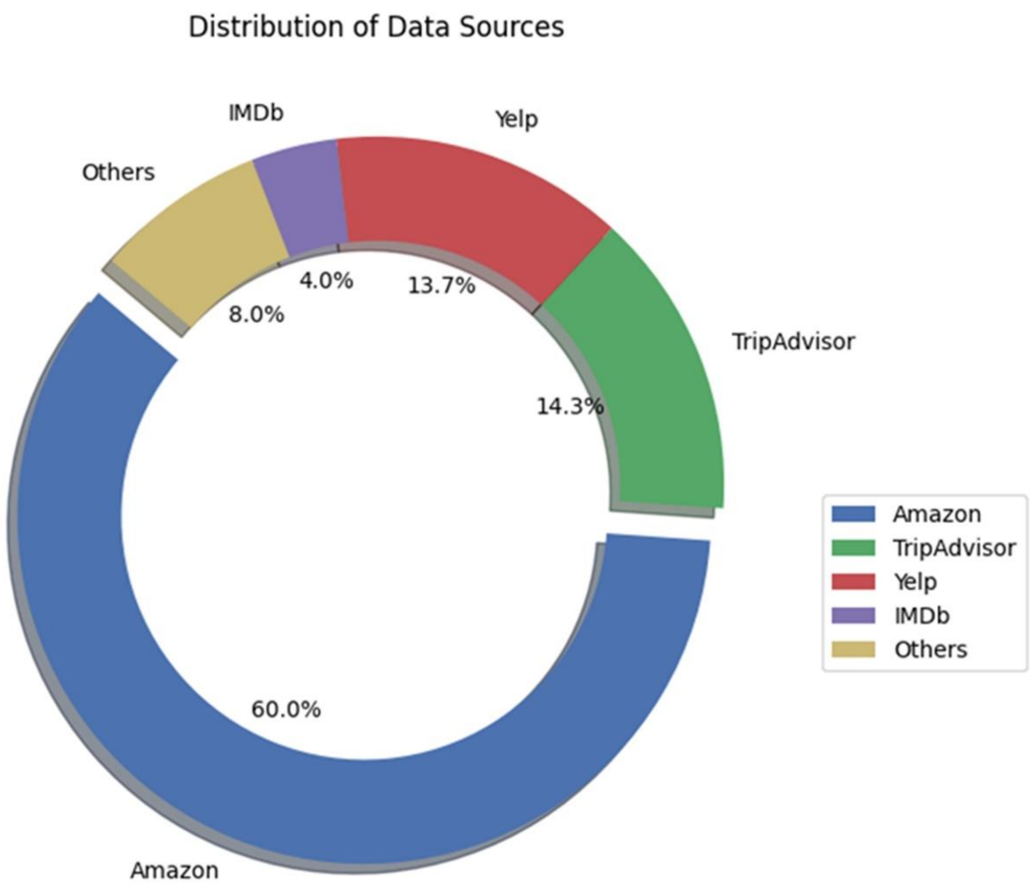


**Figure 4.** Distribution of data sources.

The changing landscape of techniques has transformed sentiment analysis. SVMs and logistic regression are still used, although sophisticated methods are becoming more popular. Deep learning models, such as deep neural networks, and cutting-edge transformers like BERT and LSTM, have shown their ability to capture complex textual relationships [14,24,29,38,39,43,110,112]. The use of cutting-edge algorithms like CNN and CF diversifies sentiment analysis [38,61]. Ensemble classifiers like Random Forest [40,48] and hybrid systems like CNN-LSTM architectures [46,101,114] demonstrate purposeful integration of several techniques for sentiment analysis. Researchers believe this technological transition is a purposeful move from conventional methodologies to advanced integration of cutting-edge technology, notably in NLP and machine learning. Advances in deep learning models and human emotion subtleties show a dedication to understanding sentiment analysis's intricacies, marking a milestone in technology and emotion comprehension [14,39,61].

NLP techniques, notably sentiment analysis, are being integrated into the e-commerce sector, with substantial consequences for platforms such as Amazon, Flipkart, and eBay. Studies show that sentiment analysis may be used to detect fake reviews and understand how cultural variables affect sentiment expression in the e-commerce business [20,45,150]. CNN-LSTM models, SVM, and hybrid techniques show the dedication to addressing these issues. Sentiment analysis, CF, and novel approaches like FusionSCF and TADO improve recommendation systems by enhancing the user experience and offer personalized information [118]. Sentiment analysis, NLP, and speech recognition are used together to derive insights from user-generated content in e-commerce [84,85]. Recent studies show that sentiment analysis is moving beyond Twitter and Facebook to embrace a wider digital environment, highlighting the increased interest in understanding attitudes across different online platforms [15].

The extent to which sentiment analysis affects e-commerce decision-making is highlighted by its effects on corporate strategy and consumer retention. Researchers study Twitter, YouTube, and

 doi:10.20944/preprints202312.2210.v2



Facebook opinions to understand user feelings in the wider online setting [59]. Sentiment analysis to identify fake reviews tackles online disinformation and emphasizes the importance of online review reliability. Studies on sentiment analysis in mobile phones, hotels, and the food business show how context affects sentiment. Domain-specific sentiment analysis algorithms are needed to capture and analyze sentiments in varied e-commerce industries [98,99]. Thus, sentiment analysis is a strategic instrument that interprets user sentiments and improves business approaches in e-commerce.

In this survey, the number of articles may not fairly reflect the volume of work done in each application area. Keyword searches may exclude papers from particular application areas, limiting their numerical representation. Rather than relying merely on the number of articles, it is preferable to investigate the specific application areas themselves. This will reduce keyword search bias by providing a deeper understanding of each application area in NLP-based review analysis.

## Open Challenges and Future Directions

The application area of NLP for analyzing online customer reviews has made significant development, as indicated by the huge number of papers included in Section 3. However, this advancement is not without its set of challenges. To guide future research and facilitate progress in this domain, the challenges and future directions are outlined below.

### *Handling Diverse Data Sources*

For reliable and adaptive sentiment analysis algorithms, varied data sources must be handled effectively. The existing research studies emphasize how important it is to understand the nuances of various social media platforms, such as Twitter and Facebook. Opinion mining in different online user data requires specialized methodologies, as [20] shows. The HABSC approach [21] uses syntactic characteristics and domain-specific information to improve sentiment analysis in TripAdvisor and Amazon datasets. Focusing on Amazon electronics product review sentiment research [35] highlights the platform's unique issues. With BERT-based models on Twitter, YouTube, Facebook, Amazon, TripAdvisor, Opera, and Personal Healthcare Agent, [45] analyses Italian corpora for sentiment. [65] mines Twitter data for sentiment analysis while [129] suggests using NLP to solve erroneous user-generated reviews in e-commerce. Furthermore, [145] proposes a Twitter reputation generating method that handles multiple sources. These studies demonstrate that sentiment analysis is complex and requires advanced methods to handle varied data sources. Federated averaging with weighting [20,35] ensures model adaptation and performance in multiple online ecosystems with varied expressions.

