FOOD DEMAND FORECASTING FOR FOOD DELIVERY COMPANY

A PROJECT REPORT ON REMOTE SUMMER INTERNSHIP PROGRAM-2020

ACKNOWLEDGEMENT

CONTENT

INTRODUCTION

Overview

The project Food Demand Forecasting helps all the food delivery companies to forecast the raw materials for the next 10 weeks. Being computerized the system will be more helpful and provide companies at ease of stocking and keeping raw materials so that they are replenished enough according to needs and there is no shortage.

Purpose

One of the biggest challenges of food and beverage manufacturers is to adjust the production and the stocks to minimize the loss of products due to its short perishability. Hence, time series analysis is very important in a wide range of applications, especially when it comes to forecasting, and it encloses many different forecasting models.

LITERATURE SURVEY

Existing Problem

The food and beverage industry is one of the most important sector of any economy, with a significant participation in GDP index. The food delivery companies have been worried about investing in planning their operations, making use, mainly, of forecasting methods in order to become more competitive in the market.

In the case of food industry, the seasonal and the short perishability factors are a limitation to the maintenance of stocks, requiring a forecast with a high accuracy level.

Proposed Solution

The solution that we are proposing to this problem is a Machine Learning Model that analyzes the information as follows:

➤ To forecast the raw materials for next ten weeks.
➤ Assembling the data about fulfillment center like area, city etc., and meal information like category of food and should also have the information about number of orders of a particular category in particular week.

The algorithms used to train the Machine Learning Model are:
1. Linear Regression
2. Decision Tree Regressor

THEORETICAL ANALYSIS

Block Diagram
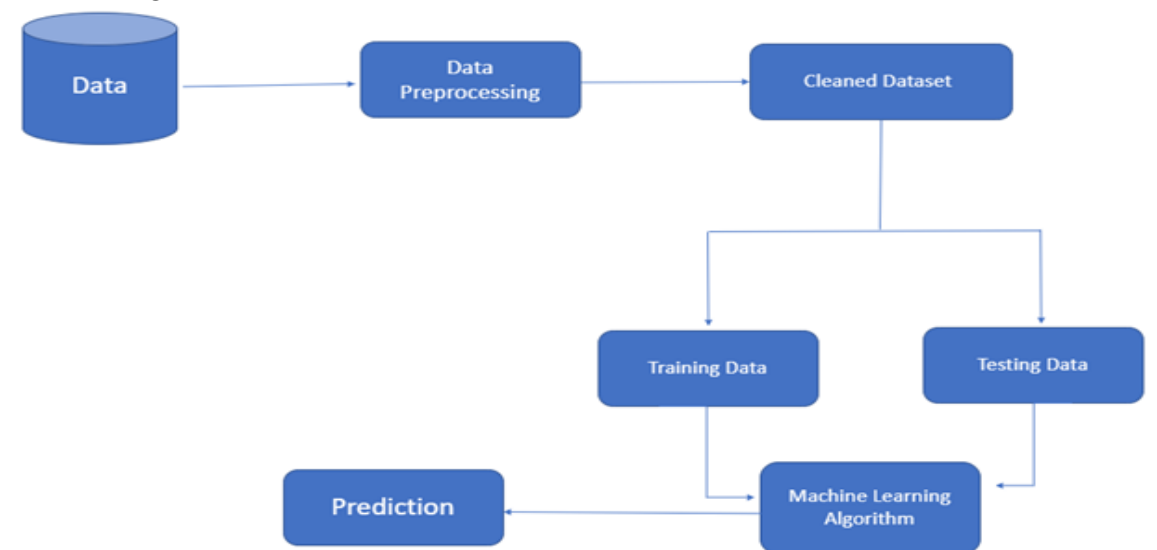
Hardware/Software Designing

The steps followed in developing the model are as follows:
1. Data Acquisition : We have downloaded the dataset from kaggle.
2. Data Visualization : Visualizing the different columns of the dataset by plotting graphs .
3. Data Preprocessing : Taking care of any missing entries and of any relations within the dataset.
4. Model Building : Developing a model with the help of different classification algorithms to predict the raw materials for the next 10 weeks.

# EXPERIMENTAL INVESTIGATIONS

Importing Libraries:
The first step of starting the analysis on the dataset is by importing the libraries numpy, pandas, matplotlib. Numpy is the numerical python library used for all sorts of mathematical calculations.Matplotlib is used for visualization of data. Pandas is used for data manipulation.

