International Journal of Advancement in Electronics and Computer Engineering (IJAECE)
Volume 14, Issue 1, January. 2025, pp. 106-112 ISSN 2278 -1412

Advanced Intelligent System For Website Review Analysis

Syed Fazal Ur Rahman[1], Jeetendra Singh Yadav[2]
[1] Research Scholar, fazalurrahman.fazal@gmail.com, CSE BERI, Bhopal, India
[2,]Assis. Prof., jeetendra2201@gmail.com, CSE BERI, Bhopal, India

---

Abstract *– The rise of digital marketplaces has transformed consumer decision-making, with online reviews playing a crucial role in shaping trust and influencing purchasing behavior. This paper examines sentiment analysis in e-commerce platforms, focusing on how review dynamics impact consumer perceptions. Our findings indicate that early reviewers tend to assign higher ratings, while previous reviewers often provide more lenient feedback, highlighting distinct rating trends over time.*

To systematically analyze consumer sentiments, we implemented sentiment classification at both the sentence and comment levels, effectively categorizing reviews into positive and negative sentiments. This approach enhances the reliability of feedback processing, enabling better organization of large volumes of user-generated content.

The results validate the effectiveness of our sentiment classification framework, demonstrating its role in improving the credibility and trustworthiness of online reviews. By addressing the challenges associated with review analysis, our research contributes to enhancing consumer trust and informed decision-making in digital marketplaces. Future work will explore advanced deep learning models to further improve sentiment prediction accuracy and capture nuanced emotions in textual feedback.

Keywords: Sentiment Analysis, Digital Marketplaces, Consumer Trust, Online Reviews, Review Dynamics, Sentiment Classification, Early Reviewers, Machine Learning, Deep Learning, Opinion Mining, Text Classification, Consumer Behavior, Trustworthiness, E-commerce Analytics.

I. INTRODUCTION

The rapid growth of e-commerce platforms has revolutionized consumer purchasing behaviors, allowing users to share their experiences through product reviews. These reviews, which contain valuable insights and opinions, play a crucial role in shaping consumer trust and influencing purchasing decisions. Among these, early reviews, posted during the initial phase of a product's launch, hold significant weight in determining the product's success. Early reviewers, despite being a small fraction of the total review base, exert considerable influence over subsequent consumer decisions. Their feedback is often perceived as credible and sets the tone for later evaluations, making them a focal point for businesses aiming to refine marketing strategies and product designs.

Given the reliance on online reviews, businesses and consumers face challenges in processing and extracting meaningful insights from the overwhelming volume of user-generated content. Traditional methods of sentiment analysis have been employed to classify reviews into positive and negative sentiments. However, existing dictionary-based approaches require extensive manual effort in constructing emotional lexicons and often fail to capture implicit opinions embedded in neutral or factual statements. Similarly, early machine learning-based sentiment classifiers, such as Naïve Bayes and Support Vector Machines, depend heavily on feature engineering, making them domain-sensitive and less adaptable to diverse datasets. To overcome these limitations, deep learning techniques have emerged as a powerful alternative for sentiment classification. Deep neural networks (DNNs) eliminate the need for manual feature engineering by learning high-level data representations, allowing more accurate sentiment predictions. Recent advancements in Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks have demonstrated superior performance in understanding context and emotion polarity in text data. These methods enable a more refined sentiment classification process, effectively handling complex user reviews and detecting subtle sentiment variations.

In this study, we propose an advanced deep learning-based framework for sentiment analysis and early reviewer prediction. Unlike conventional approaches, our model treats review ratings as weak labels, enabling it to learn meaningful embeddings that reflect the underlying sentiment distribution. By leveraging deep embedding techniques, we aim to mitigate the noise introduced by misleading sentiment labels and improve classification accuracy. Our methodology consists of two key phases: (1) constructing an embedded space where similar sentiment-based reviews are grouped, and (2) employing a ranking-based loss function to optimize the sentiment classification model.

Additionally, our study focuses on identifying the behavioral characteristics of early reviewers in large-scale e-commerce datasets. By analyzing review posting patterns, sentiment trends, and crowd behavior effects, we aim to develop a predictive model capable of ranking and classifying early reviewers based on their impact on future product evaluations. This research not only contributes to improving sentiment classification in online reviews but also provides businesses with a robust framework for leveraging early reviewer insights in product marketing and consumer engagement strategies.