### *Aspect-Based Sentiment Analysis*

Aspect-Based Sentiment Analysis offers important opportunities for research and enhancement. Despite advances like [112] and [142], approaches for collecting embedded opinions about product and service attributes remain difficult. Future research should focus on improving ABSA procedures for e-commerce and tourism, which need a deep grasp of customer attitudes. According to the research, sentiment analysis of different internet user data is challenging and requires creative methods [20,21]. Advanced natural NLP approaches, such as those described in [112], may improve Aspect-Based Sentiment Analysis accuracy and applicability across domains. As suggested in [36], data pre-processing and categorization ambiguity must be addressed. Future Aspect-Based Sentiment Analysis models should effectively recognize sentiments and respond to user views' changing subtleties and complexities to better grasp customer sentiments in different scenarios.

### *Handling Multimodal Data*

Since online reviews increasingly include visual information, multimodal techniques that analyze text and images are crucial. Previous research [85,113] have recognized the usefulness of

 doi:10.20944/preprints202312.2210.v2



multimodal techniques in sentiment analysis. These methods go beyond text analysis to picture sentiment analysis. Multimodal data is important, but various problems remain. Effective algorithms that smoothly blend textual and visual data to obtain sentiment insights are a big challenge. To fully understand visual content's emotional subtleties, image sentiment analysis requires novel methods. Standardized benchmarks and assessment measures for multimodal sentiment analysis are needed to ensure model performance. In the ever-changing world of online reviews, overcoming these problems will be essential to maximize multimodal techniques and improve sentiment analysis.

*Dealing with Sarcasm and Irony*

More complex algorithms that can recognize and analyze sarcasm and irony in customer reviews are needed to improve sentiment analysis. Sarcasm and irony are difficult to recognize and grasp, even using sentiment classification methods. Researchers in [53] introduces a method for spotting sarcastic thoughts in online communication, emphasizing the necessity for novel methods to capture subtle language subtleties. [128] employs character language model classifiers to identify irony in social media and e-commerce communications. The study shows how difficult it is to spot irony in user-generated content on Twitter and e-commerce sites. Additionally, [159] presents CF approaches using sentiment analysis on user reviews, illustrating the persistent problem to manage metaphorical language. Future sentiment analysis research must refine existing models and explore new methods to better understand sarcasm and irony in customer reviews, improving system accuracy.

*Fake Review Detection*

Even with advances in this field, maintaining the correctness of these models is vital, especially given the dynamic and changing nature of online deception. Various research, including those listed, have shown the need for new methods and improvements [69,74,75]. Because of the flexibility of fraudsters, it is necessary to investigate new ways that can keep up with evolving misleading practices. To identify more complex fake reviews, future research should use powerful machine learning and NLP techniques. To build a viable fake review detection methodology, researchers, platforms, and regulatory agencies must collaborate. These issues must be addressed to preserve online reviews and build consumer and company confidence.

*User-generated Content Challenges*

Existing research on user-generated content problems reveals crucial factors that need additional focus for sentiment analysis model improvement. Importantly, these models must be more adaptable to different language settings. Existing studies emphasize the need to address language, writing, and cultural differences. To make sentiment analysis systems more robust and successful, methods must accommodate for user-generated content's linguistic complexities [153]. This requires methods that can detect sentiment expressions in different languages, accommodate different writing styles, and delicately capture cultural subtleties in the text. Overcoming these problems will help sentiment analysis algorithms become more applicable and reliable in varied language and cultural situations as user-generated content evolves. The references provide vital insights into existing attempts to address these difficulties and establish the framework for future initiatives.

*Integration of Machine Learning Models*

The incorporation of machine learning models in sentiment analysis presents challenges and research opportunities. Despite advances in this field, thorough studies that assess the performance of different machine learning algorithms across different contexts are needed. Understanding the strengths and weaknesses of different models is essential to finding the best solutions for certain

 doi:10.20944/preprints202312.2210.v2



applications [162]. For advanced sentiment analysis machine learning model evaluation, future research should overcome this gap. Such investigations can improve models and provide new methods, advancing the field of study. For actual implementation and real-world application, these models' adaptability and scalability to multiple datasets and domains must be studied. As sentiment analysis evolves, improving machine learning model integration will improve sentiment categorization accuracy and reliability across contexts.

*Explainable and Interpretable Models*

Developing explainable and interpretable models is essential for understanding sentiment analysis model decision-making. Existing studies offer useful methods for reaching this aim [163]. However, federated models may lack interpretability, making their decision-making procedures difficult to understand. The lack of transparency in federated learning models raises questions about their dependability and trustworthiness in real-world applications. Federated learning interpretability research is needed to solve this problem. Understanding the decision processes of federated sentiment analysis models will improve their practicality and promote openness and accountability in machine learning applications. Future research should focus on building methods to understand federated models, allowing stakeholders to understand and accept sentiment analysis results across multiple domains.

*Cross-Domain Generalization*

To improve sentiment analysis model adaptation across domains, cross-domain generalization strategies must be explored. Existing sentiment analysis methods have been successful in e-commerce, tourism, and social media, but generalizing them to other domains is difficult. The findings suggest exploring ways to improve sentiment analysis model adaptability and performance in new or diversified environments. To achieve robust cross-domain generalization, domain-specific subtleties, language differences, and sentiment expressions must be overcome. Future research might focus on transfer learning, using pre-trained models to capture domain-agnostic information, and domain adaptation to fine-tune models for specific domains [164]. Addressing cross-domain generalization difficulties makes sentiment analysis models more versatile and useful in real-world settings with different and dynamic domains.

*Real-Time Sentiment Analysis*

Applications that need real-time insights require fast algorithms and models for sentiment analysis. Recent studies such as [154] have explored real-time sentiment analysis methods, but more robust algorithms and models are needed. In applications that need quick insights, sentiment analysis must be fast and accurate. Developing real-time methods to handle and analyze massive data streams and respond to changing user views and contextual details is difficult. Real-time sentiment analysis in customer feedback, social media monitoring, and online reviews requires addressing data scalability, algorithm efficiency, and model adaptation. Novel real-time data stream handling methods, model architecture optimization for quick inference, and sentiment dynamics temporal relationships may be explored in future study. Integration of edge computing and effective parallel processing might also improve real-time sentiment analysis systems [165].

