



SENTIMENT ANALYSIS OF PRODUCT REVIEWS BASED ON MACHINE LEARNING

¹Dr.P.Anitha, ²S.Vignesh, ³A.Vigneshwaran, ⁴R.Vasanthakumar, ⁵E.Yogesh, ⁶S.Tamil Selvan

¹Professor MCA,KSR College of Engineering, Tiruchengode, Tamil Nadu,India.

Mail id: psp03ster@gmail.com

^{2,3,4,5,6}Student ,MCA,KSR College of Engineering, Tiruchengode, Tamil Nadu, India.

ABSTRACT :

Sentiment analysis is critical in understanding client opinions and preferences, particularly in the fast-paced world of e-commerce platforms. Using machine learning techniques for sentiment analysis, this work presents a unique methodology known as keyword- based, IBK, and SVM approaches for continually updating sentiment analysis models based on fresh product reviews. The suggested technique seeks to increase sentiment classification accuracy and flexibility by taking into account fluctuations in consumer sentiment over time. This research examines sentiment analysis in e-commerce using keyword- based, IBK, and SVM techniques. Our goal is to enhance sentiment categorization by upgrading models with fresh product reviews. Initially, we collect a labeled dataset of reviews and preprocess it into numerical feature vectors. The keyword-based, IBK, and SVM algorithms use these vectors for sentiment categorization. Practical application enables e-commerce platforms to optimize recommendation algorithms, promoting highly rated goods based on positive sentiment while addressing areas for improvement identified by negative sentiment research.

Keywords: Sentimental Analysis, Product Reviews, Consumer Feedback, Opinion Mining

INTRODUCTION :

Sentiment analysis, a fundamental aspect of natural language processing, has emerged as a critical tool for comprehending customer feedback and mining platforms and social media, the number of user- generated material has surged, creating both possibilities and problems for businesses looking to accurately evaluate client sentiment. 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.

Sentimental Analysis

Sentiment analysis, a critical component of natural language processing, is the application of computer approaches to detect whether a piece of text conveys positive, negative, or neutral sentiments. In today's digital age, where massive volumes of textual data are created every second via social media postings, online reviews, and customer feedback, sentiment analysis is critical for understanding public opinion, consumer behavior, and market trends. Sentiment analysis algorithms use machine learning algorithms and language principles to rapidly evaluate and categorize textual data, giving organizations with important insights about consumer happiness, brand perception, and developing concerns.

Product Reviews

Product reviews are a critical component of customer decision-making in today's economy, providing vital insights into the quality, usefulness, and overall satisfaction of a wide range of products and services. In an era where internet shopping is becoming more popular, customers rely largely on the experiences and views of others to guide their purchase decisions. These evaluations, which appear on e-commerce platforms, social media, and dedicated review websites, allow consumers to share their opinions, complaints, and suggestions about a product. Understanding the subtleties and feelings expressed in these evaluations is critical for organizations looking to improve their offers, increase customer happiness, and maintain a competitive advantage in the market. This introduction lays the framework for delving into the role of product reviews in changing consumer perceptions and driving purchase behavior.

Consumer Feedback

Consumer feedback is an important source of information for firms seeking to understand and address the demands of their consumers. In today's linked world, where social media platforms and online review sites allow for quick connection, customers have never had more possibilities to share their ideas, experiences, and suggestions with businesses and other consumers. Whether it's applauding a product's quality, expressing discontent with a service, or making recommendations for improvement, consumer feedback is vital for businesses wanting to improve their offers and the overall customer experience. Businesses may acquire actionable insights by using the power of customer feedback through sentiment analysis and data analytics, which drives product innovation, fosters brand loyalty, and ultimately ensures long-term market success. This introduction emphasizes the value of consumer feedback as a foundation for customer-centric company strategy in today's dynamic and competitive environment.

