iNeuron

# HIGH LEVEL DESIGN (HLD)

# Restaurant Rating Prediction (Zomato)

# Contents

# *Abstract*

Bangalore has more demography with high number of restaurants, it is also increasing the restaurant chain with new restaurants opening each day. Restaurants from all over the world can be found in and out Bangalore surroundings, since it will be difficult to compete with established restaurant chain, key issues will be real estate costs, pricing strategy, shortage of quality manpower and over-licensing, Zomato data analysis provide an insight about different factors such as deciding their theme, cuisine, menus and also finding similarity between neighbourhoods of Bengaluru on the basis of food. The dataset also contains reviews for each of the restaurant which will help in finding overall rating for the place.

# 1 *Introduction*

## 1.1 why his high-level design document?

This document is designed to help in operational requirement and can be used as reference manual how the module interacts. It is also identifying those project parts that might be time consuming or risk taking, HLD provides a brief description about various sub-systems and components of the system fit together.

This HLD will:

- presents all of the design aspects and define them in detail.
- Describe the user interface being implemented.
- Describe the software and hardware interfaces.
- Describe the performance requirements.
- Include design features and architecture of the project.
- List and describe the non-functional attributes like:
    - Security
    - Reliability
    - Maintainability
    - Portability
    - Resource utilization

## 1.2 Scope

The High-Level Design documentation presents the structure of the system as the application/database architecture, application flow and technology architecture. High-Level Design documentation may use some non-technical terms.

## 1.3 Definitions

| Term | Description |
|---|---|
| Database | Collection of all information monitored by system |
| IDE | Integrated development environment |
| AWS | Amazon web services |

# 2 General Description

## 2.1 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.

## 2.2 Proposed Solution

It can be seen that data has been phased in two the first contains only URL, address and name of restaurant and phase two contains recorded data for each restaurant and each category was read and data for each restaurant was scraped individually and recorded in Csv file, such that will this we can find out the various factors such as approximate price of food, theme based on restaurant or not and for people who are striving to get best cuisine of the neighbourhood.

## 2.3 Further Improvements

With the help of it can get review based on the location of restaurant and neighbourhood particular famous for its own kind of food. Able to identify the most visited restaurant in particular location.

## 2.4 Data Requirements

Data requirement completely based on our problem statement

- It must contain the file in csv and should have 17 columns.
- Need to have information about 6-7 categories of restaurant serving details itself.
    - Buffet
    - Cafes
    - Delivery
    - Dine-out

- Drinks and nightlife
- Pubs and Bars

- Data scraped was entirely for educational purposes only.

## 2.5 Tools Used


- PyCharm is used as IDE.
- Aws or Azure is used as deployment of the model.
- For visualization of plots matplotlib, seaborn and plotly are used.
- Cassandra database is used to retrieve, insert, delete and update the database.
- GitHub is used as version control system.

## 2.6 Assumptions

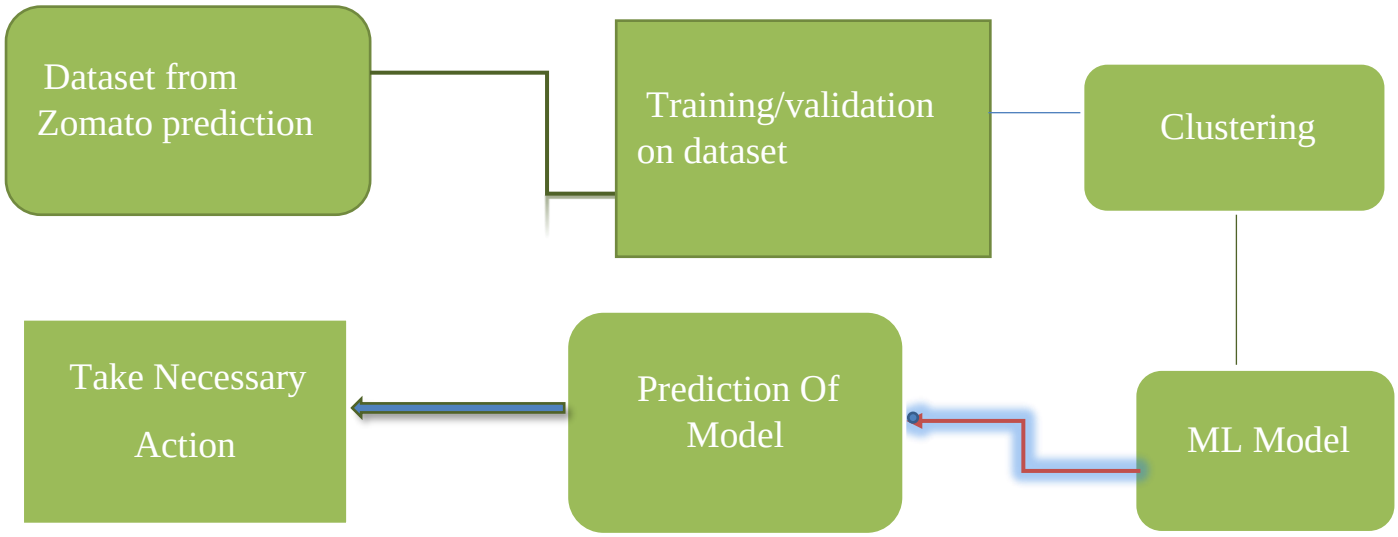
Get a fair idea about the factors affecting the establishment of different types of restaurants at different places in Bengaluru, aggregate rating of each restaurant. food chain category restaurant likely to have more customers than its counterpart.