With the increasing dependency on online reviews in digital commerce, this study presents a novel deep learning-based sentiment analysis approach that enhances predictive accuracy, reduces information overload, and facilitates a more intelligent review management system. Future advancements in this domain may integrate reinforcement learning techniques for adaptive sentiment analysis and further improve the interpretability of neural network-based models in real-world applications..

II. Background

The exponential growth of e-commerce and online marketplaces has transformed the way consumers evaluate and purchase products. With the shift from traditional brick-and-mortar shopping to digital platforms, user-generated content, particularly product reviews, has become a fundamental source of information for potential buyers. These reviews provide insights into product quality, functionality, and user satisfaction, influencing purchasing decisions and brand reputation. Given the significant impact of reviews, businesses have recognized the importance of understanding consumer sentiments to refine marketing strategies and enhance product development.

Among all reviews, early reviews, which are posted shortly after a product launch, play a crucial role in setting the perception of a product. Early reviewers often establish a product's initial credibility, thereby influencing later reviews and subsequent consumer behavior. This phenomenon aligns with herd behavior theory, where people tend to follow the opinions and actions of others, especially when making uncertain decisions. As a result, businesses actively seek to analyze early reviews to predict market responses and adapt their offerings accordingly. Platforms like Amazon have introduced initiatives such as the Early Reviewer Program to encourage initial feedback and shape consumer trust in new products.

However, managing the vast volume of user-generated reviews presents a significant challenge. Traditional sentiment analysis methods, such as dictionary-based approaches, require extensive manual effort to construct lexicons of opinion words. While effective to some extent, these approaches struggle to interpret implicit sentiments—statements where opinions are expressed indirectly. For example, a review saying, "I bought this phone last week, and the battery drains quickly," may indicate dissatisfaction, but dictionary-based methods may not classify it as negative sentiment.

Early machine learning (ML)-based sentiment classification techniques, including Naïve Bayes and Support Vector Machines (SVMs), sought to improve accuracy by using predefined features such as n-grams, part-of-speech tagging, and syntactic relationships. While these models improved text classification, they required significant feature engineering, making them less adaptable across different datasets and domains. The complexity of user reviews—ranging from short, vague statements to lengthy, detailed feedback—further challenges traditional sentiment analysis approaches.

The emergence of deep learning models, particularly Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, has revolutionized sentiment analysis. These models automatically learn representations from data, eliminating the need for manual feature engineering. Deep learning models have proven effective in capturing contextual meaning, sentiment polarity, and subtle opinion variations within textual data. However, deep learning also presents challenges, such as the need for large-scale training data and computational power.

A significant issue in sentiment classification is the misalignment between review ratings and textual sentiment. A five-star rating may still contain negative comments, while a one-star review may include positive aspects. Relying solely on binarized ratings as sentiment labels can mislead machine learning models. Recent studies propose embedding-based approaches, where models learn from sentence representations rather than explicit labels, allowing for better sentiment differentiation.

In this research, we aim to leverage deep learning techniques, particularly Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTM) networks, for sentiment classification and early reviewer analysis. Our approach treats review ratings as weak sentiment labels, optimizing classification accuracy through ranking-based learning. Furthermore, we focus on predicting early reviewers' behavior by analyzing their linguistic patterns, sentiment trends, and influence on future product evaluations.

The findings from this study will help e-commerce platforms and businesses improve product review management, sentiment prediction, and consumer behavior modeling. The adoption of intelligent sentiment analysis frameworks will enhance the reliability of online reviews, reduce misinformation, and provide consumers with more credible product insights. Future extensions of this work may explore reinforcement learning-based adaptive sentiment analysis and the integration of blockchain technology for trust verification in online reviews..

III. IMPLEMENTATION

To predict the first reviewer, we propose a novel approach by viewing the review process as a multiplayer competitive game. In this context, only the most competitive user can be considered the first reviewer for a given product. The competition process can be further simplified into pairwise comparisons between two players. In this two-tier competition, the winner outpaces the loser earlier in the timestamp. Inspired by recent advancements in learning-based dispersion representation, we first map both users and products into a shared integrated space. Then, we determine the ranking order of users who have reviewed a product based on their respective distances in this space. Our approach involves using an embedded margin-based model to represent products and predict the first reviewer.

