No	Paper name	Dataset	Findings	Architecture	Accuracy(%)	Precision(%)	Recall(%)	F1-Score (%)	Loss (%)	Research Gap	Classification Class
1	Product Sentiment Analysis For Amazon Reviews	Amazon Mobile phone Dataset	The study states, in order to deal with ambiguity and enhance opinion categorization, classifying three-star reviews on Amazon mobile phones as neutral sentiments when conducting sentiment analysis.	Logistic Regression	85.3	86.3	85.3	85.7	59.6	The relevance of its findings to different product or platform categories. The 2016 data collection date of the dataset means it may not reflect current trends and preferences. Model performance may be impacted by unbalanced classes and data quality issues. The research lacks an in-depth comparison investigation of various methodologies and does not thoroughly address model interpretability. There is a lack of comparison investigation between more sophisticated models like BERT and conventional sentiment categorization techniques. Filling up this knowledge gap would enable us to have a deeper understanding of how this method influences sentiment analysis accuracy and its applicability in real-world situations.	Positive, Negative, Neutral
				Naïve Bayes	78.4	78.5	78.4	77.9	115.9		
				Random Forest	89.9	90.1	89.9	89.7	40.3		
				Bi-LSTM	93.3	92.9	93.3	92.9	23.4		
2	Comparative Study of Sentiment Analysis with Product Reviews Using Machine Learning and Lexicon-Based Approaches	Amazon Product Reviews	By providing a thorough study of sentiment classification techniques, the area of natural language processing and sentiment analysis is advanced, enabling researchers and professionals select the best way to evaluate sentiment in text data.	Pattern Lexicon	69	88	72	79	-	The dataset is mostly made up of Amazon product reviews, which restricts the generalizability of results to other textual fields. Limitations that may effect model performance include unbalanced data, reliance on predetermined sentiment lexicons, and the use of TF-IDF for feature extraction. The absence of hyperparameter adjustment, the exclusion of emoji, and non-textual data are further restrictions. To get a deeper knowledge of sentiment analysis, future studies should investigate more varied datasets, cutting-edge vectorization techniques, and overcome these constraints.	Positive (4,5 star), Negative (1,2,3 star)
				VADER Lexicon	83	90	89	89	-		
				SentiWordNet Lexicon	80	88	88	88	-		
				Gradient Boosting	87	88	98	92	-		
				Logistic Regression	90	91	97	94	-		
3	Context-based sentiment analysis on customer reviews using machine learning linear models	Twitter Dataset	The challenges of accurately predicting sentiments from large volumes of customer reviews while mitigating issues like incoherence in polarity, model overfitting, and high data processing costs.	LSVM	90.11	90	90	-	-	Lack of diversity in datasets, little description of the demands placed on the model's resources, insufficient investigation of the effects of preprocessing, and need for in-depth hyperparameter adjustment. It might also benefit from taking ethical issues, model interpretability, real-time deployment difficulties, domain adaptability, and temporal sentiment dynamics into account. The study's usefulness and significance in the field of sentiment analysis would be further increased by benchmarking against cutting-edge algorithms and looking at user demographics.	itive Sentiment Negative Sentiment
				fastText	90.71	91	89.6	-	-		
				SA-BLSTM	77	79	74	-	-		
4	Sentiment Classification based on Machine Learning Approaches in Amazon Product	Amazon Reviews dataset	To analyze customer reviews of online stores using machine learning, emphasizing Amazon. by classifying reviews as favorable or unfavorable to aid companies in improving their goods and services in response to client input.	Logistic Regression	94	-	-	-	-	The dataset's size and concentration may constrain the findings' applicability to other product kinds in a particular product category. The research also neglects other essential performance criteria, including precision, recall, and F1-score, in favor of focusing solely on accuracy. Potential drawbacks include an unbalanced class distribution and a disregard for reviews with neutral sentiment. Future research might investigate more extensive and varied datasets, use sophisticated feature extraction techniques, and address bias and fairness issues in sentiment analysis algorithms to increase the study's application.	Positive, Negative
				Decision Tree	99	-	-	-	-		