# 3 *Design Details*

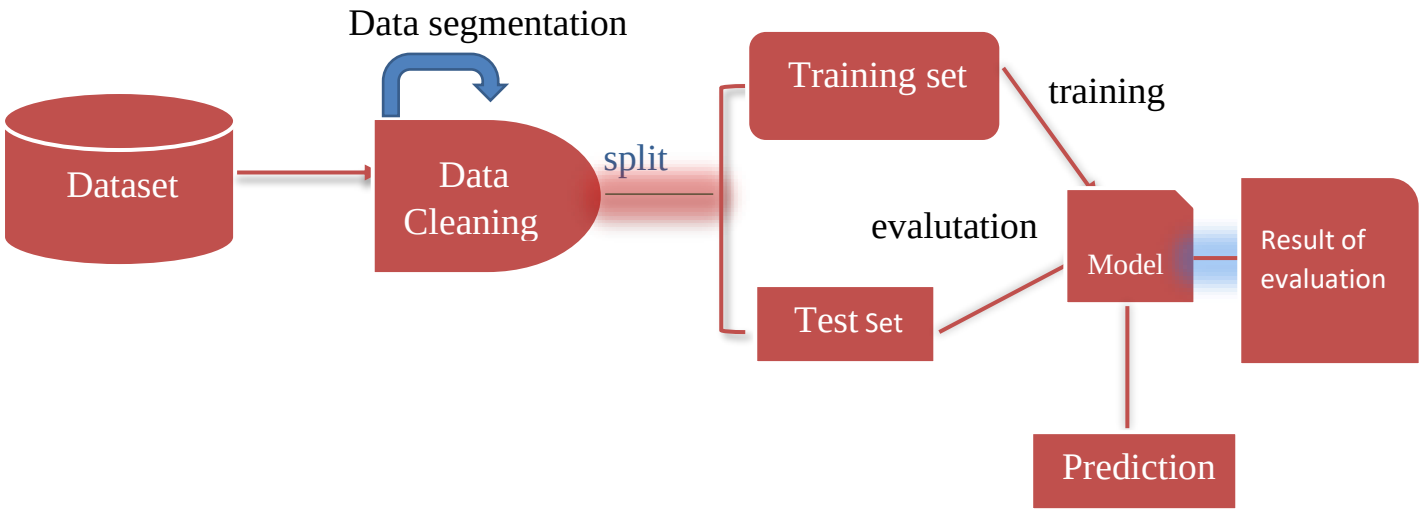
## 3.1 Process Flow

We will use machine learning model and try to analyse and use the best algorithms which it can suit and give a better idea or prediction,

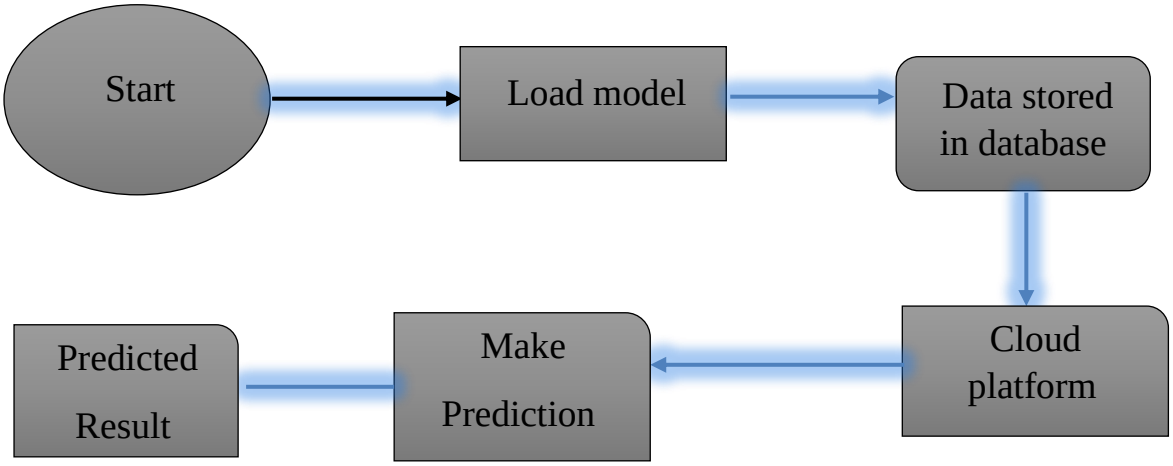
### Proposed Technology




## 3.2 Model Training and Validation



## 3.3 Deployment Process




## 3.4 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 at what step logging required.
- System must be able to log each and every system flow,
- Developer can choose logging method.
- you can choose database logging/file logging as well.
- System should not hang even after using so many loggings, logging just because We can easily debug issues so logging is mandatory to do.

## 3.5 Error Handling

Errors when encountered an explanation will be displayed what went wrong? An error will be defined as anything that fails outside the normal and intended usage.

# 4 *Performance*

ML algorithms gives an idea about the final output of the cuisine service and price tag of categories held across the cuisine.

## 4.1 Reusability

The code written and the components used should have the ability to be reused with no problems.

## 4.2 Application compatibility

The different components used for this project will be using Python as an interface between them. Each component will have its own task to perform and it is work of python to see information is transferred properly.

## 4.3 Resource Utilization

When any task is performed it will have process until the function is finished.

## 4.4 Deployment


# 5 *Dashboards*

Dashboard will be implemented to display and indicate certain and relevant indicators.



# 6 *Conclusion*

Appropriate Machine Learning Model that will help various Zomato Restaurants to predict their respective Ratings based on certain features.