Food Delivery Time Prediction

COURSE PROJECT REPORT

**18CSE398J -Machine Learning - Core Concepts with Applications**

(2018 Regulation)

**III Year/ VI Semester**

Academic Year: 2022 -2023 (EVEN) By

# Varun Khachane (RA2011026010072) Shresth Gupta (RA2011026010091) Kumar Yash (RA2011026010102)

Under the guidance of

**Dr. V. Vijayalakshmi**

**Assisstant Professor**

**Department of Data Science and Business Systems**



# DEPARTMENT OF DATA SCIENCE AND BUSINESS SYSTEMS FACULTY OF ENGINEERING AND TECHNOLOGY

**SRM INSTITUTE OF SCIENCE AND TECHNOLOGY**

# Kattankulathur, Kancheepuram

# MAY 2023

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **S. No.** | **Contents** | **Page No.** |
| **1** | Abstract | 3 |
| **2** | Introduction | 4 |
| **3** | Dataset | 5 |
| **4** | Methodology | 7 |
| **5** | Experiment and Result | 14 |
| **6** | Conclusion | 16 |
| **7** | References | 17 |

1. **ABSTRACT**

The food delivery industry has seen a significant boom in recent years, especially with the advent of food delivery apps. However, one of the biggest pain points for customers is the uncertainty around the delivery time. Delayed deliveries can be frustrating for customers and impact the overall experience of using food delivery services.

To address this issue, a food delivery time prediction project can be developed that utilizes machine learning algorithms to predict the delivery time accurately. This project aims to develop a model that takes into account various factors such as traffic, weather, distance, and historical delivery data to predict the delivery time accurately.

The project will involve collecting a dataset of delivery orders and their corresponding delivery times. This dataset will be used to train and test the machine learning model. The model will use a combination of regression and classification algorithms to predict the delivery time.

Once the model is developed, it can be integrated into food delivery apps to provide customers with accurate delivery time estimates. This can help to improve the overall customer experience and increase customer loyalty. Additionally, it can also help food delivery companies to optimize their delivery routes and improve their operational efficiency.

Overall, the food delivery time prediction project can be a valuable addition to the food delivery industry, improving the delivery experience for customers while also benefiting food delivery companies.

# INTRODUCTION

The food delivery time prediction project aims to utilize machine learning algorithms to predict the delivery time accurately.

The project will involve the following steps:

* **Data Collection:** The first step is to collect a dataset of delivery orders and their corresponding delivery times. The dataset should include information such as the delivery address, distance from the restaurant, traffic conditions, weather conditions, and historical delivery data.
* **Data Preparation:** The collected data will need to be cleaned and preprocessed before it can be used to train the machine learning model. This will involve tasks such as removing duplicates, handling missing values, and converting categorical data to numerical data.
* **Model Training:** Once the data is prepared, the next step is to train the machine learning model. The model will use a combination of regression and classification algorithms to predict the delivery time accurately. The model will be trained using various features such as distance, traffic, weather, and historical delivery data.
* **Model Evaluation:** After the model is trained, it will be evaluated using a test dataset. The evaluation metrics will be used to determine the accuracy of the model and whether it is performing as expected.
* **Integration:** Once the model is trained and evaluated, it can be integrated into food delivery apps to provide customers with accurate delivery time estimates. The model will continuously learn from new data, which will improve its accuracy over time.

The machine learning-based food delivery time prediction project has the potential to improve the overall customer experience and increase customer loyalty. Additionally, it can also help food delivery companies to optimize their delivery routes and improve their operational efficiency.

# DATASET

Food delivery is a courier service in which a restaurant, store, or independent food- delivery company delivers food to a customer. An order is typically made either through a restaurant or grocer's website or mobile app, or through a food ordering company. The delivered items can include entrees, sides, drinks, desserts, or grocery items and are typically delivered in boxes or bags.

The delivery person will normally drive a car, but in bigger cities where homes and restaurants are closer together, they may use bikes or motorized scooters.

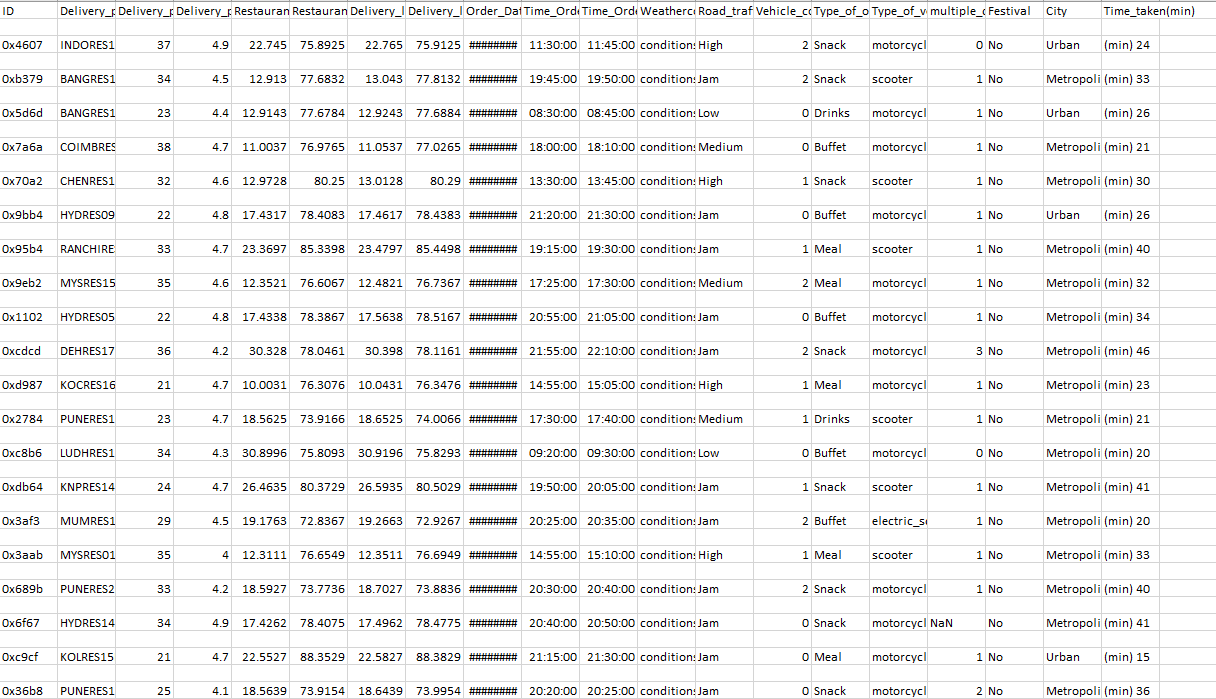
Below are all the features in the dataset:

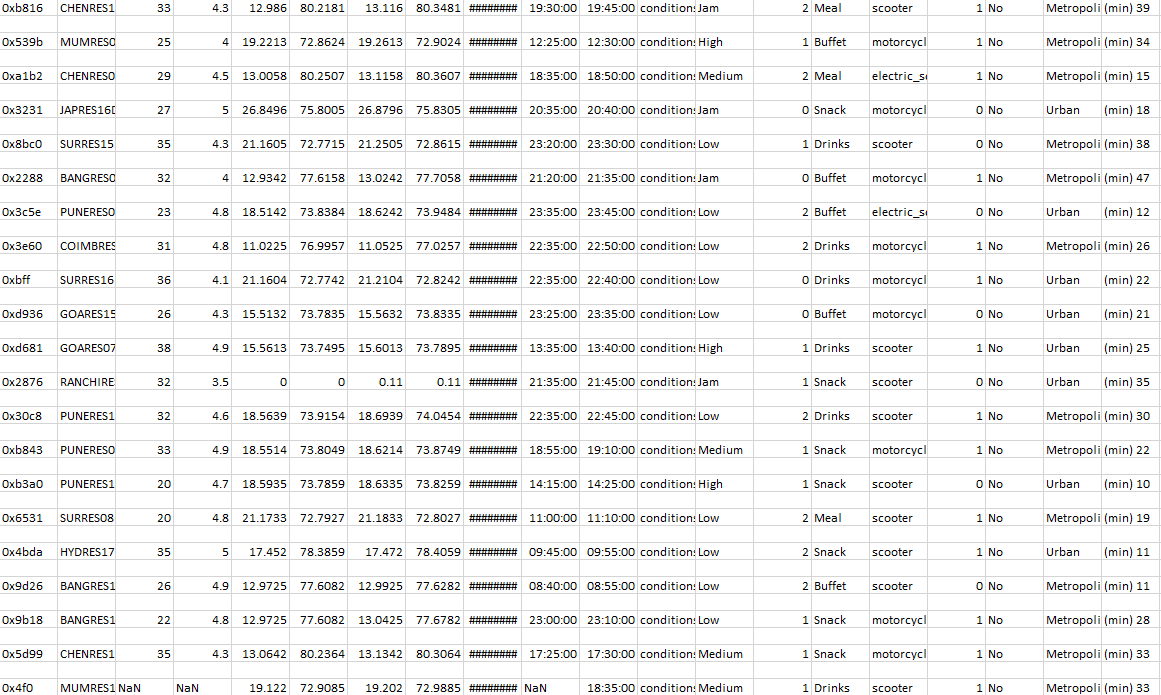
* ID: order ID number
* Delivery\_person\_ID: ID number of the delivery partner
* Delivery\_person\_Age: Age of the delivery partner
* Delivery\_person\_Ratings: ratings of the delivery partner based on past deliveries
* Restaurant\_latitude: The latitude of the restaurant
* Restaurant\_longitude: The longitude of the restaurant
* Delivery\_location\_latitude: The latitude of the delivery location
* Delivery\_location\_longitude: The longitude of the delivery location
* Type\_of\_order: The type of meal ordered by the customer.
* Type\_of\_vehicle: The type of vehicle delivery partner rides
* Time\_taken(min): The time taken by the delivery partner to complete the order

Dataset Link:

https://www.kaggle.com/datasets/gauravmalik26/food-delivery-dataset

# Dataset Snapshots:

****



1. **METHODS**

The food delivery time prediction project methodology can be broken down into several steps, as follows:

1. **Data Collection:** The first step in the project is to collect a dataset of delivery orders and their corresponding delivery times. This dataset should include information such as the delivery address, distance from the restaurant, traffic conditions, weather conditions, and historical delivery data. The dataset can be collected from various sources such as food delivery companies, restaurant chains, and public data sources.
2. **Data Preparation:** The collected data will need to be cleaned and preprocessed before it can be used to train the machine learning model. This will involve tasks such as removing duplicates, handling missing values, and converting categorical data to numerical data.
3. **Feature Selection:** The next step is to select the most relevant features for the model. This will involve analyzing the data to determine which features have the most significant impact on the delivery time.
4. **Model Training:** Once the data is prepared and the features are selected, the next step is to train the machine learning model. The model will use a combination of regression and classification algorithms to predict the delivery time accurately. The model will be trained using various features such as distance, traffic, weather, and historical delivery data.
5. **Model Evaluation:** After the model is trained, it will be evaluated using a test dataset. The evaluation metrics will be used to determine the accuracy of the model and whether it is performing as expected.

Data collection involves obtaining data on delivery times, order details, traffic patterns, and other relevant factors. This data is then preprocessed, which involves cleaning, formatting, and transforming the data to prepare it for machine learning algorithms. Feature engineering involves selecting and extracting the most relevant features from the data, such as the distance between the restaurant and the delivery location, the time of day, and the order size.

Overall, the food delivery time prediction project methodology involves collecting data, preparing it, selecting relevant features, training a machine learning model, evaluating, and tuning the model, integrating it into food delivery apps, and deploying it into a production environment.