*Ethical Considerations*

To responsibly deploy sentiment analysis models, approaches and frameworks must address ethical issues including algorithm bias, privacy, and consumer data usage. This is stressed in studies like [145,149]. These studies contribute to sentiment analysis model ethics discussions. Algorithm

 doi:10.20944/preprints202312.2210.v2



biases must be acknowledged and mitigated to avoid unforeseen outcomes and maintain fairness. Privacy problems, especially in regulated businesses, require rigorous methods and structures. Federated learning prioritizes data protection and addresses privacy challenges. Federated learning improves privacy and ethics by keeping raw data on local devices and only sharing model changes [166]. Creating and following ethical norms will help us build trust, transparency, and responsible innovation in sentiment analysis.

The literature supports NLP in online customer review analysis, and the obstacles and future research paths will develop sentiment analysis approaches and their practical applications. To address these issues and advance sentiment analysis in the digital era, researchers should work across disciplines.

## Conclusions

This study examines 154 articles published between 2013 and 2023, revealing a decade of NLP advances in online customer review analysis. The literature shows that sentiment analysis approaches have evolved to improve user opinion mining accuracy and quality. From SVM and genetic algorithms to cutting-edge methods like BERT and deep neural networks, researchers have successfully used advanced algorithms and machine learning models. The findings suggest that the review analysis is being used in e-commerce, tourism, and other industries to improve product, marketing, and decision-making. In addition, the taxonomy offered in this study provides an organized summary of the changing landscape by classifying the research based on applications. Even though there have been great advancements in sentiment classification accuracy and the exploration of new methodologies in the reviewed literature, there are still a number of open research issues. These include fake review detection and prevention, multi-modal data integration for better analysis, and cross-domain sentiment transfer learning. These problems must be addressed to produce more robust and universally applicable NLP models for online customer reviews as the industry advances. This survey shows how NLP research affects many different areas and can change how businesses interpret and exploit customer reviews in the digital age.

## References

1. Dwivedi, Y.K., et al., *Setting the future of digital and social media marketing research: Perspectives and research propositions.* International Journal of Information Management, 2021. **59**: p. 102168.
2. Thakur, R., *Customer engagement and online reviews.* Journal of Retailing and Consumer Services, 2018. **41**: p. 48-59.
3. Sezgen, E., K.J. Mason, and R. Mayer, *Voice of airline passenger: A text mining approach to understand customer satisfaction.* Journal of Air Transport Management, 2019. **77**: p. 65-74.
4. Sun, S., C. Luo, and J. Chen, *A review of natural language processing techniques for opinion mining systems.* Information fusion, 2017. **36**: p. 10-25.
5. Chauhan, G.S., et al., *Aspect based sentiment analysis using deep learning approaches: A survey.* Computer Science Review, 2023. **49**: p. 100576.
6. Mehra, P., *Unexpected surprise: Emotion analysis and aspect based sentiment analysis (ABSA) of user generated comments to study behavioral intentions of tourists.* Tourism Management Perspectives, 2023. **45**: p. 101063.
7. Gupta, H. and M. Patel. *Method of text summarization using LSA and sentence based topic modelling with Bert*. in *2021 international conference on artificial intelligence and smart systems (ICAIS)*. 2021. IEEE.
8. Marrese-Taylor, E., J.D. Velásquez, and F. Bravo-Marquez, *A novel deterministic approach for aspect-based opinion mining in tourism products reviews.* Expert Systems with Applications, 2014. **41**(17): p. 7764-7775.
9. Singla, Z., S. Randhawa, and S. Jain. *Statistical and sentiment analysis of consumer product reviews*. in *8th International Conference on Computing, Communications and Networking Technologies, ICCCNT 2017*. 2017. Institute of Electrical and Electronics Engineers Inc.
10. Fazzolari, M., et al., *A Study on Text-Score Disagreement in Online Reviews.* Cognitive Computation, 2017. **9**(5): p. 689-701.
11. Li, J.B. and L.B. Yang. *A Rule-Based Chinese Sentiment Mining System with Self-Expanding Dictionary - Taking TripAdvisor as an Example*. in *14th IEEE International Conference on E-Business Engineering, ICEBE 2017*. 2017. Institute of Electrical and Electronics Engineers Inc.
12. Magesh, G. and P. Swarnalatha, *Analyzing customer sentiments using machine learning techniques.* International Journal of Civil Engineering and Technology, 2017. **8**(10): p. 1829-1842.

