Mini Project

Project Description

The data scientists at BigMart have collected 2013 sales data for 1559 products across 10 stores in different cities. Also, certain attributes of each product and store have been defined. The aim of this data science project is to build a predictive model and find out the sales of each product at a particular store.

Using this model, BigMart will try to understand the properties of products and stores which play a key role in increasing sales.

The data has missing values as some stores do not report all the data due to technical glitches. Hence, it will be required to treat them accordingly.

We will handle this problem in a structured way. We will be following the table of content given below.

1).Problem Statement

2).Hypothesis Generation

3).Loading Packages and Data

4).Data Structure and Content

5).Exploratory Data Analysis

6).Univariate Analysis

7).Bivariate Analysis

8).Missing Value Treatment

9).Feature Engineering

10).Encoding Categorical Variables

11).Label Encoding

12).One Hot Encoding

13).PreProcessing Data

14).Modeling

15).Linear Regression

16).Regularized Linear Regression

17).RandomForest

18).XGBoost

19).Summary

Curriculum For This Project

The Business Problem Exploring

The Dataset

Exploratory Data Analysis (eda) - Outliers

Exploratory Data Analysis (eda) - Graphs

Converting Categorical To Numerical

Seperating Training And Test Data

Running The Models

Hyper Parameter Tuning XGB And GBR

Standard Scaling 06m Robust Scaling

Final Predictions On The Test Dataset

Saving The Final Model