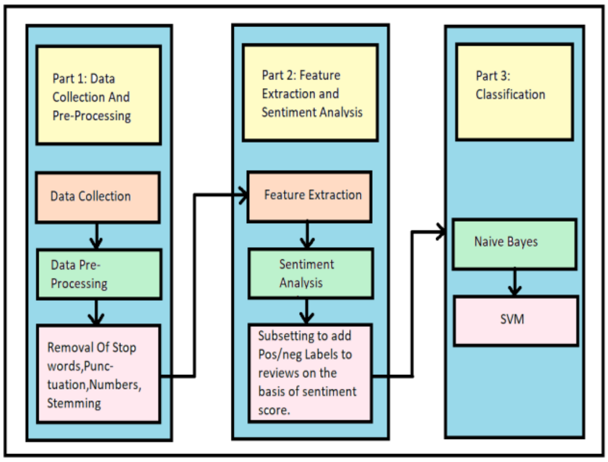


Figure 1: Proposed Framework

• Upload Products

Product uploads are overseen by the admin, with authorized personnel responsible for adding new products to the system, making them accessible to users. Each product can be uploaded along with key details such as brand, color, warranty information, and other relevant attributes. Users have the option to block or unblock products according to their preferences.

PRODUCT REVIEW BASED ORDER

The product suggestions displayed to users are based on their reviews and ratings for specific items. In this project, the Naïve Bayes algorithm is used to determine the sentiment of a given review, classifying it as either positive or negative. Based on the output of this algorithm, personalized product suggestions are then provided to users. The algorithm is applied and lists the products in user side based on the positive and negative.

•Ratings And Reviews

Ratings and reviews are main concept of the project in order to find effective product marketing. The main aim of the project is to get the user reviews based on how they purchased or whether they purchased or not. The major find out of the project is when they give the ratings and how effective it is. And this will helpful for the users who are willing to buy the same kind of product.

Data Analysis

The main part of the project is to analysis the ratings and reviews that are given by the user. The products can be analysis based on the numbers which are given by user. The user data analysis of the data can be done by charts format. The graphs may vary like pie chart, bar chart or some other charts.

• Preprocessing

In this calculation, the reviews which are unfamiliar made to information base from the reviewer API, these reviews involve trivial words, whitespaces, hyperlinks and remarkable characters. First, we need to isolate interactions by removing all unnecessary words, whitespace, hyperlinks, and special characters. This step ensures that the data is clean and ready for further processing.

The preprocessing steps mean to start the element extraction interaction and begin separating sacks of words from the examples. One of the primary center is to decrease the last measure of highlights removed. Undoubtedly, highlights decrease is significant to improve the precision of the forecast for both point displaying and conclusion investigation. Highlights are utilized to address the examples, and the more the calculation will be prepared for a particular element the more exact the outcomes will be. Consequently, if two highlights are comparable it is advantageous to consolidate them as one exceptional component. Also, if a component isn't important for the investigation, it tends to be taken out from the pack of words.

• Lower capitalized letters: The initial phase in the preprocessing is to go through all the information and change each capitalized letter to their relating lowercase letter. When handling a word, the examination will be case delicate and the program will consider "information" and "Information" as two entirely unexpected words. It is significant that, these two words are considered as similar highlights. Something else, the calculations will influence notions which may vary to these two words. For instance, on these three sentences: "information are acceptable", "Marvelous information", and "Awful Data". The first and second sentences both contain "information" and are positive, the third sentence contains "Information" and is negative. The calculation will figure that sentences containing "information" are

International Journal of Advancement in Electronics and Computer Engineering (IJAECE)
Volume 14, Issue 1, January. 2025, pp. 106-112 ISSN 2278 -1412

bound to be positive and those containing "Information" negative. In the event that the uppercases had been eliminated the calculation would have had the option to figure that the way that the sentence contains "information" isn't extremely pertinent to identify whether or the sentence is positive. This preprocessing step is much more significant since the information are recovered from Reviewer. Online media clients are regularly writing in capitalized regardless of whether it isn't needed, hence this

preprocessing step will betterly affect web-based media information than other

"traditional" information.

