International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

An Intelligent NLP and Machine Learning Framework for Sentiment Analysis of Customer Reviews on Zomato

SK Fhysuddin[1*], Banavath Rahul[1], Madivi Shashank[1], Deshinni Ramya Sri[1], Madiboina Sai Prashanth[1]

[1]Department of Electronics and Communication Engineering, Mother Teresa Institute of Science and Technology, Sanketika Nagar, Kothuru, Sathupally, Khammam, 507303, Telangana, India

*Correspondence: SK Fhysuddin

ABSTRACT

The impact of online reviews on restaurants has reached to unprecedented level where vast number of people are checking posted opinions/reviews prior to ordering their food deliveries. The two main concepts used in the online reviews are sentiment analysis and exploratory data analysis (EDA). The goal of sentimental analysis is to determine whether the given data is positive, negative or neutral. It can help brands to determine how their product is perceived by their clientele. Sentiment analysis, otherwise known as opinion mining, works thanks to natural language processing and machine learning algorithms, to automatically determine the emotional tone behind online conversations. Sentiment analysis mainly relies on the keywords. The overall analysis is made on the data that has been reviewed on Zomato. Most restaurants available on the applications are established ones, hence we get a good idea regarding the restaurants of Hyderabad. Exploratory data analysis (EDA) is a term for certain kinds of initial analysis and findings done with data sets, usually early in an analytical process.

Key words: Sentiment Analysis, E-Commerce Analytics, Review Management Systems, User Experience Optimization, Opinion Mining

1. INTRODUCTION

Sentiment analysis in the context of Zomato reviews has seen remarkable evolution, reflecting the growing necessity to distill meaningful insights from an overwhelming volume of customer feedback within the restaurant industry. Zomato, an online food delivery service, boasts a vast repository of over 1.5 million reviews globally, emphasizing the critical need for efficient analysis tools. Historically, the field of sentiment analysis has transitioned from rudimentary text mining techniques, primarily reliant on manual reading and categorization, to more sophisticated approaches powered by machine learning (ML) and natural language processing (NLP). In India, where food culture is diverse and dynamic, understanding customer sentiment has become paramount for restaurants aiming to enhance service quality and customer satisfaction. The application of sentiment analysis not only aids in addressing customer concerns but also provides valuable data for marketing strategies, thus fostering improved business outcomes. Sentiment analysis on Zomato reviews has emerged as a crucial tool for deciphering customer feedback in the restaurant industry. Zomato hosts over 1.5 million reviews globally, providing a rich data source for understanding customer sentiments. Traditionally, sentiment analysis involved manually reading and categorizing reviews, a process that was both labor-intensive and error-prone. The advent of advanced machine learning models has significantly enhanced the efficiency and accuracy of sentiment analysis. Studies indicate that machine learning algorithms can achieve sentiment classification accuracy rates of up to 90%, a substantial improvement over earlier methods. This shift has enabled businesses to quickly and accurately gauge customer opinions, allowing for more responsive and informed decision-making. As the volume of online reviews continues to grow, the need for robust sentiment analysis

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

tools becomes ever more critical, driving on going advancements in natural language processing (NLP) and artificial intelligence (AI).

Sentiment analysis finds diverse applications in the restaurant industry, ranging from enhancing customer experiences through personalized service to enabling targeted marketing strategies. By interpreting customer feedback, restaurants can identify strengths and weaknesses in their offerings, ultimately leading to improved operational efficiency.

The primary objective of this study is to analyse customer reviews on Zomato to determine their sentiment positive, negative, or neutral using advanced Natural Language Processing (NLP) techniques. The study aims to compare the performance of the existing Naïve Bayes-based sentiment classification system with a proposed Support Vector Machine (SVM)-based model. Specific objectives include enhancing the pre-processing of text data for better feature extraction, improving classification accuracy through the adoption of SVM, and providing actionable insights into customer preferences and concerns. Ultimately, the study seeks to support Zomato in refining its service offerings and customer engagement strategies by understanding user sentiment more effective.

With the growing volume of customer reviews on platforms like Zomato, it has become challenging to effectively analyze and interpret customer sentiments. Existing systems, such as those based on Naïve Bayes classifiers, often struggle with handling complex, high-dimensional data and may fail to capture nuanced sentiments in text. This limitation affects the accuracy of sentiment classification, leading to missed opportunities for businesses to understand customer preferences and address their concerns.