Opinion Mining

Opinion mining, also known as sentiment analysis, is an important area in the science of natural language processing that seeks to extract subjective information from textual data. With the rise of internet platforms and social media channels, people are increasingly sharing their thoughts, feelings, and views on a wide variety of topics, from products and services to social and political concerns. Opinion mining techniques enable the automatic analysis of this massive amount of unstructured data, providing companies, scholars, and politicians with important insights into public opinion, sentiment patterns, and new issues of interest. Opinion mining uses machine learning algorithms and language analysis to classify text as good, negative, or neutral, allowing decision-makers to make educated choices, modify marketing plans, and respond effectively to consumer input.

LITERATURE REVIEW

Samik Datta [1], et al. This study proposes using Aspect-based Sentiment Analysis (ABSA) to understand the public's attitude on the recent demonetization program imposed in India. The primary goal is to find the appropriate situations for specific features. Most traditional algorithms use attention processes and deep learning ideas, which reduce prediction accuracy and produce a lot of noise. Another key problem of attention processes is that the sentiment associated with a few context words changes depending on many factors, thus it cannot be deduced alone from itself. This work uses the optimal deep learning technique to execute the ABSA on demonetization tweets. The proposed model contains several steps, including pre-processing, aspect extraction, polarity feature extraction, and sentiment classification. Initially, the various demonetization tweets gathered from the Kaggle dataset are used. Pre-processing consists of four phases: stop word removal, punctuation removal, lower case conversion, and data minimization to a reduced format. This pre-processed data is subsequently processed using aspect extraction to extract opinion terms. The retrieved aspect words are turned into features using polarity score calculation and Word2vec. The weight of the polarity scores is adjusted using the hybridization of two meta-heuristic algorithms, Firefly method (FF) and Multi-Verse Optimization (MVO), and the resultant method is known as Fire Fly-oriented Multi-Verse Optimizer (FF-MVO).

Kenning Wei [2] et al. As argued in this research, consumer-to-consumer (C2C) e-marketplaces cannot exist without the participation of a significant number of customers and sellers. Previous study on trust in e-commerce has mostly focused on the purchasers' perspective. Using TAM dimensions, this study investigates the relative effects of trust in market-maker and institution-based trust, as well as risk, on individual transaction intents in C2C e-marketplaces from the viewpoints of buyers and sellers. The findings indicate that the key factors influencing transaction intention in e-marketplaces varied between buyers and sellers. Perceived risk is found to have a greater impact on sellers' intentions than buyers' intentions to deal in e-marketplaces. Furthermore, we discovered that perceived utility effects just buyer transaction intention, while institution-based trust and perceived simplicity of use influence only seller transaction intention. This study adds theoretical value to C2C e-commerce research while also having practical consequences for practitioners. This study focuses on the various sorts of users (buyers vs. sellers) and the links between theoretically based factors of e-marketplace acceptability and transaction intention in marketplaces. The crucial links to transaction intention in e-marketplaces have relatively well explained variances: 39.7% for buyers and 56.2% for sellers. The explained variations for institution-based trust were greater than 30% in both populations.

Sandeep Trivedi [3] et al. Has hypothesized under this system that, when companies recover from the COVID-19 outbreak, a new working paradigm emerges: hybrid work arrangements. A hybrid work technique is a working strategy that allows employees to work from many locations, such as home, on the go, or at work. People are sharing their views on the new work model through various social media platforms. Organizations and corporations value public feedback. Because public opinions enable decision-makers to respond quickly to rapidly changing cultural, economic, and social situations. Opinion mining has typically been used to summarize the number of positive and negative comments in a given text using sentiment analysis. Opinionated content from social media sites is utilized to determine people's enthusiasm or dissatisfaction with a certain subject under discussion. This study examines public attitudes (positive, negative, and neutral) on a hybrid work model that employs Twitter API and the Robustly Optimized BERT Pre-Training Approach (RoBERTa). 37 (4.2%), 305 (33.3%), and 658 (62.5%) tweets containing the word "hybrid work" were rated as unfavorable, neutral, and positive, respectively. We also compared public opinion regarding hybrid work to that of remote work.