13. Dufourq, E. and B.A. Bassett. *Automated classification of text sentiment*. in *23rd South African Institute of Computer Scientists and Information Technologists Conference, SAICSIT 2017*. 2017. Association for Computing Machinery.
14. Jain, V.K., S. Kumar, and P. Mahanti, *Sentiment recognition in customer reviews using deep learning.* International Journal of Enterprise Information Systems, 2018. **14**(2): p. 77-86.
15. Gawade, J. and L. Parthiban, *Opinion mining of amazon product data by hybrid svm.* Journal of Advanced Research in Dynamical and Control Systems, 2018. **10**(13): p. 2034-2041.
16. Aryo Prakoso, A., et al. *A Lexicon-Based Sentiment Analysis for Amazon Web Review*. in *3rd International Seminar on Application for Technology of Information and Communication, iSemantic 2018*. 2018. Institute of Electrical and Electronics Engineers Inc.
17. Panthati, J., et al. *Sentiment Analysis of Product Reviews using Deep Learning*. in *7th International Conference on Advances in Computing, Communications and Informatics, ICACCI 2018*. 2018. Institute of Electrical and Electronics Engineers Inc.
18. Wang, J., et al. *Classification of multiple affective attributes of customer reviews: Using classical machine learning and deep learning*. in *2nd International Conference on Computer Science and Application Engineering, CSAE 2018*. 2018. Association for Computing Machinery.
19. Angioni, M., et al. *An evaluation method for the performance measurement of an opinion mining system*. in *20th International Conference on Enterprise Information Systems, ICEIS 2018*. 2018. SciTePress.
20. Karthikayini, T. and N.K. Srinath. *Comparative Polarity Analysis on Amazon Product Reviews Using Existing Machine Learning Algorithms*. in *2nd International Conference on Computational Systems and Information Technology for Sustainable Solutions, CSITSS 2017*. 2018. Institute of Electrical and Electronics Engineers Inc.
21. Bansal, B. and S. Srivastava, *Hybrid attribute based sentiment classification of online reviews for consumer intelligence.* Applied Intelligence, 2019. **49**(1): p. 137-149.
22. Nakayama, M. and Y. Wan, *The cultural impact on social commerce: A sentiment analysis on Yelp ethnic restaurant reviews.* Information and Management, 2019. **56**(2): p. 271-279.
23. Chauhan, C. and S. Sehgal. *Sentiment classification for mobile reviews using KNIME*. in *2018 International Conference on Computing, Power and Communication Technologies, GUCON 2018*. 2019. Institute of Electrical and Electronics Engineers Inc.
24. Jabbar, J., et al. *Real-time sentiment analysis on E-Commerce application*. in *16th IEEE International Conference on Networking, Sensing and Control, ICNSC 2019*. 2019. Institute of Electrical and Electronics Engineers Inc.
25. Pathuri, S.K., et al., *Feature-based opinion mining for amazon product's using MLT.* International Journal of Innovative Technology and Exploring Engineering, 2019. **8**(11): p. 4105-4109.
26. Dharaiya, S., et al. *Generating Positive and Negative Sentiment Word Clouds from E-Commerce Product Reviews*. in *2020 International Conference on Computational Performance Evaluation, ComPE 2020*. 2020. Institute of Electrical and Electronics Engineers Inc.
27. Jamadi Khiabani, P., M.E. Basiri, and H. Rastegari, *An improved evidence-based aggregation method for sentiment analysis.* Journal of Information Science, 2020. **46**(3): p. 340-360.
28. Khatun, F., et al., *Sentiment analysis of Amazon book review data using lexicon based analysis*, in *3rd International Conference on Computational Vision and Bio Inspired Computing, ICCVBIC 2019*, S. Smys, et al., Editors. 2020, Springer. p. 1303-1309.
29. Chatterjee, D.P., et al., *Efficacy of Oversampling Over Machine Learning Algorithms in Case of Sentiment Analysis*, in *4th International Conference on Data Management, Analytics and Innovation, ICDMAI 2020*, N. Sharma, et al., Editors. 2021, Springer Science and Business Media Deutschland GmbH. p. 247-260.
30. Bhamare, B.R. and J. Prabhu, *A supervised scheme for aspect extraction in sentiment analysis using the hybrid feature set of word dependency relations and lemmas.* PeerJ Computer Science, 2021. **7**: p. 1-22.
31. Mukherjee, P., et al. *Effect of Negation in Sentences on Sentiment Analysis and Polarity Detection*. in *2021 Complex Adaptive Systems Conference*. 2021. Elsevier B.V.
32. Gupta, A., A. Rastogi, and A. Katal. *A Comparative Study of Amazon Product Reviews Using Sentiment Analysis*. in *7th IEEE International Conference on Advances in Computing, Communication and Control, ICAC3 2021*. 2021. Institute of Electrical and Electronics Engineers Inc.
33. Khanam, R. and A. Sharma. *Sentiment Analysis using Different Machine Learning Techniques for Product Review*. in *2021 International Conference on Computational Performance Evaluation, ComPE 2021*. 2021. Institute of Electrical and Electronics Engineers Inc.
34. Chatterjee, I., et al., *Statistics-based outlier detection and correction method for amazon customer reviews.* Entropy, 2021. **23**(12).
35. Urkude, S.V., et al., *Comparative Analysis on Machine Learning Techniques: A case study on Amazon Product Reviews.* International Journal of Mechanical Engineering, 2021. **6**(3): p. 739-744.
36. Chen, Y. and J. Yao. *Sentiment Analysis Using Part-of-Speech-Based Feature Extraction and Game-Theoretic Rough Sets*. in *21st IEEE International Conference on Data Mining Workshops, ICDMW 2021*. 2021. IEEE Computer Society.

37. Hawlader, M., et al. *Amazon Product Reviews: Sentiment Analysis Using Supervised Learning Algorithms*. in *2021 International Conference on Electronics, Communications and Information Technology, ICECIT 2021*. 2021. Institute of Electrical and Electronics Engineers Inc.
38. 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, 2021. **2**: p. 64-69.
39. Sinnasamy, T. and N.N.A. Sjaif, *Sentiment Analysis using Term based Method for Customers' Reviews in Amazon Product.* International Journal of Advanced Computer Science and Applications, 2022. **13**(7): p. 685-691.
40. Maurya, S. and V. Pratap. *Sentiment Analysis on Amazon Product Reviews*. in *2022 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing, COM-IT-CON 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
41. Naureen, A., A. Siddiqa, and P.J. Devi, *Amazon Product Alexa's Sentiment Analysis Using Machine Learning Algorithms*, in *9th International Conference on Innovations in Electronics and Communication Engineering, ICIECE 2021*, H.S. Saini, et al., Editors. 2022, Springer Science and Business Media Deutschland GmbH. p. 543-551.
42. Chu, M., et al., *Language interpretation in travel guidance platform: Text mining and sentiment analysis of TripAdvisor reviews.* Frontiers in Psychology, 2022. **13**.
43. Nasfi, R. and N. Bouguila, *Sentiment Analysis from User Reviews Using a Hybrid Generative-Discriminative HMM-SVM Approach*, in *Joint IAPR International Workshop on Structural, Syntactic, and Statistical Pattern Recognition, S+SSPR 2022*, A. Krzyzak, et al., Editors. 2022, Springer Science and Business Media Deutschland GmbH. p. 74-83.
44. Yarkareddy, S., T. Sasikala, and S. Santhanalakshmi. *Sentiment Analysis Of Amazon Fine Food Reviews*. in *4th International Conference on Smart Systems and Inventive Technology, ICSSIT 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
45. Roccabruna, G., S. Azzolin, and G. Riccardi. *Multi-source Multi-domain Sentiment Analysis with BERT-based Models*. in *13th International Conference on Language Resources and Evaluation Conference, LREC 2022*. 2022. European Language Resources Association (ELRA).
46. Juyal, P. *Sentimental analysis of amazon customers based on their review comments*. in *1st IEEE International Interdisciplinary Humanitarian Conference for Sustainability, IIHC 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
47. Uma, R., et al. *Support Vector Machine and Convolutional Neural Network Approach to Customer Review Sentiment Analysis*. in *1st International Conference on Computational Science and Technology, ICCST 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
48. Kamalesh, B. and S. Vijayalakshmi. *Sentiment Analysis on Amazon Product Review*. in *4th International Conference on Advances in Computing, Communication Control and Networking, ICAC3N 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
49. Singh, U., et al., *Towards improving e-commerce customer review analysis for sentiment detection.* Scientific Reports, 2022. **12**(1).
50. Yang, Z., P.C.I. Pang, and H.Y. Kan. *Performance Comparison of Deep Learning Text Embeddings in Sentiment Analysis Tasks with Online Consumer Reviews*. in *10th International Conference on Information Technology: IoT and Smart City, ICIT 2022*. 2022. Association for Computing Machinery.
51. Gupta, D., et al. *Food Review Analysis and Sentiment Prediction using Machine Learning Models*. in *2nd International Conference on Advances in Electrical, Computing, Communication and Sustainable Technologies, ICAECT 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
52. Hariharan, K., et al. *Analyzation of Sentiments of Product Reviews using Natural Language Processing*. in *9th International Conference on Advanced Computing and Communication Systems, ICACCS 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
53. Maheswari, S.U. and S.S. Dhenakaran, *Improved ensemble based deep learning approach for sarcastic opinion classification.* Multimedia Tools and Applications, 2023.
54. Geetha, K., et al. *Deep Learning and Sentiment Analysis Improve E-commerce Sales Prediction*. in *2023 International Conference on Data Science and Network Security, ICDSNS 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
55. Dharrao, D., et al. *E-Commerce Product Review Classification based on Supervised Machine Learning Techniques*. in *9th International Conference on Advanced Computing and Communication Systems, ICACCS 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
56. Kumar, A. *A Machine Learning-based Automated Approach for Mining Customer Opinion*. in *4th International Conference on Electronics and Sustainable Communication Systems, ICESC 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
57. Mehul, A.R., et al. *Sentiment Polarity Detection Using Machine Learning and Deep Learning*. in *3rd International Conference on Electrical, Computer and Communication Engineering, ECCE 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
58. Solairaj, A., G. Sugitha, and G. Kavitha, *Enhanced Elman spike neural network based sentiment analysis of online product recommendation.* Applied Soft Computing, 2023. **132**.

