Restaurant Rating prediction

Using Machine Learning

Dr.Siva nageshwarao[1], Addanki Venkateshwarlu [2], B Ravi teja[3], K Vamsi Krishna[4]
[1] Professor, [2, 3 & 4] Student
[1]profssnr@gmail.com,[2]addankivenky934@gmail.com,[3]ravi4323teja@gmail.com,[4]vamsikrishnakunda205@gmail.com
Department of Computer Science and Engineering,
Narasaraopeta Engineering College, Narasaraopet, Andhra Pradesh, India

***Abstract*— One of the most competitive industries nowadays is the restaurant business. Since people have loved sharing meals with one another for generations, Each passing day witnesses a burgeoning demand for dining establishments. Bangalore, renowned for its diverse array of international cuisines, stands as an epitome of delight for food enthusiasts. This study embarks on unraveling the intricacies of food culture, honing in on a specific dataset of restaurants within the locale. Employing advanced data visualization technologies, our endeavor is to dissect the trends and patterns that define Bangalore's dynamic culinary landscape. This essay offers a methodology to comprehend the variables influencing restaurant ratings. A broad spectrum of methodologies from the realms of machine learning and predictive analytics can be harnessed to forecast restaurant ratings. Leveraging an extensive toolkit encompassing various tools and techniques, we aim to anticipate and understand the intricacies of restaurant ratings, unveiling valuable insights and patterns within the data.Several regression techniques are used to build the model in this work, and the most effective approach is taken into consideration.The outcome of this study aids newly opened eateries in selecting their menu, style, theme, price, target market, etc., ultimately leading to a rise in sales.**

Keywords--Restaurant business, Food culture, Bangalore,Data visualization,Regression techniques, Predictive analytics, Restaurant ratings ,Menu selection, Pricing strategy,Target marketSales increase, Strategic decisions, Data-driven decision-making, Competitive industry, Machine learning.

I.INTRODUCTION

The restaurant industry has grown to be one of the most powerful sectors with the emergence of several restaurants, dine-out options, and meal delivery apps. The abundance of dining options compels customers to explore a wide variety of cuisines. Bengaluru, recognized as a hub of global culture, mirrors this trend within its vibrant restaurant industry.With more than 12,000 restaurants, Bangalore is known as the foodie's paradise.Due to intense industry competition, restaurants are required to offer not just high-quality food but also a variety of innovative options within their current menu.The way that consumers dineout has changed significantly over the last ten years as a result of factors like urbanization, socioeconomic expansion, demographic shifts, and exposure to a global way of life.

Consumers' dining-out habits have changed significantly in India during the last ten or so years.By employing machine learning to identify trends in data and comprehend consumerbehavior, the restaurant industry is expanding its possibilities and raising the bar on service quality.Machine learning offers a forecast for the elements influencing the business based on historical data[14].Because machine learning techniques are so effective at managing huge datasets with inconsistent properties and noisy data, they are often used in conjunction with predictive analytics.

Predictive analytics encompasses a diverse array of statistical techniques aimed at estimating or foreseeing future outcomes. These methodologies utilize statistical methods to anticipate various scenarios and trends.Using previous company data, predictive models are constructed. The outcomes generated by the model may encompass predictions regarding customer behaviors or anticipate potential market fluctuations.Through the analysis of historical data, predictive analytics assists us in comprehending potential future trends.In this work, predictive analytics is performed by regression[8].This work uses the following algorithms Bayesian Regression, Random Forest Regression, Ridge Regression, Lasso Regression, KNN Regressor, SVM Regressor, Extra Tree Regressor, and Elastic Net Regressor are among the regression methods under consideration. Determining the most effective model involves comparing these methods and gaining a deep understanding of their underlying mechanisms.

II. LITERATURE SURVEY

J. Priya et al.[8] study about served as a platform for showcasing the latest advancements in the field.In their study, they focused on utilizing machine learning techniques to forecast restaurant ratings, aiming to enhance understanding of consumer preferences and improve service quality in the restaurant industry. By comparing various regression models, the authors sought to identify the most effective approach for predicting restaurant ratings based on factors such as cuisine type, location, and customer reviews. In their study, Priya and colleagues sought to offer valuable insights into the utilization of machine learning within the hospitality industry. Their objective was to further enhance predictive analytics specifically tailored to the realm of restaurant management.

Mohan S Acharya et al.[10] conducted a study about the presentation shed light on emerging trends and innovative approaches within the field. In their research, they evaluated various regression models to forecast graduate admissions, analyzing factors such as GRE scores and undergraduate GPA. By employing machine learning techniques, the authors aimed to identify trends in data and understand consumer behavior, thereby enhancing service quality in the restaurant industry. Through comparative analysis of regression algorithms like Bayesian Regression, Random Forest Regression, and others, they sought to determine the most effective model for predicting future trends in graduate admissions, thereby contributing valuable insights to academia and admission processes.