• Remove URLs and client references: Reviewer permits client to incorporate hashtags, client references and URLs in their messages. By and large, client references and URLs are not applicable for investigating the substance of a book. Hence, this preprocessing step depends on normal articulation to discover and supplant each url by "URL" and client reference by "AT_USER", this permits to decrease the aggregate sum of highlights extricated from the corpus [2]. The hashtags are not eliminated since they frequently contain a word which is important for the investigation, and the "#" characters will be taken out during the tokenization interaction. Remove digits: Digits are not significant for breaking down the information, so they can be taken out from the sentences. Besides, now and again digits will be blended in with words, eliminating them may permit to relate two highlights which may have been considered diverse by the calculation in any case. For instance, some information may contain "iphone", when other will contain "iphone7". The tokenization cycle, which will be presented later.

• Remove stop words: In normal language preparing, stop words are regularly taken out from the example. These stop words will be words which are normally utilized in a language, and are not applicable for a few regular language preparing techniques, for example, theme demonstrating and supposition investigation [10]. Eliminating these words permits to decrease the measure of highlights extricated from the examples.

Self-Learning and word standardization System

In this calculation, first we need to instate the word reference (first accentuation dictionary).In the vocabulary generally we need to present the positive, negative fair and things. Each and every enormous datum and data mining adventures considering the pre-arranged data, without arranged data (presentation of words).So instatement of the pre-arranged data is indispensable. In oneself learning structure, we are doing word standardization, here we are not considering past, present and future status of the words, just we are pondering the word.

•Sentiment Analysis

In this calculation, pre-processed reviews are brought from the data set individually. In any case we require check individually watchword whether that expression is thing are not, if thing we will oust it from the particular audit. After that the remainder of the watchwords checked with evaluation create, whether or not those expressions are sure assessment or adverse end or unbiased inclination. The remainder of the watchwords in the review which doesn't has a spot with any of the assumption will be consigned fleeting end considering the more check of positive, negative and fair. In the subsequent cycle if the reaming word crosses the restriction of positive, negative or impartial, that watchword everlastingly included as improvement in the vocabulary.

Algorithm Step In Sentiment Analysis Step1:Get_some_sentiment_examples

As for every supervised learning problem, the algorithm needs to be trained from

labeled examples in order to generalize to new data.

.Step2:Extract_features_from_examples

Transform each example into a feature vector. The simplest way to do it is to have a vector where each dimension represents the frequency of a given word in the document.

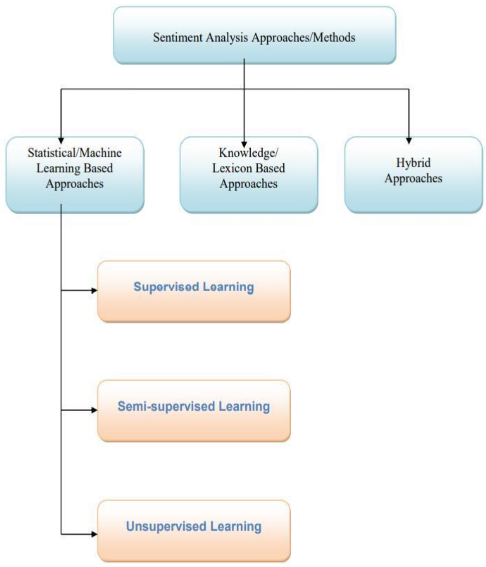


Figure 2: Sentiment Analysis Approaches/Method

Step3:Train_the_parameters

This is where your model will learn from the data. There are multiple ways of using features to generate an output, but one of the simplest algorithms is logistic regression. Other well-known algorithms are Naive Bayes. In the simplest form, each feature will be associated with a weight. Let's say the word "love" has a weight equal to +4, "hate" is -10, "the" is 0 ... For a given example, the weights corresponding to the features will be summed, and it will be considered "positive" if the total is > 0, "negative" otherwise. Our model will then try to find the optimal set of weights to maximize the number of examples in our data that are predicted correctly. If you have more than 2 output classes, for example if you want to classify between "positive", "neutral" and "negative", each feature will have as many weights as there are classes, and the class with the highest weighted feature sum wins.

Step4:Test_the_model

After we have trained the parameters to fit the training data, we have to make sure our model generalizes to new data, because it's really easy to over fit. The general way of regularizing the model is to prevent parameters from having extreme values..