Importing the Dataset:
The dataset that was downloaded from kaggle is first uploaded in a .csv format and then imported in the file.
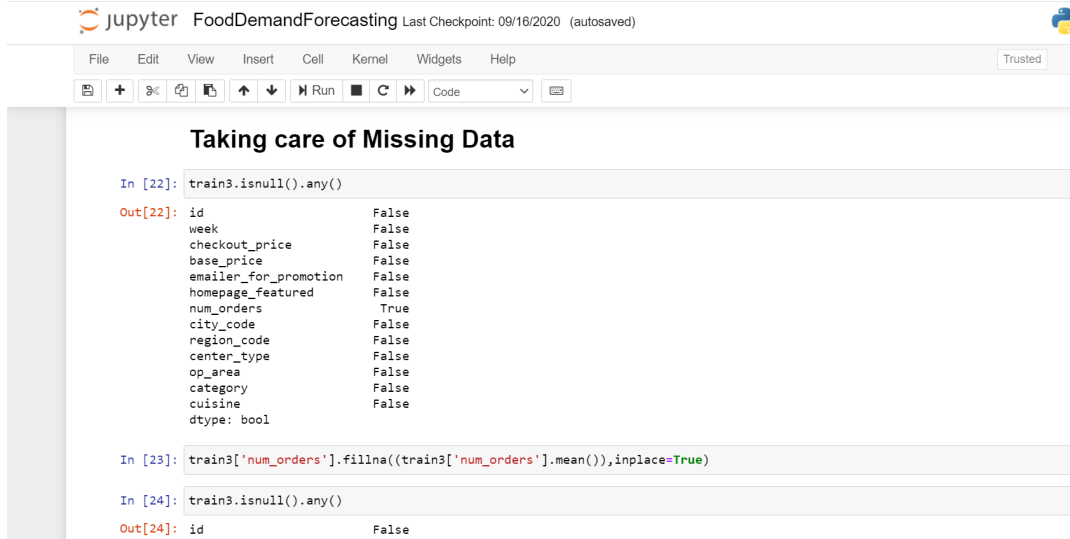
Data Visualization:
The raw data is visualized by plotting graphs of different input characteristics and the solution is proposed based on this analysis. This is an important step in analysing which category of algorithms(regression or classification) are to be applied on the model.

