

**RESTAURANT REVIEW USING NLP**

**(Natural Language Processing)**

**By DEEPRA MAZUMDER**

**ABSTRACT**

Starting a business, such as a restaurant, is not easy. Reading all the evaluations and comparing the inputs with the other restaurants become challenging and time consuming for a restaurant owner. As a result of this project, it is an important factor to consider when making a dining selection. This project provides an efficient restaurant review prediction model to predict the reviews for the restaurant based on certain collections of client reviews. Restaurant reviews are labelled as 0 or 1, with 0 denoting a negative review and 1 denoting a positive one. The findings of this project could be employed as a long-term marketing strategy for review website creators, allowing customers to sort and filter useful reviews based on their preferences.

**OBJECTIVE**

Reviews help customers choose a business over other restaurant. [Customers heavily rely on reviews](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7503372/) when deciding where to eat since they have so many options. Positive reviews let customers know that the food is delicious, safely prepared and worth choosing over other restaurants. No matter how attractive a restaurant is or how many items the menu features, customers rely on reviews to decide if it’s worth ordering from a restaurant. Studies show that [90% of consumers consider reviews](https://blogs.cornell.edu/info2040/2021/11/10/following-the-crowd-reviews-sites/) before visiting an establishment. Gathering plenty of positive reviews shows people how great your food is and persuades them to try a restaurant. Positive reviews build customers’ trust.

**INTRODUCTION**

When visiting a new restaurant, customers want to know about it. Who is better to inform them than a prior client? The more people hear about a restaurant, the more likely they are to visit it. In this day of globalisation, we are constantly on the lookout for new ideas that will save us time, simplify our tasks and eliminate the need for manual processing. Customers at restaurants rate and leave reviews based on their level of satisfaction. These can assist other customers to decide whether to visit those eateries. These ratings can also help restaurant owners make changes based on customer feedback in order to improve their business. Textual information can be found in restaurant reviews. Machine learning (ML) allows systems to learn without having to be explicitly designed and this learning can be used to solve issues. ML takes data as input and learns essential relationships from it in order to make judgments that meet the needs of the user. The learning process begins with observations such as samples and direct experience and then looks for patterns in the data to help make better judgements about how to forecast or categorise new things in the future. Natural language processing (NLP) skills are provided by machine learning for text processing. Using NLP approaches, we can simply analyse our textual datasets. Data analysts can use NLP to apply machine learning and deep learning algorithms to their textual datasets. For classifying reviews, we employ machine learning algorithms.

**METHODOLOGY**

The Dataset is a .tsv file containing 1000 reviews in text format. It is divided into two parts: 80% training data and 20% test data. There are two columns in the dataset. The first column

comprises written reviews from various people that are connected to the restaurant's food as well as an overall assessment of the establishment. The mood or whether the review is good or negative, is displayed in the second column. The number 1 denotes a good review, while the number 0 denotes a bad review. Import the dataset and transform it to a Pandas Dataframe. The model should be able to tell whether the review will be good or negative. The Dataframe is then provided as input to the CountVectorizer for further processing after the dataset is cleaned from 1000 reviews and the reviews that are not suitable are deleted from the dataset. Natural Language Processing (NLP) is a technique that allows computers to comprehend human language. It is a branch of machine learning that involves a computer's capacity to comprehend, interpret, modify, and perhaps produce human language. Virtual assistants like Google Assist, Siri, and Alexa are probably the most well-known instances of NLP in action. NLP deciphers and translates human language into numbers, such as "Hey Siri, where is the nearest restaurant?" to make it easier for robots to interpret. Chatbots are another well-known use of NLP, since they may assist you in solving problems while conducting natural language generation — in other words, conversing in plain English. The field of natural language processing (NLP) is concerned with how computers comprehend and translate human language. Machines can understand written or spoken material and execute tasks such

as translation, keyword extraction, topic categorization, and more using natural language processing (NLP).

**CODE**

*# IMPORT LIBRARIES*

import pandas as pd

*# IMPORT THE DATASET*

dataset = pd.read\_csv("Restaurant\_Reviews.tsv", delimiter = "\t" , quoting = 3)

import re

import nltk

nltk.download('stopwords')

from nltk.corpus import stopwords

from nltk.stem.porter import PorterStemmer

corpus = []

for i in range(0,1000):

review = re.sub('[^a-zA-Z]', ' ', dataset['Review'][i])

review = review.lower()

review = review.split()

ps = PorterStemmer()

all\_stopwords = stopwords.words('english')

all\_stopwords.remove('not')

review = [ps.stem(word) for word in review if not word in set(all\_stopwords)]

review = ' '.join(review)

corpus.append(review)

*# CREATE BAG OF WORDS (BOW)*

from sklearn.feature\_extraction.text import CountVectorizer

cv = CountVectorizer(max\_features = 1500)

X = cv.fit\_transform(corpus).toarray()

Y = dataset.iloc[: , -1].values

from sklearn.model\_selection import train\_test\_split

X\_train, X\_test, Y\_train, Y\_test = train\_test\_split(X, Y, test\_size = 0.2)

*# TRAIN MODEL USING NAIVE BAYES*

from sklearn.naive\_bayes import GaussianNB

gnb = GaussianNB()

gnb.fit(X\_train, Y\_train)

Y\_pred = gnb.predict(X\_test)

*# CONFUSION MATRIX*

from sklearn.metrics import confusion\_matrix, accuracy\_score

cm = confusion\_matrix(Y\_test, Y\_pred)

accuracy\_score(Y\_test, Y\_pred)

**CONCLUSION**

This project presents an efficient restaurant review prediction algorithm to predict the review for the restaurant based on 1000 customer reviews. Restaurant reviews are denoted by the numbers 0 and 1, with 0 denoting a negative review and 1 denoting a positive one. Natural Language Processing (NLP) method yielded a great prediction accuracy - (67%). This strategy will aid business owners in anticipating consumer feedback and improving the customer experience. As a result of this project, it is an important factor to consider when making a dining selection. It is an important component of the planning process when starting a business like a restaurant