

CONSUMER SENTIMENT INSIGHTS on SMART RING USING SUPERVISED CLASSIFIERS: a COMPARATIVE STUDY

*Rashmi Prabha, **Ruturaj B. Nawale, ***Sakshi P. Belhekar, ****Poonam R. Prajapati

Department of Computer Science, Changu Kana Thakur Arts, Commerce and Science College, New Panvel (Autonomous), Maharashtra, India

*Email: rashmiryan22@gmail.com

**Email: ruturajnawale888@gmail.com

***Email: belhekarsakshi@gmail.com

****Email: poonamprajapati1145@gmail.com

Abstract:

Sentiment analysis plays a crucial role in understanding customer opinions and improving business decision-making. This study focuses on analysing customer reviews of the product Smart Ring using machine learning classifiers. The dataset, collected from various e-commerce platforms like Amazon, Flipkart, etc, includes 100 records of customer feedback. This study applies text pre-processing techniques, such as tokenization, stop word removal, and feature extraction using Document-Term Matrix (DTM), to prepare the data for sentiment classification. Two machine learning models— Support Vector Machine (SVM) and Naïve Bayes (NB)—are employed to classify customer reviews into positive, negative, and neutral sentiments. The analysis reveals that SVM outperforms the NB classifier, achieving an accuracy of 55% compared to NB's 20%. The findings highlight the challenges posed by a limited dataset, inconsistencies between ratings and textual reviews, and the need for further model optimization. This research contributes to improving sentiment classification methodologies for smart wearable products and provides insights into customer satisfaction trends.

Keywords: Sentiment Analysis, Customer Reviews, Machine Learning (ML), Support Vector Machine (SVM), Naïve Bayes (NB)

Abbreviations:

BERT	Bidirectional Encoder Representations from Transformers	LR	Logistic Regression
DT	Decision Tree	LSTM	Long Short-Term Memory
GRU	Gated Recurrent Unit	RNN	Recurrent Neural Network
CNN	Convolutional Neural Network	RF	Random Forest
SVM	Support Vector Machine	NB	Naïve Bayes
ANN	Artificial Neural Network	VADER	Valence Aware Dictionary and Sentiment Reasoner
KNN	K-Nearest Neighbors	HRS	Hierarchical Random Sampling

TABLE. I ABBREVIATIONS USE

I. INTRODUCTION

In this tech era, customer reviews are essential for most of the firms to operate quickly and are useful for sentiment research. Every day, millions of reviews of different goods and services are posted online [1]. They aggregate data from online platforms like web portals, Amazon, etc. Review Categorization involves classifying fresh documents according to predefined categories based on classification patterns [2]. When multiple reviews are available for the same product, it might be difficult for customers to make informed decisions.

Analysing customer feedback enables businesses to increase the quality of their services and foster long-lasting relationships with their customers. With the help of sentiment analysis,

There are three types of sentiment analysis:

- A. **DOCUMENT LEVEL:** In this level, we categorize the entire document as positive, negative, or neutral
- B. **SENTENCE LEVEL:** Each sentence is analysed and stated as yes, no, or neutral.
- C. **FEATURE LEVEL:** Involves examining the words or phrases to determine what they are expressing either positive, negative, or neutral.

We compile data from Amazon, Flipkart, eBay, Jiomart, Gabit, boAt, and pi-ring. According to this study, Amazon was one of the most well-known internet retailers [2]

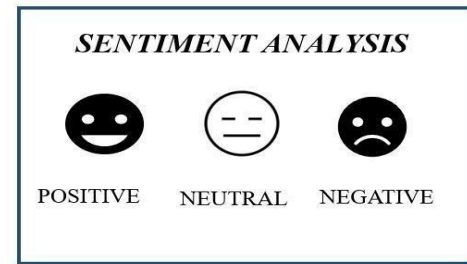


Fig. 1 Types of Sentiment Analysis

This study faced several challenges, including limited reviews due to the smart ring's recent launch in 2024, inconsistencies between ratings and review text, and a small dataset of only 100 records, affecting model accuracy.

