Zomato Project

zomato

Unsupervised ML - Zomato Restaurant Clustering

Machine Learning & GenAI with Microsoft Azure

Project Description

Business Context

Zomato is an Indian restaurant aggregator and food delivery start-up founded by Deepinder Goyal and Pankaj Chaddah in 2008. Zomato provides information, menus and user-reviews of restaurants, and also has food delivery options from partner restaurants in select cities. India is quite famous for its diverse multi cuisine available in a large number of restaurants and hotel resorts, which is reminiscent of unity in diversity. Restaurant business in India is always evolving. More Indians are warming up to the idea of eating restaurant food whether by dining outside or getting food delivered. The growing number of restaurants in every state of India has been a motivation to inspect the data to get some insights, interesting facts and figures about the Indian food industry in each city. So, this project focuses on analysing the Zomato restaurant data for each city in India.

The Project focuses on Customers and Company, you have to analyze the sentiments of the reviews given by the customer in the data and make some useful conclusions in the form of Visualizations. Also, cluster the zomato restaurants into different segments. The data is vizualized as it becomes easy to analyse data at instant. The Analysis also solves some of the business cases that can directly help the customers finding the Best restaurant in their locality and for the company to grow up and work on the fields they are currently lagging in. This could help in clustering the restaurants into segments. Also the data has valuable information around cuisine and costing which can be used in cost vs. benefit analysis Data could be used for sentiment analysis. Also the metadata of reviewers can be used for identifying the critics in the industry.

Restaurant Data

Fields	Description
Name	Name of Restaurants
Links	URL Links of Restaurants
Cost	Per person estimated cost of dining
Collection	Tagging of Restaurants w.r.t. Zomato categories
Cuisines	Cuisines served by restaurants

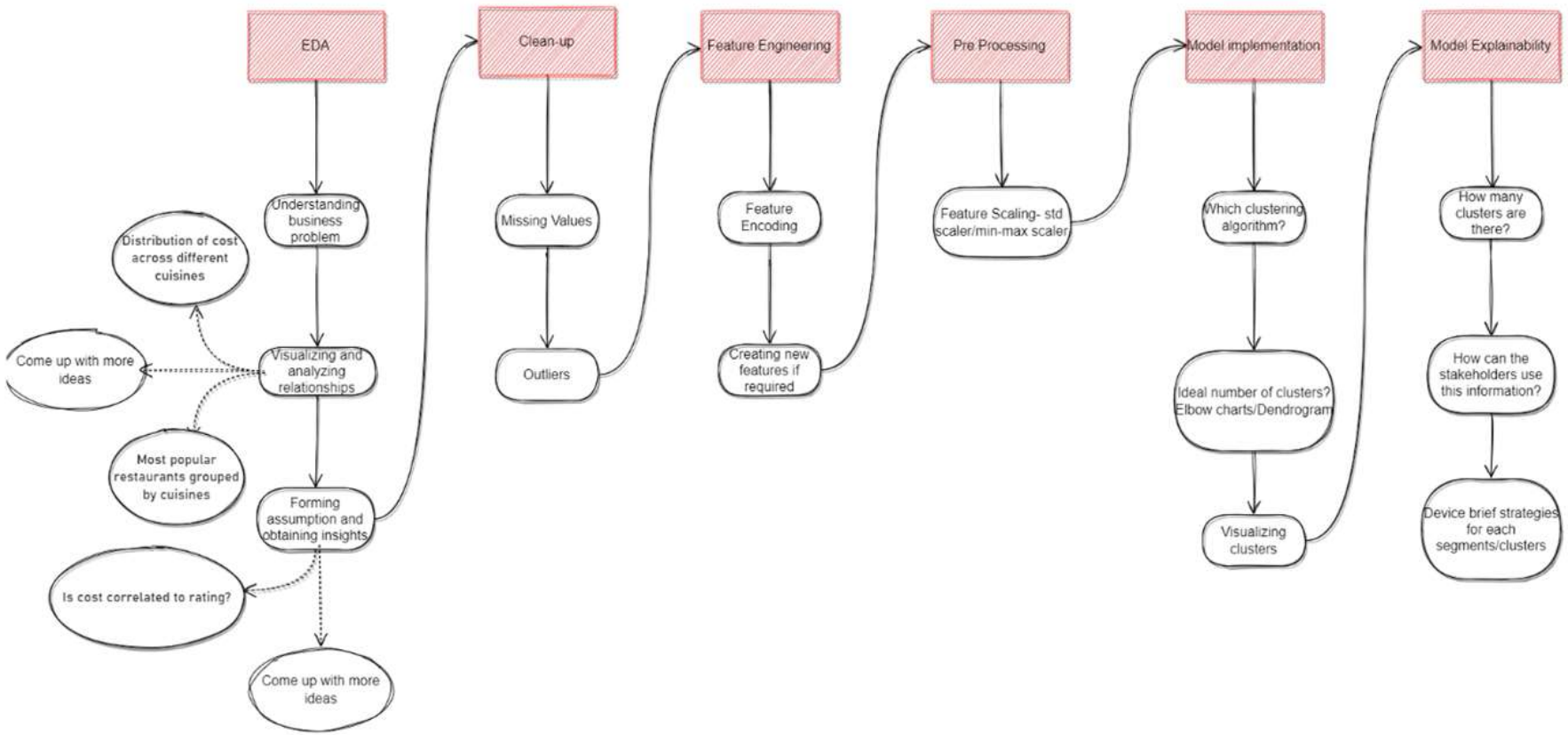
Review Data

Fields	Description
Reviewer	Name of the reviewer
review	Review text
Rating	Rating provided
MetaData	Reviewer metadata-No of reviews and followers
Time	Date and Time of Review
Pictures	No of pictures posted with review.

Main Libraries used:

- Pandas for data manipulation, aggregation
- Matplotlib and Seaborn for visualization and behavior with respect to the target variable
- NumPy for computationally efficient operations
- Scikit learn for model building

Project Architecture:



Rubrics	Weightage
Summary and Technical Documentation in Collab Notebook	10
EDA and Visualization	5
Data Cleaning and Feature Engineering	10
Making Some Hypothesis From the Data Visualized	5
Pick Appropriate Model and Train It	15
Prediction and Calculate Some Evaluation Metrics for Model	10
Final Summary of Conclusion	5
Commented Code	5
Proper Output Formatting	5
Modularity of Code	5
Video Presentation	20
Fluency and Grammatical Accuracy in Video	5