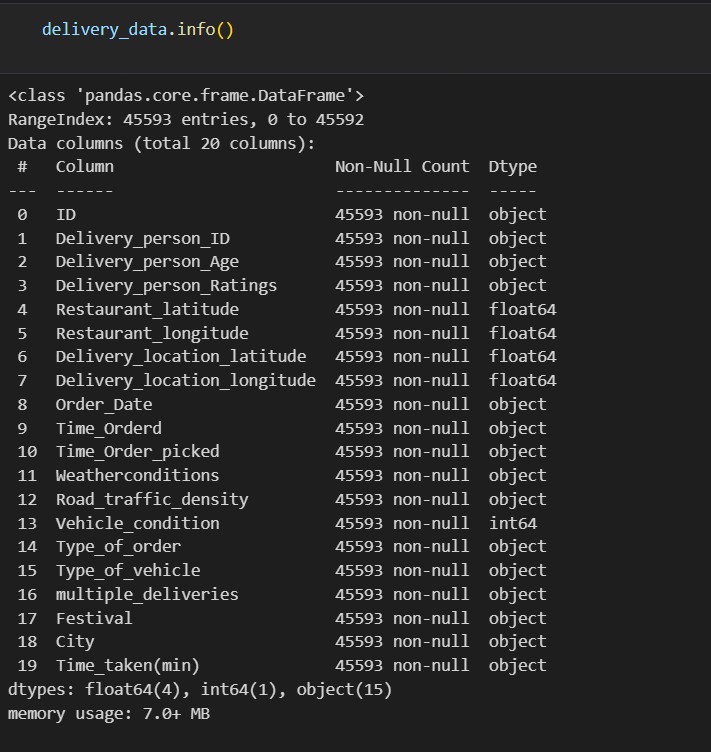
**4A. Data Collection**

This code reads a CSV file named "train.csv" using the pd.read\_csv() function from the Pandas library in Python. The CSV file is expected to be encoded in UTF-8 format. The data from the CSV file is loaded into a Pandas DataFrame object named delivery\_data.

After reading the CSV file, the head() method is called on the delivery\_data DataFrame, which displays the first few rows of the data. This allows you to quickly inspect the data and get an overview of its structure.



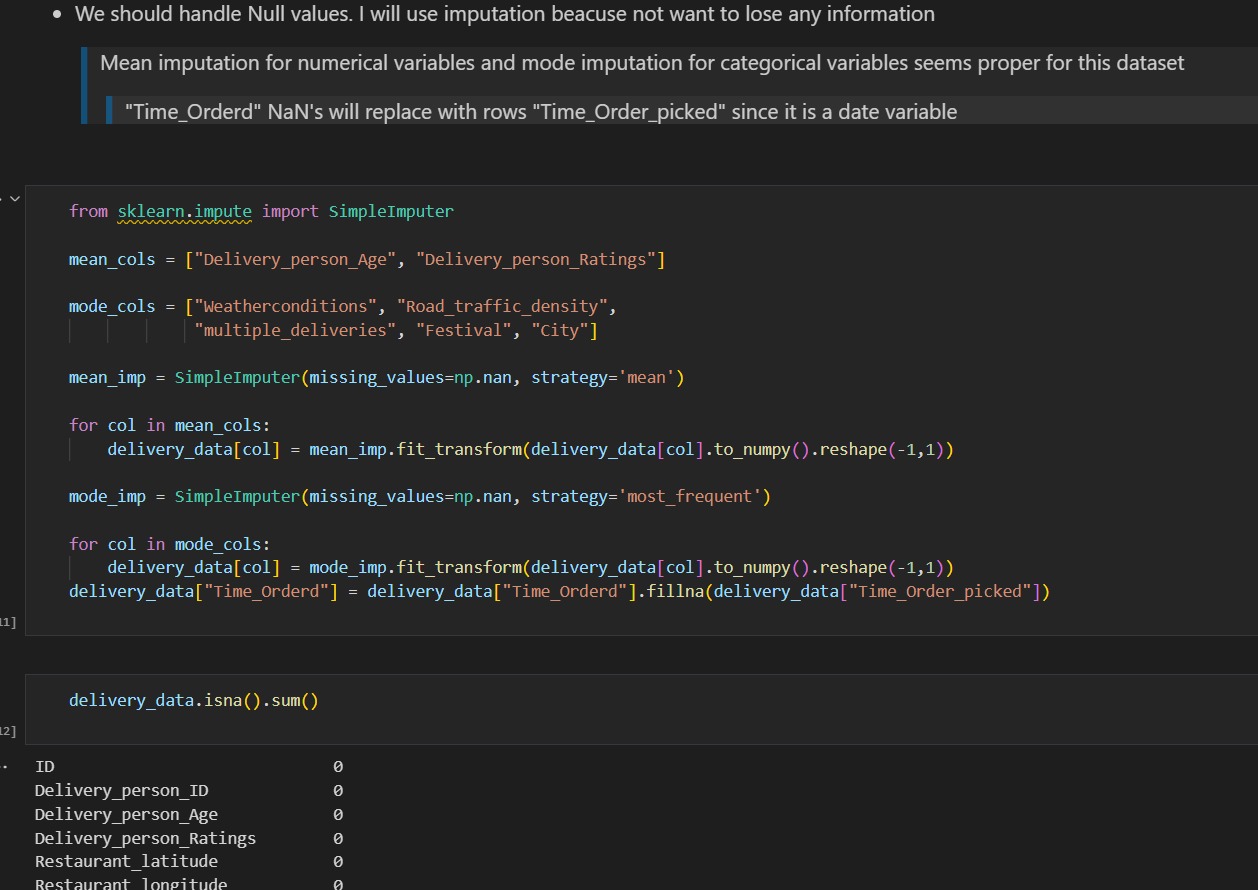
After analyze we identify the following attriubutes



