e-ISSN:2582-7219

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 7, Issue 4, April 2024

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.521

International Journal Of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704206 |

# Sentiment Analysis on Product Reviews Using Machine Learning Algorithms

**Yamunathangam D [1], Shanmugapriya R [2], Sriramkumar T [3], Subapriya B [4]**

Assistant Professor, Velalar College of Engineering and Technology, Erode, India[1]

UG Students, Velalar College of Engineering and Technology, Erode, India[2,3,4]

**ABSTRACT:** Sentiment analysis plays a vital role in understanding customer feedback and preferences, especially in the fast-paced environment of e-commerce platforms. This paper presents sentiment analysis using machine learning algorithm such as Keyword Based technique, Instance Based Learning and Support Vector Machine approaches for continually updating sentiment analysis model based on new product reviews. The suggested techniques seek to increase sentiment classification accuracy and flexibility by taking the account of customer sentiments over time. This paper provides mechanism to enhance sentiment categorization by upgrading product models with new online product reviews. Initially, the labelled dataset of reviews are collected and preprocess it into numerical feature vectors. The keyword-based technique, Instance Based Learning, and Support Vector Machine algorithms use these vectors for sentiment categorization (positive, negative, neutral). Further the new product reviews can be taken in the online and provided as an input to the system to optimize the sentiment classification.

## I. INTRODUCTION

Sentiment analysis, a fundamental aspect of natural language processing, has emerged as a critical tool for comprehending customer feedback and mining opinions in the huge terrain of online product evaluations. The increasing use of online shopping and social media means there's a lot more customer feedback for businesses to deal with. While this feedback can be helpful, it also presents challenges in accurately understanding what customers really think. Machine learning algorithms play an important role in automating this process, using techniques like text categorization and sentiment score to filter through massive volumes of textual data and extract relevant insights. In this context, sentiment research not only acts as a measure of consumer happiness, but it also influences critical business choices such as product enhancements and focused marketing initiatives. This study investigates the use of machine learning in sentiment analysis of product evaluations, emphasizing its importance in the current environment of consumer-driven businesses.

Sentiment analysis, product reviews, consumer feedback, and opinion mining are integral components of modern business strategies, driven by the influx of textual data from social media, online reviews, and customer interactions. Leveraging machine learning algorithms and language principles, organizations gain valuable insights into consumer sentiments, brand perception, and emerging trends. Product reviews, shared on e-commerce platforms and social media, play a pivotal role in shaping consumer decisions and guiding business improvements to enhance customer satisfaction and competitiveness. Likewise, consumer feedback, facilitated by digital platforms, fuels product innovation and brand loyalty when analyzed through sentiment analysis and data analytics. Opinion mining techniques further enable decision-makers to gauge public sentiment, adapt marketing strategies, and respond effectively to evolving consumer preferences in today's dynamic marketplace.

## II. LITERATURE REVIEW

In this paper, Samik Datta and Satyajit Chakrabarti propose an Aspect-based Sentiment Analysis (ABSA) approach using an optimized Recurrent Neural Network (RNN) enhanced by a novel optimization method called Fire Fly-oriented Multi-Verse Optimizer (FF-MVO) [1]. The study focuses on analyzing public opinion on demonetization in India through tweets, aiming to identify specific sentiments related to different aspects of the policy. The method involves preprocessing the tweets to extract aspect words, computing polarity scores using Word2vec, and optimizing sentiment classification with the FF-MVO algorithm to improve prediction accuracy. Comparative analysis with various machine learning algorithms demonstrates superior performance of the FF-MVO-RNN model. However, the study acknowledges limitations in real-world optimization and multi-modal search procedures, suggesting avenues for future research to enhance algorithm effectiveness and validate results on larger datasets.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704206 |

Kangning Wei, Yong Zha, and Heshan Suz examine the transaction intentions of buyers and sellers in consumer-to-consumer (C2C) e-marketplaces [2], emphasizing the distinct factors influencing each group. By integrating trust and risk perceptions into the Technology Acceptance Model (TAM) framework, they explore the impact of these constructs on transaction intentions from both perspectives. The findings reveal that perceived risk significantly influences sellers' intentions more than buyers', while perceived usefulness affects only buyers' intentions, and institution-based trust and perceived ease of use impact sellers' intentions exclusively. Notably, trust in the e-marketplace's owner influences institution-based trust for both buyers and sellers. These insights offer theoretical contributions to C2C e-commerce research and practical implications for e-marketplace practitioners, suggesting tailored strategies to address the specific concerns and motivations of buyers and sellers within these platforms.

