

# Sentiment Analysis Of Consumer Behavior

Bhoogarv Maheshwary<sup>1</sup>, Dhruv Vashishtha<sup>1</sup>, Raghav Kashyap<sup>1</sup>, Dhananjay Kumar<sup>1</sup>, Prof. Sahana M<sup>2</sup>

1. Department of Artificial Intelligence and Machine Learning, Bangalore Institute of Technology, Bangalore, India
2. Assistant Professor, Department of Artificial Intelligence and Machine Learning, Bangalore Institute of Technology, Bangalore, India

**Abstract** - In a rapidly changing digital world, consumers express themselves through vast amounts of textual data, highlighting the significance of understanding their sentiments. The paper introduces a eclectic sentiment analysis system designed to provide actionable insights to businesses by analyzing consumer reviews from e-commerce platforms such as Amazon. Leveraging an ensemble learning model comprising Random Forest, XGBoost, LightGBM, and Linear Regression algorithms, the system extracts insightful information on customer attitudes and preferences. Through a multi-stage process encompassing data acquisition, preprocessing, sentiment analysis, and recommendation generation, the system offers personalized recommendations tailored to individual user profiles. The usefulness of the suggested approach is demonstrated by experimental findings, and prospective improvements, such as real-time flexibility and sophisticated machine learning approaches, are considered. The paper also highlights the significance of sentiment analysis in informing strategic decision-making and driving business success in today's digital marketplace.

**Keywords** – Sentiment Analysis, Consumer Behavior, Customer Reviews, Business Intelligence.

## I. INTRODUCTION

"Sentiment Analysis of Consumer Behavior" isn't just software; it's a strategic asset for business owners. The web application functions as an electronic ear, analyzing customer reviews to understand precisely how customers feel about products. It then translates these sentiments into actionable insights, providing personalized recommendations tailored to each business's niche and region. With clear explanations for each suggestion and a feedback mechanism for continual improvement, businesses can refine their offerings, enhance customer satisfaction, and build brand loyalty. These results ensure businesses stay ahead of the curve, perpetually positioning them at the forefront of market dynamics. Leveraging customer reviews enables businesses to gain candid insights into consumer experiences, preferences, and grievances, facilitating informed decision-making. By comprehensively analyzing this feedback, businesses can refine their strategies, optimize product offerings, and strengthen their competitive edge. The symbiotic relationship between businesses and their clientele, fostered by sentiment analysis, underpins sustained growth and resilience in today's dynamic marketplace.

## II. LITERATURE SURVEY

Customer satisfaction is a critical metric for businesses to understand and improve. In order to assess customer satisfaction, a number of recent research investigate the use of machine learning and sentiment analysis techniques to customer review data analysis.

**Machine Learning Models:** A variety of machine learning models are used for sentiment analysis, including Random Forest [4], XGBoost [1], LightGBM [1], Support Vector Machines and deep learning architectures [3][4]. Ensemble models, which combine multiple models for improved accuracy and robustness, are also used [1].

**Data Acquisition:** Data for analysis comes from various sources, including customer reviews on e-commerce platforms [5][13], and customer satisfaction surveys [9]. Web scraping methods are frequently employed to gather data from online sources.

**Data Preprocessing:** Preprocessing steps are crucial to ensure data quality and improve model performance. Common techniques include removing outliers [1], handling missing data, and text normalization (e.g., stemming, lemmatization) which are not explicitly mentioned but likely used in many studies.

**Sentiment Analysis:** The core function of the system is sentiment analysis, which classifies text data into positive, negative, or neutral categories. Machine learning models are trained on labeled data sets to perform this task [3].

**Data Enrichment:** After sentiment analysis, data can be further enriched by categorizing it by product, service, or other relevant factors [1].

**Customer Satisfaction Scoring:** The sentiment analysis results are often used to calculate a customer satisfaction score, which reflects the overall customer sentiment towards a product, service, or brand [2].

**Recommendation Generation:** Several studies highlight the use of sentiment analysis to generate recommendations for businesses. These recommendations can be related to product development, marketing strategies, or customer service improvements [12].

In handling missing values, data pre-processing techniques typically involve removal or imputation. Removal entails discarding instances or attributes with missing data, but it's practical only for datasets with few missing values and can introduce bias. Imputation methods, widely applied across domains, replace missing values statistically or via machine learning. In the application, we'll explore various imputation methods to improve the precision of sentiment analysis. For instance, we'll consider K-Nearest Neighbors imputation, shown to outperform other methods regardless of missing value percentages, especially in numerical datasets.[2]

Mhd Ridwan AlHabbal [5] explored various models, including Decision Trees, Logistic Regression, Naïve Bayes, and Artificial Neural Networks (ANN), highlights the importance of employing a range of techniques to achieve accurate predictions. However, the identification of Random Forest as the most effective predictor particularly caught our attention.