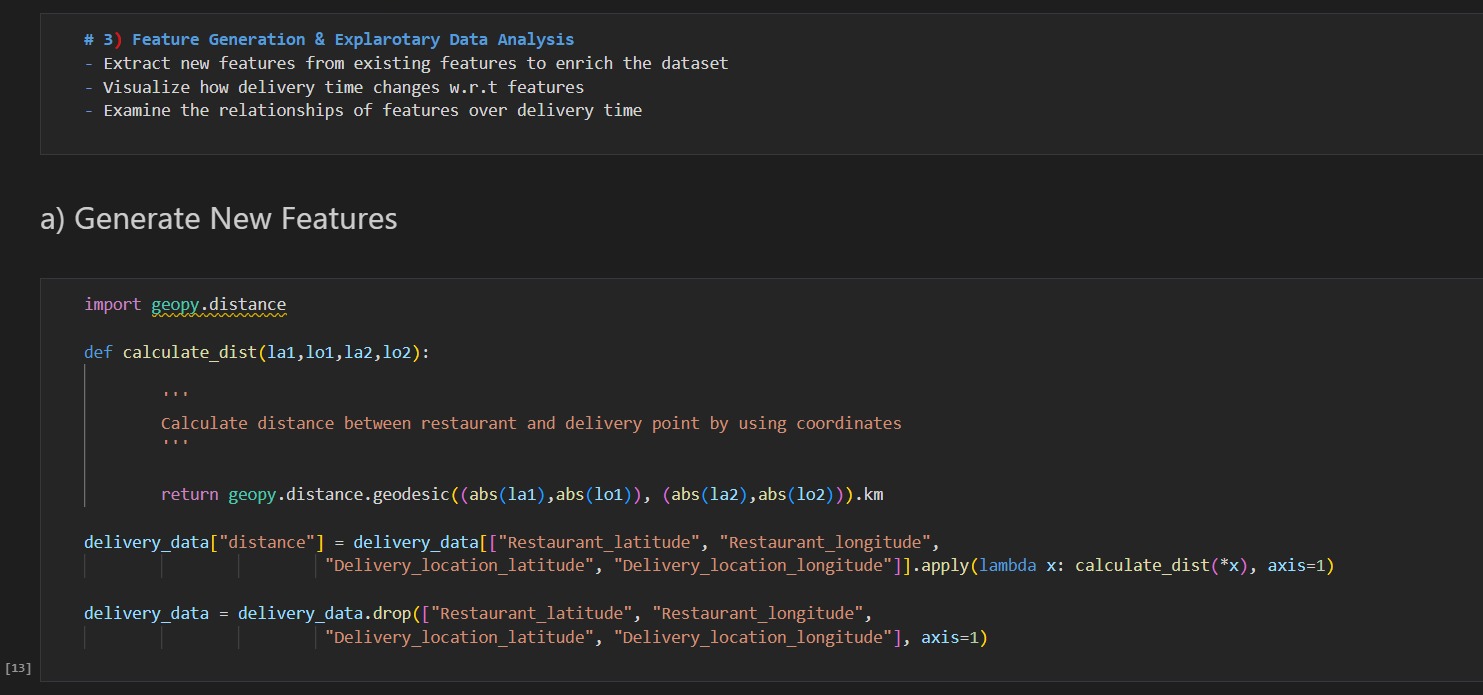
**4B. Data Preparation**

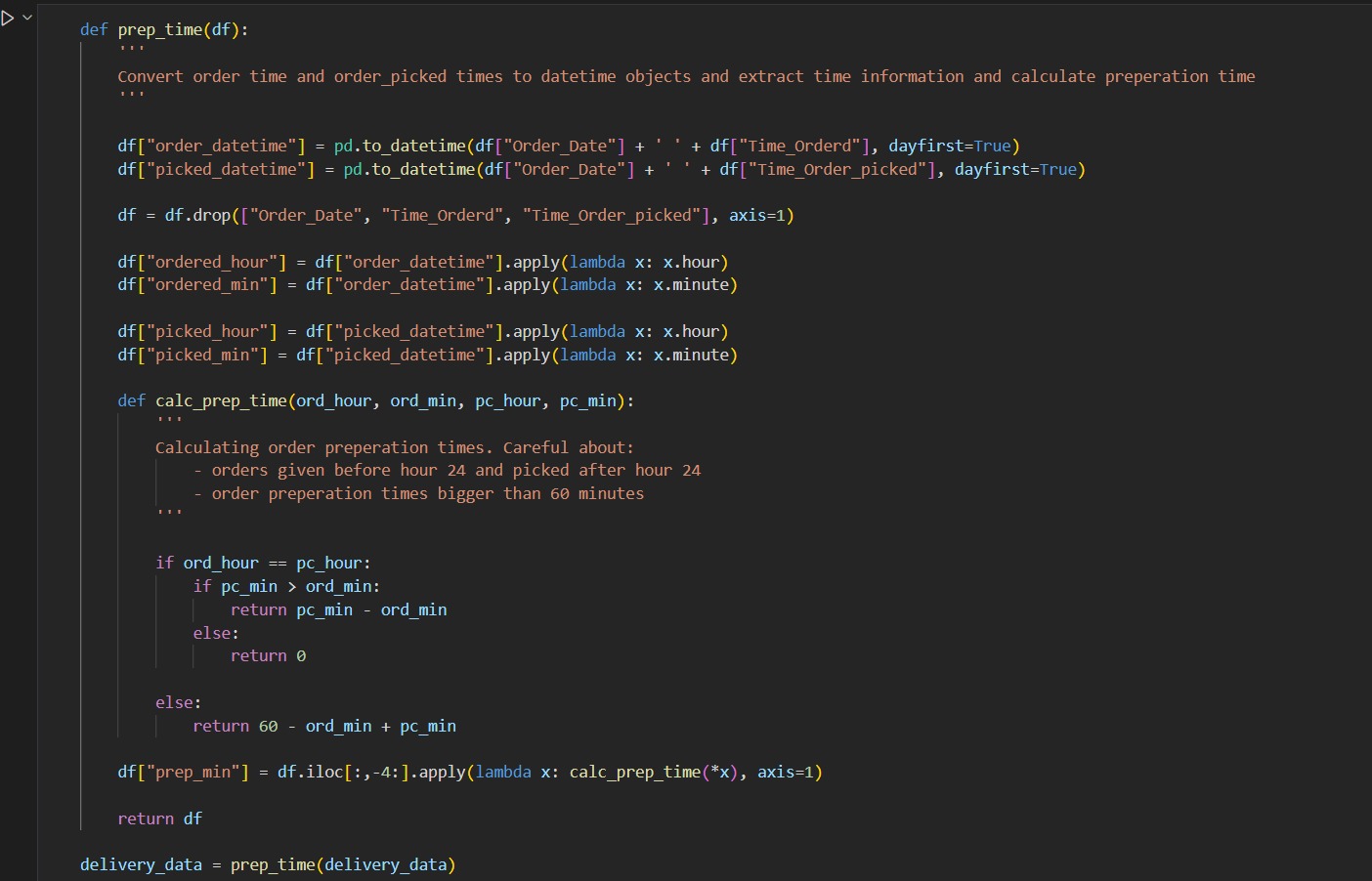
This code snippet is using the SimpleImputer class from scikit-learn library to perform imputation on a DataFrame called delivery\_data. The DataFrame has columns with missing values that need to be imputed.

