JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# Beyond the Stars : Unveiling Customer Sentiment in Restaurant Reviews

[1]Freny Babaria, [2]Darshan Kolhe, [3]Samiksha Raut, [4]Soham Pal
Vidyalankar Institute of Technology , Mumbai -Wadala , India

*Abstract :* In the digital age, internet reviews are critical to restaurant success. However, manual analysis of these reviews is ineffective and prejudiced. This research investigates sentiment analysis, an automated technique used to overcome these restrictions and get deeper customer insights. We propose a sentiment analysis project to determine customer sentiment in restaurant reviews. This research has substantial advantages over previous methods due to its use of Natural Language Processing (NLP) and machine learning. By automating sentiment analysis, restaurants can acquire important insights from customer comments, allowing them to make data-driven decisions to improve customer happiness and gain a competitiveadvantage.

***IndexTerms – Sentimental Analysis, Restaurant reviews, Natural Language Processing, Text Classification***

## I INTRODUCTION

In the modern digital world, feedback from customers are incredibly effective for businesses, particularly in the competitive restaurant market [1]. They serve as a virtual word-of-mouth, influencing potential customers and developing a restaurant's image. Restaurants frequently rely on manual searches of review websites, social media platforms, and even physical comment cards [2,3]. However, the conventional method of manually gathering and assessing these reviews has shortcomings. This fragmented approach not only takes valuable time away from essential operations, but also risks missing out on critical feedback distributed over the web [4]. Sifting through innumerable written reviews is atime-consuming and arduous operation, which frequentlyresults in inadequate data sets [5]. Furthermore, human analysis is prone to prejudice and inconsistency. For example, reviewers may employ sarcasm or comedy that a human reader might miss, resulting in misinterpretations of the meaning [6]. Human reviews may also unintentionally focuson keywords or favor bad events, creating an imperfect picture of consumer happiness. Furthermore, the sheer volume of reviews can make it difficult to spot repeating themes or trends in the data [7-8]. This lack of efficiency and accuracy prevents restaurants from utilizing vital consumer feedback to improve their services and overall customer experience [9]. The lack of easily accessible information from manual analysis impedes effective decision-making.

Restaurants require actionable intelligence to improve their services, handle client problems, and maximize their offerings. Manual techniques frequently fail to produce clear and succinct reports, making it difficult to turn client feedback into actionable improvements. The static nature of manual review analysis restricts its usefulness in a changing sector. Customers' preferences and trends change frequently,and restaurants require real-time feedback to adjust. Manual approaches are simply too sluggish to keep up with the market's constantly changing needs. The current state ofmanual review gathering, and analysis is a huge challenge forrestaurants looking to harness the power of customer feedback. By embracing digitization, restaurants can gain valuable insights from the large ocean of online evaluations. Automated systems may speed data collecting, eliminate human bias in analysis, and provide actionable data that can catapult a restaurant to success. As the business evolves, restaurants' ability to leverage the power of digital sentimentanalysis will become an important distinction in thecompetitive landscape.

In today's digital age, online restaurant evaluations have become an essential component of the food service industry. Customers are increasingly turning to networks such as Yelp, Zomato, and Google Reviews to discuss their positive and negative experiences. This amount of data provides an excellent opportunity for restaurants to better understand consumer sentiment, identify areas for development, and ultimately improve the dining experience [10]. Exploratory data analysis (EDA) is critical to uncovering the insights concealed within these reviews. EDAapproaches can help restaurants obtain a better understandingof their customers, the language they use to describe their experiences, and the elements that determine their overall satisfaction. This research goes beyond merely looking at star ratings and delves into the written content of the reviews to identify repeating themes and trends [11]. Beyond internal changes, EDA on restaurant evaluations might provide usefulinformation for industry research. Trends in consumer expectations, dining habits, and the evolving terminology used to describe food and service can be detected by evaluating a larger dataset that includes multiple locations. This information can help to inform industry-wide best practices and the creation of better customer service models in the restaurant business.