IV. SIMULATION RESULT

The Ratings and reviews are main concept of the project in order to find effective product marketing. The main aim of the project is to get the user reviews based on how they purchased or whether they purchased or not. The major find out of the project is when they give the ratings and how effective it is. And this will helpful for the users who are willing to buy the same kind of product.The main part of the project is to analysis the ratings and reviews that are given by the user. The products can be analysis based on the numbers which are given by user. The user data analysis of the data can be done by charts format. The graphs may vary like pie chart, bar chart or some other charts.

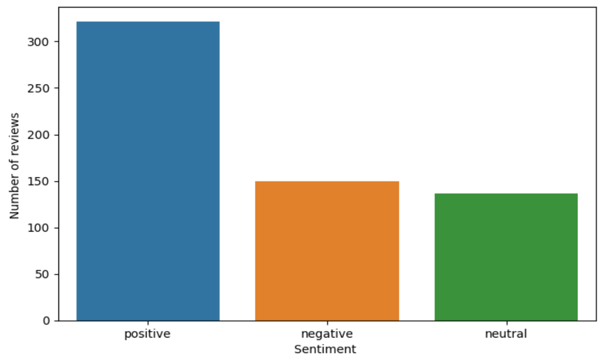


Figure 3: Polarity analysis

Figure 4: Accuracy of proposed Algorithm

Figure 4 shows the Polarity Analysis of the Reviews given by users. Out of total 607 reviews, 315 are positive, 152 are negative and 140 are neutral reviews. Next figure is the table that shows the accuracy of our classifier. SVM has performed very good with an accuracy of 87.6664%.

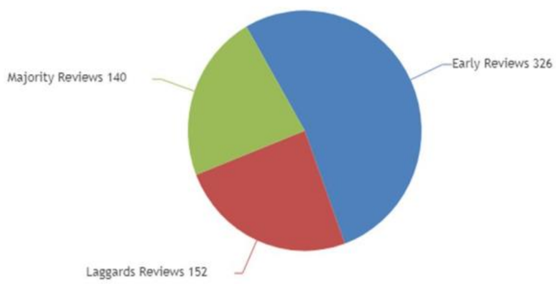
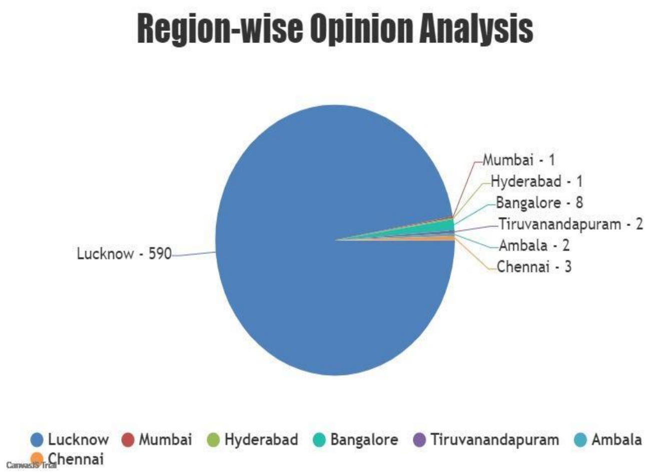


Figure 5: reviews classified into three types

Figure 5 represents three types of reviews which we have classified on the basis of product life time. Out of 607 reviews we can see that 326 are early reviews, 140 are majority reviews while the remaining 152 are laggard reviews.

The below table displays the average rating and average text letters per Review given by each category of reviews.

	Early	**Majority**	**Laggard**
Ave. Rating	3.97	3.56	3.42
Ave. Text per Review	69.2	51.9	32.4



International Journal of Advancement in Electronics and Computer Engineering (IJAECE)
Volume 14, Issue 1, January. 2025, pp. 106-112 ISSN 2278 -1412