The code imputes the missing values in the "Time\_Ordered" column with the values from the "Time\_Order\_picked" column using the fillna method of the DataFrame.’



**4C. Feature Selection**

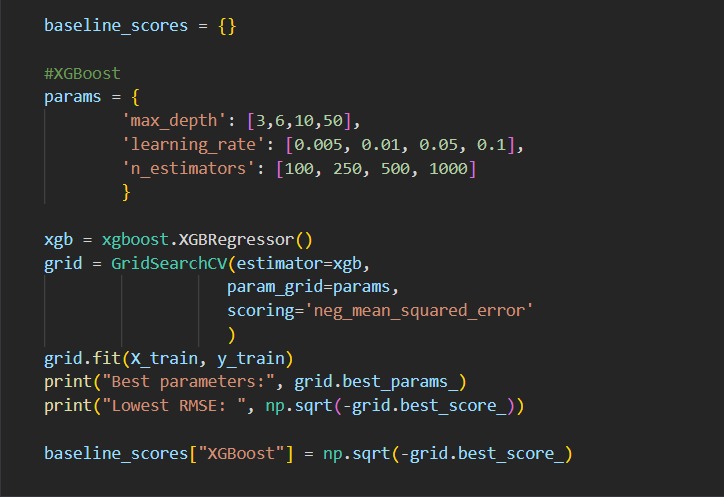
This code calculates the distance between a restaurant and a delivery point using their respective latitude and longitude coordinates. It uses the geopy.distance module to calculate the geodesic distance in kilometers. The calculate\_dist function takes four arguments: la1 and lo1 which represent the latitude and longitude of the restaurant, and la2 and lo2 which represent the latitude and longitude of the delivery location.

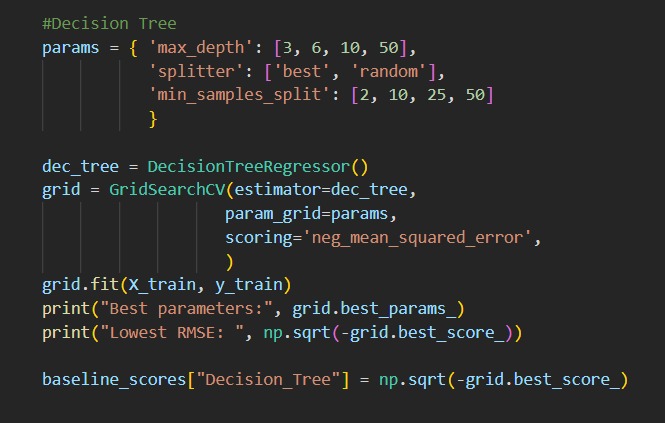
The code defines a function prep\_time(df) that takes a DataFrame df as input and performs various operations to calculate preparation time for orders. It converts the "Order\_Date" and "Time\_Orderd" columns in the DataFrame to datetime objects, extracts hour and minute information, and creates new columns for them. It also does the same for the "Time\_Order\_picked" column. Then, it defines an inner function calc\_prep\_time to calculate the preparation time for each order based on the hour and minute information.

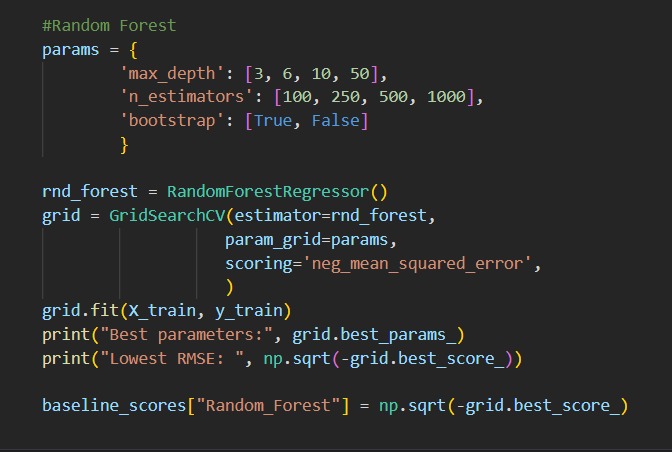
**4D. Model Training**

Train regression-based machine learning models and make predictions

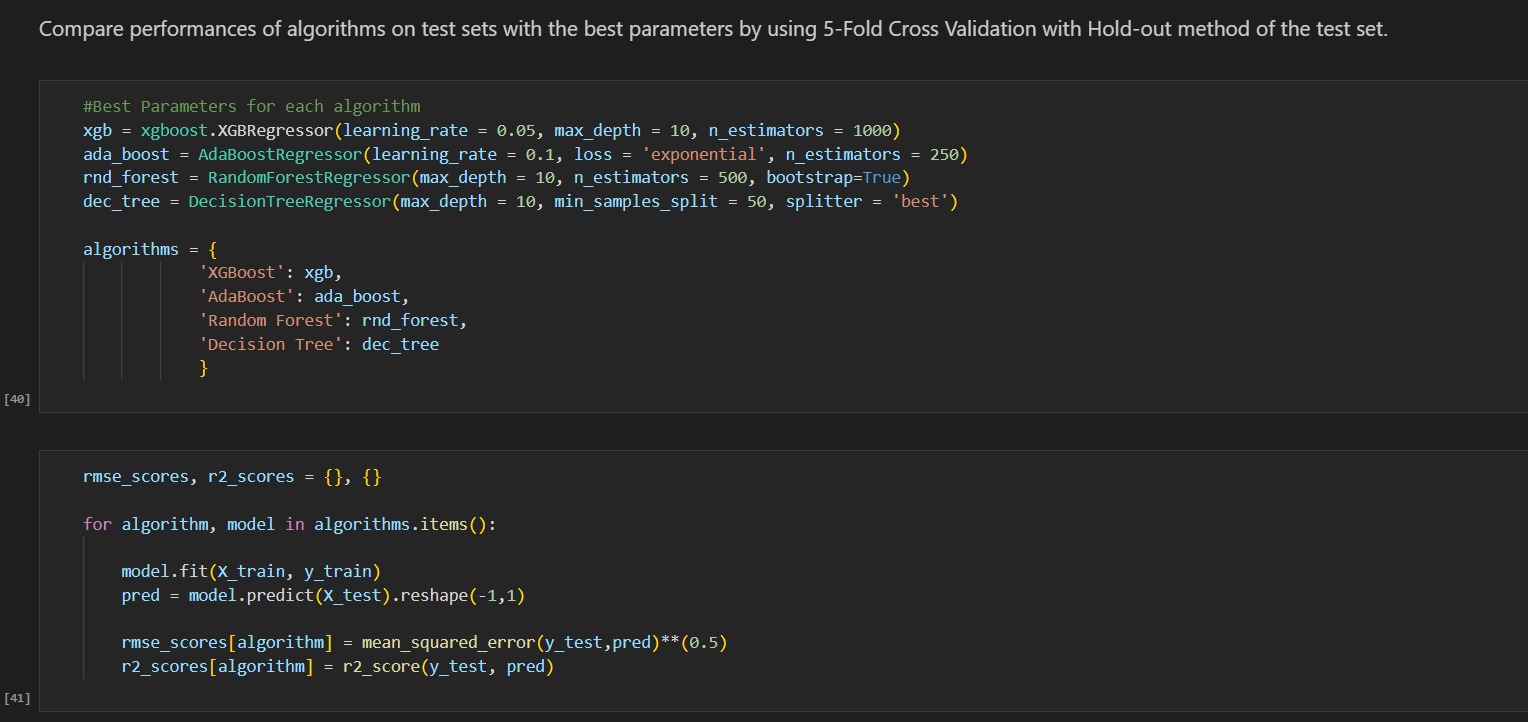
* Train deep learning models
* Make predictions and compare results
* Apply:
* Standardization
* Grid search cross-validaton for hyperparameter tuning
* Examine feature importances

