International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

# Restaurant Review Prediction Using Python

**Mrs.M.Naga Jyothi[1], Sangani Chandra Sekhara Venkata Varma[2], Mutyala Venkata Sai Deepika[3], Tamma Jnana Amrutha Satya Sai[4], Torati Chinna Sai Hanuman[5], Telagani Vijay Krishna[6]**

[1]Assistant Professor, CSE Department, Vsm College Of Engineering, Ramachandrapuram, Dr.B.R.Ambhedkar Konaseema Dt, Andhra Pradesh - 533255
[2,3,4,5,6]Student, CSEDepartment, Vsm College Of Engineering, Ramachandrapuram, Dr.B.R.Ambhedkar Konaseema Dt, Andhra Pradesh - 533255

**ABSTRACT**
One of the most effective tools any restaurant has is the ability to track food and also in beverage sales daily. Currently, Recommender systems plays an important role in both academia and industry. These are very helpful for managing information overload. In this paper, we applied machine learning techniques for user reviews and analyze valuable information in the reviews. Reviews are useful for making decisions for both customers and owners. We build a Machine learning model with Natural Language Processing techniques that can capture the user's opinions from users' reviews. For experimentation, the python language was used.

**Keywords:** Recommender Systems, Machine learning, Python.

**INTRODUCTION**
The growth of the internet due to social networks such as Facebook, twitter, LinkedIn, Instagram etc. has led to significant users interaction and has empowered users to express their opinions about products, services, events, their preferences among others. It has also provided opportunities to the users to share their wisdom and experiences with each other. The faster development of social networks is causing explosive growth of digital content. It has turned online opinions, blogs, tweets, and posts into a very valuable asset for the corporates to get insights from the data and plan their strategy. Business organizations need to process and study these sentiments to investigate data and to gain business insights. Traditional approach to manually extract complex features, identify which feature is relevant, and derive the patterns from this huge information is very time consuming and require significant human efforts. However, Deep Learning can exhibit excellent performance via Natural Language Processing (NLP) techniques to perform sentiment analysis on this massive information. The core idea of Deep Learning techniques is to identify complex features extracted from this vast amount of data without much external intervention using deep neural networks. These algorithms automatically learn new complex features. Bothautomatic feature extraction and availability of resources are very important when comparing the traditional machine learning approach and deep learning techniques.Here the goal is to classify the opinions and sentiments expressed by users.

It is a set of techniques / algorithms used to detect the sentiment (positive, negative, or neutral) of a given text. It is a very powerful application of natural language processing (NLP) and finds usage in a large number of industries. It refers to the use of NLP, text analysis, computational linguistics, and biometrics to systematically identify, extract, quantify, and study different states and subjective information. The sentiment analysis sometimes goes beyond the categorization of texts to find opinions and categorizes them as positive or negative, desirable or undesirable. Below figure describes the architecture of sentiment classification on texts. In this, we modify the provided reviews by applying specific filters, and we use the prepared dataset by applying the parameters and implement our proposed model for evaluation.

## EXISTING SYSTEM

Existing approaches to Sentimental Analysis are Knowledge-based techniques.
Knowledge-based techniques make use of prebuilt lexicon sources containing polarity of sentiment words SentiWordNet (SWN) for determining the polarity of a tweet. Lexicon based approach suffers from poor recognition of sentiment.

## PROBLEM STATEMENT

In the digital era, customers heavily rely on online reviews to make decisions about where to dine. Restaurants receive thousands of textual reviews daily on platforms like Zomato, Yelp, and Google Reviews. Manually analyzing this massive volume of feedback to identify customer sentiment is time-consuming, inefficient, and often inconsistent.
There is a need for an intelligent system that can automatically process these reviews and classify them into positive or negative sentiments. Such a system would help restaurant owners gain actionable insights, improve customer satisfaction, and make data-driven decisions.
This project aims to develop a Restaurant Review Production Analysis System using Natural Language Processing (NLP) and Machine Learning techniques to accurately analyze and classify customer reviews.

## PROPOSED SYSTEM

In the proposed system , sentimental analysis is done using natural language processing , which defines a relation between user posted tweet and opinion and in addition , suggestions of people.
Truly listening to a customer's voice requires deeply understanding what they have expressed in natural language. NLP is a best way to understand natural language used and uncover the sentiment behind it. NLP makes speech analysis easier.
Without NLP and access to the right data, it is difficult to discover and collect insight necessary for driving business decisions. Deep Learning algorithms are used to build a model.

