**🍽️ Project Title:**

**“Smart Restaurant Insights & Rating Prediction App”**

**✅ Objective:**

Build an end-to-end system that:

1. Predicts **average item rating or price**
2. Recommends **high-rated, value-for-money restaurants**
3. Visualizes **restaurant trends** by location, cuisine, and popularity
4. Allows interactive exploration via **Streamlit web app**

**🔄 Project Pipeline (End-to-End)**

**1. Data Collection (Given)**

* You already have the dataset (CSV or Excel) - [Zomato Restaurant Dataset](https://www.kaggle.com/datasets/gauravkumar2525/zomato-restaurant-dataset)
* Columns include rating, votes, prices, city, cuisine, bestseller flag, etc.

**2. Data Cleaning & Feature Engineering**

📌 Tools: pandas, numpy

* Handle missing values and incorrect formats
* Encode:
  + Categorical: Cuisine, City, Best\_Seller, etc.
  + Binary: Is\_Highly\_Rated, Is\_Expensive, etc.
* Feature ideas:
  + Rating\_Gap = Dining\_Rating - Delivery\_Rating
  + Value\_Score = Average\_Rating / Price
  + Popularity = Total\_Votes / Price

**3. Exploratory Data Analysis (EDA)**

📌 Tools: matplotlib, seaborn, plotly

* Bar plots of average rating per cuisine or city
* Price vs rating scatter plots
* Heatmap of correlations
* Bestseller trends

**4. Model Training**

📌 Tools: scikit-learn, xgboost, lightgbm, joblib

**🎯 Model 1: Regression**

* **Goal**: Predict Average\_Rating or Prices
* Models: Linear Regression, Random Forest, XGBoost

**🎯 Model 2: Classification**

* **Goal**: Predict Is\_Highly\_Rated, Is\_Expensive
* Models: Logistic Regression, LightGBM, CatBoost

✔️ Evaluate with metrics: RMSE (regression), F1, Accuracy (classification)

**5. Model Serialization**

📌 Tool: joblib

* Save trained model and encoders:

python

CopyEdit

import joblib

joblib.dump(model, "rating\_predictor.pkl")

**6. Web App (Streamlit)**

📌 Tool: streamlit

**🧠 Features:**

* Form for user input: Cuisine, City, Delivery Rating, Votes, etc.
* Display:
  + Predicted Rating or Price
  + Recommended Restaurants
  + Filter options by City/Cuisine/Price
* Show:
  + Bestseller items
  + Value-for-money items
  + Trend visualizations

**7. Dashboard (Optional)**

📌 Tool: Power BI, Tableau, or Plotly Dash

* Create a BI dashboard for stakeholders:
  + Top 10 cities/cuisines by rating
  + Distribution of prices
  + Popularity heatmaps

**8. Deployment**

📌 Options:

* Deploy Streamlit app on **Streamlit Cloud** or **Render**
* Use **GitHub** for code + model repo
* Write a clean **README** with:
  + Project overview
  + Dataset description
  + Model metrics
  + App screenshots
  + How to run locally and online

**🔥 Final Deliverables**

| **Component** | **Tool/Output** |
| --- | --- |
| Data Cleaning & EDA | Jupyter Notebooks |
| Feature Engineering | Script/module |
| ML Model | Trained .pkl model |
| Streamlit App | app.py with UI and prediction |
| Dashboard (Optional) | Tableau or Power BI |
| GitHub Repo | Code + Model + Docs |
| Deployment | Streamlit Cloud/Render link |

**🎁 BONUS: Add-ons for Extra Impact**

* 📈 Trend Analysis: Show change in rating vs price
* 🎯 NLP: If you have text reviews (not in this dataset), apply sentiment analysis
* 🤖 AutoML: Try PyCaret to benchmark models quickly