In this paper, Sandeep Trivedi and Nikhil Patel investigate public sentiment surrounding the emerging paradigm of hybrid work arrangements post-COVID-19. Using sentiment analysis techniques with the RoBERTa model and Twitter API, they analyze attitudes toward hybrid work, finding that a majority of individuals express positive sentiments. Comparisons with remote work sentiments reveal an affinity between the two concepts. The study suggests several approaches for organizations to develop effective hybrid work environments, including adapting the physical workspace, implementing flexible office policies, fostering virtual collaboration, and ensuring equitable benefits for hybrid workers. These strategies aim to promote efficiency, engagement, and seamless communication among employees working from various locations [3].

Sulaiman Ainin, Ali Feizollah, Nor Badrul Anuar, and Nor Aniza Abdullah conducted a study analyzing 85,259 multilingual tweets on halal tourism [4] using 19 keywords in English and Bahasa Malaysia. The research identified Japan as the most-tweeted-about halal tourist destination, followed by Malaysia and Indonesia. The findings provide insights for stakeholders in the tourism industry, highlighting the growing popularity of halal tourism beyond Muslim-majority countries, particularly in non-Muslim nations like the United Kingdom, Canada, and Spain. Tweet activity spikes during holiday seasons, reflecting increasing interest in international halal travel. The study, spanning a 10-year period, utilized Twitter's concise format to efficiently analyze sentiments and opinions related to halal tourism trends.
Jing-Rong Chang, Hsin-Ying Liang, Long-Sheng Chen, and Chia-Wei Chang address the challenge of sentiment classification in text-based social media, where the volume of online reviews leads to high-dimensional text data affecting machine learning performance. They introduce two novel feature selection methods: the Modified Categorical Proportional Difference (MCPD) approach and the Balance Category Feature (BCF) strategy. These methods aim to enhance sentiment classification by improving attribute selection without bias toward majority sentiments. Experimental results demonstrate that combining BCF with MCPD significantly reduces feature space while improving classification performance. The study emphasizes the importance of efficient feature selection to handle the increasing volume of online text data and highlights the potential for further research in applying different term weighting methods to enhance sentiment analysis and text classification tasks [5].

In summary, the work presented in this paper is built on previous research to explore how the product reviews are classified based on the analysis of their sentiments as positive, negative, both and undefined. While earlier work focused on review classifications, we focus on the comparative review classification based on various machine learning algorithms with aspect-based technique and providing product aspect ranking for the offline reviews.

## III. EXISTING SYSTEM

The existing system faces challenges in managing large volumes of unstructured data on internet platforms. Traditional methods for analysis are time-consuming and inefficient. Sentiment analysis, a subset of NLP using machine learning techniques, aims to extract insights from product reviews by predicting sentiment polarity. This approach enhances the understanding of user opinions in textual data. Challenges in sentiment analysis include accurately capturing sarcasm and irony and addressing context-dependent language nuances. Additionally, there are difficulties with acquiring labelled training data and effectively handling multilingual or dialectal expressions.