The sentiment analysis project for restaurant reviews seeks to automatically identify the emotional tone and opinions offered by customers. This entails processing the reviews' textual content using Natural Language Processing (NLP) techniques. NLP activities such as tokenization, stemming/lemmatization, and negation handling will be used to prepare the text data for analysis. Machine learning algorithms, like Naïve Bayes or Support Vector Machines (SVM), can categorize reviews as good, negative, or neutral depending on their sentiment. This classification enables the automated study of enormous datasets of evaluations, providing restaurants with important information about customer satisfaction and areas for improvement [12]

## II LITERATURE SURVEY

Since the early 2000s, analyzing consumer sentiment in restaurant reviews has been a major area of research in Natural Language Processing (NLP). This sectiondigs into the research timeline, outlining major breakthroughs and approaches during the last two decades. Pang and Lee (2008) were among the first to publish in this field. Their research focused on developing a lexicon-based sentiment analysis system for film reviews. This method used pre- defined sentiment dictionaries comprising positive and negative words. They used this method to assess online restaurant evaluations, demonstrating the utility of sentiment classification for customer input in the food service business[13].

Building on this foundation, Thet et al. (2010) investigated the use of supervised machine learning algorithms to sentiment analysis of restaurant reviews. The study compared the performance of classifiers like Naïve Bayes, Support Vector Machines (SVM), and MaximumEntropy models. Their findings indicated that SVM was the most accurate in categorizing reviews as favorable, negative, or neutral. This study highlighted the potential of machine learning for automated sentiment analysis, opening the way for future developments [14].

As social media platforms grew in popularity in thelate 2000s and early 2010s, researchers began looking into sentiment analysis on restaurant reviews obtained through these platforms. Hu and Liu (2004) published a landmark work on opinion mining and sentiment analysis in social media content. Their work addressed the issues raised by informal language used on these platforms, such as slang, acronyms, and emoticons. By including these elements into their sentiment analysis model, they improved the accuracy of sentiment classification for social media evaluations, particularly those about restaurants [15].The mid-2010s saw an increase in the use of deep learning algorithms for sentiment analysis. Wang et al. (2016) suggested a Convolutional Neural Network (CNN)-based technique for sentiment analysis in brief texts, such as onlinereviews. Their methodology outperformed established machine learning methods, especially for brief and informal evaluations commonly found on social media networks. Thisstudy demonstrated the power of deep learning to handle theintricacies of restaurant review language [16].

More recently, academics have concentrated on improving sentiment analysis tools to capture subtle aspects of the consumer experience that go beyond simple positive, negative, or neutral categories. Liu et al. (2020) developed a multi-aspect sentiment analysis (MASA) framework for restaurant reviews. Their methodology went beyond general sentiment categorization to identify sentiment related to specific characteristics of the eating experience, such as foodquality, service, and ambiance. This fine-grained analysis gives restaurants more specific information into areas for improvement [17].The contemporary research landscape is constantly exploring new approaches. Studies such as Luo et al. (2023) investigate the use of transformers, a sophisticated deep learning architecture, for sentiment analysis in restaurant reviews. These upgraded models show much more accuracy in understanding the sentiment represented in reviews [18].

## III. METHODOLOGY

The methodology applied for the working of the code and the project can be summarized as shown in Figure 1.

