PRODUCT DEMAND PREDICTION WITH MACHINE LEARNING

Problem Statement

The primary problem is to develop a machine learning model that can accurately predict product demand based on historical sales data and external factors. The goal is to assist businesses in optimizing inventory management and production planning to efficiently meet customer needs.

Design Thinking:

Data Collection: Gathering historical sales data and external factors that influence product demand, such as marketing campaigns, holidays, economic indicators, etc.

Data Preprocessing: Cleaning and preprocessing the collected data, handling missing values, and converting categorical features into numerical representations.

Feature Engineering: Creating additional features that capture seasonal patterns, trends, and external influences on product demand.

Model Selection: Choosing suitable regression algorithms (e.g., Linear Regression, Random Forest, XGBoost) for demand forecasting.

Model Training: Training the selected model using the preprocessed data.

Evaluation: Evaluating the model's performance using appropriate regression metrics (e.g., Mean Absolute Error, Root Mean Squared Error).

Dataset

The dataset required for this project is available on Kaggle at the following link: https://www.kaggle.com/datasets/chakradharmattapalli/product-demand-prediction-with-machine-learning

Design Thinking

1. Data Collection

To address the problem, we will start by collecting relevant data. This involves gathering historical sales data and external factors that influence product demand. The data will be obtained from the provided dataset on Kaggle.

2. Data Preprocessing

Once the data is collected, we will perform data preprocessing tasks. This includes cleaning the data, handling missing values, and converting categorical features into numerical representations. Proper data preprocessing is crucial for building accurate machine learning models.

3. Feature Engineering

Feature engineering is an essential step to create meaningful features that can capture seasonal patterns, trends, and external influences on product demand. We will explore techniques to extract valuable information from the data.

4. Model Selection

Choosing the right machine learning algorithm is critical for demand forecasting. We will consider various regression algorithms, such as Linear Regression, Random Forest, and XGBoost, to determine which one best suits our dataset and problem.

5. Model Training

Once the algorithm is selected, we will train the machine learning model using the preprocessed data. This step involves splitting the data into training and testing sets and training the model on the training data.

6. Evaluation

To assess the model's performance, we will employ appropriate regression metrics such as Mean Absolute Error (MAE) and Root Mean Squared Error (RMSE). The model will be evaluated using the testing dataset to ensure its accuracy in predicting product demand.

Conclusion

This document provides a clear understanding of the problem statement and outlines the design thinking process for solving the problem of product demand forecasting. The subsequent phases of data collection, preprocessing, feature engineering, model selection, training, and evaluation will be executed systematically to achieve the project's objectives. The provided dataset on Kaggle will serve as the foundation for our analysis and model development.