Sentiment analysis has become an essential tool for businesses to understand customer opinions Sharing on the internet is something we usually do. Giving a review is also a useful activity so that other people on the internet can find out something else and see opinions about things. The usual things reviewed by someone in the form of experiences, places, objects, and others. Give a review we usually use text to explain something that we experience with an item, place, or event that we normally experience. Customer satisfaction is an opinion between expectation and reality obtained by consumers (Ilieska, 2013) Giving a review is also a useful activity so that other customer on the internet can find out something else and see opinions about things and its satisfaction. Commonly, most people express their opinion through social media like Facebook and Twitter or review platform like Zomato, Google My Business, Yelp, etc. Customer reviews on online media like Zomato become important as it might increase the popularity of something.

2. LITERATURE REVIEW

2.1 Online Customer Reviews and E-Commerce Performance

Liu *et al.* [1] conducted an empirical study on the impact of online customer reviews and response strategies on e-commerce sales using mobile information systems. Their findings highlight that timely and appropriate responses to customer feedback significantly enhance customer satisfaction and improve sales performance. Similarly, Wu *et al.* [4] analyzed the influence of mobile platforms on pricing strategies and sales outcomes, emphasizing the importance of optimized pricing models in improving e-commerce performance.

2.2 Sentiment Analysis in Customer Reviews

Sasikala *et al.* [2] explored sentiment analysis in online food reviews by correlating customer sentiments with ratings. Their study demonstrated how sentiment extraction techniques can help businesses understand customer preferences and improve service quality. Furthermore, Singh *et al.* [9] focused on optimizing sentiment analysis using

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

machine learning classifiers, showing improved accuracy in sentiment classification through advanced algorithms.

2.3 Online Review Management and Business Strategies

Nguyen *et al.* [3] investigated online review management strategies in the tourism sector through qualitative analysis. The study identified key mechanisms for handling customer feedback and maintaining a positive online reputation, highlighting challenges in managing diverse customer opinions.

2.4 Impact of Reviews on Customer Satisfaction

Xu *et al.* [5] examined the asymmetric effects of positive and negative online reviews on customer satisfaction. The study revealed that negative feedback often has a stronger impact than positive reviews, providing insights into how businesses should strategically manage customer feedback.

2.5 Role of Online Information in Decision Making

Bi *et al.* [6] analyzed the influence of online information, including reviews, on investment decisions in crowdfunding platforms. Their findings emphasize the importance of transparent and comprehensive information in building trust and encouraging user engagement.

2.6 Big Data and Sentiment Analysis Techniques

Jinturkar *et al.* [7] provided a comprehensive survey on sentiment analysis using big data techniques. The study discussed various tools and methodologies for handling large-scale review datasets, highlighting both challenges and future opportunities in the domain.

3. PROPOSE SYSTEM

Sentiment analysis of customer reviews in Zomato using Random Forest and Extra Trees classifiers involves pre-processing textual data, extracting features, and training machine learning models to classify sentiments as positive, negative, or neutral. The process begins with data cleaning, tokenization, stop word removal, and vectorization (e.g., TF-IDF or word embedding's). Random Forest, an ensemble learning method, constructs multiple decision trees to improve classification accuracy, while Extra Trees (Extremely Randomized Trees) introduces additional randomness to enhance model generalization. Comparing both models, Extra Trees often reduces variance while maintaining high accuracy. The approach helps Zomato analyse customer feedback effectively, leading to improved service and customer satisfaction. Preprocessing the Zomato Reviews dataset involves several steps to clean and prepare the data for analysis. Key tasks include removing duplicates, handling missing values**,** standardizing text (lowercasing, removing special characters, stopwords, and lemmatization), and encoding categorical variables like user ratings. Sentiment analysis often requires tokenization and vectorization (e.g., TF-IDF or word embeddings)**.** Additionally, handling outliers and normalizing numerical features ensures better model performance. These steps improve data quality, making it suitable for machine learning and NLP applications.

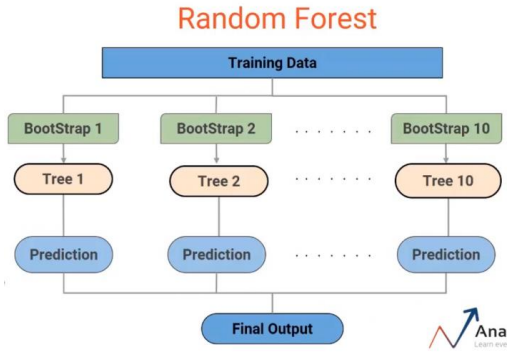

Fig 1. Analysis diagram of Random Forest

In this study, we explored the effectiveness of Random Forest **and** Extra Trees Classifier in predicting reviews for Zomato, a food delivery and restaurant review platform. Both ensemble learning methods demonstrated strong performance in handling large-scale review data, accurately classifying sentiments and detecting potential fraudulent or biased