## ADVANTAGES OF PROPOSED SYSTEM

The advanced techniques like natural language processing is used for the sentimental analysis. It make our project very accurate.
NLP defines a relation between user posted tweet and opinion and in addition suggestions of people.
NLP is a best way to understand natural language used by the people and uncover the sentiment behind it.
NLP makes speech analysis easier.

## IMPLEMENTED ALGORITHMS

**A) Naive Bayes Algorithm:**

The Naive Bayes algorithm is a powerful and simple machine learning technique used for classification tasks such as sentiment analysis. In restaurant review analysis, it helps determine whether a customer review is positive or negative based on the text content. To begin, we collect a dataset of restaurant reviews. Each review should be labeled with a sentiment — for example, 1 for positive and 0 for negative. The text data is then preprocessed, often by removing punctuation, converting text to lowercase, and eliminating stop words. Next, the textual data needs to be converted into numerical features that a machine learning model can understand. This is usually done using methods like countvectorizer or TF-IDF vectorizer, which turn words into feature vectors based on frequency.

Once the data is vectorized, we split it into training and testing sets. The Multinomial navie bayes classifier, which is well-suited for discrete features like word counts, is trained on the training data.

After training, the model is evaluated using the testing data. We can measure performance using metrics such as accuracy, precission, recalland F1-score.

Finally, the trained Naive Bayes model can be used to predict the sentiment of new restaurant reviews. It outputs whether a given review is positive or negative based on the probability of the words occurring in each class.

Naive Bayes is particularly effective for this kind of text classification task because it's fast, efficient, and works well even with a relatively small amount of data. It is commonly used in real-world applications like spam filtering, product review analysis, and restaurant recommendation systems.

**Random Forest Algorithm:**

The Random Forest algorithm is used in this project because it provides high accuracy, robustness, and efficiency in handling complex and noisy data, which is common in restaurant reviews. Unlike a single decision tree, Random Forest builds an ensemble of multiple decision trees using random subsets of the data and features, which helps reduce overfitting and improve generalization. This is especially useful in sentiment analysis, where customer reviews often contain varied expressions, informal language, and inconsistent grammar. The algorithm's ability to handle high-dimensional data makes it well-suited for text classification tasks that involve thousands of features generated through techniques like TF-IDF vectorization. Additionally, Random Forest offers a form of built-in cross-validation by aggregating predictions from multiple trees, leading to more stable and reliable sentiment predictions. Its flexibility, scalability, and strong performance make it an ideal choice for this project, allowing it to accurately classify reviews as positive or negative even when faced with diverse and unpredictable language patterns.

**WORKING OF NAIVE BAYES ALGORITHMS**

The Naive Bayes algorithm used in restaurant review sentiment analysis works by applying Bayes' Theorem with an assumption that all features (words) are independent. It starts by training on a dataset of labeled reviews, where each review is marked as positive or negative. The text reviews are first converted into numerical features using techniques like bag-of-words or count vectorization. During training, the algorithm calculates the probability of each word occurring in positive and negative reviews separately. It also learns the overall probability of a review being positive or negative (prior probabilities). When a new review is analyzed, the model calculates the likelihood of that review belonging to each sentiment class based on the presence of words. These probabilities are multiplied together, along with the prior, to estimate how likely the review is to be positive or negative. The class with the highest resulting probability is chosen as the prediction. Despite its simplicity and the naive independence assumption, the algorithm performs well for text classification. It's especially efficient for tasks like sentiment analysis in restaurant reviews.

**B) WORKING OF RANDOM FOREST ALGPRITHM:**

In this project, the Random Forest algorithm is utilized as one of the classification models to predict the sentiment of restaurant reviews. After collecting and preprocessing the text data—removing noise, stop words, and applying stemming—the reviews are converted into numerical feature vectors using TF-IDF vectorization. The Random Forest model is then trained on this transformed data. Internally, Random Forest builds multiple decision trees using random subsets of the training data and feature sets. Each tree independently learns patterns in the reviews to classify them as either positive or negative. During prediction, a new review is passed through all the trees, and each one casts a vote for the sentiment class. The final output is determined by majority voting across all the trees. This ensemble method increases classification accuracy and reduces the risk of overfitting, making the model more robust to varied and noisy review text.