Out[13]: <matplotlib.axes._subplots.AxesSubplot at 0x1eda15f5ee0>

Out[14]: <matplotlib.axes._subplots.AxesSubplot at 0x1eda2ab56a0>
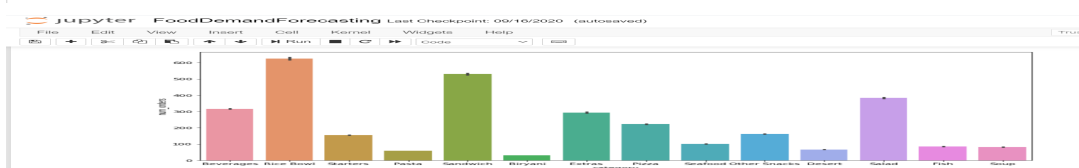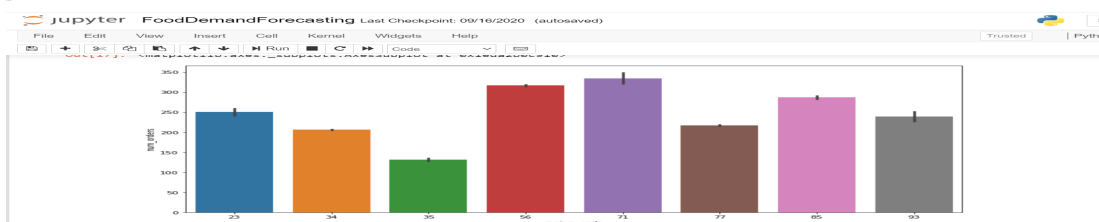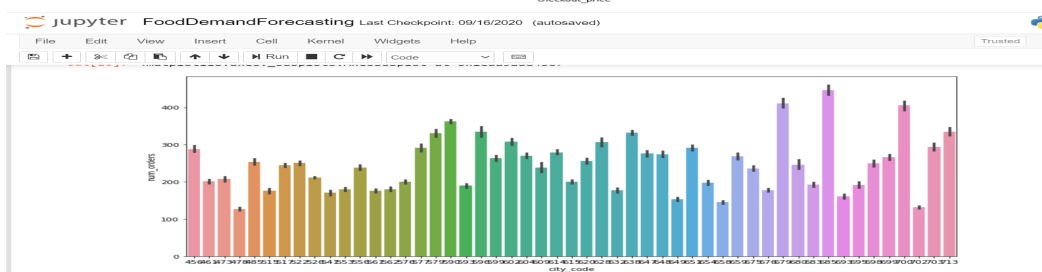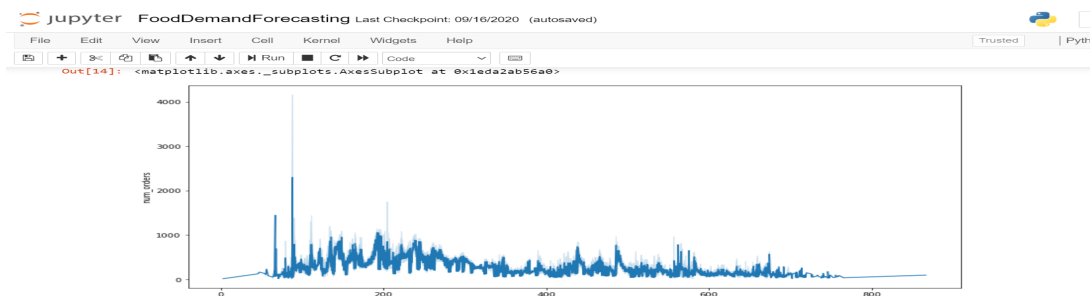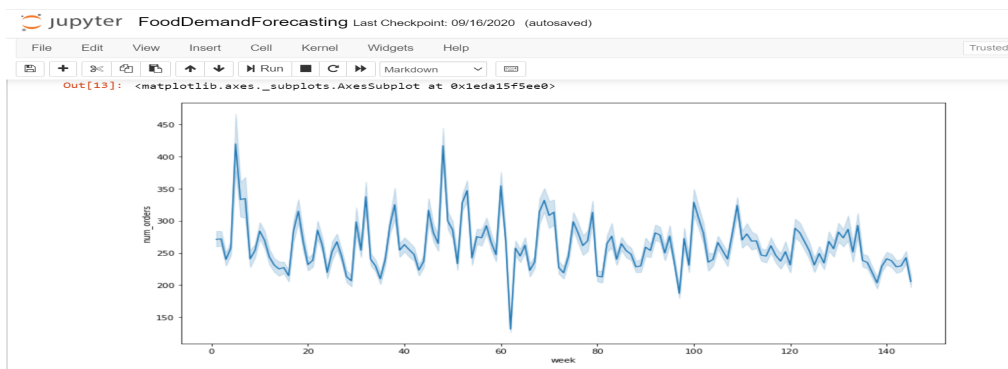
Taking care of missing data:

The data after visualization is then checked to find out the missing data in the dataset , and then filling them by using different mode and mean methods as per requirement. For numerical values we use mean and for categorical values , mode is used.

Applying One Hot Encoder and Label Encoder:

It mainly did the work of converting categorical columns into binary elements for easy acceptibility by the algorithm.

Feature Scaling:

The main purpose of feature scaling is to convert all varying data in one format lying between 0-1.

Splitting the dataset:

The dataset is then divided into two sets:
1. Training Data- It mainly comprises of 80% of the data. It is used to train the data to acquire desired output.
2. Testing Data- It comprises of the other 20% of the data that is tested on the model and on the basis of which accuracy of the model is predicted.

Model Building:

The main work here is of applying the different algorithms on the training set and checking the accuracy of the model on the basis of these algorithms. The evaluation is done using RMSE values.

The two algorithms that we used to build the model are:

1. Linear Regression
2. Decision Tree Regressor

Application Building:
The webpage is created using HTML and python using Flask framework.

RESULT

The Food Demand Forecasting Model is analysed using the Linear Regression and the predicted RMSE value is 334.3732. On using the Decision Tree Regressor algorithm the predicted RMSLE value is 69.565.

ADVANTAGES AND DISADVANTAGES

The advantages that the model proposes are:
1. User friendly interface.
2. Easy results with one click.
3. Accuracy in results.

The disadvantage is that the precision is not so high using these algorithms , so there is requirement of advanced algorithms.

APPLICATIONS

The application of the model is that it can be used by the food delivery companies in predicting the amount of raw materials that will be needed to keep in store for the next 10 weeks. The wastage and loss issues can thus be solved.

CONCLUSION

In this project , we have proposed methods for predicting the raw materials used in food delivery companies using Machine Learning techniques. The two methods that we used are Linear Regression and Decision Tree Regressor.The system was implemented using the models and their performance was evaluated.A GUI was made based on the model for use in the medical industry.

REFERENCE

1.https://www.kaggle.com/kannanaikkal/food-demand-forecasting
2.https://github.com/CallMeAmartya/Food-Demand-Forecasting