Shina et al. [12] conducted a study about served as the backdrop for their investigation. Their study delved into the utilization of tree-based machine learning algorithms to dissect and interpret restaurant reviews,aiming to uncover patterns and insights that could improve decision-making in the hospitality industry. In their endeavor, the authors aimed to discern the optimal approach for sentiment analysis and review classification by contrasting various tree-based models. This involved a comparative analysis of decision trees, random forests, and gradient boosting trees to ascertain their effectiveness in the task. Through their investigation, Shina et al. sought to contribute to the advancement of machine learning techniques for restaurant management and customer satisfaction assessment.

Vishnu Tharavath et al. [14] study Their presentation unveiled insights into the intricate dynamics influencing cuisine choices in the local dining landscape. The study aimed to understand the preferences and decision-making processes of Indian consumers regarding dining out, particularly focusing on the selection of cuisines. Tharavath et al. examined various factors such as cultural influences, socio-economic backgrounds, and personal preferences to elucidate the determinants behind consumers' choices. Through a comprehensive examination of surveys and interviews, the authors endeavored to furnish the restaurant industry with valuable insights. Their aim was to bolster strategies geared towards enhancing customer satisfaction within the dynamic Indian dining-out market.

III.PROPOSED METHODOLOGY

In Figure 1, the methodology for interpreting data obtained from online reviews unfolds through a structured process. Initially, the data gathering and pre-processing phases are pivotal. This involves not only the collection of data but also meticulous steps such as outlier removal to ensure the integrity of the dataset. After completing the preparatory phase, the next step involves the application of a Machine Learning (ML) algorithm to the refined dataset.

The chosen algorithms for classifying the overall class based on the review parameters are the Random Forest Algorithm and Decision Tree.These algorithms are selected for their proven efficacy in handling classification tasks and their suitability for the complexities inherent in restaurant rating prediction. Once the algorithms are applied, the resultant outcomes are meticulously analyzed and visualized to facilitate comprehension and informed decision-making. This visualization step is crucial as it provides stakeholders with a tangible representation of the data trends and model predictions.

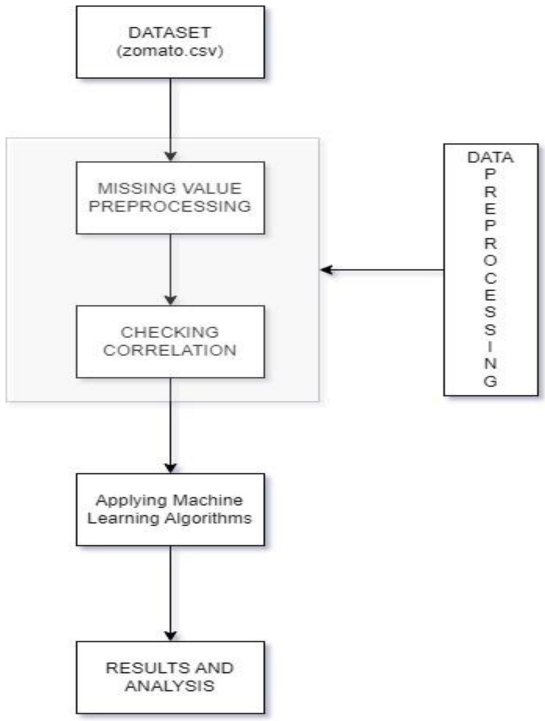


Fig-1:Proposed methodology

Moreover, the comparative analysis extends to evaluating the performance of the machine learning algorithms. The evaluation process primarily revolves around gauging the precision of the algorithms in forecasting the previously mentioned restaurant class within the dataset.By scrutinizing the accuracy metrics, researchers can discern which algorithm exhibits superior performance characteristics tailored to the nuances of the dataset at hand. This process of algorithm evaluation serves as a critical benchmarking exercise, guiding the selection of the most suitable algorithm for the restaurant rating prediction task.

IV. DATASET DESCRIPTION

Since Bangalore is a restaurant hotspot, the Zomato website is used to gather information on different restaurants[15].Gaining a general understanding of the elements influencing each restaurant's overall rating and the establishment of various restaurant kinds in various locations is the main goal of the Zomato dataset analysis.

The dataset encompasses 51715 rows and 17 columns, each containing valuable information about the restaurants. Among these columns, one describes the restaurant's URL from the Zomato website. Another column, labeled address, provides the location details of the eatery in Bangalore. The column named name holds the unique names of each restaurant. Additionally, the fields online_order and book_table signify whether the establishment offers online ordering and table booking options, respectively. rate: The restaurant has a five-star rating overall.

The dataset provides a comprehensive overview of each restaurant's details. Within the dataset, the column labeled votes displays all ratings recorded up to the specified date.

