Product Demand Prediction with Machine Learnings

Problem Statement:

The problem is to create a machine learning model that forecasts product demand based on historical sales data and external factors. The goal of this project is to predict product demand using machine learning techniques. Accurate demand forecasting is crucial for optimizing inventory management, production planning, and overall business efficiency.

Steps to be followed:

I.Data Collection

1. Collect Historical Sales Data:
   - Gather historical data on product sales, including timestamps, quantities sold, and product identifiers.
   - Include data from multiple time periods to capture seasonal variations.

2. External Factors:
   - Collect external factors that influence demand:
     - Marketing campaigns: Record the timing and impact of marketing efforts.
     - Holidays: Include data on holidays and their effects on demand.
     - Economic Indicators: Gather relevant economic data such as GDP, unemployment rates, and consumer sentiment.

II. Data Preprocessing

1. Data Cleaning:
   - Remove duplicate records and outliers that can distort predictions.
   - Handle missing values through imputation or data removal.

2. Categorical to Numerical Conversion:
   - Convert categorical features (e.g., product categories) into numerical representations using techniques like one-hot encoding or label encoding.

III. Feature Engineering

1. Time-based Features:
   - Create features that capture temporal patterns, such as day of the week, month, or season.

2. Lag Features:
   - Generate lag features to incorporate historical data, e.g., sales from previous months.

3. External Influence Features:
   - Incorporate external factors like marketing campaigns and holidays as features.

IV. Model Selection

1. Regression Algorithms:
   - Choose suitable regression algorithms for demand forecasting, such as:
     - Linear Regression
     - Random Forest Regression
     - XGBoost Regression
     - Support Vector Regression (SVR)

2. Hyperparameter Tuning:
   - Fine-tune model hyperparameters for optimal performance using techniques like cross-validation.

V. Model Training

1. Data Split:
   - Split the preprocessed data into training, validation, and test sets. Typically, an 80-10-10 split is used.

2. Model Training:
   - Train the selected regression model on the training data.

VI. Evaluation

1. Regression Metrics:
   - Evaluate the model's performance using appropriate regression metrics, including:
     - Mean Absolute Error (MAE)
     - Root Mean Squared Error (RMSE)
     - Mean Absolute Percentage Error (MAPE)
     - R-squared (R^2)

2. Visualization:
   - Visualize predicted vs. actual demand to gain insights into model performance.

Conclusion

In this document, we have outlined the steps involved in a product demand prediction project using machine learning. Accurate demand forecasting can lead to improved business decision-making and resource allocation. Ensure to iterate and refine your model as new data becomes available and business conditions change.