59. Sumathi, K. and K. Santharam, *Sentiment Analysis on Feedback Data of E-commerce Products Based on NLP*, in *7th International Conference on Inventive Systems and Control, ICISC 2023*, V. Suma, P. Lorenz, and Z. Baig, Editors. 2023, Springer Science and Business Media Deutschland GmbH. p. 387-401.
60. Qorich, M. and R. El Ouazzani, *Text sentiment classification of Amazon reviews using word embeddings and convolutional neural networks.* Journal of Supercomputing, 2023. **79**(10): p. 11029-11054.
61. Elangovan, D. and V. Subedha, *Adaptive Particle Grey Wolf Optimizer with Deep Learning-based Sentiment Analysis on Online Product Reviews.* Engineering, Technology and Applied Science Research, 2023. **13**(3): p. 10989-10993.
62. Kaur, G. and A. Sharma, *A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis.* Journal of Big Data, 2023. **10**(1).
63. Venkataraman, T.K. and A. Jadhav. *Classifying the Sentiment Polarity of Amazon Mobile Phone Reviews and their Ratings*. in *4th International e-Conference on Recent Advancement in Mechanical Engineering and Technology, ICRAMET 2021*. 2023. American Institute of Physics Inc.
64. Meghana, M.S., et al. *Sentiment Analysis on Amazon Product Reviews using LSTM and Naive Bayes*. in *7th International Conference on Computing Methodologies and Communication, ICCMC 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
65. Kshirsagar, P.R., et al. *A Scalable Platform to Collect, Store, Visualize and Analyze Big Data in Real- Time*. in *3rd International Conference on Innovative Practices in Technology and Management, ICIPTM 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
66. Wimmer, H. and V. Yoon. *Leveraging technology to improve intent to purchase*. in *17th Annual International Conference on Electronic Commerce, ICEC 2015*. 2015. Association for Computing Machinery.
67. Ata Ur, R., et al. *Intelligent interface for fake product review monitoring and removal*. in *16th International Conference on Electrical Engineering, Computing Science and Automatic Control, CCE 2019*. 2019. Institute of Electrical and Electronics Engineers Inc.
68. Lutz, B., N. Pröllochs, and D. Neumann. *The longer the better? The interplay between review length and line of argumentation in online consumer reviews*. in *40th International Conference on Information Systems, ICIS 2019*. 2019. Association for Information Systems.
69. Jin, K., X. Zhang, and J. Zhang. *Learning to Generate Diverse and Authentic Reviews via an Encoder-Decoder Model with Transformer and GRU*. in *2019 IEEE International Conference on Big Data, Big Data 2019*. 2019. Institute of Electrical and Electronics Engineers Inc.
70. Sadman, N., et al. *Detect Review Manipulation by Leveraging Reviewer Historical Stylometrics in Amazon, Yelp, Facebook and Google Reviews*. in *6th International Conference on E-Business and Applications, ICEBA 2020*. 2020. Association for Computing Machinery.
71. Xu, S., S.E. Barbosa, and D. Hong, *BERT Feature Based Model for Predicting the Helpfulness Scores of Online Customers Reviews*, in *Future of Information and Communication Conference, FICC 2020*, K. Arai, S. Kapoor, and R. Bhatia, Editors. 2020, Springer. p. 270-281.
72. Bilal, M., et al., *Profiling reviewers' social network strength and predicting the "Helpfulness" of online customer reviews.* Electronic Commerce Research and Applications, 2021. **45**: p. 101026.
73. Jain, P.K., R. Pamula, and S. Ansari, *A Supervised Machine Learning Approach for the Credibility Assessment of User-Generated Content.* Wireless Personal Communications, 2021. **118**(4): p. 2469-2485.
74. Anas, S.M. and S. Kumari. *Opinion Mining based Fake Product review Monitoring and Removal System*. in *6th International Conference on Inventive Computation Technologies, ICICT 2021*. 2021. Institute of Electrical and Electronics Engineers Inc.
75. Birim, Ş.Ö., et al., *Detecting fake reviews through topic modelling.* Journal of Business Research, 2022. **149**: p. 884-900.
76. Aishwarya, R. and K.N. Prashanth Kumar, *An Application for Genuine Reviewer Identification Using Deep Learning Techniques*, in *7th International Conference on Information and Communication Technology for Competitive Strategies , ICTCS 2022*, A. Joshi, M. Mahmud, and R.G. Ragel, Editors. 2023, Springer Science and Business Media Deutschland GmbH. p. 257-270.
77. Thilagavathy, A., et al. *Fake Product Review Detection and Elimination using Opinion Mining*. in *2023 IEEE World Conference on Communication and Computing, WCONF 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
78. Iliev, A.I., et al., *Fake Review Recognition Using an SVM Model*, in *Proceedings of the Computing Conference 2023*, K. Arai, Editor. 2023, Springer Science and Business Media Deutschland GmbH. p. 885-893.
79. Deshai, N. and B. Bhaskara Rao, *Unmasking deception: a CNN and adaptive PSO approach to detecting fake online reviews.* Soft Computing, 2023. **27**(16): p. 11357-11378.
80. Bilal, M. and A.A. Almazroi, *Effectiveness of fine-tuned BERT model in classification of helpful and unhelpful online customer reviews.* Electronic Commerce Research, 2023. **23**(4): p. 2737-2757.
81. Marrese-Taylor, E., et al. *Identifying customer preferences about tourism products using an aspect-based opinion mining approach*. in *17th International Conference in Knowledge Based and Intelligent Information and Engineering Systems, KES 2013*. 2013. Kitakyushu: Elsevier B.V.

