BIG MART SALES PREDICTION USING MACHINE LEARNING

NAME: MADHURA M V

COLLEGE: EAST POINT COLLEGE OF ENGINEERING AND TECHNOLOGY

DOMAIN: DATA SCIENCE WITH AIML

PROJECT NAME: BIG MART SALES PREDICTION USING ML

ABSTRACT

Nowadays shopping malls and Big Marts keep the track of their sales data of each and every individual item for predicting future demand of the customer and update the inventory management as well. These data stores basically contain a large number of customer data and individual item attributes in a data warehouse. Further, anomalies and frequent patterns are detected by mining the data store from the data warehouse. The resultant data can be used for predicting future sales volume with the help of different machine learning techniques for the retailers like Big Mart. In this paper, we propose a predictive model using XG boost Regressor technique for predicting the sales of a company like Big Mart and found that the model produces better performance as compared to existing models.

INTRODUCTION

Big Mart is a big supermarket chain, with stores all around the country and its current board set out a challenge to all Data Scientist out there to help them create a model that can predict the sales, per product, for each store to give accurate results. Big Mart has collected sales data from the year 2013, for 1559 products across 10 stores in different cities. Big Mart has collected sales data from Kaggle, for various products across different stores in different cities. With this information the corporation hopes we can identify the products and stores which play a key role in their sales and use that information to take the correct measures to ensure success of their business.

OBJECTIVE

The objective of this framework is to predict the future sales from given data of the previous year's using Machine Learning Techniques. Another objective is to conclude the best model which is more efficient and gives fast and accurate result by using XG Boost Regressor. To find out key factors that can increase their sales and what changes could be made to the product or store's characteristics. Objectives of these project are: a) Predicting future sales from a given dataset. b) To understand the key features that are responsible for the sale of a particular product. c) Find the best algorithm that will predict sales with the greatest accuracy.

PROBLEM STATEMENT

Problem Statement: To find out what role certain items play and how they affect their sales by understanding Big Mart Sales. To find out what role certain properties of an item play and how they affect their sales by understanding Big Mart sales." In order to help Big Mart achieve this goal, a predictive model. The aim is to build a predictive model and find out the sales of each product at a particular store. Using this model, Big Mart will try to understand the properties of products and stores which play a key role in increasing sales. This model is used to predict both classification and regression types of problems

IMPLIMENTATION

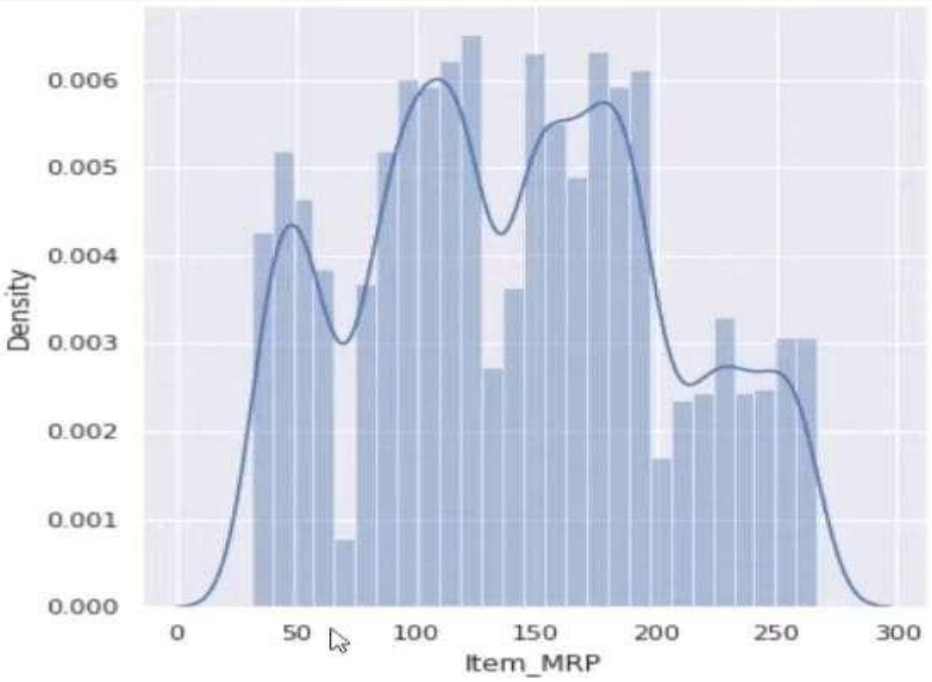
predicting sales in a retail environment, like Big Mart, can be done using various machine learning algorithms. I'll provide a high-level outline of the steps involved in implementing a sales prediction model and we are using here pandas, Numpy and seaborn, and Matplot. It includes the steps like Importing necessary Libraries, Load and Preprocess the data, Model selection and training, and visualization.

Importing Libraries:

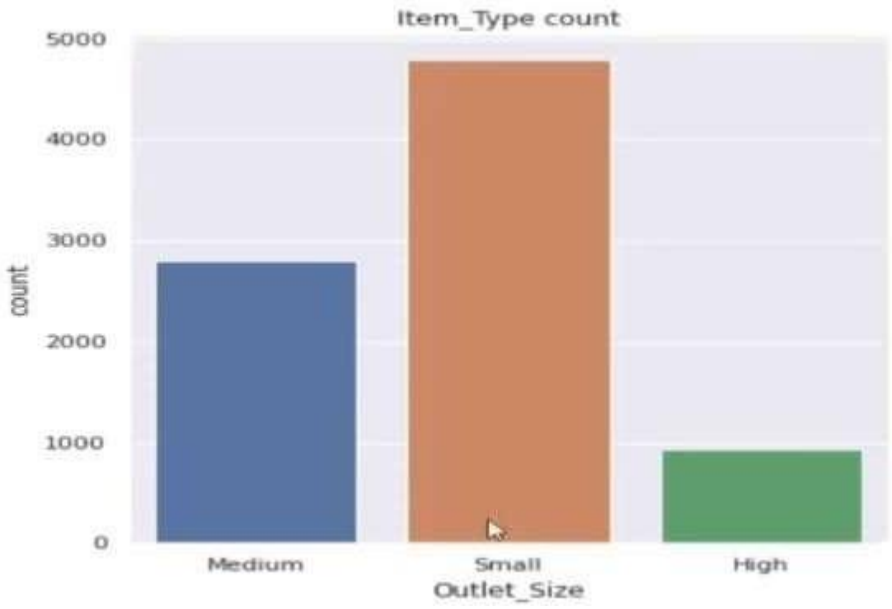
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns
from sklearn.preprocessing import LabelEncoder
from sklearn.model_selection import train_test_split
from xgboost import XGBRegressor
from sklearn import metrics
```

Results

```
# Item MRP distribution
plt.figure(figsize=(6,6))
sns.distplot(big_mart_data['Item_MRP'])
plt.show()
```




```
plt.figure(figsize=(6,6))
sns.countplot(x='Outlet_Size', data=big_mart_data)
plt.show()
```



CONCLUSION

Our predictions help big marts to refine their methodologies and strategies which in turn helps them to increase their profit. The results predicted will be very useful for the executives of the company to know about their sales and profits. This will also give them the idea for their new locations or Centre’s of Big-mart.