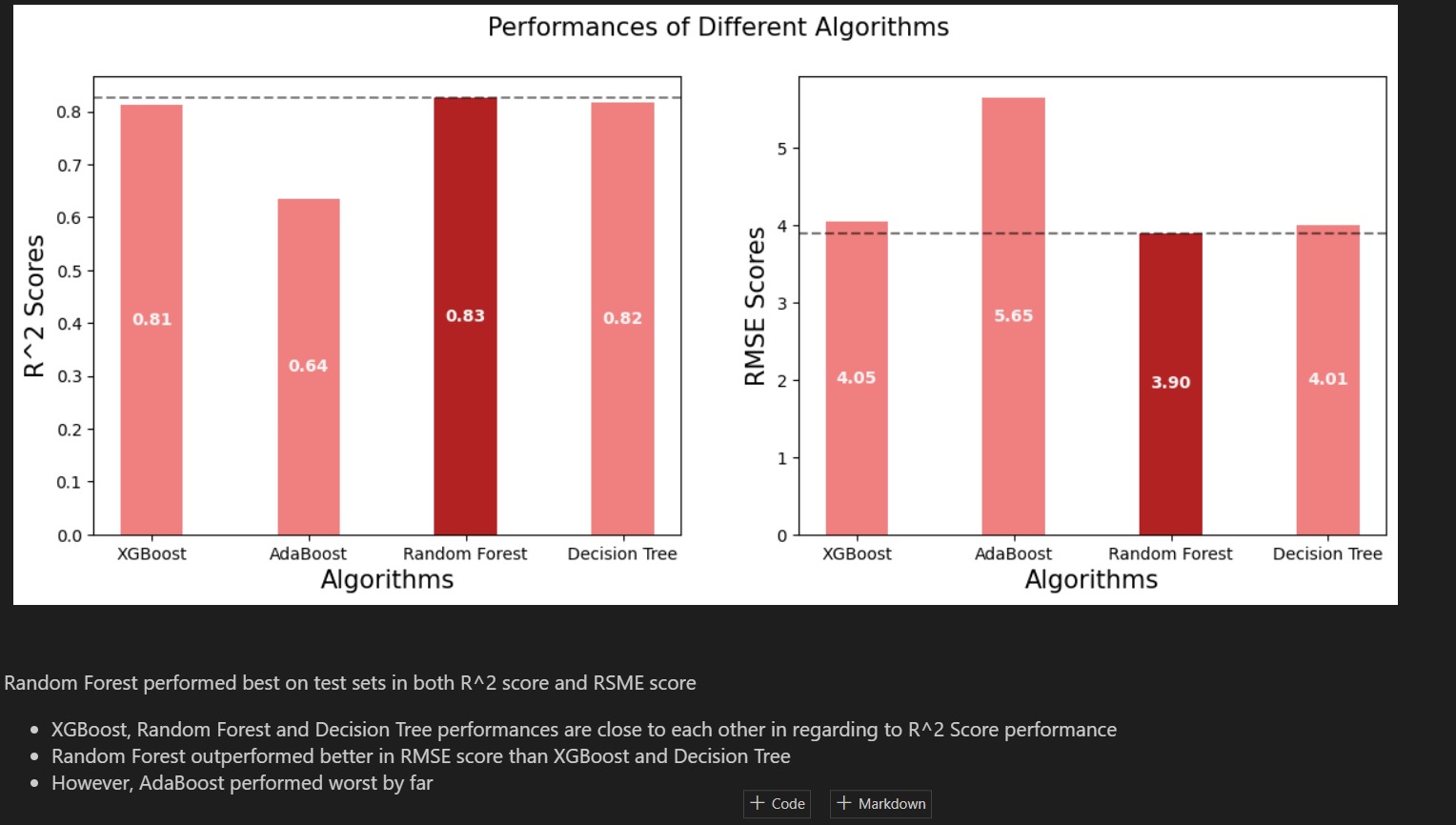




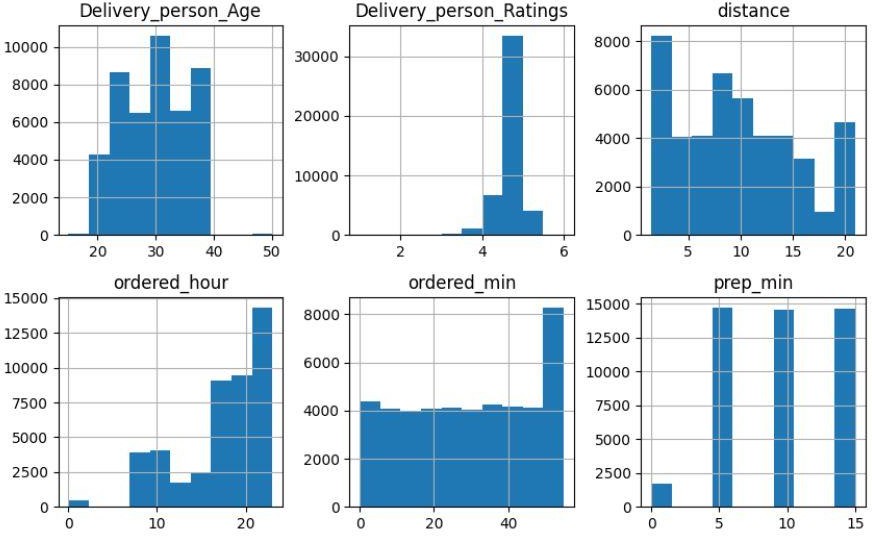


**4E. Data Evaluation**

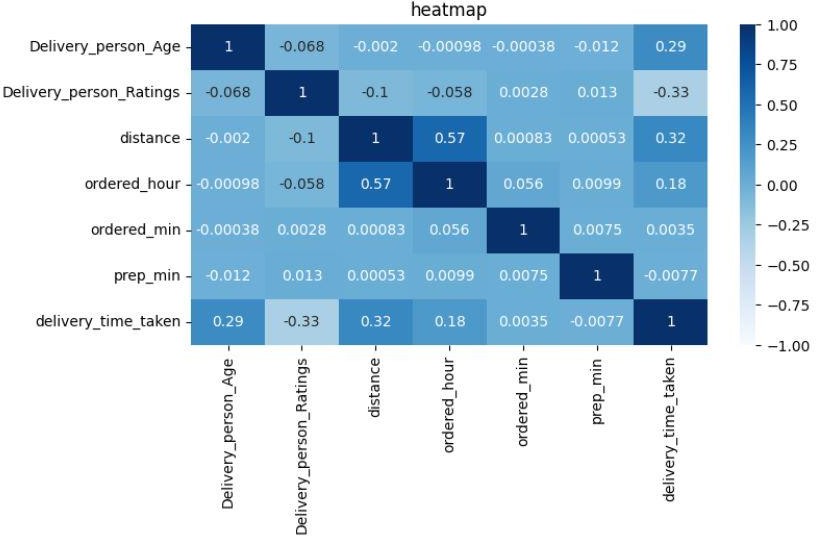
RMSE (Root Mean Squared Error) is a commonly used metric to measure the accuracy of a prediction or a model's performance. It is a measure of the square root of the average of the squared differences between the predicted values and the actual values. RMSE is often used in machine learning and statistics to evaluate the performance of regression models, where the goal is to predict a continuous target variable.

