

IBM NaanMudhalvan

Domain : Applied Data Science

Project 10: Product Demand Prediction with Machine Learnings

Agenda :

Introduction

Problem Definition

Design Thinking

Dataset

Conclusion



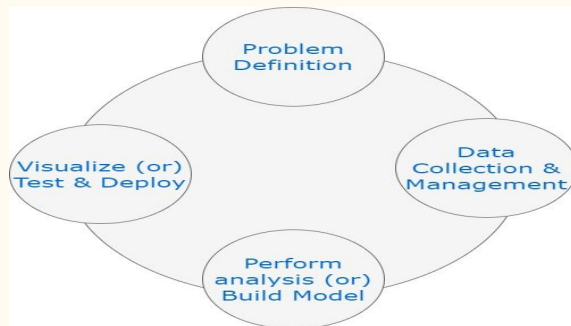
Introduction :

Data science is getting useful information from large amounts of unstructured data using software, data mining and statistical methods, algorithms, and machine learning principles.

According to our project we already have studied that the demand for a product varies with the change in its price. If you take real-world examples, you will see if the product is not a necessity, then its demand decreases with the increase in its price and the demand increases with the decrease in its price. In this project, I will walk you through the task of product demand prediction with machine learning using Python.

Problem Definition :

The problem is to create a machine learning model that forecasts product demand based on historical sales data and external factors. The goal is to help businesses optimize inventory management and production planning to efficiently meet customer needs. This project involves data collection, data preprocessing, feature engineering, model selection, training, and evaluation. We'll start by importing the necessary Python libraries and the dataset we need for the task of product demand prediction.

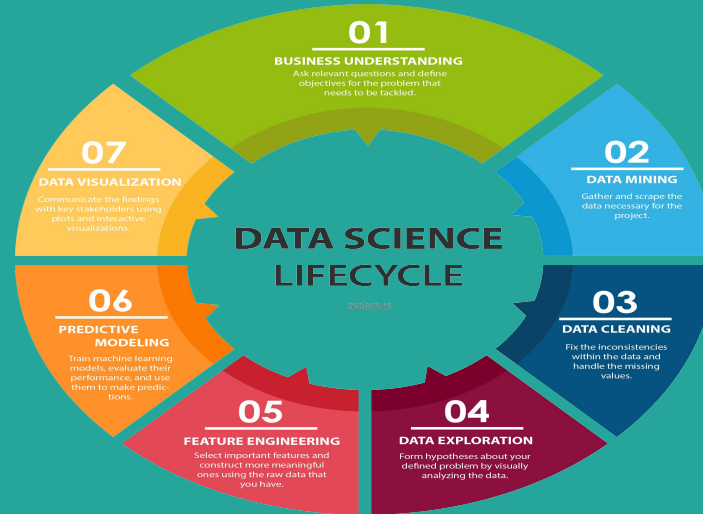


Design Thinking :

In this phase we are going to explain how we crack the problem for creating a machine learning model that forecast product demand based on historical sales data and external factors.

By achieving the following 6 steps we have create a model for product demand prediction. They are:

- Data Collection
- Data Preprocessing
- Feature Engineering
- Model Selection
- Model Training
- Evaluation



Data Collection :

This process involves collecting historical sales data and external factors that influence demand, such as marketing campaigns, holidays, economic indicators, etc..But in our case we are already provided with datasets for product demand prediction from KAGGLE platform.



Data Preprocessing :

1. Clean and preprocess the data:

Here , we are going to clear outliers that perform distinct significant from the other observation

2. Handle the missing values:

Here, we are working with large amount of data,therefore every is need to be consistent and every datas are need to be verified before the process.