By extracting information from these evaluations, customers can understand the main characteristics of the product. To solve the sentiment classification issue, we use machine learning algorithms. The primary objective of this investigation was to identify customer evaluations, both favorable and unfavorable, and assess their applicability. This survey simulated support vector machine and naïve bayes two machine learning algorithms and came up with an analysis in R tools. Several libraries were used, such as tidyverse, readx, ggplot2, dplyr, tidytext, wordcloud2, snowball, e10701, tm, RColorBrewer. The accuracy of the Support Vector Machine achieves a higher rate of 55 % than Naïve Bayes, 20%.

II. LITERATURE REVIEW

The table 2 highlights the summarized information of reviewed papers based on Customer feedback analysis using Sentiment Analysis.

Ref. No.	Authors	Title	Methods	Research Gap	Results
[2],2023	Gagan Kumar Patra, Chandrababu Kuraku	A Sentiment Analysis of Customer Product Review Based on Machine Learning Techniques in ECommerce	ML, BERT, LR, DT	The model is less efficient for real-time applications	BERT- 89% LR- 83.1% DT- 75%
[4],2022	Roobaea Alroobaea	Sentiment Analysis on Amazon Product Reviews using the Recurrent Neural Network	ML, LSTM, GRU, RNN, CNN	Limited research in Arabic language	RNN -85% CNN -83% GRU -80% LSTM -76%

[3],2024	Hashir Ali, Ehtesham Hashmi	Analysing Amazon Products Sentiment: A Comparative Study of Machine and Deep Learning, and Transformer-Based Techniques	ML, LR, RF, BERT	Difficulty in categorizing between similar reviews	LR -86 % RF -84% BERT – 89%
[1], 2021	Behrooz Noori	Classification of Customer Reviews Using Machine Learning Algorithms	ML, SVM, NB, DT, ANN	Challenging task for studying the predictive accuracy of sentiment analytics	DT – 98.9% ANN -74.5% SVM -73% NB -70.3 %
[6] ,2024	Koushik Barik, Sanjay Mishra	Analysis of customer reviews with an improved VADER lexicon classifier	VADER lexicon classifier	Not able to understand the full meaning of the word in certain situations	VADER – 98.64%
[7],2024	Ohud Alsemaree , Atm S. Alam , Sukhpal Singh Gill	Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions	SVM, RF, KNN	Still, Improvement is required for further analysis	SVM -95% RF – 94% KNN -74%
[5],2020	Shanshan Yi, Xiaofang Liu	Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review	ML, HRS	Limited scope of data source	HRS-98%

TABLE II SUMMARIZED INFORMATION OF REVIEWED PAPERS

III. OBJECTIVES

The purpose of the study is to examine the multiple elements of the smart ring, a recently developed device. By examining customer opinion, businesses can improve their products and potentially boost marketing.

IV. METHODOLOGY

Research methodology consists of methods, resources, and processes utilized to gather and analyse data to answer a research question or hypothesis. Data processing techniques, feature extraction, and data collection methods via many platforms such as Amazon, Flipkart, and others are all detailed in the research methodology. It ensures the reliability, validity, and repeatability of the study. This section summarises the intended step-by-step approach for conducting Sentiment

Analysis on Customer Reviews. The process is illustrated in figure 2.

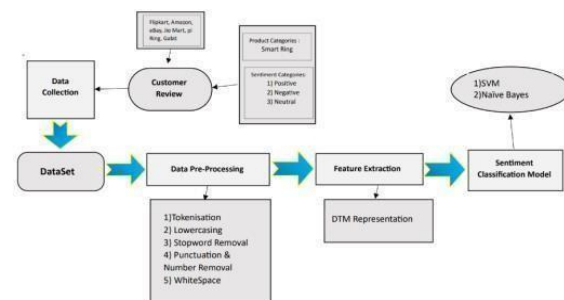


Fig. 2 Smart Ring Customer Feedback Architecture

The first steps are vectorization and then pre-processing of the collected data. Techniques for feature extraction are then used, and machine learning algorithm training comes next. Lastly, this model's output is assessed [3].

A. DATA COLLECTION

We gathered customer reviews and feedback from a collection of e-commerce platforms, including Amazon, Flipkart, eBay, Pi Ring, Gabit, Jio Mart, and Croma, for our study. We gathered product reviews from websites that sell smart wearables devices, such as Smart Rings and etc. Three categories were used to classify the sentiments: positive, negative, and neutral.