## IV. PROPOSED SYSTEM

The proposed system revolutionizes e-commerce sentiment analysis with continuous model updates based on new reviews. Using machine learning (keyword-based, Instance Based Learning, Support Vector Machine), it identifies product aspects and categorizes sentiments for effective analysis. By ranking reviews based on sentiment and aspect weights, it optimizes product recommendations and addresses customer feedback dynamically, ensuring responsiveness

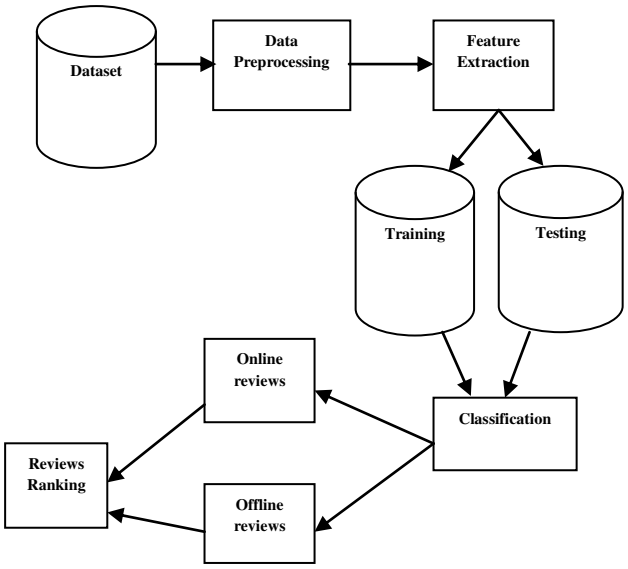
| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704206 |

to evolving sentiments over time. The system enhances e-commerce understanding of customer preferences and feedback, improving sentiment classification accuracy over time.

## V. SYSTEM ARCHITECTURE




## VI. METHODOLOGY OF PROPOSED SYSTEM

### LOAD DATA:

In the load data process, collect e-commerce product reviews and item information from various channels. Structure and format the data systematically for clarity ad consistency. Ready the formatted data for preprocessing and analysis, setting the stage for sentiment and review analysis.

### DATA PREPROCESSING:

Clean text by removing special characters and unnecessary elements. Standardize text by making everything lowercase and reconstruct sentences. Enhance clarity by addressing misspelled words, removing stop words, and applying lemmatization. Identify and extract crucial aspects or features relevant to sentiment analysis.

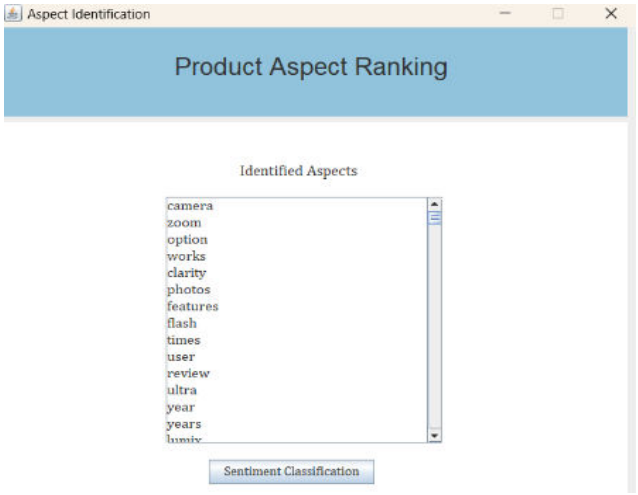
### ASPECT BASED TECHNIQUE:

Aspect-based sentiment analysis (ABSA) is a technique used in natural language processing (NLP) and text mining to analyze sentiment at a more granular level by focusing on specific aspects or attributes within text data. In ABSA, the goal is to identify and extract opinions and sentiments expressed towards different aspects or features of a product, service, or entity mentioned in text data, such as product reviews or social media posts. In this paper, the aspect from the reviews are identified through the external files consisting keywords of positive text and negative text. Using these external text files the aspects from the product reviews are categorized and classified for analysis.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704206 |




**KEYWORD BASED TECHNIQUE:**

The keyword-based technique collects a dataset of product reviews containing text and corresponding sentiment labels. Create lists of sentiment-bearing keywords and phrases that indicate positive or negative sentiment based on product reviews. Converts each review into a feature vector representation based on the presence or absence of sentiment keywords. Apply a classification algorithm to the feature vectors to predict the sentiment of each review.