82. Tian, X., et al. *Mining online hotel reviews: A case study from hotels in China*. in *22nd Americas Conference on Information Systems: Surfing the IT Innovation Wave, AMCIS 2016*. 2016. Association for Information Systems.
83. Hu, Y.H., Y.L. Chen, and H.L. Chou, *Opinion mining from online hotel reviews – A text summarization approach.* Information Processing and Management, 2017. **53**(2): p. 436-449.
84. Rogojanu, I., et al. *Smart shopping technologies for indoor markets*. in *21st IEEE International Conference on Computational Science and Engineering, CSE 2018*. 2018. Institute of Electrical and Electronics Engineers Inc.
85. Limsopatham, N., O. Rokhlenko, and D. Carmel. *Research Challenges in Building a Voice-based Artificial Personal Shopper - Position Paper*. in *2nd International Workshop on Search-Oriented Conversational AI, SCAI 2018*. 2018. Association for Computational Linguistics (ACL).
86. Ghani, U., I.S. Bajwa, and A. Ashfaq, *A fuzzy logic based intelligent system for measuring customer loyalty and decision making.* Symmetry, 2018. **10**(12).
87. John, D.L., et al. *Topic Modeling to Extract Information from Nutraceutical Product Reviews*. in *16th IEEE Annual Consumer Communications and Networking Conference, CCNC 2019*. 2019. Institute of Electrical and Electronics Engineers Inc.
88. Ashfaq, A. and M. Kausar, *A Fuzzy Logic Model for Evaluating Customer Loyalty in e-Commerce*, in *1st International Conference on Intelligent Technologies and Applications, INTAP 2018*, I.S. Bajwa, A. Costa, and F. Kamareddine, Editors. 2019, Springer Verlag. p. 216-227.
89. Ching, M.R.D. and R. De Dios Bulos. *Improving restaurants' business performance using yelp data sets through sentiment analysis*. in *3rd International Conference on E-Commerce, E-Business and E-Government, ICEEG 2019*. 2019. Association for Computing Machinery.
90. Chang, Y.C., C.H. Ku, and C.H. Chen, *Social media analytics: Extracting and visualizing Hilton hotel ratings and reviews from TripAdvisor.* International Journal of Information Management, 2019. **48**: p. 263-279.
91. Srivastava, V. and A.D. Kalro, *Enhancing the Helpfulness of Online Consumer Reviews: The Role of Latent (Content) Factors.* Journal of Interactive Marketing, 2019. **48**: p. 33-50.
92. Sann, R. and P.C. Lai, *Understanding homophily of service failure within the hotel guest cycle: Applying NLP-aspect-based sentiment analysis to the hospitality industry.* International Journal of Hospitality Management, 2020. **91**.
93. Christodoulou, E., A. Gregoriades, and S. Papapanayides. *A data analytics approach to online tourists' reviews evaluation*. in *22nd International Conference on Enterprise Information Systems, ICEIS 2020*. 2020. SciTePress.
94. van Vliet, M., et al., *Identifying and Classifying User Requirements in Online Feedback via Crowdsourcing*, in *26th International Working Conference on Requirements Engineering: Foundation for Software Quality, REFSQ 2020*, N. Madhavji and L. Pasquale, Editors. 2020, Springer. p. 143-159.
95. Herbst, C.M., et al., *What do parents value in a child care provider? Evidence from Yelp consumer reviews.* Early Childhood Research Quarterly, 2020. **51**: p. 288-306.
96. Bhatia, S., *A Comparative Study of Opinion Summarization Techniques.* IEEE Transactions on Computational Social Systems, 2021. **8**(1): p. 110-117.
97. Moharkar, K., et al. *Responding to customer queries automatically by customer reviews' based Question Answering*. in *6th International Conference on Natural Language Processing and Information Retrieval, NLPIR 2022*. 2022. Association for Computing Machinery.
98. Sergiacomi, C., et al., *Exploring National Park Visitors' Judgements from Social Media: The Case Study of Plitvice Lakes National Park.* Forests, 2022. **13**(5).
99. Barbierato, E., I. Bernetti, and I. Capecchi, *Analyzing TripAdvisor reviews of wine tours: an approach based on text mining and sentiment analysis.* International Journal of Wine Business Research, 2022. **34**(2): p. 212-236.
100. Ratmele, A. and R. Thakur, *OpExHAN: opinion extraction using hierarchical attention network from unstructured reviews.* Social Network Analysis and Mining, 2022. **12**(1).
101. Suresh, P. and K. Gurumoorthy, *Mining of Customer Review Feedback Using Sentiment Analysis for Smart Phone Product*, in *International Conference on Computing, Communication, Electrical and Biomedical Systems, ICCCEBS 2021*, A. Ramu, C. Chee Onn, and M.G. Sumithra, Editors. 2022, Springer Science and Business Media Deutschland GmbH. p. 247-259.
102. Hsu, T.Y., Y. Suhara, and X. Wang. *Summarizing Community-based Question-Answer Pairs*. in *2022 Conference on Empirical Methods in Natural Language Processing, EMNLP 2022*. 2022. Association for Computational Linguistics (ACL).
103. Alves, M.D., A.G. Lobo, and A.M. Reis. *Assessing the use of pre-trained transformers to classify customer reviews*. in *5th International Conference on Quality Engineering and Management: A Better World with Quality! Quality in the Digital Transformation, ICQEM 2022*. 2022. Universidade do Minho.
104. Nainwal, H., et al. *Text Summarization of Amazon Customer Reviews using NLP*. in *17th INDIACom; 2023 10th International Conference on Computing for Sustainable Global Development, INDIACom 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
105. Wang, H., et al. *Extracting Fine-Grained Service Value Features and Distributions for Accurate Service Recommendation*. in *24th IEEE International Conference on Web Services, ICWS 2017*. 2017. Institute of Electrical and Electronics Engineers Inc.

