

HIGH LEVEL DOCUMENTATION

Restaurant Rating Prediction

Basawanayya Hiremath

Contents

1.	Introduction.....	4
	What is High-Level Design document?	4
	Scope.....	4
2.	General Description.....	5
	Problem statement.....	5
	Proposed Solution.....	5
	Further improvements.....	5
	Data Requirements.....	5
	Tools used.....	6
	Constraints.....	6
	Assumptions.....	6
3.	Design Details.....	7
	Process flow.....	7
	Model training and evaluation.....	7
	Event log.....	8
	Error Handling.....	8
4	Performance.....	8
	Reusability.....	8
	Application compatibility.....	8
	Resource Utilization.....	8
	Deployment.....	8
5.	Conclusion.....	9

Abstract

The restaurant industry is growing rapidly as earning of current generation increased as economy increased and it was found that current generation people are spending more on food than previous generation. From this lot of competition is also increased. In the generation of social media if service or food is not good then rating may fall. Bangalore is hub of IT sector in India and it is increasing from last 25 years and we can find wide range of cuisines from all over the world. The zomato dataset captured the features like online order, book table, votes, rating, location etc. from this dataset we explore the features by visualizing and comparing them. Finally to build the appropriate ML model with high R2 score.

1. Introduction

Why this high level Design Documentation

The purpose of this High-Level Design (HLD) Document is to add the necessary detail to the current project description to represent a suitable model for coding. This document is also intended to help detect contradictions prior to coding, and can be used as a reference manual for how the modules interact at a high level.

High Level Design (HLD) is the overall system design - covering the system architecture and database design. It describes the relation between various modules and functions of the system. Data flow, flow charts and data structures are covered under HLD.

The HLD will:

- Present all of the design aspects and define them in detail
- Describe the user interface being implemented
- Describe the hardware and requirements
- Include design features and the architecture of the project
- List and describe the non-functional attributes like:
 - Security
 - Reliability
 - Maintainability
 - Portability
 - Reusability
 - Application compatibility
 - Resource utilization
 - Serviceability
 - Cost limitations
 - Degree of performance
 - Degree of scalability

Scope

The scope of the project is to predict restaurant rating and visualize, compare the data with other restaurant to adapt and to build UI to show the Rating as

a output based on certain features like location, votes, cost for two people, restaurant type etc.

2. General Description

Product perspective

Restaurant rating system in Bangalore based on machine learning models which predicts rating with only few input features in developed UI.

Problem statement

The main goal of this project is to perform extensive Exploratory Data Analysis (EDA) on the Zomato Dataset and build an appropriate Machine Learning Model that will help various Zomato Restaurants to predict their respective Ratings based on certain features.

Proposed Solution

This model/UI makes prediction based on features/services provided by restaurant like table booking, Online orders facility, type of food they serve, location, cost for two person etc. some unnecessary or unique features like phone number, e-mail id etc. are dropped.

Further improvements

There are so many improvements are to be done in dataset like example in menu column there are so many unordered items in just one row, if we split them for additional features it rises some issues like. If first row is dosa, idly and second row is idly, dosa. It will consider it them different or unique feature although they are same. To improve this we have to list menu items in ordered way.

The data should contain some more features like 'Bestselling dish', 'Service', 'Hygiene' etc. and The data is small to make best optimized solution.

Data Requirements

Data is requirement completely depends on problem statement.

Tools used



- IDE used are Jupyter notebook and VS code
- For visualization Matplotlib and Seaborn library are used
- Sklearn is used to train the model
- For front end HTML is used
- For web development Flask is used
- Github is used as version control

Hardware requirement

None

Constraints

The data only limited to Bangalore so it will be biased if we use it for other cities.

Assumptions

The main objective of the project is extensive Exploratory Data Analysis (EDA) on the Zomato Dataset and build an appropriate Machine Learning Model that will help various Zomato Restaurants to predict their respective Ratings in Bangalore region. We assume that the system will provide accurate results.