**INSTANCE BASED LEARNING:**

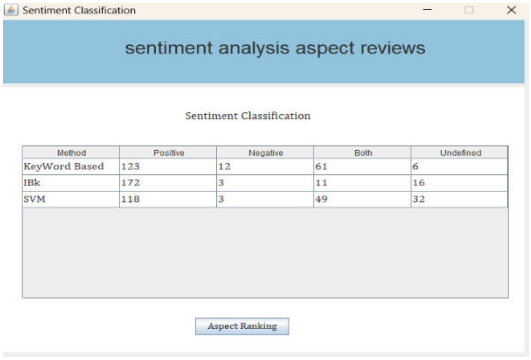
Each product review is represented as a data instance containing text and its corresponding sentiment label. All labeled reviews (instances) are stored in memory without explicit model training. The stored instances serve as a reference for making predictions later. To predict the sentiment of a new review, the similarity between the new review and each stored instance is computed. Once the nearest neighbors are identified, the predicted sentiment for the new review is determined. Finally, the predicted sentiment (positive, negative, or neutral) for the new reviews is assigned.

**SUPPORT VECTOR MACHINE:**

Convert product reviews into numerical feature representations. Train an SVM classifier using labeled data, optimizing hyperparameters. Assess the performance of the SVM classifier using evaluation metrics (e.g., accuracy, precision, recall) on a validation dataset. Use the trained SVM model to predict sentiment for new product reviews based on the feature representations.

**SENTIMENT ANALYSIS CLASSIFICATION:**

This paper categorizes comparative study for the sentiment classification of the product reviews using various machine learning algorithms such as Keyword based, Instance Based Learning, and Support Vector Machine. This comparative study provides the relatively accurate classification on positive, negative, neutral and undefined reviews. The results are tabulated in such a way that how many reviews from the overall product review dataset are positive, negative, both and undefined are differentiated from various algorithms.



| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704206 |

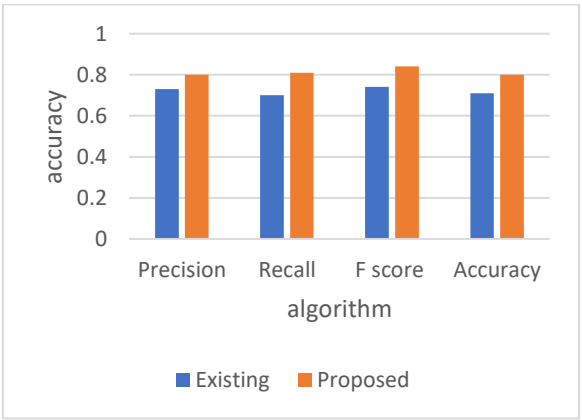
**PRODUCT ASPECT RANKING:**

After sentiment analysis, aspects are ranked based on their importance or impact on overall customer satisfaction or opinion. This ranking can be determined by aggregating sentiment scores across reviews or by considering other metrics such as frequency of aspects with its overall product reviews occurrence. The goal of product aspect ranking is to extract meaningful insights about which aspects of a product are most positively or negatively perceived by customers.


**COMPARSION TABLE:**

These are important metrics used to evaluate the performance of classification models in machine learning and sentiment analysis. Each metric serves a specific purpose in assessing how well a model is performing based on its predictions compared to the actual ground truth labels:

| *Algorithm* | *Precision* | *Recall* | *F score* | *Accuracy* |
|---|---|---|---|---|
| *Existing* | 0.73 | 0.7 | 0.74 | 0.71 |
| *Proposed* | 0.8 | 0.81 | 0.84 | 0.8 |




## VII.CONCLUSION AND FUTURE WORK

The development and deployment of the sentiment analysis system described represent a significant advancement in understanding customer feedback within the evolving landscape of e-commerce platforms. Leveraging machine learning techniques like keyword-based, Instance Based K, and Support Vector Machine approaches, this system offers a comprehensive method for sentiment classification, enabling reliable detection of product attributes and sentiments expressed in both online and offline reviews. Through rigorous testing and iterative refinement, the system