106. Wang, C., M. Niepert, and H. Li. *LRMM: Learning to recommend with missing modalities*. in *2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018*. 2018. Association for Computational Linguistics.
107. Neeraj, S., C. Oswald, and B. Sivaselvan. *A Novel Gamification Approach to Recomendation Based Mobile Applications*. in *9th International Conference on Advanced Computing, ICoAC 2017*. 2018. Institute of Electrical and Electronics Engineers Inc.
108. Hüsson, D., A. Holland, and R.A. Sánchez, *Intelligent Personal Assistant in Business-Context: Key-feature Evaluation for User Acceptance.* Business Systems Research, 2020. **11**(3): p. 147-166.
109. Xia, X., W. Fang, and X. Shi, *Multi-interest Network Based on Double Attention for Click-Through Rate Prediction*, in *5th International Joint Conference on Asia-Pacific Web and Web-Age Information Management, APWeb-WAIM 2021*, L.H. U, et al., Editors. 2021, Springer Science and Business Media Deutschland GmbH. p. 310-322.
110. Zhuang, Y. and J. Kim, *A bert-based multi-criteria recommender system for hotel promotion management.* Sustainability (Switzerland), 2021. **13**(14).
111. Tahir, R. and M. Asif Naeem. *A Machine Learning based Approach to Identify User Interests from Social Data*. in *24th International Multitopic Conference, INMIC 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
112. Patidar, R. and S. Patel. *Design & Implementation of Product Recommendation Solution using Sentiment Analysis*. in *2022 International Conference on Edge Computing and Applications, ICECAA 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
113. Lee, C., et al., *Improving Graph-Based Movie Recommender System Using Cinematic Experience.* Applied Sciences (Switzerland), 2022. **12**(3).
114. Acharya, N., A.M. Sassenberg, and J. Soar, *Effects of cognitive absorption on continuous use intention of AI-driven recommender systems in e-commerce.* Foresight, 2023. **25**(2): p. 194-208.
115. Noorian, A., A. Harounabadi, and M. Hazratifard, *A sequential neural recommendation system exploiting BERT and LSTM on social media posts.* Complex and Intelligent Systems, 2023.
116. Paul, P. and R.P. Singh. *A weighted hybrid recommendation approach for user's contentment using natural language processing*. in *1st International Conference on Applied Computational Intelligence and Analytics, ACIA 2022*. 2023. American Institute of Physics Inc.
117. Ananth, G.S., K. Raghuveer, and S. Vasanth Kumar. *Fusion of Various Sentiment Analysis Techniques for an Effective Contextual Recommender System*. in *16th Innovations in Software Engineering Conference, ISEC 2023*. 2023. Association for Computing Machinery.
118. Cui, L., et al., *Mining e-commerce feedback comments for dimension rating profiles*, in *9th International Conference on Advanced Data Mining and Applications, ADMA 2013*. 2013: Hangzhou. p. 1-12.
119. Zhang, X., L. Cui, and Y. Wang, *Commtrust: Computing multi-dimensional trust by mining e-commerce feedback comments.* IEEE Transactions on Knowledge and Data Engineering, 2014. **26**(7): p. 1631-1643.
120. Angioni, M., M.L. Clemente, and F. Tuveri. *Combining Opinion Mining with Collaborative Filtering*. in *11th International Conference on Web Information Systems and Technologies, WEBIST 2015*. 2015. SciTePress.
121. Bhargava, K., et al. *Comment based Seller Trust model for E-commerce*. in *2016 International Conference on Computational Techniques in Information and Communication Technologies, ICCTICT 2016*. 2016. Institute of Electrical and Electronics Engineers Inc.
122. Kokkodis, M. and T. Lappas. *The relationship between disclosing purchase information and reputation systems in electronic markets*. in *2016 International Conference on Information Systems, ICIS 2016*. 2016. Association for Information Systems.
123. Sorokina, D. and E. Cantú-Paz. *Amazon search: The joy of ranking products*. in *39th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR 2016*. 2016. Association for Computing Machinery, Inc.
124. Anjum, B. and C.L. Sabharwal. *An Entropy based product ranking algorithm using reviews and Q&A data*. in *22nd International Conference on Distributed Multimedia Systems, DMS 2016*. 2016. Knowledge Systems Institute Graduate School.
125. Adams, D.Z., R. Gruss, and A.S. Abrahams, *Automated discovery of safety and efficacy concerns for joint & muscle pain relief treatments from online reviews.* International Journal of Medical Informatics, 2017. **100**: p. 108-120.
126. Megasari, M., et al. *Can models learned from a dataset reflect acquisition of procedural knowledge? An experiment with automatic measurement of online review quality*. in *20th International Workshop on Design, Optimization, Languages and Analytical Processing of Big Data, DOLAP 2018*. 2018. CEUR-WS.
127. Kumar, G. *A multi-criteria decision making approach for recommending a product using sentiment analysis*. in *12th International Conference on Research Challenges in Information Science, RCIS 2018*. 2018. IEEE Computer Society.
128. Clavel Quintero, Y. and L. Arco García, *Irony detection based on character language model classifiers*, in *6th International Workshop on Artificial Intelligence and Pattern Recognition, IWAIPR 2018*, Y.H. Heredia, V.M. Nunez, and J.R. Shulcloper, Editors. 2018, Springer Verlag. p. 158-165.