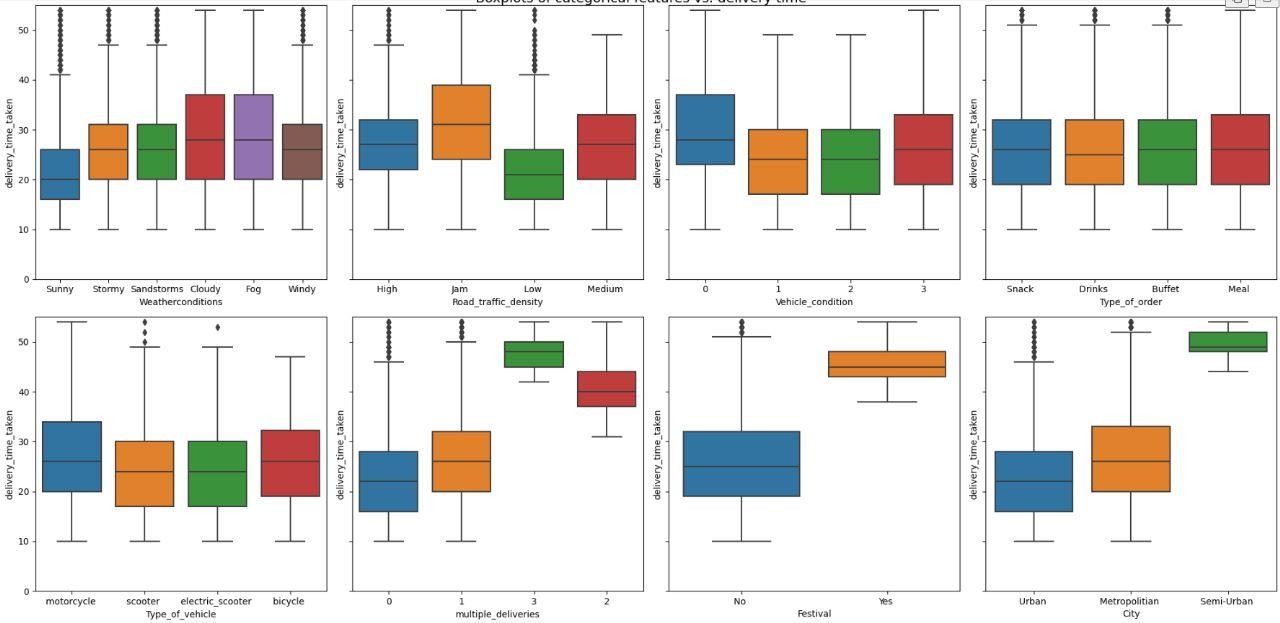


1. **EXPERIMENT AND RESULTS**

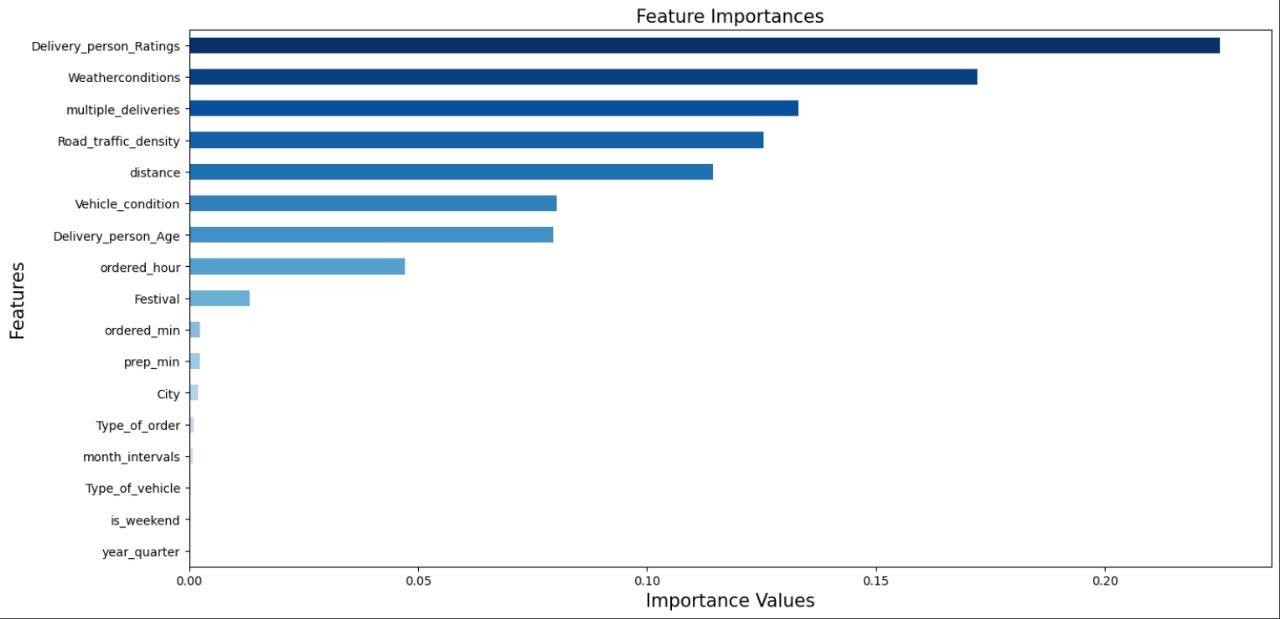


Histogram of Numerical Features in our dataset

Heatmap of the Numerical Columns



Boxplot of Categorical Columns vs delivery time



Feature Importance Graph for the given attribute

# CONCLUSIONS AND FUTURE WORKS

The food delivery industry has grown tremendously in recent years, with the advent of food delivery services that bring restaurant-quality meals straight to people's doors. However, one common problem that plagues this industry is the issue of delivery time estimation. Customers expect their food to arrive within a reasonable timeframe, and delivery drivers need to know how long it will take to deliver an order. Accurate food delivery time prediction is, therefore, essential for improving customer satisfaction and operational efficiency.

Machine learning techniques can be used to solve this problem by predicting delivery times based on various factors such as traffic, distance, and order volume. The methodology for a food delivery time prediction project involves several steps, including data collection, data preprocessing, feature engineering, model training, and model evaluation.

Several machine learning algorithms can be used to train a predictive model, including linear regression, decision trees, and neural networks. These algorithms use the features selected during feature engineering to make predictions about delivery times. The model is then evaluated using metrics such as mean absolute error and mean squared error to determine its accuracy and effectiveness.

In conclusion, food delivery time prediction is an essential problem in the food delivery industry, and machine learning techniques can be used to solve it. A food delivery time prediction project involves several steps, including data collection, data preprocessing, feature engineering, model training, and model evaluation. Several machine learning algorithms can be used to train predictive models, and there are many resources available for learning machine learning concepts and techniques. Accurate food delivery time prediction can improve customer satisfaction and operational efficiency, making it a valuable application of machine learning in the food delivery industry.

# REFERENCES

* **Research Papers:**

1. **Bhandari, R., Jain, A., & Tiwari, R. (2021).** Predicting delivery time of food orders using machine learning algorithms. Journal of Ambient Intelligence and Humanized Computing, 12(7), 6431-6443.
2. **Singhal, P., & Rastogi, A. (2020).** A survey on food delivery time prediction. International Journal of Advanced Science and Technology, 29(7), 4263-4273.