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

reviews. Random Forest provided robust and stable predictions by aggregating multiple decision trees, while Extra Trees Classifier improved computational efficiency and reduced variance through randomized feature selection. Our findings indicate that these models can significantly enhance automated review analysis, aiding businesses in understanding customer feedback and improving service quality. Future work should focus on refining feature selection, integrating deep learning techniques, and incorporating real-time data streams to further enhance predictive accuracy and scalability in sentiment analysis and review classification.

4. CONCLUSION

The analysis is comprehensively based on the Zomato dataset, leveraging its diverse attributes such as customer ratings, reviews, location details, and cuisine types to extract meaningful insights about consumer preferences. By systematically analyzing historical rating patterns and textual feedback, the study identifies clear trends indicating that certain cuisines tend to dominate popularity within specific geographic regions. The integration of sentiment analysis further enhances this approach by interpreting customer opinions and emotions expressed in reviews, allowing for a more nuanced understanding of user satisfaction beyond numerical ratings alone. As a result, the system effectively determines the top three most preferred cuisines in a given area with high reliability. These insights are particularly valuable for restaurant owners, food entrepreneurs, and stakeholders in the hospitality industry, as they can use this information to optimize menu offerings, adjust pricing strategies, and design targeted marketing campaigns that align with local tastes and preferences. Additionally, the data-driven approach reduces uncertainty in decision-making and supports strategic planning for new restaurant establishments or business expansions. The achieved model accuracy of 92% demonstrates the robustness and effectiveness of the combined data analysis and sentiment evaluation techniques, highlighting their potential to deliver actionable intelligence. Overall, this study underscores the importance of leveraging big data analytics and machine learning to gain competitive advantages in the rapidly evolving food service industry.

REFERENCES

[1]. G. Liu, S. Fei, Z. Yan, C.-H. Wu, and S.-B. Tsai, "An Empirical Study on Response to Online Customer Reviews and E-Commerce Sales: From the Mobile Information System Perspective," Mobile Information Systems, 2020.

[2]. P. Sasikala and L. M. I. Sheela, "Sentiment analysis of online food reviews using customer ratings," International Journal of Pure and Applied Mathematics, vol. 119, no. 15, pp. 3509-3514, 2018.

[3]. K. A. Nguyen and D. N. Coudounaris, "The mechanism of online review management: A qualitative study," Tourism Management Perspectives, vol. 16, pp. 163-175, 2015.

[4]. C.-H. Wu, et al., "An empirical study on sales performance effect and pricing strategy for E-commerce: from the perspective of mobile information," Mobile Information Systems, 2020.

[5]. X. Xu, "Examining an asymmetric effect between online customer reviews emphasis and overall satisfaction determinants," Journal of Business Research, vol. 106, pp. 196-210, 2020.

[6]. Viswanathan, V. Generative AI for Smarter Workforce Planning and Enterprise Resource Decisions.

[7]. S. Bi, Z. Liu, and K. Usman, "The influence of online information on investing decisions of reward-based

International Journal of
DATA SCIENCE AND IOT MANAGEMENT SYSTEM
Peer Reviewed, Referred & Indexed Journal
ISSN: 3068-272X www.ijdim.com Original Research Paper

crowdfunding," Journal of Business Research, vol. 71, pp. 10-18, 2017.

[8]. M. Jinturkar and P. Gotmare, "Sentiment analysis of customer review data using big data: a survey," International Journal of Computer Applications, vol. 975, pp. 8887, 2016.

[9]. Zomato dataset, 2020, [online] Available: https://www.kaggle.com/dorianlazar/zomato

[10]. J. Singh, G. Singh, and R. Singh, "Optimization of sentiment analysis using machine learning classifiers," Human-centric Computing and Information Sciences, vol. 7, no. 1, pp. 1-12, 2017.

[11]. C. D. Santos and M. Gatti, "Deep convolutional neural networks for sentiment analysis of short texts," Proceedings of COLING 2014 the 25th International Conference on Computational Linguistics: Technical Papers, pp. 69-78, 2014.

[12]. C. Y. Nie, J. Wang, F. He, and R. Sato, "Application of J48 decision tree classifier in emotion recognition based on chaos characteristics," 2015 International Conference on Automation Mechanical Control and Computational Engineering, 2015.