Figure 6: Region-wise opinion analysis

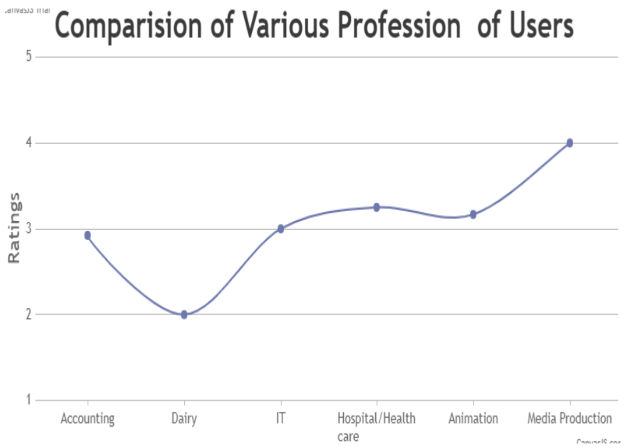


Figure 7: Comparison of various profession of users

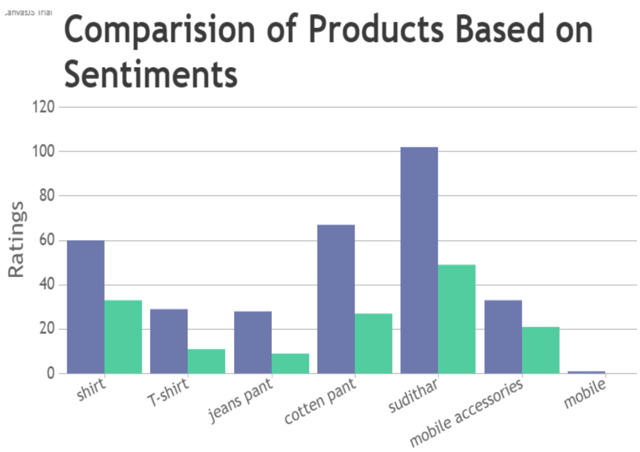


Figure 8: Comparison of product based on sentiments

Table 1: Accuracy Comparison of Existing and Proposed System

	Classifier	Number of Review	Accuracy (%)
1	NB (PROPOSED)	200	65%
2	SVM (PROPOSED)	200	87.6664
3	Random Forest	200	81%
4	RNN	200	85.42%
5	Linear Regression	200	78.60%
6	Cosine Similarity	200	73.23%

The above table shows the accuracy comparision of existing and Proposed system. In this table we can see that Support Vector Machine has clearly outperformed other techniques.

V. Conclusion

This paper presents examines sentiment analysis in digital marketplaces, highlighting how review dynamics influence consumer trust. Our findings show that early reviewers tend to give higher ratings, while previous reviewers often provide more lenient assessments. By implementing sentiment classification at both the sentence and comment levels, we effectively categorized opinions into positive and negative sentiments.

With the increasing reliance on online reviews, our research aids in managing large volumes of feedback, ensuring credible and trustworthy insights for consumers. The results demonstrate the effectiveness of our approach in sentiment classification, reinforcing its importance in shaping consumer behavior. Future work can explore advanced deep learning models to enhance accuracy and capture more nuanced sentiment patterns.

REFERENCES

[1] Najma Sultana , Pintu Kumar , Monika Rani Patra , Sourabh Chandra And S.K. Safikul Alam" Sentiment Analysis For Product Review" ICTACT Journal On Soft Computing, April 2019, Volume: 09, Issue: 03

[2] Alpna P. , Arvind K. T. , “Sentiment Analysis by using Recurrent Neural Network” proceedings of 2nd ICACSE, 2019

[3] Yoon-Joo Park "Envisioning the Helpfulness of Online Customer Reviews across Different Product Types" MDPI, Sustainability 2018, 10, 1735

[4] Sunil Saumya, Jyoti Prakash Singh, Nripendra P. Rana,Yogesh k. Dwivedi, Swansea University Bay Campus, Swansea "Situating Online Consumer Reviews" Article in Electronic Commerce Research and Applications, March 2018

[5] Liao, Shiyang, Junbo Wang, Ruiyun Yu, Koichi Sato, and Zixue Cheng, “CNN for situations understanding based on sentiment analysis of twitter data”, Procedia Computer Science, vol. 111, pp.376-381, 2017.

[6] Xing Fang* and Justin Zhan "Sentiment Analysis using product review data” Journal of Big Data , 2015

[7] D. Imamori and K. Tajima, "Foreseeing notoriety of twitter accounts through the revelation of link-propagating early adopters," in CoRR, 2015, p. 1512.

[8] Yeole V., P.V. Chavan, and M.C. Nikose, "Opinion mining for emotions determination", ICIIECS 2015-2015 IEEE Int. Conf. Information, Embed. Commun. Syst., 2015.

[9] F. Luo, C. Li, and Z. Cao, Affective-feature-based sentiment analysis using SVM classifier, 2016 IEEE 20th Int. Conf. Comput. Support. Coop. Work Des., pp.276281, 2016.