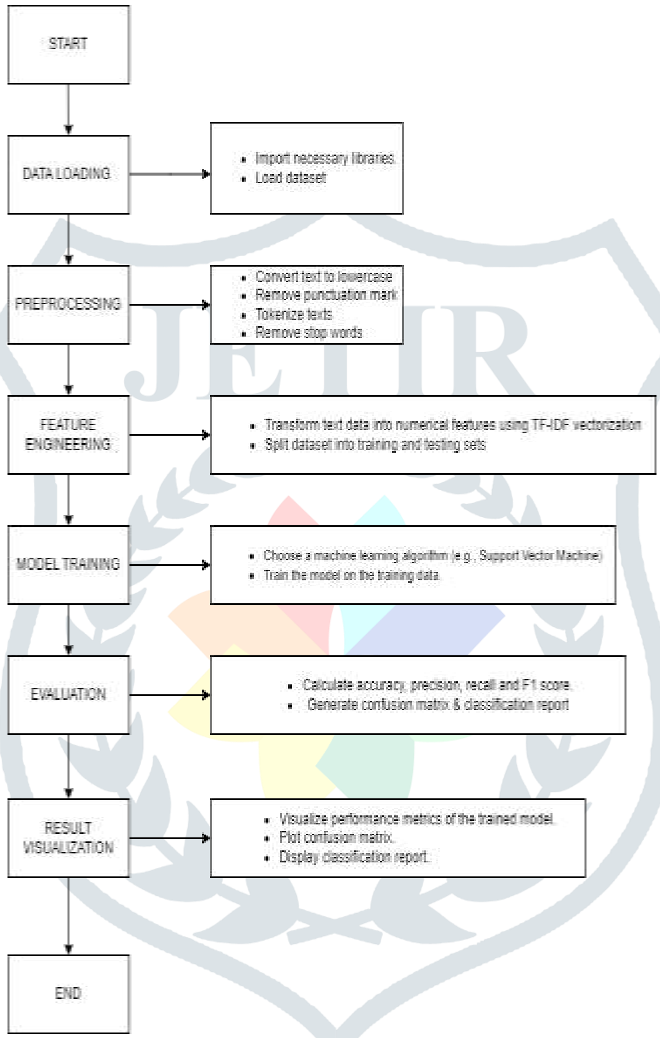


Figure 1: Working of the project

1. Data Loading and Preprocessing:

- The code begins by importing necessary libraries and loading the dataset from using Pandas.
- Next, it preprocesses the data by converting text to lowercase, removing punctuation, and applying tokenization.
- It also removes stop words to focus on meaningful words that contribute to sentiment analysis.

2. Feature Engineering:

- The text data is transformed into numerical features using the TF-IDF vectorization technique, which converts text documents into vectors of numerical values.
- This process helps in preparing the data for training machine learning models by representing textual information in a format suitable for analysis.

3. Model Training and Evaluation:

- The code splits the dataset into training and testing sets using a standard train-test split.
- It then trains a Support Vector Machine (SVM) classifier and a Naïve Bayes classifier on the training data to predict the sentiment labels (positive, negative, or neutral) associated with each review.
- After training, the model is evaluated on the testing data to assess its performance using metrics such as accuracy, precision, recall, and F1-score.

4. Results Visualization:

- Finally, the code visualizes the performance metrics of the SVM classifier & the Naïve Bayes classifier using a confusion matrix and classification report.
- These visualizations provide insights into the model's ability to correctly classify reviews into their respective sentiment categories.

## IV. RESULTS

| | Review | | | |
|---|---|---|---|---|
| | count | unique | top | freq |
| Liked | | | | |
| 0 | 500 | 497 | The food was terrible. | 2 |
| 1 | 500 | 499 | I love this place. | 2 |

Figure 2. Reviews after grouping

| | Review | Liked | Length |
|---|---|---|---|
| 0 | Wow... Loved this place. | 1 | 24 |
| 1 | Crust is not good. | 0 | 18 |
| 2 | Not tasty and the texture was just nasty. | 0 | 41 |
| 3 | Stopped by during the late May bank holiday of... | 1 | 87 |
| 4 | The selection on the menu was great and so wer... | 1 | 59 |

Figure 3. Printing first 5 rows of messages Data Frame

```
count     1000.000000
mean        58.315000
std         32.360052
min         11.000000
25%         33.000000
50%         51.000000
75%         80.000000
max        149.000000
Name: Length, dtype: float64
```

Figure 4. Generating summary statistics for “Messages”