5	Sentiment Analysis on Product Reviews Using Machine Learning Techniques	Amazon Camera, Laptops, Mobile phones, tablets, TVs, video surveillance Dataset	Analyze various Amazon product reviews for cameras, laptops, mobile phones, tablets, TVs, and video surveillance. Use practical sentiment analysis and opinion mining tools—machine learning techniques to categorize these reviews as favorable or unfavorable. Improve firms' decision-making processes, give them valuable insights from consumer feedback, and increase customer happiness.	Naïve Bayes(Camera)	98.17	98.3	-	99.03	-	It only uses Amazon reviews for particular product categories, which may introduce bias. The data pretreatment methods are briefly described, and the decision to use a lexicon-based approach may need to be made to capture nuanced feelings. Other evaluation metrics should be present, even when accurate data are supplied. The possibility for aspect-level analysis is only briefly mentioned in the text; it must be substantially investigated. There needs to be more debate about the potential overfitting, subjectivity, and generalization to different settings inherent in sentiment analysis, as well as the neglect of temporal analysis and generalization. It would be more complete and credible if these shortcomings were acknowledged and addressed.	Positive, Negative
				Naïve Bayes(Laptops)	90.22	90.01	-	94.74	-		
				Naïve Bayes(Mobile Phones)	92.85	91.64	-	95.64	-		
				Naïve Bayes(Tablets)	97.17	98.73	-	98.31	-		
				Naïve Bayes(TVs)	90.16	90.17	-	94.72	-		
				Naïve Bayes(Video Surveillance)	91.13	89.95	-	94.71	-		
				SVM(Camera)	93.54	93.58	-	96.66	-		
				SVM(Laptops)	88.16	88.52	-	93.71	-		
				SVM(Mobile Phones)	92.85	91.64	-	95.64	-		
				SVM(Tablets)	84.12	84.31	-	91.37	-		
				SVM(TVs)	88.49	85.56	-	93.89	-		
				SVM(Video Surveillance)	79.43	84.25	-	88.53	-		
6	Customer sentiment analysis with more sensibility	Korean motor company in 2017	We are examining the feedback left by customers in online product reviews. Existing techniques that rely on general sentiment dictionaries need help to evaluate and capture sentiments expressed in non-standard languages. To better understand customer dissatisfaction, improve sentiment analysis accuracy, incorporate contextually comparable terms, and identify the main complaint subjects.	ANN(Automobile)	72.4 AUC				-	The method can only be applied to specific languages because of its language dependence. The vehicle reviews' industry-specific focus might need to be revised in other sectors. Furthermore, the sentiment lexicon's completeness and associated biases are not considered. The use of data from particular online communities could lead to selection bias, and privacy issues about data sources are noted but have yet to be further explained. The application of semi-supervised learning is not explicitly described, and the generalizability of the results beyond the dataset and the particular car models utilized is still being determined. Underemphasized are parameter sensitivity and practical applications. Although the research is encouraging, these constraints should be considered when evaluating its findings and their relevance to more general circumstances.	"Positive, Negative"
				ANN(Movie)	68.3 AUC				-		
				ANN(Game)	72.9 AUC				-		
				SVM(Automobile)	97.4 AUC				-		
				SVM(Movie)	90.5 AUC				-		
				SVM(Game)	91.2 AUC				-		
				GSSL(Automobile)	98.1 AUC				-		
				GSSL(Movie)	93.4 AUC				-		
				GSSL(Game)	97.4 AUC				-		
7	Sentilyzer: Aspect-Oriented Sentiment Analysis of Product Reviews	Online Retailer Amazon	Sentilyzer is a method for aspect-oriented sentiment analysis that uses customer evaluations on Amazon to create sentiment dictionaries and aspect lexicons automatically. The evaluation's findings show that both English and German evaluations may be classified highly in terms of sentiment. The system's capacity to automate the construction of sentiment dictionaries is an impressive accomplishment, and its domain-specific lexicon pairs give it flexibility in answering the needs of sentiment analysis for various products. The research's findings do, however, also emphasize the significance of carrying out more thorough assessments to guarantee the system's scalability and generalizability to various datasets and languages.	Sentilyzer (English)	80	97.37	78.72	87.06	-	It uses more extensive and varied sets of product reviews to assess Sentilyzer's performance, particularly when combined with machine learning techniques like BERT and NLP. Sentilyzer has potential for sentiment analysis, but to assure its flexibility and scalability to various data sources and languages, it has to be thoroughly evaluated with a machine learning framework like BERT. To comprehend Sentilyzer's competitive advantage and room for progress, it is crucial to compare its performance against cutting-edge NLP methods. Further study should concentrate on user experience, real-world applications, and how Sentilyzer interfaces with sophisticated machine learning algorithms for improved sentiment analysis. Closing these gaps may result in sentiment analysis methods that are more reliable and adaptable.	Positive, Negative
				Sentilyzer (Germany)	80.43	97.14	80.95	88.31	-		

