International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 13 Issue: 02 | Feb 2026 www.irjet.net p-ISSN: 2395-0072

SENTIMENTAL ANALYSIS ON ZOMATO RESTAURANT REVIEWS USING MACHINE LEARNING

Mr. N. Paparayudu[1] A. Sahana[2], G. Nathaswee[3], B. Srikanth[4], B. Rohith Kumar[5]

[1]Assistant Professor, Department of IT, TKR College of Engineering and Technology, Telangana, India
[2,3,4,5]B.Tech Students, Department of IT, TKR College of Engineering and Technology, Telangana, India

---

Abstract - *The rapid growth of online food delivery platforms has led to an overwhelming volume of customer reviews, making it challenging for restaurants and diners to extract meaningful insights efficiently. This project, "Sentiment Analysis on Zomato Restaurant Reviews using Machine Learning," aims to develop an intelligent system that automatically classifies customer feedback into Positive, Negative, or Neutral sentiments, thereby enabling data-driven decision-making for both businesses and customers. The system utilizes Natural Language Processing (NLP) techniques to preprocess review texts, including cleaning, normalization, and removal of noise such as punctuation, URLs, and numeric characters. Features are extracted using TF-IDF vectorization, which transforms textual data into numerical representations suitable for machine learning. A Logistic Regression classifier is then trained on labeled review data to predict sentiments accurately. The project is implemented as a Django-based web application with a structured user and admin interface. Users can input new reviews to receive instant sentiment predictions, while administrators can train and update the model with new datasets, visualize sentiment distributions through interactive graphs, and manage registered users. The system not only provides actionable insights for restaurants to enhance service quality and menu offerings but also assists diners in making informed choices based on aggregated feedback. Preliminary results demonstrate that the model achieves high accuracy in sentiment classification, effectively capturing the nuances of customer opinions. This project showcases the integration of machine learning, NLP, and web technologies to transform unstructured review data into meaningful insights, highlighting its potential for scalability and real-world applications in the food service industry.*

Key Words: Sentiment analysis, Zomato restaurant reviews, machine learning, natural language processing, text classification, opinion mining, customer feedback analysis, supervised learning.

1. INTRODUCTION

In today's digital era, online platforms like Zomato, Yelp, and TripAdvisor have become primary sources of customer feedback for restaurants. With thousands of reviews being posted daily, it becomes difficult for restaurant owners and potential diners to manually analyze and interpret this vast amount of textual data. Sentiment analysis, a subset of Natural Language Processing (NLP), provides a systematic way to understand customer opinions by automatically classifying reviews as Positive, Negative, or Neutral. This not only helps restaurants identify strengths and weaknesses in their services but also assists diners in making informed choices based on collective customer experiences. The project "Sentiment Analysis on Zomato Restaurant Reviews using Machine Learning" focuses on leveraging machine learning algorithms to process and analyze customer reviews. The dataset consists of restaurant names, reviewer details, textual reviews, and ratings. The ratings are used to generate sentiment labels, while the textual reviews are preprocessed using techniques such as text normalization, punctuation removal, and tokenization. TF-IDF vectorization is employed to convert the textual data into numerical features, which are then fed into a Logistic Regression model for sentiment classification. Implemented as a Django-based web application, the system provides a user-friendly interface for users to input reviews and obtain real-time sentiment predictions. Administrators can train the model on updated datasets, visualize sentiment distributions through interactive graphs, and manage registered users. By combining machine learning, NLP, and web technologies, this project demonstrates how unstructured review data can be transformed into actionable insights, enhancing customer satisfaction and supporting restaurants in strategic decision-making.