Contact information is available in the phone column, while the location column describes the area surrounding the restaurant. Additionally, the rest_type field categorizes the type of eatery, and dish_liked highlights popular dishes among diners. Cuisines are listed in groups separated by commas, and the approx_cost (for two people) variable estimates the cost of a meal for two individuals. Reviews and ratings from patrons are stored in the reviews_list, with each tuple containing the review and corresponding rating. Furthermore, the menu_item column lists the menus provided by the restaurant. Finally, listed_in(type) categorizes the meals offered, while listed_in(city) specifies the city where the restaurant is listed.

The data set attribute description is displayed in the table.

TABLE1:DATASET BEFORE PREPROCESSESING

```
Data columns (total 17 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   url                          51717 non-null  object
 1   address                      51717 non-null  object
 2   name                         51717 non-null  object
 3   online_order                 51717 non-null  object
 4   book_table                   51717 non-null  object
 5   rate                         43942 non-null  object
 6   votes                        51717 non-null  int64
 7   phone                        50509 non-null  object
 8   location                     51696 non-null  object
 9   rest_type                    51490 non-null  object
 10  dish_liked                   23639 non-null  object
 11  cuisines                     51672 non-null  object
 12  approx_cost(for two people)  51371 non-null  object
 13  reviews_list                 51717 non-null  object
 14  menu_item                    51717 non-null  object
 15  listed_in(type)              51717 non-null  object
 16  listed_in(city)              51717 non-null  object
dtypes: int64(1), object(16)
memory usage: 6.7+ MB
```

V.DATA PREPROCESSING

In order to transform raw data into a format that can be used, data preparation—also known as data pretreatment—is essential[9].It addresses problems such as null values, transforming category data into numerical form (online orders, table reservations, etc.), and fixing inconsistent data. The dataset is streamlined for analysis by removing unnecessary features. Preprocessing improves the dataset's clarity and usefulness and makes it ready for analysis. This methodical procedure guarantees that the data is effectively structured and formatted, providing a strong basis for further analysis and insights extraction.

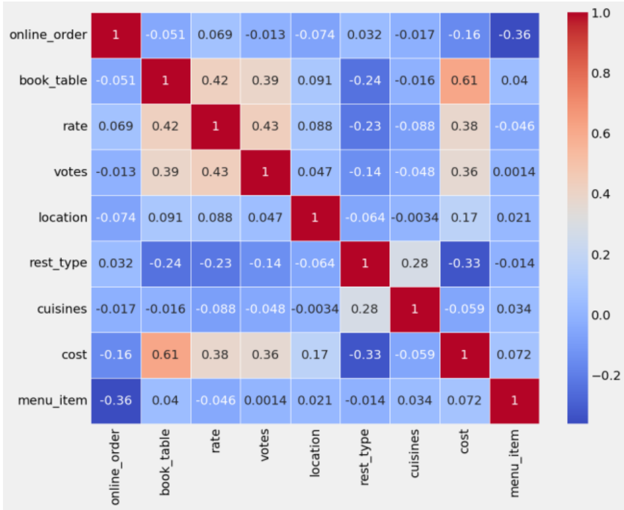
The Data in dataset will be reduced to some what less in the previous data of before preprocessing.The preprocessing takes place a vital role in the project with data cleaning and removal of unwanted data .The dataset contains information about restaurants and café in Bangalore city with around 5000 null values. We used mean for numerical data and mode for the categorical data. With this methods the null values are removed. Also in this dataset contains duplicate data which may reduce the accuracy. So we also removed the duplicate data from the dataset which increased the accuracy of the model.

We also removed two columns named url and phone number because this features are not useful to predict the restaurant rating.

TABLE2:DATASET AFTER PREPROCESSING

```
Index: 51674 entries, 0 to 51716
Data columns (total 15 columns):
 #   Column                       Non-Null Count  Dtype
---  ------                       --------------  -----
 0   address                      51674 non-null  object
 1   name                         51674 non-null  object
 2   online_order                 51674 non-null  object
 3   book_table                   51674 non-null  object
 4   rate                         51674 non-null  object
 5   votes                        51674 non-null  int64
 6   location                     51674 non-null  object
 7   rest_type                    51674 non-null  object
 8   dish_liked                   51674 non-null  object
 9   cuisines                     51674 non-null  object
 10  approx_cost(for two people)  51674 non-null  object
 11  reviews_list                 51674 non-null  object
 12  menu_item                    51674 non-null  object
 13  listed_in(type)              51674 non-null  object
 14  listed_in(city)              51674 non-null  object
dtypes: int64(1), object(14)
memory usage: 6.3+ MB
```

VI.CORREALTION COEFFICIENT



The correlation matrix serves as a foundational tool in the prediction of restaurant ratings, providing a comprehensive overview of the interrelationships between numerous factors. Among these factors are cuisine type, location, availability of online ordering and table booking services, diversity of menu items, restaurant type, and price range. By scrutinizing the correlation coefficients derived from this matrix, analysts can discern the strength and direction of associations between each factor and the overall ratings.