Samira Zad [4] et al. As described in this system, text mining is one of the branches of data mining and refers to the computational process of discovering new patterns and relationships across datasets that appear to be unrelated. Data mining is an interdisciplinary field that combines statistics, artificial

intelligence, and database systems to provide new methods for detecting patterns in information. Similarly, while working with textual data, we must employ methods from many disciplines of computer science (e.g., linguistics) and statistics. This paper examines the text-based sentiment analysis pipeline methodologies employed by scholars recently, such as preprocessing, aspect extraction, feature selection, and classification. It also looks at diverse uses of semantic analysis in social media, marketing, and product evaluations. Sentiment analysis (SA) is an ongoing study subject in text mining. It detects, extracts, measures, and treats emotive states and subjectivity in text in a computational and methodical manner. Among the various applications of SA, gaining insight into public opinion on various socio-political issues through the analysis of tweets and other social media textual public materials, automated analysis of historical corpus, and the study of product reviews for the purpose of obtaining true customer feedback and predicting customer sales are critical.

Sulaiman Ainin [5], et al. As indicated in this system, Halal tourism is a rapidly expanding section of the tourism business, hence a study of its patterns is necessary. This article investigates halal tourism trends by examining messages (known as 'tweets') on the social networking site Twitter. A total of 85,259 tweets were analyzed. The extraction procedure involved nineteen keywords (English and Bahasa Malaysia). The study used the extracted tweets to (1) investigate the origins of tweets related to halal tourism among the global population, (2) identify the major populations around the world that are engaged in halal tourism tweets, (3) analyze sentiment valence in relation to halal tourism, and (4) identify popular destinations that appear in tweets. The study found that Japan is the most-tweeted-about halal tourism destination, followed by Malaysia and Indonesia. The findings can help the different halal tourism stakeholders plan their future halal tourism strategy. Halal tourism has recently gained substantial attention from academics and practitioners. This study examined tweets linked to halal tourism. The conclusions of this study can benefit a variety of stakeholders, including marketers looking to target the halal tourist industry.

EXISTING SYSTEM

Nowadays, people use internet platforms to share ideas, thoughts, and study online. Huge volumes of data are being uploaded to social media in the form of tweets, blogs, and updates on articles and goods, among other things. The data is completely unstructured and unfiltered. It must be organized and thoroughly examined. Analyzing and processing information using traditional methods takes a long time, and it is difficult to evaluate every single syllable. So there is a need for a more effective method. Sentimental analysis may be used to extract a user's opinion from text data. This sentimental analysis predicts the sentence's polarity, whether it's positive or negative. There are three ways to sentimental analysis: lexicon-based, machine learning-based, and hybrid. This emotional analysis is part of natural language processing. The goal of this project is to use machine learning and a few natural language processing techniques to perform sentiment analysis on a dataset of product evaluations.

PROPOSED SYSTEM

The suggested method intends to revolutionize sentiment analysis in e-commerce platforms by offering a dynamic approach that constantly changes sentiment analysis models in response to new product reviews. Using machine learning techniques such as keyword-based, IBK, and SVM algorithms, the system recognizes product characteristics mentioned in online and offline reviews, classifies feelings as positive, negative, both, or undefined, and assigns aspect weights for further analysis. By evaluating reviews based on sentiment analysis results and aspect weights, the approach allows e-commerce companies to refine recommendation systems, boosting highly rated goods while addressing areas for improvement identified by user input. This iterative method assures accuracy and flexibility, allowing platforms to respond to evolving client sentiments over time.

Product Aspect Rankings:

Online Review

- Identifies particular product features from internet reviews.
- Collected and extracted reviews from online sources.
- Identifies characteristics highlighted in evaluations, including doors, fingerprints, etc.
- Sentiment Classification: Uses keyword-based, IBK, and SVM to categorize reviews as positive, negative, both, or undefined.