This finding motivated us to incorporate ensemble learning techniques, such as Random Forest, into the sentiment analysis system. By leveraging the strengths of multiple algorithms, as demonstrated in the paper, we aim to improve the precision and robustness of the sentiment analysis model.

### III. EXISTING SYSTEM

Existing sentiment analysis systems that are used by platforms like Hootsuite, Brandwatch, and Sprout Social, employ machine learning algorithms and natural language processing techniques to analyze customer feedback across digital platforms and surveys. However, these systems face challenges with limited accuracy due to difficulties in interpreting nuanced language, sarcasm, and complex sentiments, often resulting in misclassifications. Additionally, reliance on keyword-based approaches can lead to a lack of context comprehension in longer text, while handling mixed sentiments within a single piece of text remains a significant challenge. These drawbacks underscore the need for advancements in sentiment analysis to accurately capture and interpret diverse and nuanced customer expressions across various digital channels.

### IV. PROBLEM STATEMENT

To build a comprehensive sentiment analysis system that deciphers consumer sentiments from extensive data sources, providing actionable insights on consumer behavior through a user-friendly interface.

### V. PROPOSED METHODOLOGY

The proposed interactive web application provides an innovative solution in the realm of sentiment analysis by implementing an ensemble learning model that amalgamates the strengths of four distinct algorithms: XGBoost, LightGBM, Linear Regression, and Random Forest [1][5]. This ensemble approach is devised to significantly enhance the accuracy and robustness of sentiment analysis while simultaneously mitigating the limitations inherent in individual algorithms, such as overfitting or bias. By leveraging the diverse capabilities of these algorithms, the solution aims to provide a more comprehensive and nuanced understanding of consumer sentiment, thereby enabling businesses to make more informed decisions.

One of the paramount challenges in analyzing online consumer behavior is its dynamic nature, characterized by rapid shifts and evolving trends. To address this challenge, the ensemble model is intended to be highly adaptable and responsive. It achieves this by continuously learning from newly scraped data sourced from the web, allowing it to stay abreast of emerging sentiments in real-time. This adaptability guarantees that the sentiment analysis remains relevant and accurate, even in the face of constantly changing consumer preferences and market dynamics.

Moreover, the application encompasses the creation of a user-friendly web application, tailored specifically for business owners seeking to harness the meaning of sentiment analysis. The application provides an intuitive interface, allowing users to effortlessly access real-time consumer sentiment insights and actionable recommendations. By democratizing access to this valuable information, the web application empowers businesses to make data-driven decisions, optimize their strategies, and potentially increase profitability. In essence, the

solution represents a holistic approach to sentiment analysis, integrating advanced algorithms with user-friendly technology to deliver tangible benefits for businesses in today's dynamic marketplace.

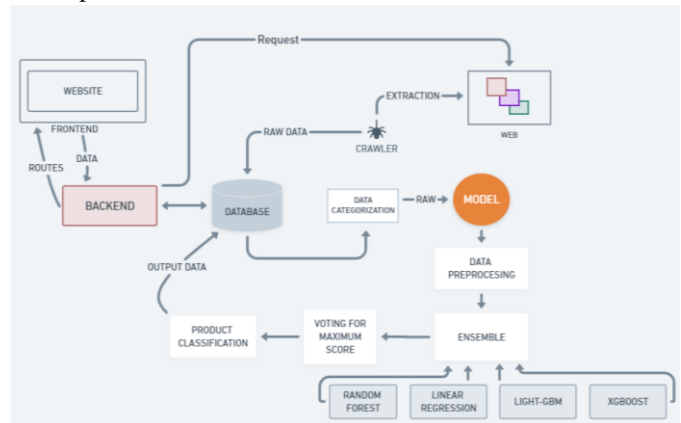


Fig 1. Architecture of the proposed system

#### A. DATA ACQUISITION

In the data acquisition phase, the system gathers essential inputs crucial for the sentiment analysis process. User profile data and product data serve as foundational components, providing insights into the preferences and characteristics of the target audience. Moreover, to enrich the dataset and capture real-time consumer sentiments, a web crawler is deployed. This crawler scours the web, meticulously extracting relevant data from a multitude of sources, including e-commerce platforms like Amazon and various social media channels. For instance, the system can ingest the URL of a specific product from Amazon, enabling it to retrieve and analyze reviews, comments, and discussions related to that product. This comprehensive approach ensures that the sentiment analysis is based on a diverse and representative dataset, reflecting the varied perspectives and opinions of consumers across different platforms.