This examination enables the identification of variables that exert a significant influence on ratings, as well as those that have minimal impact.

Such insights gleaned from the correlation matrix are invaluable for refining predictive models in restaurant rating prediction tasks. By prioritizing variables with stronger correlations to ratings, analysts can optimize model performance and enhance prediction accuracy. Additionally, identifying variables with weaker correlations allows for a more nuanced understanding of the multifaceted nature of restaurant ratings. Ultimately, this analytical approach empowers stakeholders in the restaurant industry to make informed decisions and implement targeted strategies aimed at improving customer satisfaction and overall dining experiences.

VII. DATA VISUALIZATION

The visual display of data as figures, graphs, or charts is known as data visualization.Massive data sets are analyzed using it in order to identify patterns and trends. Data from eateries is represented visually[6].

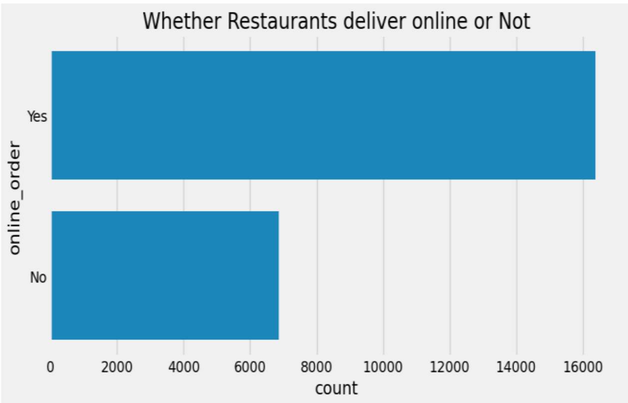


Fig.2:Online order

The fig.2 shows the distribution of restaurants that deliver online or not[11]. The y-axis shows the count, while the x-axis shows whether the restaurant delivers online or not. There are two bars, one for online orders and one for restaurants that don't deliver online.Overall, the graph seems to indicate that there are more restaurants that deliver online than those that don't.

There are many restaurant where most of them are offering online order and services which may increase their reputation and restaurant or café rating. While the restaurants which do not offer online are somewhat less popular particurarly in Bangalore because most of the people work in office and do not have time to go the restaurants or café for either lunch or dinner. This may impact the rating or reputation of the restaurant or café.

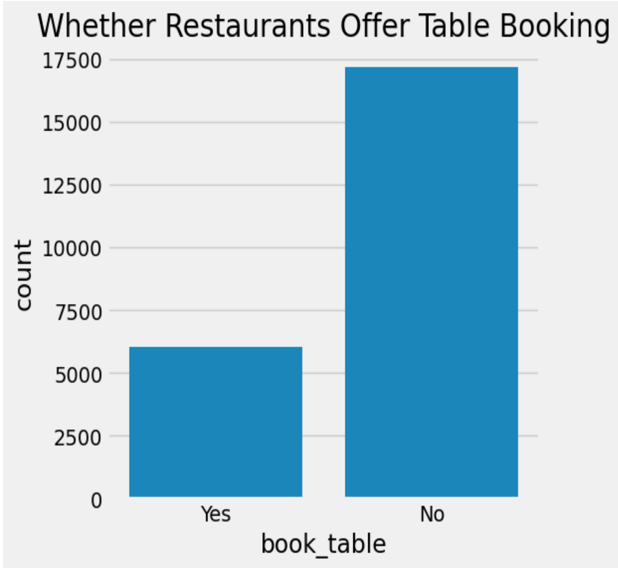


Fig.3:Booking table

The Fig.3 shows that distribution of restaurants that offer table booking. The y-axis shows the count of restaurants, and the x-axis shows whether a restaurant offers table booking ("Yes") or not ("No"). The more number of people do not prefer the table booking in a restaurant but, some may be prefer also.

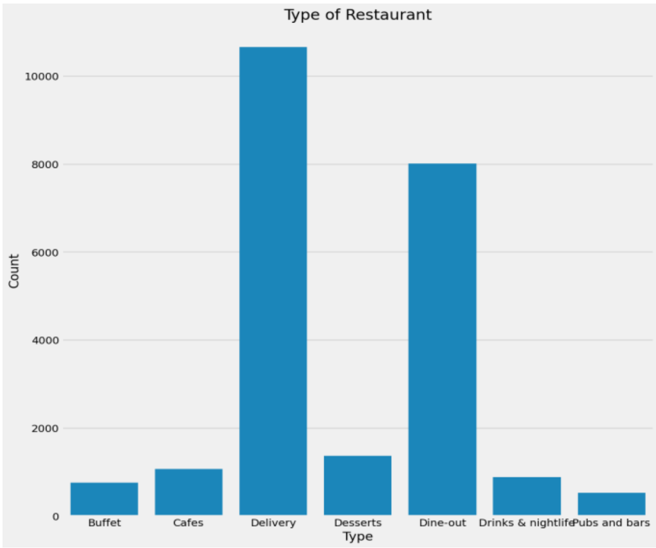


Fig.4: Type of restaurants

The Fig.4 shows that number of restaurants in the Banglore city by type of restaurant. Delivery restaurants are the most popular, with around 10,000 restaurants. Dine-out are the second most popular type of restaurant, with around 8,000 restaurants. Cafes and dessert restaurants follow in popularity, with 1,300 and 2,000 restaurants respectively. It is vital to note that the data labels for the y-axis are count for the restaurants.

