**PRODUCT DEMAND PREDICTION WITH MACHINE LEARNING**

**INTRODUCTION:**

In demand forecasting, machine learning algorithms can analyze historical sales patterns and predict future trends. The first step is collecting data about past sales, such as:

* Product type
* Quantity purchased
* Purchase frequency
* Seasonality sold
* Discounts, and more.

Demand forecasting is the process of using predictive analysis of historical data to estimate and predict customers' future demand for a product or service. It helps businesses make better-informed supply decisions and estimate total sales and revenue for future periods.

The limitations of traditional linear regression models in capturing complex relationships emphasize the need for advanced regression techniques like Gradient Boosting and XG Boost to enhance prediction accuracy.

**Content for Project Phase 2:**

Consider exploring advanced regression techniques like Gradient Boosting or XG Boost for improved prediction accuracy.

**Data Source:**

Data source for product demand price prediction takewn from kaggle source to perform a prediction.

**Dataset Link:** <https://www.kaggle.com/datasets/chakradharmattapalli/product-demand-prediction-with-machine-learning>

**Details about column:**

1. **ID**: A unique identifier for each data record, serving as a reference point for the dataset.
2. **Store ID**: An identifier specifying the store associated with the transaction, facilitating store-specific analysis.
3. **Total Price**: The cumulative cost of items sold in a transaction, typically in the transaction's currency.
4. **Base Price**: The initial price of a product or service before any discounts or extra charges are applied, providing insights into pricing dynamics.
5. **Units Sold**: The quantity of products sold in a transaction, crucial for understanding sales volume.

**Data Collection:**

Collect historic sales data and external factors that influence demand, such as marketing campaigns, holidays, economic indicators, etc.

Data collection is the process of gathering and measuring information from different sources. Data must be collected and stored in a way that makes sense for the business problem and machine learning.

**Data Preprocessing:**

* Clean and preprocess the data, handle missing values, and convert categorical features into numerical representations.
* Data Processing is the task of converting data into a more usable form, making it more meaningful and informative for machine learning.

**Feature Engineering:**

Feature engineering is the pre-processing step of machine learning, transforming raw data into features used for creating predictive models. It aims to improve model performance by creating features that capture seasonal patterns, trends, and external influences on product demand.

**Advanced Regression Techniques:**

* Ridge Regression: Introduce L2 regularization to mitigate multicollinearity and overfitting.
* Lasso Regression: Employ L1 regularization for feature selection and model simplification.
* Elastic Net Regression: Combine L1 and L2 regularization.
* Random Forest Regression: Implement an ensemble technique to handle nonlinearity.
* Gradient Boosting Regressors (e.g., XG Boost, Light GBM): Utilize gradient boosting algorithms.

**Model Selection:**

Choose suitable regression algorithms (e.g., Linear Regression, Random Forest, XGBoost) for demand forecasting. Model selection involves assessing and contrasting various models to identify the one that best fits the data and produces the best results.

**LIBRARIES USED:**

* Pandas: Python library for working with data sets.
* NumPy: Python library for working with arrays.
* Matplotlib: Comprehensive library for creating visualizations in Python.

**To download and use pandas , numpy and matplot lib use command:**

pip install pandas

pip install numpy

pip install matplotlib

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

**Model Training**

**Model training is the phase in the data science development lifecycle where practitioners try to fit the best combination of weights and bias to a machine learning algorithm to minimize a loss function over the prediction range.**

PROGRAM:

**from s k learn import datasets  
from s k learn . model \_ selection import train \_ test \_ split  
from s k learn. metrics import accuracy \_ score  
from s k learn .tree import Decision Tree Classifier**

**data = datasets .load \_ wine(as \_ frame = True)  
X = data . data  
y = data . target**

**X \_ train, X \_ test, y \_ train, y \_ test = train \_ test \_ split(X, y, test \_ size = 0.25, random \_ state = 22)**

**D tree = Decision Tree Classifier (random \_ state= 22)  
d tree . fit (X \_ train , y \_ train)**

OUTPUT:

**Decision Tree Classifier (random \_ state=22)**

**Conclusion:**

Machine learning enables accurate product demand predictions by analyzing historical data. This enhances inventory management and customer satisfaction. Data quality, model choice, and continuous refinement are critical to success.