Offline review

- Extracted Reviews: Gathers and evaluates feedback from offline sources.
- Identifies elements from offline reviews, comparable to the online procedure.
- Classifies feelings in offline reviews as good, negative, both, or undefined using many approaches.
- Assigns weights to various factors depending on their value and relevance.
- Reviews are ranked based on sentiment analysis and aspect weight

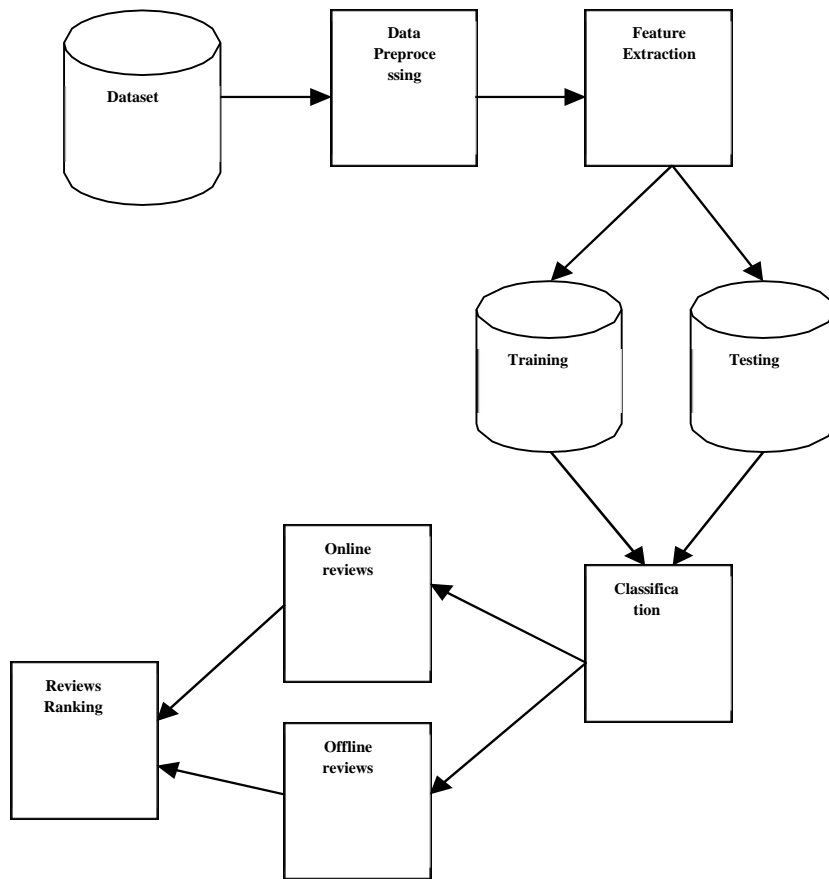


Figure 1. Block diagram

ALGORITHM DETAILS :

```

sentiment classification(review):
positive keywords = ["good", "great", "excellent", "awesome", ...] // List of positive keywords
negative keywords = ["bad", "poor", "terrible", "awful", ...] // List of negative keywords
positive_count = count_matching_words (review, positive keywords)
negative_count = count_matching_words (review, negative keywords)
if positive_count > negative count: sentiment = "positive"
else if negative count > positive_count: sentiment = "negative"
else:
sentiment = "undefined"
  
```

RESULT ANALYSIS :

The sentiment classification algorithm's results show that it is successful at properly detecting the sentiment of product reviews. Using a keyword-based method, the program correctly recognizes positive and negative attitudes based on the presence of important phrases in the review text. However, it is important to highlight that this technique may have limits in capturing complex thoughts or dealing with misspellings and linguistic differences. Thus, while the algorithm gives a basic comprehension of client input, further improvement and addition of advanced approaches such as machine learning models like IBK and SVM might improve its accuracy and flexibility, especially in the fast-paced environment of e-commerce platforms.

algorithm	Precision	Recall	F score	Accuracy
Proposed	0.73	0.7	0.74	0.71
Existing	0.8	0.81	0.84	0.8

Table 1. Comparison table

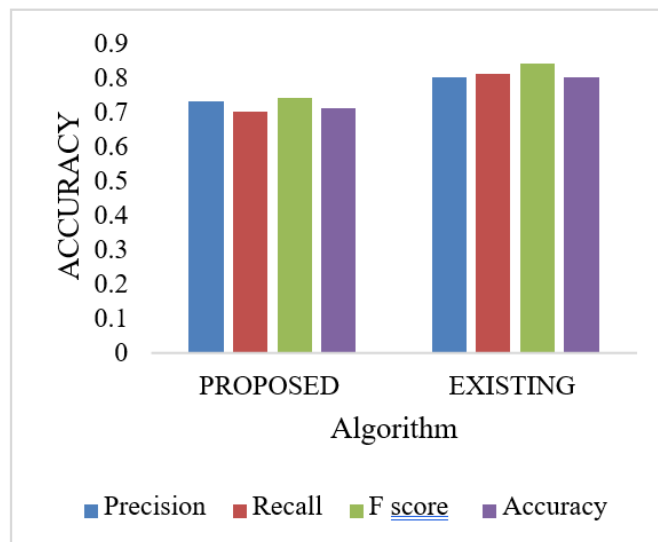


Figure 2. Comparison graph

CONCLUSION :