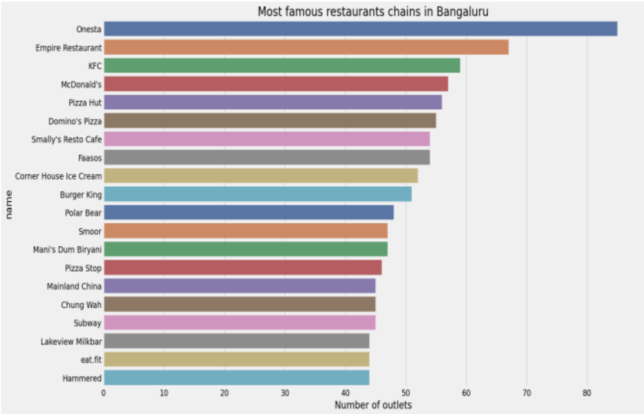


Fig.5:Most Famous restaurants

The Fig.5 shows that Most Famous restaurants in the Banglore city with their names.On the x-axis it shows the number of outlets for the restaurants where as on the y-axis its names should be displaying.The Omesta, Empire restaurants are the more number of outlets in this by the count.

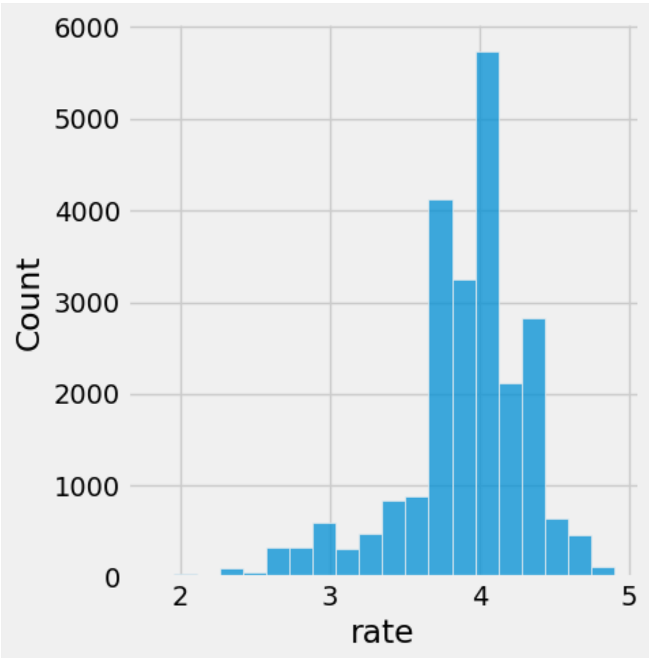


Fig.6:Rating Distribution

The Fig.6 shows that Rating Distribution for the Restaurant. The x-axis as rating column from 0 to 5 with responding of giving,where as the y-axis has the count of rating giving with corresponding restaurant. Here are some possibilities for what the x-axis might represent:Customer satisfaction rating (with 0 being the least satisfied and 5 being the most satisfied)Product quality rating (with 0 being the lowest quality and 5 being the highest quality)Difficulty level (with 0 being the easiest and 5 being the hardest).The rating will be displaying in a bar graph with having highest rating is 4 having more than 6000 restaurants for the dataset.The less number of restaurants having lessthan rating 2 and more average rating 3 to 4 in the rating Distribution.

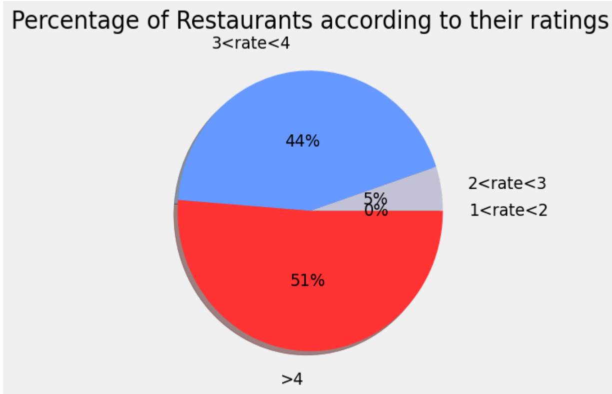


Fig.7:Restaurant Rating in percentages

The Fig.7 shows the distribution of restaurants according to their customer ratings. The slices of the pie chart are labeled with ranges of ratings, but the specific range for each slice is not provided. The largest slice, colored blue, represents the highest percentage of restaurants with a rating that falls somewhere between 3 and 4. The red slice represents the highest percentage of restaurants, with ratings greater than 4.