We have extracted a dataset from web scraping that contains approximately 100 records. This dataset includes a set of attributes: Product_ID, Category, Brand, Price, Platform, Rating, Review_Title, Review_Text, Reviewer_Name, Review_Date, Verified_Purchase, Review_Likes, and Product URL. These qualities offer a complete basis for carrying out research as part of this undertaking.

B. DATA PRE-PROCESSING

A crucial phase in Natural Language Processing activities is data pre-processing, which maximizes the effectiveness of knowledge discovery. To prepare and enhance the dataset for more efficient analysis, it uses methods including stop word and punctuation removal, root word discovery, tokenization, lemmatization, post-tagging, and word clouds [3].



Fig. 3 Step for Data Pre-processing

Below table 3 highlights how the data appears pre- (raw data) and post-processing.

Raw Data	Processed Data
Boat smart ring gen 1 - it's not charging even after 4 hours.	boat smart ring gen 1 charging even 4 hour
Packaging and all is good but 200rs for rings of plastic is overpriced	packaging good 200rs ring plastic overpriced

TABLE 3 RAW DATA VS PROCESSED DATA

1 Tokenization

The process of tokenization is breaking up a text sequence into distinct units referred to as tokens. Single words, phrases, or even complete sentences may be included in these tokens [3]. After that, these tokens are used as inputs for a number of procedures, including text mining and parsing.

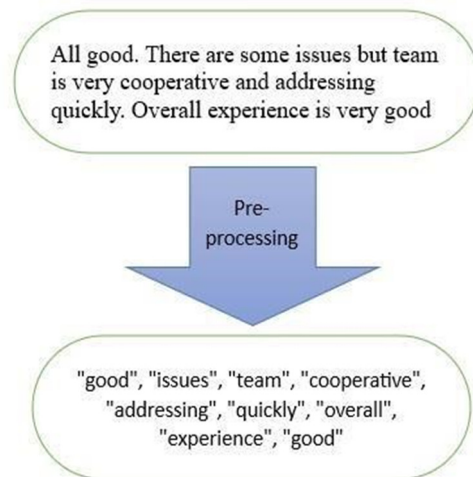


Fig. 4 Reviews before and after processing

2 Lowercasing

At this step, all of the review words are converted to lowercase. In this case, "Great" and "Amazing" are substituted by "great" and "amazing." Through case-insensitive word treatment, lowercasing reduces the dimensionality of the data and aids in text normalization [3].

3 Removal of Stop Words

Examples of stop words that are frequently used in a language are the, a, an, is, and are. Since these terms don't provide important information for the model, they have been removed from the review's content [2].

4 Punctuation Removal

This dataset excluded all punctuation, disqualified reviews that contained only one word, and eliminated one-word reviews [2].

5 Word Cloud

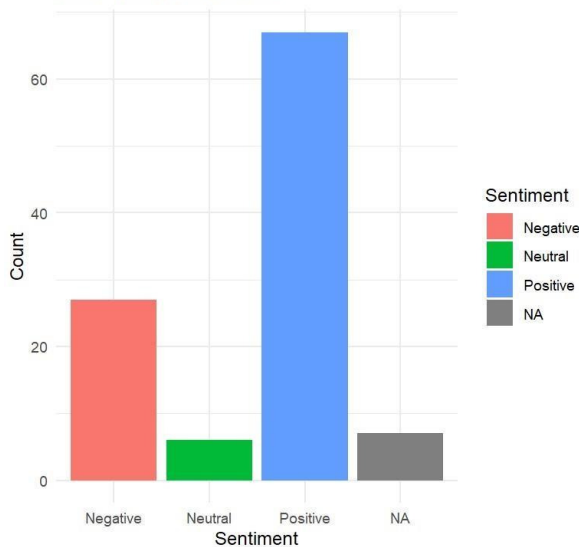
When it comes to analysis, a word cloud is a helpful tool for showing the most common themes and topics in client feedback [9]. The word cloud gives quick insights by identifying the root terms in a text without having to read through it in detail.

The most common words highlighting various sentiments in a given text are represented visually in figure 5.



The largest words like ring, sleep, health, good, data, app, product, tracking, battery, quality, and days indicate that customers frequently mention this aspect in their feedback.

Figure 6, represents the distribution of sentiments on the smart ring. It is divided into three categories: neutral, negative, and positive sentiments. Positive reviews constitute 65% of the dataset, indicating a high level of customer satisfaction with the smart ring. Around 25% of reviews are negative, reflecting customer dissatisfaction with the product. Another small portion, 5% of reviews, are neutral, suggesting a mixed experience. This sentiment classification involves overall satisfaction with the smart ring.



