

Machine Learning Internship Project

Company: Cognizify Technologies

Project Title: Predicting Restaurant Ratings using Machine Learning

Intern: Ruthika Nalajala

1. Introduction

This project is part of the Machine Learning internship at Cognizify Technologies. The objective is to build a predictive model that estimates the aggregate ratings of restaurants based on various features like cuisines, location, delivery options, and others.

2. Problem Statement

The aim is to predict the restaurant ratings using machine learning models by analyzing multiple attributes of the restaurant data. Accurate prediction can help businesses and customers better understand the factors influencing restaurant quality.

3. Dataset Description

The dataset contains information about restaurants including features such as:

- City
- Cuisines
- Price Range
- Votes
- Online Delivery and Table Booking Availability
- And other related features.

The target variable is 'Aggregate rating'.

4. Methodology

The following steps were followed to complete the project:

- Data Preprocessing: Handling missing values and encoding categorical variables using Label Encoding.
- Feature Selection: Dropped irrelevant columns like Restaurant Name, Address, Longitude, Latitude, etc.
- Model Building: Two models were trained - Linear Regression and Decision Tree Regressor.
- Evaluation: Models were evaluated based on Mean Squared Error (MSE) and R^2 Score.
- Feature Importance: Important features influencing the restaurant ratings were identified.

5. Results

- Linear Regression Model:
 - Achieved a reasonable R^2 score indicating moderate prediction ability.
- Decision Tree Regression Model:
 - Performed slightly better, showing important non-linear relationships between features.
- Most influential features included:
 - Price Range
 - Number of Votes
 - Online Delivery Availabilit

- Cuisines

6. Conclusion

The machine learning models successfully predicted the restaurant ratings with acceptable accuracy. Decision Tree Regression performed better compared to Linear Regression. Important features influencing restaurant ratings were identified, providing actionable insights for restaurant businesses.

7. Future Work

- Implement more advanced models like Random Forest or Gradient Boosting for improved accuracy.
- Fine-tune model hyperparameters.
- Use feature engineering to extract more meaningful features from textual data.