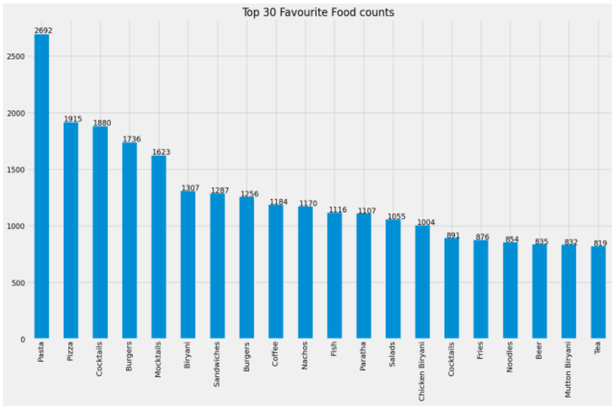


Fig.8:Most favourite food

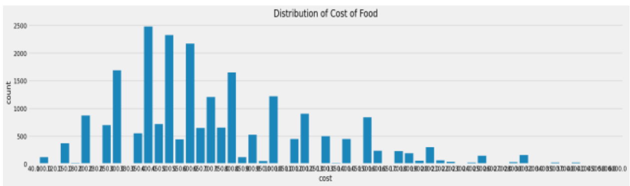


Fig.9:Distribution of cost of food

Restaurants with meals priced between Rs. 450 and Rs. 500 are highly regarded. Restaurant ratings rise in proportion to expense. The majority of customers favor Chinese, North Indian, and Indian cuisine, rolls, memos, and so forth.

VIII. MODEL BUILDING

Using the available set of instances, a predictive model is formulated. In this process, 30% of the dataset is allocated for testing purposes, whereas the remaining 70% is utilized for training the model.The goal variable for restaurant rates is fixed. One method of selecting independent variables is through standardization.A variety of regression procedures are used to build the model[4].

(a)Linear Regression

Linear regression stands as the fundamental form of regression analysis. Its objective lies in establishing a meaningful relationship, or model, between the independent variable (also known as the exploratory variable) and the dependent variable (termed as the predicted variable). In the realm of simple linear regression, there exists only one independent variable and one dependent variable. Conversely, multilinear regression accommodates multiple independent and dependent variables, adding complexity to the analysis[2].
The formula for a linear regression line is

$$Y=a+bX.$$

where X and Y are respectively, represent the independent and dependent variables.
The value of y at x=0 is the intercept, while b is the line's slope.

(b) Random Forest Regression algorithm

A highly effective machine learning model frequently utilized in predictive analytics is the Random Forest regression technique. This method employs an ensemble approach, where the judgments of multiple decision trees are amalgamated to form a robust predictive model. The Random Forest model operates by aggregating the outputs of numerous individual decision trees, thereby enhancing prediction accuracy and reliability.

$$f0(x)+f1(x)+f2(x)+...=g(x).$$

In this case, g is the total of all basic models fi that are created independently from separate subsamples of data[12]. Model esembling involves using many models to improve prediction.

(c) Ridge regression

One kind of linear regression is ridge regression. Multicollinearity in multiple regression data is resolved by the use of ridge regression. Least squares estimates are unbiased when multicollinearity is present, but because of their high variances, they could not be very close to the actual value. Introducing a degree of bias to regression estimates, Ridge regression effectively reduces standard errors, thereby improving the accuracy of projections. This adjustment enhances the model's ability to generalize and make reliable predictions. The equation representing Ridge regression is formulated as follows:

$$Y = XB + e,$$

In the equation, X denotes the set of independent variables, while B represents the regression coefficient to be computed, and e signifies the error term.

(d)Lasso Regressor

Lasso regression is a shrinkage-based variant of linear regression.In this case, shrinking is accomplished by bringing each point to itscenter.It reduces the quantative response variable's prediction error.For some variables, it results in regression coefficients that are closer to zero.Regression coefficient 0 variables are removed from the model.Response variables and variables with non-zerocoefficents have a strong correlation.It provides accurate predictions. Reduce residual sum of squares is the purpose of it[7].

(e)KNN Regression

The K Nearest Neighbors (KNN) algorithm operates by assessing the proximity between each neighbor, which forms the basis of its mapping. Leveraging the independent variable, KNN predicts the target variable. A fundamental application of KNN regression involves calculating the average of the numerical target among the K nearest neighbors. Alternatively, an inverse distance weighted average method can be employed. KNN utilizes various distance functions such as Euclidean, Manhattan, and Minkowski distances to determine neighbor proximity.

(f)Support Vector Regression

Support Vector Regression (SVR) and Support Vector Machine (SVM) exhibit minimal variance, rendering them comparable in terms of performance and stability[13].It is exceedingly difficult to estimate the result because there are endless alternatives and the output is genuine.As a result, a tolerance margin is chosen to approximate SVM. By customizing the margin, it eventually lowers error[1].