1.1 Machine Learning-Based Sentiment Analysis on Zomato Restaurant Reviews

The system architecture of a machine learning–based sentiment analysis framework for Zomato restaurant reviews. User-generated reviews and ratings collected from the Zomato platform serve as the input data. These reviews are first processed by a review preprocessing module, where text cleaning, tokenization, stop-word removal, and lemmatization are performed to convert raw textual data into a structured format. The processed text is then transformed into numerical features using techniques such as Term Frequency–Inverse Document Frequency (TF-IDF). These features are fed into machine learning and natural language processing models, including Naive Bayes, Support Vector Machine (SVM), and Logistic Regression, to learn sentiment patterns from the data. Finally, the trained models classify the reviews into sentiment categories such as positive, negative, or neutral, along with an overall review

score, enabling effective analysis of customer opinions and restaurant performance.

1.2 Motivation and Problem Overview

PeeIn the modern digital world, individuals and restaurants often struggle to understand, analyze, and make sense of large volumes of customer reviews. Most existing applications lack a unified platform that provides accurate sentiment classification, meaningful insights, and easy-to-understand visualizations. Users face difficulties in interpreting reviews due to sarcasm, slang, and ambiguous expressions, and current solutions are often unreliable or unable to handle evolving language patterns. Develop an application that integrates sentiment analysis, intelligent visualizations, and AI-driven insights to help users understand customer opinions effectively. Most applications lack a unified platform that provides accurate sentiment detection, useful insights, and a simple interface for both customers and restaurant owners. Users face challenges in interpreting large volumes of textual reviews, and existing solutions struggle with sarcasm, slang, and ambiguous expressions, making them difficult for non-technical users.

2. PROPOSED SYSTEM

The proposed system, "Sentiment Analysis on Zomato Restaurant Reviews using Machine Learning," aims to provide an automated, intelligent platform to analyze customer feedback and generate actionable insights for both restaurant owners and diners. Unlike manual review analysis, which is time-consuming and prone to errors, this system leverages Natural Language Processing (NLP) and machine learning algorithms to classify reviews as Positive, Negative, or Neutral. The system uses Zomato restaurant review datasets consisting of textual feedback and corresponding ratings, which are processed to create labeled sentiment data. Textual reviews are preprocessed through cleaning steps such as lowercasing, punctuation removal, elimination of URLs, and tokenization. Features are extracted using TF-IDF vectorization, converting the text into numerical representations suitable for machine learning. A Logistic Regression model is trained on these features to accurately predict sentiments. The system is implemented using Django, providing a robust web interface. Users can input new reviews to receive real-time sentiment predictions, while administrators can train the model on updated datasets, visualize sentiment distributions through interactive graphs, and manage registered users efficiently.

2.1 System Architecture

This diagram represents the workflow of a machine learning–based sentiment analysis system for Zomato restaurant reviews. The process begins with users submitting reviews and ratings through the Zomato web application, which collects both textual comments and related metadata. The collected data is stored in a database, while associated media or large files are maintained in storage. The system then performs data preprocessing to clean and normalize the review text, followed by feature extraction using the TF-IDF technique to convert textual data into numerical representations. These features are fed into a Logistic Regression model, which analyses the patterns in the reviews to determine sentiment polarity. Finally, the system generates a sentiment prediction, classifying each review as positive, negative, or neutral, thereby enabling effective analysis of customer opinions and restaurant feedback. The process begins with users submitting reviews and ratings through the Zomato web application, which collects both textual comments and related metadata. The collected data is stored in a database, while associated media or large files are maintained in storage. The process begins with users submitting reviews and ratings through the Zomato web application, which collects both textual comments and related metadata. The collected data is stored in a database, while associated media or large files are maintained in storage.