3. **Zhao, H., Cheng, Y., & Li, L. (2018).** A prediction model for food delivery time based on machine learning algorithm. In Proceedings of the 2018 3rd International Conference on Frontiers of Image Processing (ICFIP 2018) (pp. 225-230). Atlantis Press.
4. **Wang, X., Zhang, X., & Liu, S. (2019).** Delivery time prediction for food ordering and delivery platform based on time series analysis. In Proceedings of the 2019 IEEE 19th International Conference on Communication Technology (ICCT) (pp. 154-158). IEEE.

# Books:

1. "The Hundred-Page Machine Learning Book" by Andriy Burkov.
2. "Data Science for Business" by Foster Provost and Tom Fawcett
3. "An Introduction to Machine Learning" by Alpaydin Ethem
4. "Python Machine Learning" by Sebastian Raschka and Vahid Mirjalili

These references provide insights into the different approaches and techniques used for food delivery time prediction using machine learning. They can be useful resources for understanding the project and developing the model.

# Online Resources:

1. **Kaggle:** Kaggle is a platform for data science competitions and provides a wealth of datasets, tutorials, and forums for machine learning practitioners.
2. **Scikit-learn:** Scikit-learn is a popular machine learning library for Python, providing a range of tools for supervised and unsupervised learning, including classification, regression, and clustering.
3. **PyTorch:** PyTorch is an open-source machine learning library for Python, providing a range of tools for deep learning, including neural networks, convolutional networks, and recurrent networks.
4. **Machine Learning Mastery:** Machine Learning Mastery is a website that provides a range of tutorials, courses, and resources for machine learning practitioners, covering topics such as data preparation, feature selection, and model evaluation.
5. **Towards Data Science:** Towards Data Science is a platform for sharing knowledge and ideas in data science, providing a range of articles and tutorials on machine learning, deep learning, and data analysis.

These websites provide a range of resources for machine learning practitioners, from tutorials and examples to libraries and datasets.

# Food Delivery Time Prediction Services

1. DoorDash: https://[www.doordash.com/](http://www.doordash.com/)
2. Postmates: https://postmates.com/
3. Zomato: https://[www.zomato.com/](http://www.zomato.com/)
4. Swiggy: https://[www.swiggy.com/](http://www.swiggy.com/)
5. Deliveroo: https://deliveroo.co.uk/
6. Talabat: https://[www.talabat.com/](http://www.talabat.com/)

These websites and platforms can provide examples of how food delivery time prediction is implemented in real-world scenarios. Studying these websites can also help identify best practices for predicting delivery times accurately and efficiently.