(g)Elastic Net Regression

Elastic Net includes regularization terms with both L-1 and L-2 norms, which causes coefficients to either shrink (as in ridge regression) or set to zero (as in LASSO).As a result, Lasso and Ridge regression is inferior to Elastic Net Regression. Elastic regression offers more accurate feature selection and prediction performance.

(h)Bayesian Regression

Bayesian regression presents a departure from traditional linear regression by incorporating probability distributions rather than point estimates. In this approach, the output variable y is determined from a probability distribution rather than a single value, offering a more nuanced understanding of uncertainty. The construction of the Bayes Model follows this framework:

$$P(\beta|y,X) = \frac{P(y|\beta,X) * P(\beta|X)}{P(y|X)}$$

Additionally shown as:

$$Posterior = \frac{Likelihood * Prior}{Normalization}$$

(i) ExtraTree Regressor

The ExtraTree Regressor is a well-known ensemble learning technique used in machine learning to address regression problems. ExtraTrees are a decision tree algorithm family member that differs from Random Forests in that they use a different approach. They apply randomness to the process of choosing feature subsets for individual nodes as well as the splitting threshold selection.ExtraTrees stochastic component sets it apart from other approaches by giving the model's predictions more variability and resilience. ExtraTree Regressors are a powerful tool for predictive modeling across a wide range of domains, and they excel at tackling complicated regression tasks by utilizing this randomized technique.

IX. RESULTS AND ANALYSICS

ExtraTree Regression is clearly the best performer in Regression Score, according to Table 3's results[10]. Compared to alternative models, it exhibits remarkably low absolute, mean squared, and root mean squared errors, distinguishing its performance in predictive accuracy[5]. The higher performance of ExtraTree Regression is demonstrated by all indicators, demonstrating its effectiveness in regression problems.

Its unmatched precision and accuracy set it alone as the best option out of all the models that were assessed. These findings highlight the ExtraTree Regression's stability and dependability and highlight its potential for a variety of applications where accurate regression analysis is crucial. As a result, stakeholders can count on ExtraTree Regression to continually produce better outcomes, providing a dependable solution for a range of regression difficulties.

TABLE.3:EVALUATION OF ALL REGRESSION ALGORITHMS

Type of regression	Regression Score	Absolute error	MeanSquaredError	Root mean square error
Linear Regression[8]	0.30	0.26	0.17	0.48
Random Forest Regression[8]	0.88	0.070	0.023	0.288
Ridge Regressor[8]	0.38	0.265	0.147	0.407
Lasso Regressor[8]	0.38	0.252	0.168	0.515
KNN Regressor[8]	0.70	0.1701	0.0704	0.446
SVM Regressor[8]	0.70	0.243	0.1387	0.536
Elastic net Regressor[8]	0.60	0.2451	0.168	0.518
Bayesian Regressor[8]	0.72	0.267	0.138	0.518
Extratree Regressor	**0.93**	**0.033**	**0.014**	**0.178**

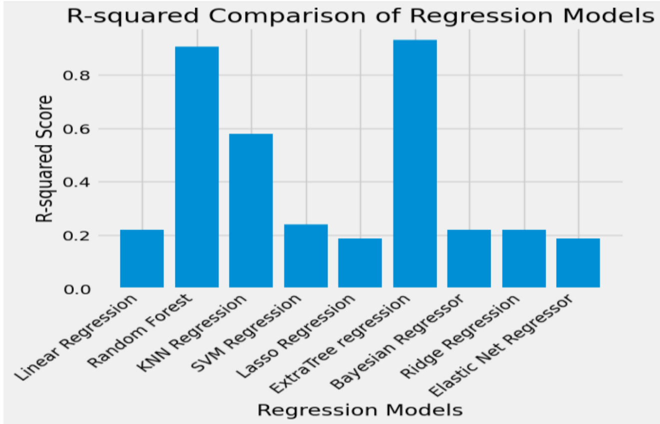


Fig.10:Regression score of models

Fig.11.Absolute Error for Models

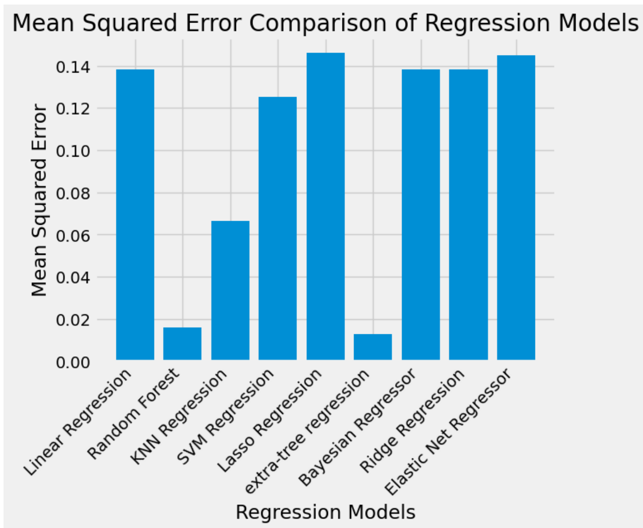


Fig.12.Mean Squared Error of Models

X.CONCLUSION

Machine learning techniques are becoming more and more integrated into the restaurant business to forecast consumer behavior and optimize operations. In order to build models, this study explores 9 regression algorithms and looks at their construction processes and behaviors.A comprehensive analysis of performance metrics unmistakably highlights the ExtraTree Regressor as the top performer, showcasing superior regression scores and lower error rates when compared to its counterparts.