7 Feature Extraction

extracted using statistical techniques or machine learning algorithms. We have used the Bag of Words (BoW) approach through Document-Term Matrix (DTM) in the feature extraction process.

It is a structured version of Bag of Words (BoW) that organizes words per review. It creates a matrix where each row is a review and each column is a word, which shows word count per review. Table 4, shows a simple Document-Term Matrix (DTM) representation.

TABLE 4 DOCUMENT-TERM MATRIX REPRESENTATION.

Terms	Amazing	Battery	Design	Drains	Good	Great	Life	Long	Product	Quickly	With
Review 1	1	1	1	0	0	1	1	0	0	0	0
Review 2	0	1	1	1	1	0	0	0	0	1	0
Review 3	1	0	0	0	0	0	1	1	1	0	1

V. RESULTS AND DISCUSSIONS

A. Support Vector Machine

Figure 7 displays the confusion matrix for the SVM model, which shows how well the model performs in sentiment classification.



Fig. 7 SVM Confusion Matrix

B. Naïve Bayes

Naïve bayes is probabilistic machine learning model. All features are assumed to be independent of one another. It labels the review as the sentiment with the highest probability.

The number of examples that the model identified properly and incorrectly is displayed in figure 8. Naïve Bayes accurately recognizes positive reviews and incorrectly classifies negative reviews as neutral, as the figure illustrates.

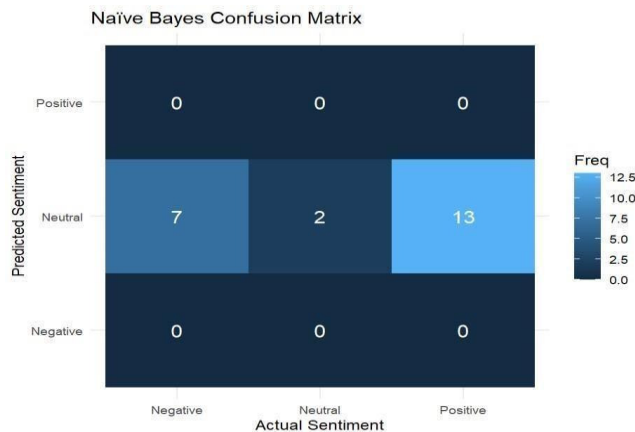


Fig. 8 Naïve Bayes Confusion matrix

C. Comparison between SVM and Naïve Bayes

Figure 9 below shows a comparison between Support Vector Machine and Naïve Bayes models. The bar chart shows a significant gap in accuracy:

Naïve Bayes at 20%, Support Vector Machine at 55%. This shows that the customer is classified by the Support Vector Machine.

reviews more accurately than naïve Bayes. This significant difference suggests that SVM performs better on sentiment classification tasks, most likely as a result of its capacity to locate the best decision boundary in high-dimensional feature spaces. Conversely, Naïve Bayes, which relies on feature independence, had trouble classifying data, which resulted in decreased accuracy. As a result, Support Vector Machine is better for separating positive, negative, and neutral sentiments than Naïve Bayes.

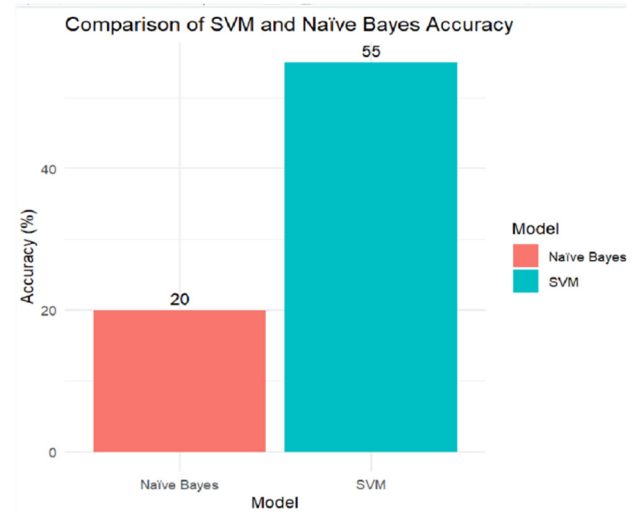


Fig. 9 SVM vs Naïve Bayes Accuracy

Table 6, highlights the comparison of results on the same data set on two different models SVM and NB with their accuracy.