#### B. DATA PREPROCESSING

Once the data is collected, it undergoes a rigorous preprocessing stage to ensure its quality and reliability [11][13]. This involves a series of data cleaning and refinement techniques aimed at filtering out noise, removing outliers, and standardizing the format of the dataset. Irrelevant or erroneous data points are identified and eliminated, while missing values are imputed or interpolated where necessary. Additionally, text data undergoes preprocessing steps such as tokenization, stop-word removal, and stemming to extract meaningful features and enhance the accuracy of sentiment analysis. By meticulously preparing the dataset, the system lays the foundation for robust and insightful analyses in subsequent stages.

#### C. SENTIMENT ANALYSIS

The system's central lies in its sophisticated sentiment analysis capabilities, driven by a machine learning model that uses the collective power of multiple algorithms. This ensemble approach, integrating RandomForest, XGBoost, LightGBM [1][4], and Linear Regression algorithms, is specifically designed to improve the precision and robustness

of sentiment analysis. Each algorithm contributes unique strengths, enabling the system to effectively capture subtle nuances in consumer sentiment while mitigating individual limitations such as overfitting or bias [1]. The sentiment analysis process involves classifying the data that has been prepped into distinct sentiment categories, including positive, negative, or neutral. By systematically analyzing the sentiments expressed in reviews, comments, and social media posts, the system generates valuable insights into consumer perceptions and attitudes towards the products or user profiles under scrutiny. Following sentiment analysis, the system enriches the dataset by categorizing the analyzed data based on the sentiments expressed. This categorization adds depth and granularity to the dataset, facilitating a nuanced understanding of consumer sentiment. By discerning patterns and trends within each sentiment category, the system provides actionable insights that can inform strategic decision-making and drive business outcomes. The culmination of the sentiment analysis process is the delivery of actionable insights to stakeholders. This involves calculating a consumer satisfaction score, which serves as a holistic measure of the overall sentiment towards the products or user profiles. Derived from sentiments expressed in reviews, comments, and social media posts, this score provides a comprehensive overview of consumer sentiment, ranging from positive to negative. Additionally, the system leverages the analyzed data to further refine and enhance the sentiment analysis model over time, contributing to its ongoing improvement and effectiveness.

#### D. RECOMMENDATION GENERATION.

Considering the consumer satisfaction score and sentiment analysis results, the system generates personalized recommendations tailored to the specific needs and preferences of the intended target audience [5][12]. These recommendations highlight areas of improvement for the products or user profiles, offering actionable insights derived from the analysis of consumer feedback. By identifying strengths, weaknesses, and opportunities for enhancement, these recommendations empower businesses to refine their products, optimize their marketing strategies, and ultimately enhance customer satisfaction and loyalty [5][12]. Through a continuous feedback loop, the system iteratively refines its recommendations, ensuring relevance and effectiveness in dynamic market conditions.

#### VI. FUTURE ENHANCEMENTS

While the current iteration of the system represents a significant advancement in sentiment analysis and recommendation generation, several avenues for future enhancements and refinements exist:

1. **Advanced Machine Learning Techniques:** Modern machine learning approaches can be further explored and integrated to improve sentiment analysis's efficiency and accuracy. Research on deep learning models, advances in natural language processing, and reinforcement learning algorithms may lead to the discovery of new information and functionalities.

2. **Multimodal Analysis:** Incorporating multimodal data sources, such as images, videos, and audio, alongside textual data, can provide a more thorough comprehension of consumer sentiment. Integrating multimodal analysis techniques can enrich the sentiment analysis process and offer deeper insights into consumer perceptions.
3. **Real-time Adaptability:** Implementing mechanisms for real-time adaptation and learning can enable the system to dynamically adjust its recommendations based on evolving consumer sentiments and market trends. Incorporating streaming data processing techniques and adaptive learning algorithms can ensure that the system remains responsive and relevant in dynamic environments.
4. **Continuous Feedback Loop:** Establishing a robust system of feedback that lets users to provide feedback on the recommendations produced by the apparatus can facilitate continuous improvement and refinement. By incorporating user feedback into the recommendation generation process, the model can iteratively enhance its accuracy and effectiveness over time.

#### VII. RESULT

The outcome demonstrates a novel approach for analyzing consumer sentiment and providing actionable insights to business owners based on product reviews from the e-commerce platform Amazon. The system takes the link of the product from Amazon as input and outputs the sentiment score of the product, derived from the sentiments expressed in the comments and reviews. Additionally, smart recommendations are also inculcated, suggesting changes in the product that would assist business owners in making improvements.

To demonstrate the effectiveness of the system, experiments were conducted on a variety of products across different categories on Amazon. The system successfully extracted and analyzed the sentiment of the products, providing sentiment scores that accurately reflected the overall consumer perception. The sentiment scores ranged from 0 to 100 percentage, providing a comprehensive understanding of how consumers perceive the product.

In addition, the section with intelligent suggestions offered insightful suggestions on how to make the products better. These recommendations, which emphasize certain features or components of the product that should be improved to better suit customer wants and preferences, were developed from the analysis of consumer feedback. This information can be used by business owners to customize their offerings and marketing plans, which will ultimately increase client happiness and revenues.

