Project Documentation:

Dominos - Predictive Purchase Order System

Sai Vaishnav J

Overview

The aim of this project is to enhance Domino's operational efficiency by predicting future pizza sales, thereby optimizing the process of ordering ingredients. By leveraging historical sales data and ingredient information, we develop a predictive model that facilitates the generation of a purchase order system, ensuring that the right amount of ingredients is stocked, thus minimizing waste and preventing stockouts.

Skills Acquired from This Project

- **Data Cleaning and Preprocessing**: Techniques to handle missing values, outliers, and data inconsistencies.
- **Exploratory Data Analysis (EDA)**: Methods to visualize and analyze sales trends and patterns in historical data.
- **Time Series Forecasting**: Applying forecasting techniques to predict future sales.
- **Predictive Modeling**: Using statistical models to make informed business decisions.
- **Business Decision Making**: Utilizing data-driven insights for strategic planning.
- **Real-world Application of Data Science**: Implementing data science techniques in a practical, industry-relevant context.

Domain

Food Service Industry

Problem Statement

Dominos aims to optimize its ingredient ordering process by accurately predicting future sales. This will allow them to maintain optimal stock levels, reduce waste, and prevent stockouts. The project utilizes historical sales data and ingredient information to create a predictive model that generates an efficient purchase order system.

Business Use Cases

1. **Inventory Management**: Maintaining optimal stock levels to meet future demand without overstocking.
2. **Cost Reduction**: Minimizing waste and reducing costs associated with expired or excess inventory.
3. **Sales Forecasting**: Accurately predicting sales trends to inform business strategies and promotions.
4. **Supply Chain Optimization**: Streamlining the ordering process to align with predicted sales and avoid disruptions.

Approach

1. Data Preprocessing and Exploration

- **Data Cleaning**:
    - Removed missing or inconsistent data entries.
    - Handled outliers using the Interquartile Range method.
    - Standardized entries in categorical columns.
- **Exploratory Data Analysis (EDA)**:
    - Analyzed historical sales data to identify trends, seasonality, and patterns.
    - Visualized data distributions to highlight significant features.

2. Sales Prediction

- **Feature Engineering**:
    - Created features such as day of the week, month, and holiday effects to enhance predictive power.
- **Model Selection**: Chose the Prophet model for time series forecasting based on its robustness and ease of use.
- **Model Training**: Trained the model using an 80/20 split of the data for training and testing.

- **Model Evaluation**:
    - Used Mean Absolute Percentage Error (MAPE) to assess model performance.

3. Purchase Order Generation

- **Sales Forecasting**: Predicted pizza sales for the next week using the trained model.
- **Ingredient Calculation**:
    - Calculated the required quantities of ingredients based on predicted sales and ingredient dataset.
- **Purchase Order Creation**:
    - Generated a detailed purchase order listing quantities of each ingredient needed for the forecasted sales period.

Results

- **Accurate Sales Predictions**: Successfully forecasted pizza sales with a MAPE score indicating good predictive performance.
- **Comprehensive Purchase Order**: Generated detailed ingredient requirements for the upcoming sales period.

Data Set

- **Sales Dataset**: Contains historical sales records (Date, Pizza Type, Quantity Sold, Price, Category, Ingredients).
- **Ingredients Dataset**: Contains ingredient requirements for each pizza type (Pizza Type, Ingredient, Quantity Needed).

Data Set Explanation

- **Sales Data**:
    - Comprises a total of 48,620 entries and 12 columns.
    - Key columns include pizza_id, order_id, pizza_name_id, quantity, order_date, order_time, unit_price, total_price, pizza_size, pizza_category, pizza_ingredients, and pizza_name.
    - Initial analysis revealed missing values in columns such as pizza_name_id and total_price, with a significant range of unit prices and total prices.
- **Ingredients Data**:
    - Details the ingredient requirements for various pizza types, providing crucial information for ingredient calculations.

Detailed Program Descriptions

Program 1: Data Loading and Initial Analysis

- Loaded the sales dataset and performed an initial analysis.
- Identified missing values and unique value counts for key columns, highlighting the dataset's diversity and complexity.

Program 2: Data Cleaning and Standardization

- Cleaned the dataset by filling missing values and standardizing entries.
- Removed outliers and visualized the distribution of sales quantities, preparing the dataset for deeper analysis.

Program 3: Sales Categorization and Visualization

- Grouped daily sales data, computed median sales thresholds, and categorized sales into 'Holiday', 'Promotional Period', and 'Normal Sales'.
- Visualizations illustrated sales distributions and categories, aiding in understanding sales dynamics.

Program 4: Weekly Sales Analysis

- Analyzed sales performance for specific weeks, identifying top-selling pizzas and visualizing trends over time.
- Conducted day-of-week breakdowns and seasonal analysis, contributing to sales insights.

Program 5: Sales Forecasting with Prophet

- Trained the Prophet model on historical sales data to predict future sales.
- Evaluated model performance using MAPE, providing a score indicative of forecast accuracy.

Program 6: Extended Sales Predictions

- Expanded on earlier predictions by generating forecasts for the next four weeks and visualizing actual vs. predicted sales data.

Program 7: Final Sales Forecasting and Ingredient Requirements

- Generated predictions for the upcoming week, calculated ingredient requirements based on forecasted sales, and created a detailed purchase order.
- Ensured alignment of inventory levels with expected sales, enhancing operational efficiency.

Conclusion

The **Dominos - Predictive Purchase Order System** project successfully demonstrated the application of data science techniques in the food service industry. By accurately predicting pizza sales and generating a corresponding purchase order, the project addresses critical operational challenges such as inventory management and cost reduction. The implementation of the Prophet model for time series forecasting provided reliable sales predictions, allowing Domino's to make informed decisions regarding ingredient stock levels. This project not only streamlines the supply chain but also contributes to minimizing waste and optimizing operational efficiency, ultimately enhancing customer satisfaction and business performance. Through this endeavor, we showcased the importance of data-driven approaches in driving strategic decision-making within the food service domain.