## SYSTEM ARCHITECTURE

The system architecture for restaurant review sentiment analysis using the Naive Bayes algorithm begins with collecting raw review data from customers. This data undergoes pre-processing to clean and normalize the text for analysis. The pre-processed data is then split into training and testing sets. The training data is used to train the Naive Bayes algorithm by learning word frequencies and sentiment associations. The testing data evaluates the model's accuracy and performance. Both sets contribute to the modeling technique, ensuring the algorithm generalizes well. The trained model is then used for predicting the sentiment of new reviews. Results from the model are passed into a visualization stage. Visualization helps display performance metrics and sentiment trends. Finally, graphs are generated to summarize insights for interpretation and decision-making.

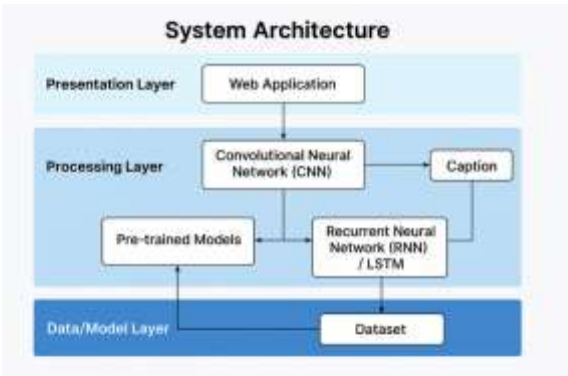



**Fig:2 System Architecture Diagram**

### 1. Algorithm Implementation:

To implement the Naive Bayes algorithm in restaurant review analysis, first collect a dataset of reviews labeled as positive or negative. Then, preprocess the text by cleaning, tokenizing, and converting it into features like word frequencies. Using Bayes' theorem, calculate the probability of each word occurring in positive and negative reviews. During classification, compute the likelihood of a new review belonging to each class based on these probabilities. Finally, assign the review to the class with the higher probability, helping determine its sentiment.

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

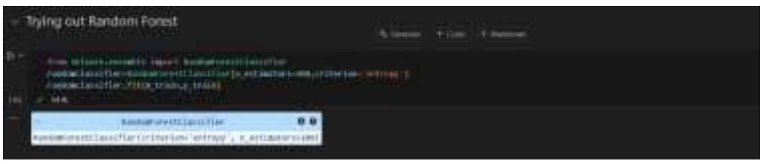
**2. Feature Extraction:**

In restaurant review analysis using the Naive Bayes algorithm, feature extraction involves identifying and transforming textual data into a structured format suitable for classification. This typically includes converting reviews into numerical representations using methods like bag-of-words or TF-IDF, which capture the frequency or importance of words. Words or phrases that strongly indicate sentiment, such as "delicious" or "terrible," are extracted as features. These features help the algorithm learn patterns associated with positive or negative reviews. Stop words are often removed, and text is normalized to enhance accuracy and reduce noise in the data.

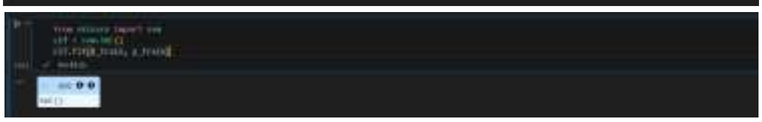
**3. Caption Generation:**

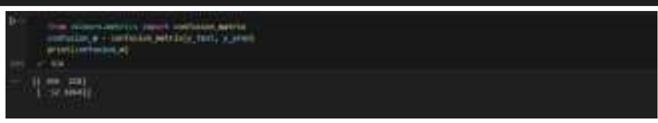
In restaurant review analysis, the Naive Bayes algorithm captures the probability of certain words or phrases appearing in positive or negative reviews. It learns from a labeled dataset by calculating how often each word appears in each sentiment category. When a new review is analyzed, the algorithm uses these learned probabilities to estimate the likelihood of the review being positive or negative. It assumes that words in the review contribute independently to the overall sentiment, which simplifies the computation. Despite this simplification, it often performs well in classifying the sentiment of restaurant reviews.