The web application demonstrates the feasibility and effectiveness of using sentiment analysis on e-commerce platforms to provide valuable insights to business owners. By leveraging consumer feedback, businesses can make informed decisions to enhance their products and better meet the needs of their target audience, ultimately driving success in the competitive marketplace.



### VIII. CONCLUSION

The "Sentimental Analysis of Consumer Behavior" has achieved significant progress in developing a comprehensive solution for businesses seeking data-driven insights, focusing on creating a user-friendly web application. The backend system incorporates a dynamic online learning machine learning (ML) model utilizing a sophisticated ensemble of algorithms to extract and process reviews from ecommerce platforms and consumer mentions from social media using a web scraper. The sentiment analysis tool provides nuanced ratings for every product, forming the basis for actionable recommendations such as product prioritization, avoidance, and identifying similar products with higher sentiment scores in specific regions. This initial phase lays the groundwork for a cutting-edge solution empowering businesses to make informed decisions based on real-time consumer sentiment, with further refinement and enhancement anticipated to deliver a powerful tool for boosting profitability through data-driven strategies.

**Competing Interests:** Not Applicable

**Funding Information:** Not Applicable

#### Author Contribution:

Bhoogarv Maheshwary: Conceptualization, Methodology  
 Dhruv Vashishtha: Data Curation, Software, Validation  
 Raghav Kashyap: Formal Analysis, Investigation  
 Dhananjay Kumar: Visualization, Project Administration  
 Prof. Sahana M: Supervision, Writing - Review & Editing

#### Data Availability Statement:

The data that support the findings of this study are available from the corresponding author upon reasonable request.

#### Research Involving Human and/or Animals:

Not Applicable

**Informed Consent:** Not Applicable

### IX. REFERENCES

- [1] Fei Zeng, Yuqing He, Chengqin Yang, Xinkai Hu, Yining Yuan (2023) Mobile Customer Satisfaction Scoring Research Based on Quadratic Dimension Reduction and Machine Learning Integration. *Applied Sciences* 13(17):9681
- [2] CARL F.W. HÖGGREN: Predicting Customer Satisfaction in the Context of Last-Mile Delivery using Supervised and Automatic Machine Learning *Mathematics* 2023, 11(6), 1482
- [3] Majed A. Alshamari (2023): Evaluating User Satisfaction Using Deep-Learning-Based Sentiment Analysis for Social Media Data in Saudi Arabia's Telecommunication Sector *Computers* 2023, 12(9), 170
- [4] Amjad Iqbal, Rashid Amin, Javed Iqbal, Roobaea Alroobaea, Ahmed Binmahfoudh, and Mudassar Hussain (2022): Sentiment Analysis of Consumer Reviews Using Deep Learning. *Sustainability* 2022, 14(17), 10844
- [5] Md. Ridwan AlHabbal (2022): Predicting & Optimizing Airlines Customer Satisfaction Using Classification Research paper from Rochester Institute of Technology.
- [6] Nisreen Ameen, Ali Tarhini, Alexander Reppel: Customer experiences in the age of artificial intelligence. *Computers in Human Behaviour* (Volume 114, January 2021, 106548)
- [7] Samer Arqawi (2022): Customer Satisfaction Prediction using Artificial Intelligence *Mathematics*, vol. 2022
- [8] Yu-Che Wang, Yu-Cheng Lee, Yi-Fang Hsieh (2021) Empirical research on customer satisfaction study: a consideration of different levels of performance. *SpringerPlus* 5, 1577
- [9] Tyler Doll, Matthew Bussing, Kai Nichols, Sidney Johnson: A Machine Learning Approach to Automated Customer Satisfaction Surveys Towards Data Science

- [10] Yolande Piris, Anne-Cécile Gay (2021): Customer satisfaction and natural language processing *Journal of Business Research* Volume 124, January 2021
- [11] Zita Bošnjak, Olivera Grljevic (2020): Sentiment analysis of customer data: *Strategic Management* (2020):38-49
- [12] Mohamed Zaki, Janet R. McColl-Kennedy, Andy Neely (2019): Gaining Customer Experience Insights That Matter *Harvard Business Journal*
- [13] B. Rajeswari, S. Madhavan, Ramakrishnan Venkatesakumar and S. Riasudeen (2020) Sentiment analysis of consumer reviews – a comparison of organic and regular food products usage - *Rajagiri Management Journal* Vol. 14 No. 2, 2020
- [14] Sachin Kumar, Mikhail Zymbler (2019). A machine learning approach to analyze customer satisfaction from airline tweets. *Big Data* volume 6.
- [15] Dr.M.Kotteeswari (2019): A Study on Consumer Behavior in Selection of Personal Care Products Based on their Financial status. (2021) *IJCRT* Volume 9, Issue 4.