129. Mastan Rao, T., et al., *A framework for generating rankings to E-commerce products based on reviews using NLP.* International Journal of Mechanical Engineering and Technology, 2018. **9**(1): p. 44-52.
130. Panigrahi, N. and T. Asha, *RHALSA: Ranking hotels using aspect level sentiment analysis.* Journal of Computer Science, 2018. **14**(11): p. 1512-1520.
131. Kauffmann, E., et al., *Managing marketing decision-making with sentiment analysis: An evaluation of the main product features using text data mining.* Sustainability (Switzerland), 2019. **11**(15).
132. Galli, A., et al., *A Business Reputation Methodology Using Social Network Analysis*, in *16th International Symposium on Pervasive Systems, Algorithms and Networks, I-SPAN 2019*, C. Esposito, J. Hong, and K.R. Choo, Editors. 2019, Springer. p. 96-106.
133. Koneru, A., et al., *FBP recommendation system through sentiment analysis.* International Journal of Advanced Science and Technology, 2019. **29**(5): p. 896-907.
134. Li, J., et al. *The Establishment of Commodity Success Model Based on Text Sentiment Analysis*. in *6th International Conference on Big Data and Information Analytics, BigDIA 2020*. 2020. Institute of Electrical and Electronics Engineers Inc.
135. Mitra, S. and M. Jenamani, *OBIM: A computational model to estimate brand image from online consumer review.* Journal of Business Research, 2020. **114**: p. 213-226.
136. Yang, J., et al. *Optimization analysis of online marketing strategy based on the regression prediction model*. in *2020 2nd International Conference on Advances in Civil Engineering, Energy Resources and Environment Engineering, ACCESE 2020*. 2020. Institute of Physics Publishing.
137. Konjengbam, A., N. Kumar, and M. Singh, *Unsupervised tag recommendation for popular and cold products.* Journal of Intelligent Information Systems, 2020. **54**(3): p. 545-566.
138. Benlahbib, A. and E.H. Nfaoui. *AmazonRep: A Reputation System to Support Amazon's Customers Purchase Decision Making Process based on Mining Product Reviews*. in *4th International Conference on Intelligent Computing in Data Sciences, ICDS 2020*. 2020. Institute of Electrical and Electronics Engineers Inc.
139. Benlahbib, A. and E.H. Nfaoui, *Aggregating customer review attributes for online reputation generation.* IEEE Access, 2020. **8**: p. 96550-96564.
140. El-Dehaibi, N., T. Liao, and E.F. MacDonald. *Validating perceived sustainable design features using a novel collage approach*. in *33rd International Conference on Design Theory and Methodology, DTM 2021, Held as Part of the ASME 2021 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, IDETC-CIE 2021*. 2021. American Society of Mechanical Engineers (ASME).
141. Liu, J. *Data Wealth Mining Based on Text and Time*. in *2020 7th International Conference on Computer-Aided Design, Manufacturing, Modeling and Simulation, CDMMS 2020*. 2021. IOP Publishing Ltd.
142. Nellutla, A.P., et al., *Online product reviews and their impact on third party sellers using natural language processing.* International Journal of Business Intelligence Research, 2021. **12**(1): p. 26-47.
143. Boumhidi, A., A. Benlahbib, and E.H. Nfaoui, *Mining online opinions and reviews using bi-lstm for reputation generation*, in *Lecture Notes in Networks and Systems*. 2021, Springer. p. 135-144.
144. Christodoulou, E., et al. *Evaluating the Effect of Weather on Tourist Revisit Intention using Natural Language Processing and Classification Techniques*. in *2021 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2021*. 2021. Institute of Electrical and Electronics Engineers Inc.
145. Boumhidi, A. and E.H. Nfaoui, *Leveraging Lexicon-Based and Sentiment Analysis Techniques for Online Reputation Generation.* International Journal of Intelligent Engineering and Systems, 2021. **14**(6): p. 274-289.
146. Asseo, K. and M.Y. Niv, *Harnessing Food Product Reviews for Personalizing Sweetness Levels.* Foods, 2022. **11**(13).
147. Lutz, B., N. Pröllochs, and D. Neumann, *Are longer reviews always more helpful? Disentangling the interplay between review length and line of argumentation.* Journal of Business Research, 2022. **144**: p. 888-901.
148. Honig, S., et al., *Using Online Customer Reviews to Classify, Predict, and Learn About Domestic Robot Failures.* International Journal of Social Robotics, 2022.
149. Sánchez, F.T. *Comparison of Memorable Tourist Experiences Based on Collections of Reviews from Trip Advisor: Acropolis of Athens and Royal Alcazar of Seville*. in *8th International Conference of the International Association of Cultural and Digital Tourism, IACuDiT 2021*. 2022. Springer Science and Business Media B.V.
150. Arora, L., P. Srivastava, and S. Ananda Kumar. *Evaluation of Product Reviews Using Deep Learning Classifier Models*. in *13th International Conference on Computing Communication and Networking Technologies, ICCCNT 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
151. Tang, J. and J. Guo. *An Analysis Method for Online Shopping Platform Comments Based on NLP-AHP: Taking Amazon as An Example*. in *3rd International Conference on Education, Knowledge and Information Management, ICEKIM 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.
152. Hossain, M.S., N. Nayla, and A.A. Rassel. *PRODUCT MARKET DEMAND ANALYSIS USING NLP IN BANGLISH TEXT WITH SENTIMENT ANALYSIS AND NAMED ENTITY RECOGNITION*. in *56th Annual Conference on Information Sciences and Systems, CISS 2022*. 2022. Institute of Electrical and Electronics Engineers Inc.

153. Ravichandran, T. and C. Deng, *Effects of Managerial Response to Negative Reviews on Future Review Valence and Complaints.* Information Systems Research, 2023. **34**(1): p. 319-341.
154. Mantilla-Saltos, G., et al. *Assessing User and Manufacturer Perceptions of Fitness Trackers through Amazon Review Analysis*. in *18th Iberian Conference on Information Systems and Technologies, CISTI 2023*. 2023. IEEE Computer Society.
155. Christodoulou, E. and A. Gregoriades, *Leveraging Natural Language Processing in Persuasive Marketing*, in *Proceedings of the 15th Asian Conference on Intelligent Information and Database Systems, ACIIDS 2023*, N.T. Nguyen, et al., Editors. 2023, Springer Science and Business Media Deutschland GmbH. p. 197-209.
156. Ullah, A., et al., *Understanding Quality of Products from Customers' Attitude Using Advanced Machine Learning Methods.* Computers, 2023. **12**(3).
157. Harth, P., et al., *From Data to Human-Readable Requirements: Advancing Requirements Elicitation through Language-Transformer-Enhanced Opportunity Mining.* Algorithms, 2023. **16**(9).
158. Vollero, A., D. Sardanelli, and A. Siano, *Exploring the role of the Amazon effect on customer expectations: An analysis of user-generated content in consumer electronics retailing.* Journal of Consumer Behaviour, 2023. **22**(5): p. 1062-1073.
159. Al-Ghuribi, S., S.A.M. Noah, and M. Mohammed, *An experimental study on the performance of collaborative filtering based on user reviews for large-scale datasets.* PeerJ Computer Science, 2023. **9**.
160. Li, Z., et al., *Improving rating predictions with time-varying attention and dual-optimizer.* Applied Intelligence, 2023. **53**(21): p. 26098-26109.
161. Harsha, K., et al. *Empirical evaluation of Amazon fine food reviews using Text Mining*. in *8th IEEE International Conference for Convergence in Technology, I2CT 2023*. 2023. Institute of Electrical and Electronics Engineers Inc.
162. Araque, O., et al., *Enhancing deep learning sentiment analysis with ensemble techniques in social applications.* Expert Systems with Applications, 2017. **77**: p. 236-246.
163. Linardatos, P., V. Papastefanopoulos, and S. Kotsiantis, *Explainable ai: A review of machine learning interpretability methods.* Entropy, 2020. **23**(1): p. 18.
164. Lai, Z., et al. *Empowering Unsupervised Domain Adaptation With Large-Scale Pre-Trained Vision-Language Models*. in *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*. 2024.
165. Qiao, Z., *Fifth-generation edge computing-oriented speech recognition system applied in Japanese education and social sentiment classification.* Soft Computing, 2023: p. 1-16.
166. Rahman, A., et al., *Federated learning-based AI approaches in smart healthcare: concepts, taxonomies, challenges and open issues.* Cluster computing, 2023. **26**(4): p. 2271-2311.