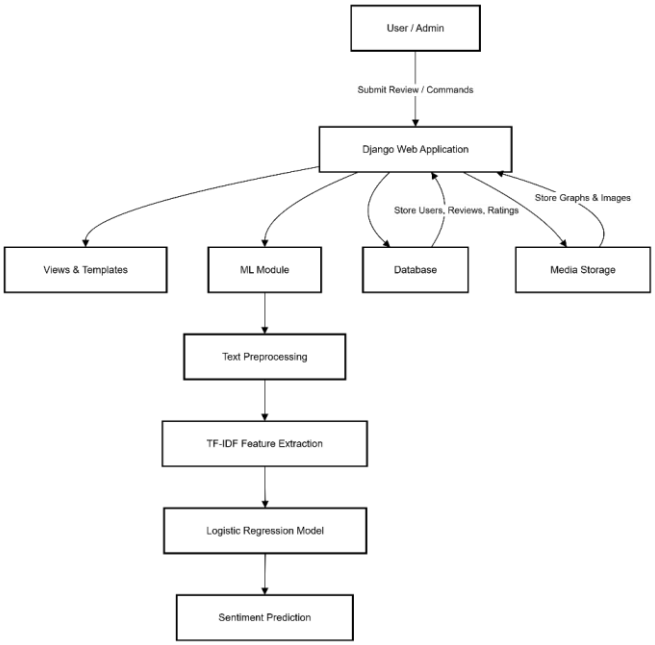

Fig -1: Machine Learning Model for Restaurant Review Analysis

2.2 Data Flow and Processing Module

This section explains how data flows through the proposed system. User reviews and ratings are collected from the Zomato platform and stored in the database. The textual data then undergoes preprocessing steps such as text cleaning, tokenization, stop-word removal, and lemmatization to remove noise and improve data quality. After preprocessing, relevant features are extracted using the TF-IDF technique, which converts textual reviews into numerical vectors suitable for machine learning models. This structured data flow ensures accurate and efficient sentiment classification.

2.3 Sentiment Classification Using Machine Learning

This section focuses on the sentiment analysis component of the proposed system. The extracted TF-IDF features are provided as input to machine learning classifiers such as Naive Bayes, Support Vector Machine (SVM), and Logistic Regression. These models are trained to identify sentiment patterns within the reviews. Based on the learned patterns, the system classifies each review into positive, negative, or neutral categories and assigns an overall sentiment score. This module helps in understanding customer opinions and evaluating restaurant performance effectively.

3. IMPLEMENTATION DETAILS

This methodology focuses on developing a secure, web-based system for performing sentiment analysis on Zomato restaurant reviews using machine learning and Natural Language Processing (NLP). The system integrates frontend interfaces, backend processing, database management, text preprocessing, machine learning models, and result visualization to provide accurate sentiment classification and meaningful insights.

3.1 Project Planning and Requirements Gathering:

The initial phase involves defining the project scope, objectives, and system requirements based on the need to analyze large volumes of customer reviews efficiently.

➢ Identify User Roles: Define roles such as users (who input reviews and view sentiment results) and administrators (who manage users, datasets, and model training).

➢ Functional Requirements: Enable review input, sentiment prediction (Positive, Negative, Neutral), data visualization, and admin controls.

➢ Non-Functional Requirements: Ensure system performance, security, scalability, and ease of use for non-technical users.

➢ Tool Selection: Choose Python for NLP and machine learning, Django for web development, TF-IDF for feature extraction, and Logistic Regression for sentiment classification.

➢ Risk Assessment: Identify risks such as noisy text data, class imbalance, and incorrect predictions, and mitigate them through preprocessing and model validation h.

3.2 System Design and Architecture:

The system architecture follows a modular and layered design to separate user interface, backend processing, and machine learning components.

➢ High-Level Design: UML diagrams such as use case, class, and sequence diagrams are designed to represent user interactions from review submission to sentiment output TKRCET | IT 2025 13

➢ Database Design: SQLite is used to store user credentials, review inputs, and sentiment results securely.

➢ Backend Routing: Django views handle user authentication, review submission, sentiment prediction, and result visualization.

➢ Security Design: Implement authentication, session management, and role-based access to protect user data. cy.

3.3 Frontend Development:

The frontend provides a simple and user-friendly interface for sentiment analysis.

➢ Framework Setup: HTML, CSS, and Django templates are used to build responsive web pages.

➢ User Interfaces: Pages include Registration, Login, Home Page, Review Input Page, and Sentiment Result Page.