8	Sentiment Analysis of Customer Product Reviews Using Machine Learning	Amazon Mobile phone Dataset	Use sentiment analysis to categorize many online product evaluations' positive and negative aspects. The difficulty lies in effectively assessing this data, given the impact of online reviews on customer decisions.	Naïve Bayes	66.95	-	-	-	-	It ignores the complex sentiment observed in real-world reviews and concentrates on binary sentiment classification (positive and negative). Dataset bias and inherent imbalances in the data could constrain the generalizability of the results. The choice of particular machine learning models needs a thorough reason, and the evaluation is based solely on accuracy without considering other pertinent metrics. The study also ignores human-centric elements like sarcasm and the time-sensitivity of sentiment in reviews. Additionally, there needs to be more discussion of the scalability or moral ramifications of using these models in practical applications. It would be easier to comprehend the research's extent and potential future work directions if these limitations were acknowledged.	Positive, Negative
				SVM	81.77	-	-	-	-		
				Decision Tree	74.75	-	-	-	-		
9	Using Machine Learning to Predict the Sentiment of Online Reviews: A New Framework for Comparative Analysis	Yelp Dataset Challenge Round 9 in 2017	The three best single classifiers for online review sentiment analysis are Linear Kernel Support Vector Machine, Logistic Regression, and Multilayer Perceptron. It illustrates the significance of neutral sentiment detection and the efficiency of text pre-processing, TF-IDF feature extraction, and sentiment lexicons in delivering consistent outcomes across various datasets.	LSVM	89.29	89.13	89.26	89.14	-	It is still challenging to categorize neutral assessments, particularly when those who provide 3-star ratings frequently have a favorable bias. Suggests that bio-inspired optimization algorithms and better feature selection methods are required to improve the performance of machine learning models, particularly in reliably categorizing neutral sentiment.	"Positive (3,4,5 star), Negative (1,2 star)"
				LR	90.17	89.94	90.17	90	-		
				MLP	90.38	90.33	90.38	90.35	-		
10	Sentiment Analysis: Predicting Product Reviews' Ratings using Online Customer Reviews	product reviews of Cell phones and Accessories (Amazon.com)	Establishes a strong link between the opinions expressed in text reviews and the star ratings for Amazon products. Polarity and review length are important factors in machine learning models like Logistic Regression's ability to predict ratings using information from text. It is simpler to anticipate extreme ratings (1-star and 5-star) than neutral ones (3-star).	Multinomial Naive Bayes	49.2	-	-	-	-	It depends on a particular Amazon dataset and corrects for class imbalances, which may limit generalizability. It only emphasizes a few features, doesn't prove causation, and doesn't delve into the subtleties of sentiment analysis. The peculiarity of the dataset and temporal factors must be considered when evaluating the results.	"Positive(4,5 star), Neutral(3 star) and negative(1,2 star)"
				Logistic Regression	54.1	-	-	-	-		
				Linear SVC	51.1	-	-	-	-		
11	Sentiment Analysis of Consumer Reviews Using Deep Learning	Amazon-Fine-Food-Reviews, Cell Phones and Accessories, Amazon-Products, Yi-Jen Mon,Yelp	The creation and assessment of models for consumer feedback and product reviews that use deep learning to analyze sentiment. To attain high accuracy, precision, recall, and F1-score in sentiment classification, the focus is on improving model architecture, feature encoding techniques, and data pretreatment approaches. Give businesses the resources they need to analyze user-generated material for valuable insights and decide based on data in light of customer feedback and online reviews.	LSTM(Dataset-Amazon-Fine-Food-Reviews)	87	78	55	47	-	Reliance on benchmark datasets, potential overfitting from sparse data, and feature engineering decisions that impact model performance are just a few examples. Scalability, model interpretability, and real-time analysis still need to be addressed. Domain-specific difficulties, various measurements, and hyperparameter tuning should all be considered in future studies. Despite these limitations, this study provides insightful tips for sentiment analysis, with room for further development.	Positive Instances, Negative Instances
				LSTM(Dataset-Cell Phones and Accessories)	87	99	70	78	-		
				LSTM(Dtase-Amazon-Products)	97	70	62	64	-		
				LSTM(Dtase-IMDB)	75	76	75	75	-		
				LSTM(Dtaset-Yelp)	87	79	59	61	-		
12	Deep learning sentiment analysis of Amazon.com reviews and rating.	Product reviews(Amazon.com)	To address the problem of inconsistent user evaluations and ratings on sites like Amazon.com. A deep learning model must be developed to analyze review sentiment successfully. The system offers user comments when there are differences between the sentiment and rating to enhance review consistency and dependability.	SVM	81.29	58.61	40.85		-	Its reliance on machine learning models, potential problems with the quality of the training data, and the presumption that review-rating discrepancies are purely the result of user mistakes. Additionally, it does not consider other elements that may contribute to differences, such as product modifications or fake reviews, and it may perform differently depending on the product category or language.	"Positive(4,5 star), Neutral(3 star) and negative(1,2 star)"
				RNN	81.82	59.45	42.52		-		