[10] Kalaivani A., Thenmozhi D, Sentiment Analysis using Deep Learning Techniques. , IJRTE,2018.

[11] Bingwei Liu, Erik Blasch, Yu Chen, Dan Shen, and Genshe Chen. Scalable sentiment classification for big data analysis using naïve bayes classifier. In big data, 2013 IEEE International Conference on, pages 99-104. IEEE,2013

[12] D. N. Devi, C. K. Kumar, and S. Prasad, "A feature based approach for sentiment analysis by using support vector machine," in Advanced Computing (IACC), 2016
IEEE 6thInterna-tional Conference on. IEEE, 2016, pp. 3–8.

[13] McAuley, J.; Leskovec, J. Hidden factors and hidden topics: Understanding rating dimensions with review text. In Proceedings of the 7th ACM Conference on Recommender Systems, RecSys', Hong Kong, China,12–16 October 2013; pp. 165–172.

[14] Ghose, A.; Ipeirotis, P.G. Estimating the helpfulness and economic impact of product reviews: Mining text and reviewer characteristics. IEEE Trans. Knowl. Data Eng. 2011, 23, 1498–1512.

[15] V. G. D. W. Shih-Lun Tseng, Shuya Lu, "The effect of herding behavior on online review voting participation," in AMCIS, 2017.

[16] S. M. Mudambi and D. Schuff, "What makes a helpful online review? a study of customer reviews on amazon.com," in MIS Quarterly, 2010, pp. 185–200.

[17] B. Pang and L. Lee, "Opinion mining and sentiment analysis," Foundations and Trends in Information Retrieval, vol. 2, no. 1-2, pp. 1–135, 2007.

[18] K. Dave, S. Lawrence, and D.M. Pennock. Mining the peanut gallery: Opinion extraction and semantic classification of product reviews. In Proceedings of the 12th international conference on World Wide Web, pages 519_528. ACM, 2003. ISBN
1581136803.

[19] J.C. Na, H. Sui, C. Khoo, S. Chan, and Y. Zhou. Effectiveness of simple linguistic processing in automatic sentiment classification of product reviews. AD- VANCES IN KNOWLEDGE ORGANIZATION, 9:49_54, 2004. ISSN 0938- 5495.

[20] Qiang Ye, Ziqiong Zhang, and Rob Law. Sentiment classification of online reviews to travel destinations by supervised machine learning approaches. Expert systems with applications, 36(3):6527–6535, 2019.

[21] M. M. Altawaier & S. Tiun "Comparison of Machine Learning Approaches on Arabic Twitter Sentiment Analysis". International Journal on Advanced Science, Engineering and Information Technology, Vol. 6 (2016) No. 6, pages: 1067-1073.

[22]K. Dave, S. Lawrence and D. M. Pennock. "Mining the peanut gallery: Opinion extraction and semantic classification of product reviews". In Proceedings of the 12th international conference on World Wide Web, pp. 519-528.

[23]A. Balahur, R. Mihalcea and A. Montoyo "Computational approaches to subjectivity and sentiment analysis": Present and envisaged methods and applications. Computer Speech & Language, 28(1): 1-6, 2014.

[24]E. Boiy and M. F. Moens "A machine learning approach to sentiment analysis in multilingual Web texts". Information retrieval, 12(5): 526- 558, 2009.

[25]M. Ghiassi, J. Skinner and D. Zimbra. "Twitter brand sentiment analysis: A hybrid system using n-gram analysis and dynamic artificial neural network". Expert Systems with applications, 40(16): 6266-6282, 2013.

[26] Tan LK-W, Na J-C, Theng Y-L, Chang K (2011) Sentence-level sentiment polarity classification using a linguistic approach. In: Digital Libraries: For Cultural Heritage, Knowledge Dissemination, and Future Creation. Springer, Heidelberg, Germany. pp 77–87

[27] Liu B (2012) Sentiment Analysis and Opinion Mining. Synthesis Lectures on
Human Language Technologies. Morgan & Claypool Publishers.

[28] Hu M, Liu B (2004) Mining and summarizing customer reviews. In: Proceedings of the tenth ACM SIGKDD international conference on Knowledge discovery and data mining. ACM, New York, NY, USA. pp 168–177