The results highlight how machine learning may improve decision-making in the restaurant industry by providing useful information on client preferences and operational efficiency. Consequently, the Extratree Regressor becomes an effective instrument for promoting advancements and creativity in the field of restaurant administration.Hence we concluded that previous paper have the highest accuracy or regression score is 88% using the Random Forest Method .But this time i used ExtraTree Regression method for the accuracy or Regression score with 93% having the highest Regression Score.This is the best accuracy method for the regression score and model implementation.

XI. FUTURE SCOPE

In order to forecast Zomato restaurant evaluations, a machine learning model is presented in this work. The extraction of features from independent variables has traditionally been done by hand. On the other hand, this work is the first to incorporate Deep Learning methods for automating feature extraction, which eliminates the need for labor-intensive manual engineering. Moreover, Deep Learning models are skillfully applied to manage a dataset that is growing at an exponential rate. This novel method improves the model's scalability and adaptability while streamlining the prediction process. This research marks a significant improvement in the field of restaurant rating prediction by utilizing Deep Learning, with the potential to yield more accurate and efficient results.

REFERENCES

[1] A. Kulkarni, D. Bhandari and S. Bhoite, "Restaurants rating prediction using machine learning algorithms", International Journal of Computer Applications Technology and Research, 2019.

[2] saedsayad,“A Tutorial on Support Vector Machine” – Regression,https://www.saedsayad.com/support_vector_machine_reg. html

[3] Bern Jonathan, Jay Idoan Sihotang, Stanley Martin 2019, “Sentiment Analysis of Customer Reviews in Zomato Bangalore Restaurants Using Random Forest Classifier”.

[4] B. Nithya,V. Ilango “Predictive Analytics in Health Care Using Machine Learning Tools and Techniques” International Conference on Intelligent Computing and Control Systems ICICCS 2017

[5] Hideo Hirose, Yusuke SoejimaKeiHirose ,“NNRMLR: A Combined Method of Nearest Neighbor Regression and Multiple Linear Regression”

[6] Huang Li, Shiaofen Fang, Snehasis Mukhopadhyay, Andrew J. Saykin, Li Shen “Interactive Machine Learning by Visualization: A Small Data Solution”

[7] I. F. Shihab, M. M. Oishi, S. Islam, K. Banik and H. Arif, "A Machine Learning Approach to Suggest Ideal Geographical Location for New Restaurant Establishment", 2018 IEEE Region 10 Humanitarian Technology Conference (R10-HTC), pp. 1-5, 2018

[8] J. Priya, "Predicting restaurant rating using machine learning and comparison of regression models", 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), pp. 1-5, 2020.

[9] J. Z. A. Wang, W. Zeng, “Predicting New Restaurant Success and Rating with Yelp,” 2016

[10] Mohan S Acharya, Asfia Armaan, Aneeta S Antony “A Comparison of Regression Models for PredictionofGraduate Admissions” Second International Conference on Computational Intelligence in Data Science (ICCIDS-2019)

[11]Qiwei Gan and Yang Yu, “Restaurant Rating: Industrial Standard and Word-of-Mouth A Text Mining and Multi-dimensional Sentiment Analysis”, Hawaii International Conference on System Sciences, pp. 1332-1340, 2015

[12] Shina, S. Sharma and A. Singha, "A study of tree based machine learning Machine Learning Techniques for Restaurant review", 2018 4th International Conference on Computing Communication and Automation (ICCCA), 2018.

[13] Vishal Lokam, Vaishnavi Shinde, Abhishek Raikar, Vivek Kate, Gaurav Jumnake 2021, “Restaurant and Cuisine Review Sentiment Analysis using SVM”.

[14] Vishnu Tharavath, Dr. Deepak Gupta, Dr. Sangeetha Gunasekar,“Factors Influencing Choice of Cuisines While Indian Consumers Eat Out” 2017 International Conference on Data Management, Analytics and Innovation (ICDMAI) Zeal Education Society, Pune, India, Feb 24-26, 2017.

[15] “Zomato Bangalore Dataset”, Predicting Restaurant rate from important paramaters, Kaggle, March 2019. https://www.kaggle.com/himanshupoddar/zomato-bangalore-restaurants.