Figure 5. Word cloud of combined messages

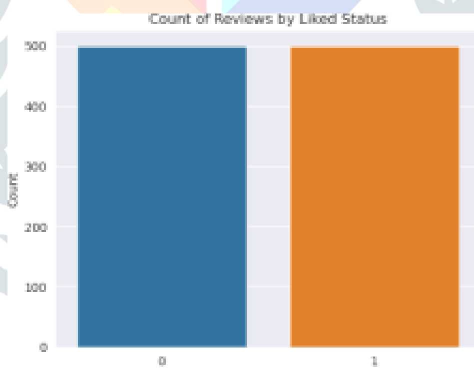


Figure 6.Word cloud of combined messages

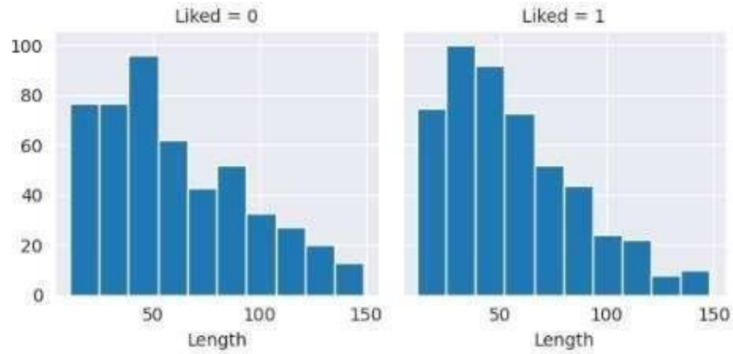


Figure 7. Histogram of liked status

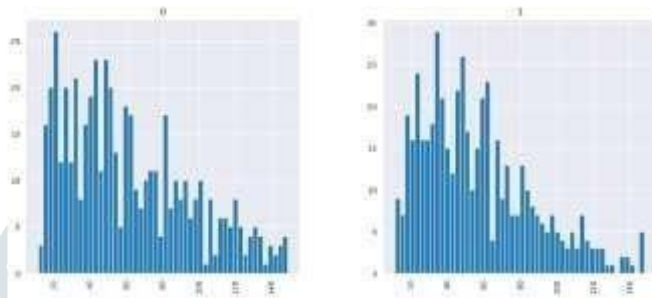

Figure 8. Histogram of length of messages as per category

| | Liked | Length |
|---|---|---|
| Liked | 1.000000 | -0.075285 |
| Length | -0.075285 | 1.000000 |

Table 1. Correlation values

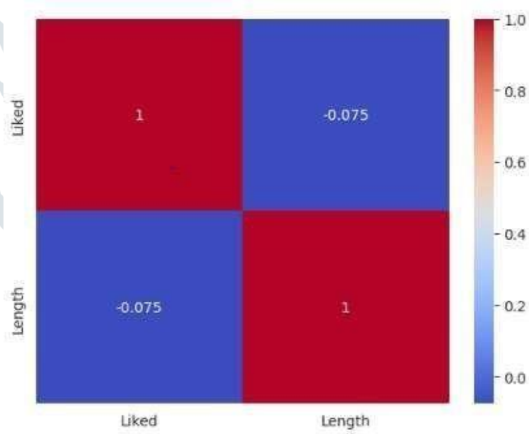


Figure 9.Heatmap of correlation values

| | |
|---|---|
| 115 | 35 |
| 30 | 120 |

Table 2.Confusion Matrix

| | precision | recall | f1-score | support |
|---|---|---|---|---|
| 0 | 0.79 | 0.77 | 0.78 | 150 |
| 1 | 0.77 | 0.00 | 0.79 | 150 |
| accuracy | | | 0.78 | 300 |
| macro avg | 0.78 | 0.78 | 0.78 | 300 |
| weighted avg | 0.78 | 0.78 | 0.78 | 300 |

Table 3. Classification Report

## V. CONCLUSION

In conclusion, this research used sentiment analysis to acquire a thorough grasp of client comments on restaurant experiences. Restaurants may boost customer happiness and loyalty by automatically categorizing reviews and assessing sentiment toward key components. Furthermore, detecting repeating patterns and comparing sentiment with competitors enables strategic decision-making to optimize strengths, rectify flaws, and achieve a competitive advantage in the marketplace. Finally, tracking sentiment changes over time allows restaurants to assess the impact of changes and constantly improve their products to meet changing client preferences.