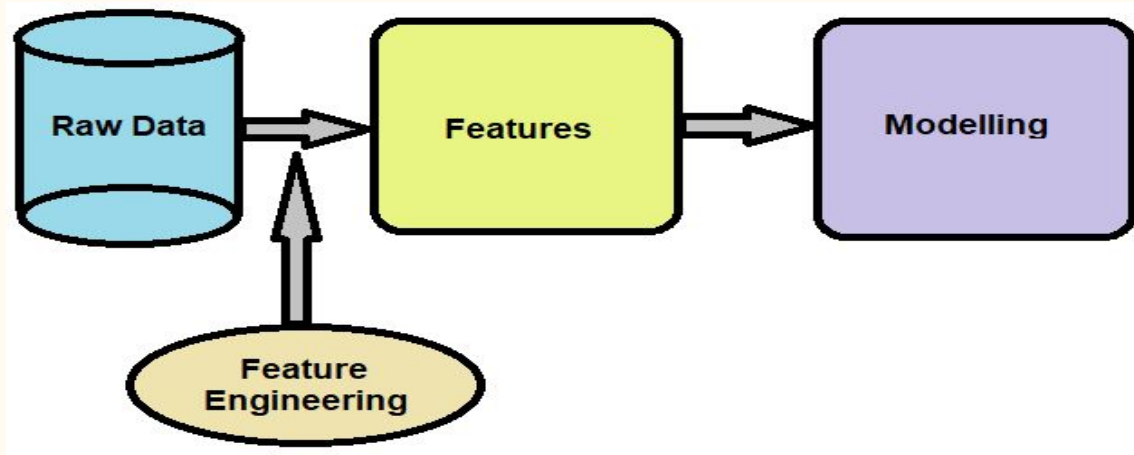
So if we found any entity with missing values we have to replace a ZN() function replaces a Null value for the data field that is placed inside the brackets with a zero.

3.convert categorical features into numerical values:

Here, we are converting the data values into numerical values for example ,most of the price values are in INR and some of them are in \$,so here we have to convert them(\$) into INR.

Feature Engineering :

Create additional features that capture seasonal patterns, trends and external influences on product demand.



Model Selection :

To choose suitable regression algorithms (e.g., Linear Regression, Random Forest, XGBoost) for demand forecasting. We split the data into training and test sets and use the decision tree regression algorithm to train our model.

Machine Learning Models



Classification Models

Clustering

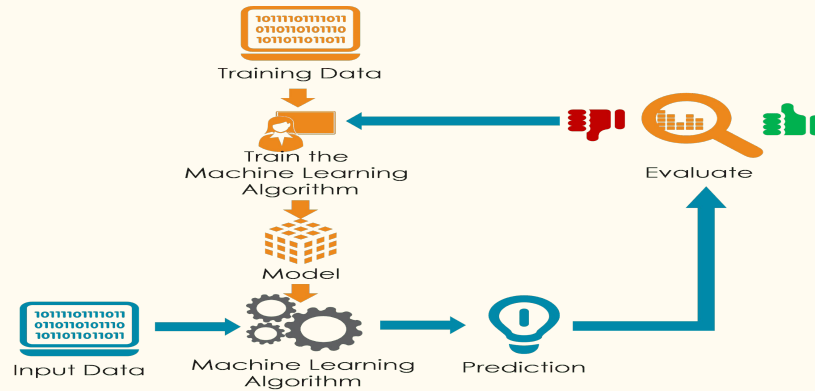
Regression Models

Dimensionality Reduction

Deep Learning etc.

Model Training :

To train the selected model using the preprocessed data. We need to train a model that can predict the demand for the product in the market with different price segments. The task of training a machine learning model to predict the demand for the product at different prices, I will choose the Total Price and the Base Price column as the features to train the model, and the Units Sold column as labels for the model.



Evaluation :

To Evaluate the model's performance using appropriate regression metrics like Mean Absolute Error, Root Mean Squared Error.



Dataset :

We will going to perform all the above processes using the data set given below

Link:

<https://www.kaggle.com/datasets/chakradharmattapalli/product-demand-prediction-with-machine-learning>.The dataset that we have for this task contains data about:

1. ID
2. Store ID
3. Total price
4. Base price
5. Unitsold



In this dataset there are 5 columns and 1750150 rows.

Conclusion :

Data science projects can be immensely valuable for beginners, providing us with practical experience and enhancing our skills in data analysis, machine learning, and data visualization.

With sincere gratitude we are very happy to take part in this IBM NaanMudhalvan data science project along with our Mentor's support.

Thank You!