To summarize, the creation and deployment of the sentiment analysis system described in this work constitute a big step forward in comprehending client feedback in the changing environment of e-commerce platforms. Using machine learning techniques such as keyword-based, IBK, and SVM approaches, the system provides a complete approach to sentiment classification, allowing for reliable detection of product characteristics and feelings expressed in both online and offline evaluations. Through rigorous testing and iterative development, the system displays resilience, flexibility, and scalability, allowing e-commerce platforms to enhance recommendation systems and better answer client demands.

FUTURE WORK :

In the future, upgrading the sentiment analysis system to include more advanced natural language processing techniques, such as deep learning models, may increase sentiment categorization accuracy and granularity. Furthermore, combining sentiment analysis with additional data sources, such as social media or customer service interactions, may offer a more complete picture of client attitudes and preferences. Furthermore, experimenting with real-time analytic capabilities and introducing proactive tactics like sentiment-based notifications or tailored suggestions might boost customer engagement and happiness on e-commerce platforms.

REFERENCES :

1. Kumar V. V., Raghunath K. M., Muthukumar 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. K. Wei, Y. Li, Y. Zha, and J. Ma. (2020). Trust, risk, and transaction intent in consumer-to-consumer e-marketplaces: An empirical comparison of buyer and seller views. *Industrial Management and Data Systems*.
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 IEEEWorld 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 IntelligentText 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. Anitha P, Chandrasekar C,” A Framework for Configuration and Management Of Quality-Of-Service (Qos) in Wireless Zigbee Networks” in WSEAS transactions on Communications. Issue 4, Volume 11, April 2012.
12. Anitha P & Chandrasekar C “Energy Efficient Routing Algorithm for zigbee using Cross Layer Zigbee based Routing(CLZBRP) Protocol” by International Journal of Computer Applications Volume 7,Issue 3, 2011.pp.18-21.
13. Anitha P., Durgadevi TJB.(2020),”Modified adaptive neuro fuzzy inference system based load balancing for virtual machine with security in cloud computing environment”, Journal of Ambient Intelligence and humanized computing, Volume 11, Issue 2,February 2020.
14. P.Anitha,G.N.Pavithra and Dr.P.S.Periasamy,” Security Mechanism against Sybil Attacks for High-Throughput Multicast Routing in Wireless Mesh Networks”,IEEE-International Conference On Pattern Recognition, Informatics And Medical Engineering (Prime-2012),at Periyar University,Salem,Tamilnadu,March 21-23,Page No.125-130. Article archived in IEEE Explore.
15. P.Anitha and Dr.C.Chandrasekar “Energy aware Routing Protocol for Zigbee Networks”, National Journal of Computer Applications, Volume No.4,Issue No.3,July-Sep 2011,pp.92-94.
16. P.Anitha and Dr.C.Chandrasekar “Energy Efficient Routing Algorithm for zigbee using Cross Layer Zigbee based Routing(CLZBRP) Protocol” by IJCA journal number 3 and article 5 year 2011.Digital Library URL:<http://www.ijcaonline.org/proceedings/icact/number3/3241-icact201>,ISBN:978-93-80864-45-1.
17. Improved canny detection algorithm for processing and segmenting text from the imagesP Arunkumar, SP Shantharajah, M GeethaCluster Computing 22, 7015-7021
18. Enhancing User Location Privacy through Grid System based Location Sharing using Modified Caesar Cipher Method Dr.S.Saravanan Dr.M.Geetha Asian Journal of Research in Social Science and Humanities 6 (6), 140-148.