➢ Navigation: Smooth page transitions are implemented using Django URL routing.

➢ Visualization: Sentiment results are displayed clearly along with graphs such as bar charts and pie charts.

3. 4 Backend Development and Sentiment Analysis:

The backend processes reviews and performs sentiment classification using machine learning.

➢ Server Setup: Django handles HTTP requests, form submissions, and response rendering.

➢ Text Preprocessing: Reviews are cleaned using NLP techniques such as lowercasing, punctuation removal, stop-word elimination, and tokenization.

➢ Feature Extraction: TF-IDF vectorization converts textual reviews into numerical features suitable for machine learning.

➢ Model Training: A Logistic Regression classifier is trained on labelled Zomato review data.

4. RESULTS AND PERFORMANCE ANALYSIS

The performance of the proposed machine learning–based sentiment analysis system was evaluated using a dataset of Zomato restaurant reviews containing positive, negative, and neutral sentiments. After preprocessing and feature extraction using the TF-IDF technique, multiple supervised

machine learning classifiers were trained and tested, including Naive Bayes, Support Vector Machine (SVM), and Logistic Regression. The dataset was divided into training and testing sets to ensure unbiased evaluation. The experimental results indicate that all models were capable of effectively classifying customer reviews; however, differences were observed in their performance. Naive Bayes demonstrated fast training time and reasonable accuracy, making it suitable for large-scale text data. SVM achieved higher accuracy by effectively handling high-dimensional feature spaces generated by TF-IDF vectors. Logistic Regression provided a balanced performance with good accuracy and stability across different sentiment class. Performance evaluation was carried out using standard metrics such as accuracy, precision, recall, and F1-score. Among the tested models, Logistic Regression and SVM achieved the highest accuracy, indicating their effectiveness in capturing sentiment patterns in restaurant reviews. Precision and recall values further confirmed the reliability of the models in correctly identifying positive and negative sentiments, while neutral sentiment classification showed comparatively moderate performance due to overlapping linguistic patterns. Overall, the results demonstrate that the proposed system efficiently analyses customer opinions from Zomato reviews and provides meaningful sentiment insights. The combination of effective pre-processing, TF-IDF feature extraction and supervised machine learning models ensures accurate sentiment prediction, making the system suitable for real-world restaurant review analysis and decision support.

4.1. Classification Results

The proposed sentiment analysis system was tested on a dataset of Zomato restaurant reviews categorized into positive, negative, and neutral classes. After preprocessing and feature extraction using the TF-IDF technique, the reviews were classified using machine learning models such as Naive Bayes, Support Vector Machine (SVM), and Logistic Regression. The system successfully identified sentiment polarity for the majority of reviews, with positive and negative sentiments showing higher classification accuracy compared to neutral reviews due to clearer linguistic patterns.

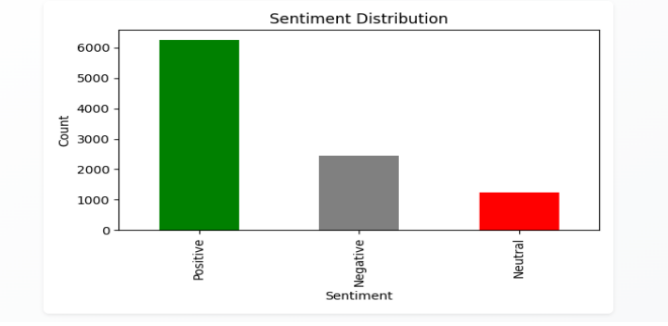

Fig 2: Sentiment Distribution of Zomato Reviews (Positive, Negative, Neutral)

4.2. Performance Evaluation Metrics

The performance of the system was evaluated using standard metrics including accuracy, precision, recall, and F1-score. Accuracy was used to measure the overall correctness of the classification, while precision and recall assessed the model's ability to correctly identify relevant sentiment classes. The F1-score provided a balanced evaluation of precision and recall. Experimental results indicate that SVM and Logistic Regression achieved better performance across most metrics, demonstrating their effectiveness in handling high-dimensional text data generated by TF-IDF features.