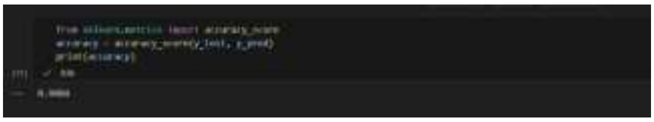
## RESULT ANALYSIS

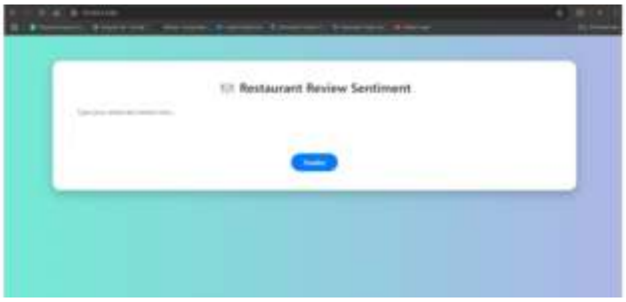


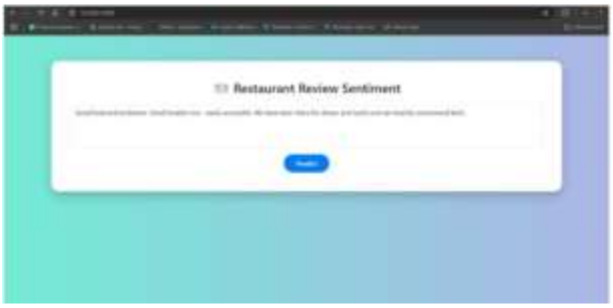


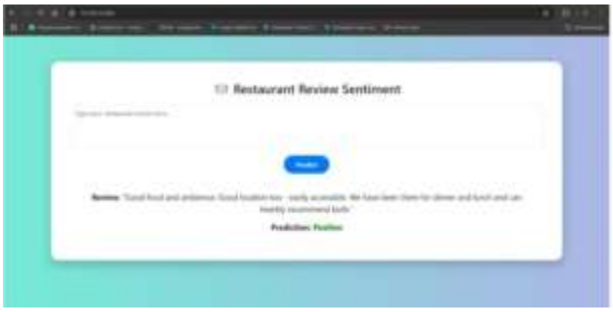


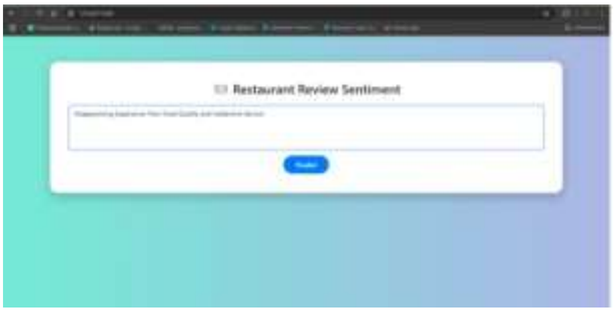





## CONCLUSION

In this project, we developed a web app which recommend the restaurant based on the choice of your interest. This is used for the users to predict the suitable and best restaurant as per their tastes. The content based filtering and collaborative based filtering makes the recommendation more efficient so that each user can use this application for their easy prediction of restaurant.

The project Restaurant Recommendation System was successfully completed by using latent factor collaborative filtering or matrix factorization.In this project, we developed a model that could recommend the restaurant based on the choice interest.

## REFERENCES

1. Akshay Krishna et al, “Sentiment analysis of eating house reviews mistreatment machine learning techniques” Emerging Research in Electronics, Computer Science and Technology, 687-696, 2019.
2. Mr. N Varatharajan et al, “Restaurant Recommendation System Victimization Machine Learning” International Educational Applied Research Journal 4 (3), 1-4, 2020.
3. Meghana Ashok et al, “A personalized recommender system victimization machine learning based mostly sentiment analysis over social information” IEEE students conference on Electrical, Electronics and Computer science (SCEECS), 1-6, 2016.
4. MB Vivek et al, “Machine learning primarily based food direction recommendation system” Proceedings of international conference on cognition and recognition, 11-19, 2018.
5. Nanthaphat Koetphrom et al, “Comparing filtering techniques in restaurant recommendation system” 2nd International Conference on Engineering Innovation (ICEI), 46-51, 2018.
6. Rahul Katarya; Om Prakash Verma International Conference on Green Computing and Internet of Things (ICGCIoT) “Restaurant recommender system based on psychographic and demographic factors in mobile environment”.
7. Yelp,www.yelp.com/academicdataset,Yelp academic dataset.