Food Delivery Time Prediction

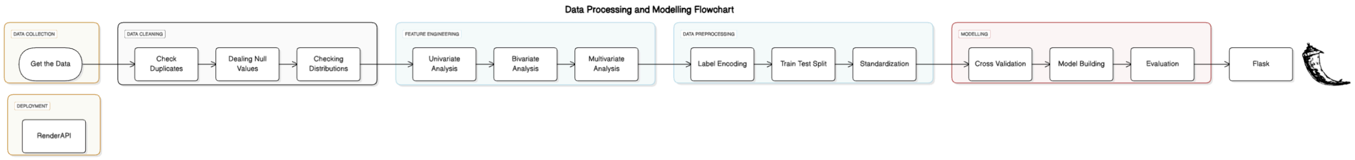

Table of Contents

1. Project Overview
2. Data Source
3. Implementation Details
   3.1 Methods Used
   3.2 Technologies
   3.3 Python Packages Used
4. Steps Followed
   4.1 Data Collection
   4.2 Data Preprocessing
   4.3 Model Development
   4.4 Model Evaluation
   4.5 Deployment
5. Results and Evaluation Criteria
6. Future Improvements

1. Project Overview

This project focuses on developing a **food delivery time prediction model**. The primary goal is to accurately estimate the time it takes for food to be delivered to customers. By providing precise delivery time predictions, food delivery platforms can:

- Enhance customer experience,
- Optimize delivery logistics, and
- Improve overall operational efficiency.

2. Data Source

The dataset used for this project contains detailed information, including:

- **Order Details**: Order time, order type, etc.

- **Location**: Pickup and delivery locations.
- **City Details**: Urban or suburban areas.
- **Delivery Personnel Information**: Experience, ratings, and delivery history.
- **Weather Conditions**: Rain, temperature, etc.
- **Actual Delivery Times**: Recorded times for analysis and validation.

3. Implementation Details

3.1 Methods Used

- Machine Learning
- Data Cleaning
- Feature Engineering
- Regression Algorithms

3.2 Technologies

- **Python** for implementation.
- **Jupyter Notebook** for analysis and model development.
- **Flask** for building the web application.
- **Render API** for deploying the application.

3.3 Python Packages Used

- **Pandas** and **NumPy** for data manipulation.
- **Scikit-learn** for model training and evaluation.
- **Matplotlib** and **Seaborn** for data visualization.
- **XGBoost** for building the best-performing model.

4. Steps Followed

4.1 Data Collection

Gathered the food delivery dataset from the provided data source.

4.2 Data Preprocessing

- **Data Cleaning**: Handled missing values, outliers, and inconsistencies.
- **Feature Engineering**: Extracted and transformed relevant features for the prediction model.

4.3 Model Development

- **Algorithms Explored**:
  - Linear Regression
  - Decision Trees
  - Random Forest
  - XGBoost
- The **XGBoost model** was identified as the best-performing model.

4.4 Model Evaluation

Evaluated the performance of the models using the following metrics:

- **Mean Squared Error (MSE)**
- **Root Mean Squared Error (RMSE)**
- **R-squared (R^2)** Score

4.5 Deployment

- The final model was deployed as a **Flask application** on **Render API** to provide real-time delivery time predictions.

5. Results and Evaluation Criteria

The best-performing model was **XGBoost**, achieving:

- **R-squared (R^2)** score: **0.82**

6. Future Improvements

1. **Feature Expansion**:
   Include additional features such as:
   - Delivery partner characteristics (e.g., mode of transport).
   - Traffic patterns or real-time GPS data.
2. **Comprehensive Data Analysis**:
   Perform deeper exploratory analysis to uncover new insights and correlations.

3. **Model Optimization**:
   Fine-tune hyperparameters to further improve model accuracy.