## REFERENCES

[1] Stabiner, K. (2017) To Survive, Restaurants Turn to Data -Mining. The New York Times Company. https://www.nytimes.com/2017/08/25/dining/restaurant-softwareanalytics-data-mining.html
[2] Kumar, Dhiraj & Gopesh, & Choubey, Avinash & Singh, Pratibha. (2020). Restaurant Review Classification and Analysis.10.13140/RG.2.2.32874.34249.
[3] "Restaurant Review Using Sentiment Analysis in Social Media", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.10, Issue 5, page no.p72-p77, May-2023, Available:http://www.jetir.org/papers/JETIR2305G10.pdf
[4] Alhubaishy, Abdulaziz & Aljuhani, Abdulmajeed.(2021). The Influence of Information Sharingthrough Social Network Sites on Customers'Attitudes during the Epidemic Crisis of COVID-19.Journal of Theoretical and Applied ElectronicCommerce Research. 16. 1390-1403. 10.3390/jtaer16050078.
[5] Alday, Ramnhell & Rosas, Maryli. (2019). BusinessIntelligence Solution for Bikers Haven Restaurant.1204- 1210. 10.1109/UEMCON47517.2019.8992956.
[6] Liu, B. (2012). Sentiment analysis and opinion mining. Synthesis Lectures on Human Language Technologies3(1),1-167
[7] Westerman, George, Didier Bonnet, and Andrew McAfee. Leading digital: Turning technology into business transformation. Harvard Business Press, 2014.
[8] Westerman, George, Didier Bonnet, and Andrew McAfee. "The nine elements of digital transformation." MIT Sloan Management Review 55, no. 3 (2014): 1-6.
[9] Watson, Hugh & Goodhue, Dale & Wixom, Barb. (2002). The benefits of data warehousing: Why some organizations realize exceptional payoffs. Information & Management. 39. 491-502. 10.1016/S0378- 7206(01)00120-3.
[10] Liu, Jun, Yunyun Yu, Fuad Mehraliyev, Sike Hu, and Jiaqi Chen. "What affects the online ratings of restaurant consumers: a research perspective on text-mining big data analysis." International Journal of Contemporary Hospitality Management 34, no. 10 (2022): 3607-3633.
[11] M. Huang, H. Xie, Y. Rao, Y. Liu, L. K. M. Poon and F. L. Wang, "Lexicon-Based Sentiment Convolutional Neural Networks for Online Review Analysis," in IEEE Transactions on AffectiveComputing, vol. 13, no. 3, pp. 1337-1348, 1 July-Sept. 2022.
[12] Reddy, Kothapally Nithesh, and P. Indira Reddy. "Restaurant Review Classification Using Naives Bayes Model." J Univ Shanghai Sci Technology (2021).
[13] Pang, B., & Lee, L. (2008). Opinion mining and sentiment analysis.
[14] [14] Thet, T. M. T., Aung, H. L., & Myo, K. T. (2010). Sentiment analysis of customer reviews for hotel services using support vector machine.
[15] [15] Hu, M., & Liu, (2004). Mining and summarizing customer reviews.
[16] [16] Wang, Y., Huang, M., Lv, L., & Zhu, X. (2016). Attention-based LSTM for aspect-level sentiment analysis.
[17] Liu, B., Yu, J., Zhang, Y., & Huang, X. (2020). A hierarchical attention networks for aspect-based sentiment analysis in food reviews.
[18] Li, Hengyun & Yu, Bruce & Li, Gang & Gao, Huicai. (2023). Restaurant survival prediction using customer- generated content: An aspect-based sentiment analysis of online reviews.Tourism Management