3. Design Details

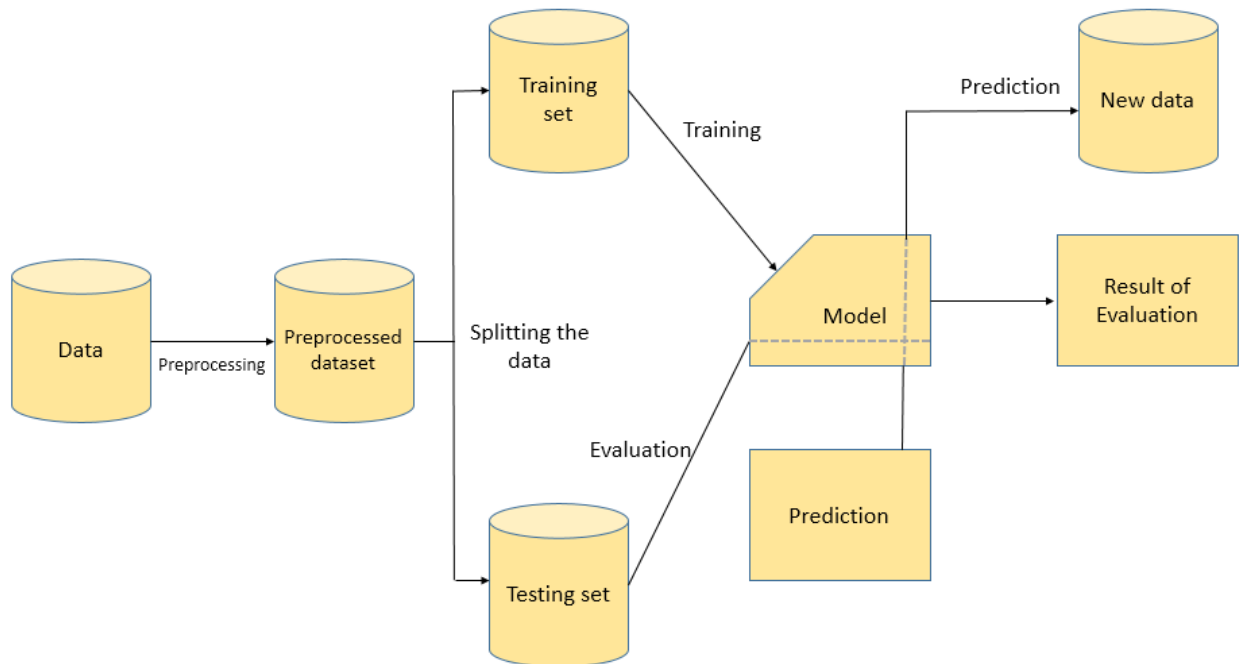
Process flow eve

To train the ML model to get high Score. The process flow is shown in below diagram.

Proposed methodology



Model training and evaluation



Event log

- The system should log every event so that the user will know what process is running internally
- Initial step by step description
- The system identifies as what step logging required
- The system should be able to log each and every system flow
- System should not hang even after using so many logging, logging just because we can easily debug issues so logging is mandatory to do

Error Handling

Should errors be encountered, an explanation will be displayed as to what went wrong? An error will be defined as anything that falls outside the normal and intended usage.

4. Performance

The system should run smoothly if we feed the inputs and get output without lagging or hanging.

Reusability

The code written and the components used should have the ability to be reused **with** no problem, if there is similar problem statement.

Application compatibility

The different components for this project will be using Python as an interface between them. Each component will have its own task to perform, and it is the job of the Python to ensure transfer of information.

Resource Utilization

When any task performed, it will likely to use all the processing power available until that is finished.

Deployment

Local host

5. Conclusion

The system/UI will predict Rating of the restaurant with only few feature inputs and there are extensive Exploratory Data Analysis (EDA) on the Dataset from this we can compare from which parameter restaurant rating will increase and to adapt the changes to make maximum profit.

There is also Dominos V/s Pizza Hut comparison.