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.521| Monthly Peer Reviewed & Referred Journal |

| Volume 7, Issue 4, April 2024 |

| DOI:10.15680/IJMRSET.2024.0704206 |

demonstrates resilience, flexibility, and scalability, empowering e-commerce platforms to enhance recommendation systems and better address customer needs. Looking ahead, augmenting the sentiment analysis system with advanced natural language processing, such as deep learning models, could elevate sentiment categorization accuracy and granularity. Moreover, integrating sentiment analysis with additional data sources like social media or customer service interactions may provide a holistic view of customer attitudes and preferences. Exploring real-time analytics and implementing proactive strategies such as sentiment-based notifications or personalized recommendations could further elevate customer engagement and satisfaction across e-commerce platforms.

## REFERENCES

[1] Kumar V. V., Raghunath K. M., Muthukumaran V., Joseph R. B., Beschi I. S., and Uday A. K. (2022). Aspect-based sentiment analysis and intelligent categorization in an unpredictable feedback pool. International Journal of System Assurance Engineering and Management, 13(1), 252–262.
[2] Kumar, S., Gahalawat, M., Roy, P. P., Dogra, D. P., & Kim, B. G. (2020). Exploring impact of age and gender on sentiment analysis using machine learning. Electronics, 9(2), 374.
[3] Trivedi, S., and Patel, N. (2022). Mining Public Opinion on Hybrid Collaboration with RoBERTa. Empirical Quests for Management Essences, 2(1), 31–44.
[4] Zad S., Heidari M., Jones J. H., and Uzuner O. (2021, May). A study of concept-level sentiment analysis algorithms for textual data. In the 2021 IEEE World AI IoT Congress (AIIoT) (pp. 0285–0291). IEEE.
[5] Ainin S., Feizollah A., Anuar N. B., and Abdullah N. A. (2020). Sentiment analysis of multilingual tweets about halal tourism. Tourism Management Perspectives, 34: 100658.
[6] Keikhosrokiani, P.; Pourya Asl, M. (Eds.). (2022). Handbook of Research on Opinion Mining and Text Analytics in Literature and Social Media. IGI Global.
[7] Chen X., Xue Y., Zhao H., Lu X., Hu X., and Ma Z. (2019). A unique feature extraction approach for sentiment analysis in product reviews. Neural Computing and Applications, 31(10): 6625–6642.
[8] S. Mukherjee, and P. Bhattacharyya (2020, March). Sentiment analysis based on certain product features. In International Conference on Intelligent Text Processing and Computational Linguistics (pp. 475–487). Springer, Berlin and Heidelberg.
[9] Vayansky, I., and Kumar, S. A. (2020). A survey of subject modeling techniques. Information Systems, 94: 101582.
[10] Nagarajan, S. M., and Gandhi, U. D. (2019). Using hybridization to classify streamed Twitter data based on sentiment analysis. Neural Computing and Applications, 31(5), 1425–1433.
[11] J.-R. Chang, H.-Y. Liang, L.-S. Chen, C.-W. Chang. (2020). Novel feature selection approaches for improving the performance of sentiment classification, Journal of Ambient Intelligence and Humanized Computing.
[12] Datta, S., Chakrabarti, S. (2021). Aspect-Based Sentiment Analysis for Demonetization Tweets by Optimized Recurrent Neural Network using Fire Fly-oriented Multi-Verse Optimizer. publication in journal Sadhana 46(2), 79.
[13] Maheswari, S & S.S, Dhenakaran. (2021). Detection of fake and Genuine Reviews with Hybridization of Fuzzy and Neural Networks Techniques.
[14] Chunduru, A. Veerappampalayam Easwaramoorthy, S., Kanchana, S., Bharath Kumar, S. (2023). Sentimental Analysis on Product Reviews Using Support Vector Machine and Nave Bayes. Applied and Computational Engineering, 2, 66-72.
[15] Zhang, Wei, Sui-xi Kong, and Yan-chun Zhu. "Sentiment classification and computing for online reviews by a hybrid SVM and LSA based approach." Cluster Computing, Volume 22, No. 5, 2019, pp. 12619-12632.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

# INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

## IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com