Sr. No.	Dataset	Classifier	Accuracy (%)
1	Ring Dataset	SVM	55%
2		NB	20%

TABLE 6 ACCURACY PERCENTAGE

VI. RESEARCH LIMITATIONS

In order to improve sentiment analysis's accuracy, future research could look into collecting additional data as more people purchase smart rings and share their opinions. Because of ambiguous data, advanced algorithms can be used to determine the text and numerical scores and their correlations. Additionally, there is also an NA category due to missing sentiment labels or ambiguous text. Negative predictions of zero show that naïve Bayes has trouble distinguish negative reviews from sentiments.

VII. CONCLUSIONS

In this Technological world, sentiment analysis is a powerful tool to understand what customers feel about products. Every

day, thousands of people are doing online shopping and posting their feedback about the purchased product. Analysing these reviews helps other customers make better decisions when buying products. To achieve this goal, we applied the machine

REFERENCES

learning algorithms Support Vector Machine and Naïve Bayes to accurately classify sentiments on the Smart Ring. The proposed support vector machine has succeeded in achieving an accuracy of 55%.

Since the smart ring was only introduced in 2024, there weren't many reviews available at the time of our analysis, which presented some difficulties. The customer's rating are not clear, which could make it difficult for the model to effectively train the system. There were just 100 records in the dataset that was utilized to train the model. As a result, the model has trouble accurately identifying various sentiment kinds.

AUTHOR CONTRIBUTION DETAILS

Ruturaj Nawale: Formal analysis, Investigation, Data curation, Writing – review & editing, Implementation. All authors have read and agreed to the published version of the manuscript. **Sakshi Belhekar:** Writing – original draft, Conceptualization, Methodology, Validation. All authors have read and agreed to the published version of the manuscript. **Poonam Prajapati:** Visualization, Review & editing, Resources, Writing – original draft. All authors have read and agreed to the published version of the manuscript. **Rashmi Prabha:** Supervision and Guidance. The published version of the manuscript has been read and approved by all authors.

- [1] Behrooz Noori (2021) *Classification of Customer Reviews Using Machine Learning Algorithms*, *Applied Artificial Intelligence*, 35:8, 567-588, DOI:10.1080/08839514.2021.1922843
- [2] Gagan Kumar Patra, Chandrababu Kuraku, Siddharth Konkimalla, Venkata Nagesh Boddapati, Manikanth Sarisa, et al. (2023) *Sentiment Analysis of Customer Product Review Based on Machine Learning Techniques in E-Commerce*. *Journal of Artificial Intelligence & Cloud Computing*. SRC/JAICC-408. DOI: [doi.org/10.47363/JAICC/2023\(2\)389](https://doi.org/10.47363/JAICC/2023(2)389)
- [3] Ali, H.; Hashmi, E.; Yayilgan Yildirim, S.; Shaikh, S. *Analyzing Amazon Products Sentiment: A Comparative Study of Machine and Deep Learning, and Transformer-Based Techniques*. *Electronics* 2024,13, 1305. <https://doi.org/10.3390/electronics13071305>
- [4] Alroobaea, Roobaea. "Sentiment analysis on amazon product reviews using the recurrent neural network (rnn)." *International Journal of Advanced Computer Science and Applications* 13.4 (2022).
- [5] Yi, Shanshan, and Xiaofang Liu. "Machine learning based customer sentiment analysis for recommending shoppers, shops based on customers' review." *Complex & Intelligent Systems* 6.3 (2020): 621-634.
- [6] Barik, K., Misra, S. *Analysis of customer reviews with an improved VADER lexicon classifier*. *J Big Data* 11, 10 (2024).
- [7] Alsemaree, Ohud, et al. "Sentiment analysis of Arabic social media texts: A machine learning approach to deciphering customer perceptions." *Heliyon* 10.9 (2024).
- [8] Tian, T., Chen, X., Liu, Z., Huang, Z., & Tang, Y. (2024). *Enhancing Organizational Performance: Harnessing AI and NLP for User Feedback Analysis in Product Development*. *Innovations in Applied Engineering and Technology*, 3(1), 1–15. <https://doi.org/10.62836/iaet.v3i1.203>