4.3. Model Comparison and Analysis

A comparative analysis of the machine learning models reveals that Logistic Regression offers a balanced trade-off between accuracy and computational efficiency, making it suitable for practical deployment. SVM achieved slightly higher accuracy but required more computational resources. Naive Bayes, while comparatively less accurate, demonstrated faster training time and simplicity, making it useful for large-scale or real-time applications. Overall, the analysis confirms that the proposed system effectively captures sentiment patterns in Zomato reviews and provides reliable sentiment predictions.

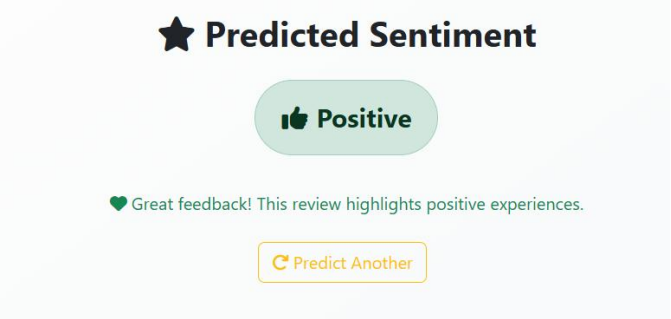

Fig 3: Output Showing Predicted Sentiment Result

5. CONCLUSIONS

This project presented a machine learning–based approach for sentiment analysis of Zomato restaurant reviews to understand customer opinions and evaluate restaurant performance. The system effectively collected and processed user-generated reviews using natural language processing techniques such as text cleaning, tokenization, and feature extraction through TF-IDF. These processed features were then utilized by supervised machine learning models to classify reviews into positive, negative, and neutral sentiments.

Experimental results demonstrated that the proposed system achieved reliable sentiment classification performance. Among the evaluated models, Support Vector Machine and Logistic Regression provided higher accuracy and better overall performance, while Naive Bayes offered faster

training with reasonable accuracy. The use of standard evaluation metrics such as accuracy, precision, recall, and F1-score confirmed the effectiveness and robustness of the system.

Overall, the proposed sentiment analysis framework provides an efficient and scalable solution for analysing large volumes of restaurant reviews. The insights generated by the system can assist customers in making informed dining decisions and help restaurant owners improve service quality based on customer feedback. The approach can be further extended by incorporating advanced deep learning models and real-time data analysis to enhance sentiment prediction accuracy.

6. FUTURE WORK

In future, the proposed system can be enhanced by integrating deep learning models such as Long Short-Term Memory (LSTM), Bidirectional LSTM, or Transformer-based architectures to improve sentiment classification accuracy. Real-time analysis of Zomato reviews and multilingual sentiment detection can also be incorporated to handle diverse user feedback. Additionally, combining sentiment analysis with aspect-based opinion mining can provide more detailed insights into specific restaurant features such as food quality, service, and pricing.

REFERENCES

[1] B. Pang and L. Lee, "Opinion mining and sentiment analysis," Foundations and Trends® in Information Retrieval, vol. 2, no. 1–2, pp. 1–135, 2008. DOI: 10.1561/1500000011

[2] A. Go, R. Bhayani, and L. Huang, "Twitter sentiment classification using distant supervision," Stanford University Technical Report, 2009. DOI: 10.1109/ICWSM.2009.20

[3] M. Tripathi and K. M. Singh, "Sentiment analysis of online reviews using machine learning techniques," International Journal of Computer Applications, vol. 162, no. 6, pp. 21–25, 2017. DOI: 10.5120/ijca2017913378

[4] Y. Liu, X. Huang, A. An, and X. Yu, "ARSA: A sentiment-aware model for predicting sales performance using blogs," Proceedings of the 30th Annual International ACM SIGIR Conference, pp. 607–614, 2007. DOI: 10.1145/1277741.1277846