Future Cart: AI-Driven Demand Prediction for Smarter Retail

Presented By:

1. Abhilash
2. Abhigna Kumar
3. Aravindhan
4. Saket Chaudhary

Project Statement

In the realm of E-Commerce, demand forecasting plays a pivotal role in ensuring business success. This project aims to develop a demand forecasting model in an E-commerce business that predicts future product demand leveraging time series analysis and multivariate regression based on historical sales data, along with Google Analytics KPIs such as Google clicks and Facebook impressions, which are valuable indicators of customer interest.

Time Series Analysis

- **Definition:** Time series analysis is a method of analyzing data points collected at regular intervals over time, aiming to understand underlying patterns, relationships, and behaviors in the data.
- **Purpose:** It focuses on identifying key components such as trends (long-term movement), seasonality (recurring patterns within specific intervals), and noise (random fluctuations), which help in explaining the data's dynamics and variability.
- **Goal:** The primary objective is to use historical data to forecast future values, detect anomalies, and provide actionable insights for decision-making across various domains like finance, marketing, and environmental sciences.

Expected Outcomes

- **Improved Inventory Management:** More accurate demand forecasts lead to better inventory decisions, potentially reducing stock-outs and excess inventory.
- **Enhanced Marketing Efficiency:** Identify periods of high demand for targeted marketing campaigns, optimizing resource allocation.
- **Data-Driven Decision Making:** Reliable forecasts provide a basis for business decisions, such as pricing adjustments or product promotions.
- **Accurate Demand Predictions:** Implement a forecasting model that achieves high accuracy in predicting future demands thereby improving customer service levels.
- **Scalable Solution:** Develop a solution that can scale to handle large datasets and varying demand patterns across multiple products.

Modules

Module 1: **Data Collection**
Gathering relevant and accurate data from various sources to use in the analysis and forecasting process.

Module 2: **EDA and Data Preprocessing**
Exploring and cleaning the data to understand its structure, handle missing values, and prepare it for modeling.

Module 3: **Time Series Modeling**
Analyzing historical data to identify trends and patterns, and using models like ARIMA or SARIMAX to make forecasts.

Module 4: **Dynamic Multivariate Regression**
Using regression models to forecast by considering multiple variables and their time-dependent relationships.

Metrics Used

- **MAE (Mean Absolute Error):** Measures the average magnitude of errors in predictions, without considering their direction. Lower MAE indicates better accuracy.
- **RMSE (Root Mean Squared Error):** Calculates the square root of the average squared differences between predicted and actual values, penalizing larger errors more heavily.
- **MSE (Mean Squared Error):** Averages the squared differences between predicted and actual values, providing a measure sensitive to larger deviations.
- **PD (Percentage Deviation):** Expresses the average deviation of predictions from actual values as a percentage, offering a normalized view of forecast accuracy.

Evaluation Metrics After Tuning

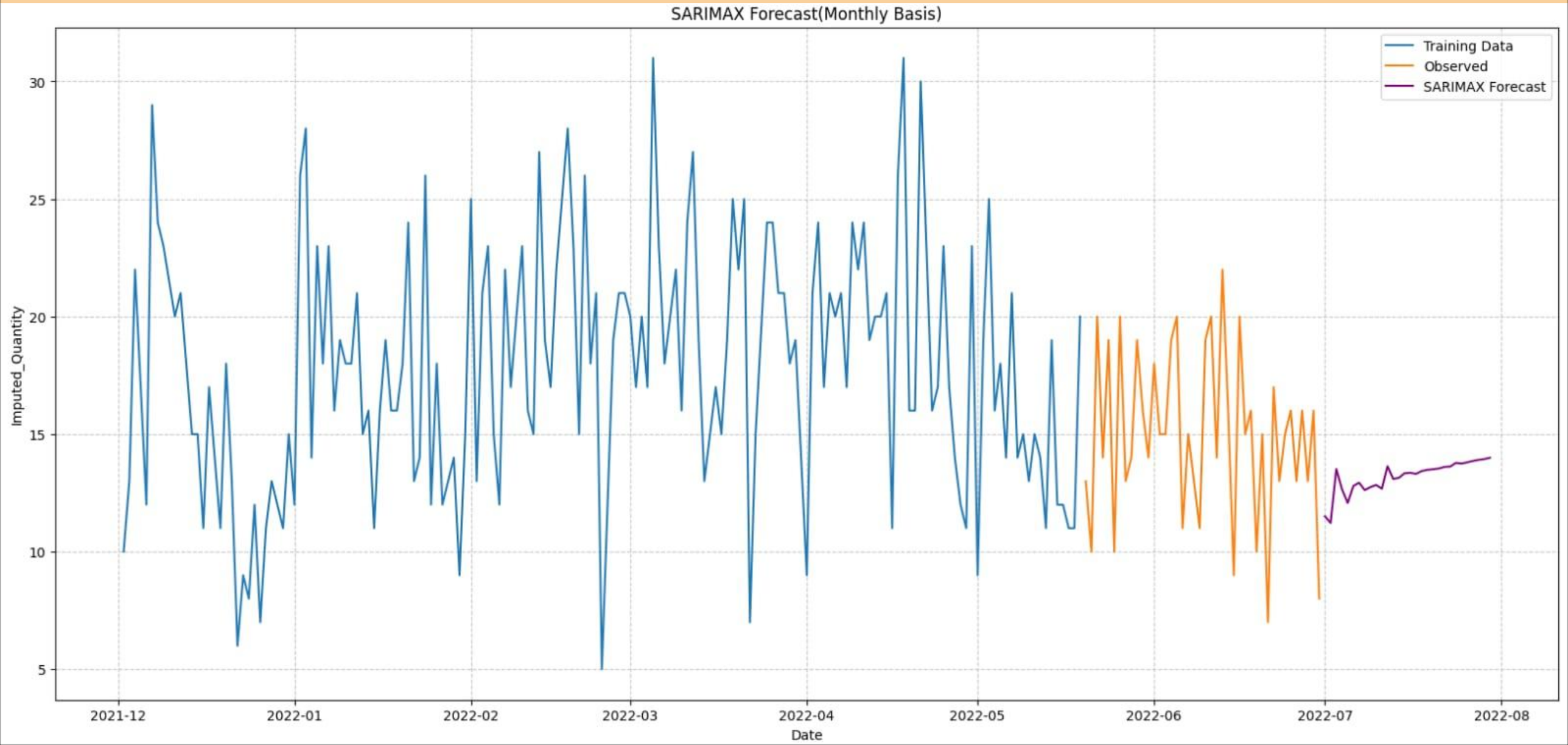
Metrics After Tuning:

	Model	MAE	RMSE	MSE	PD
0	AR	4.029659	4.762827	22.684517	32.988444
1	MA	4.032763	4.769063	22.743960	33.050780
2	ARIMA	3.089326	3.854395	14.856360	22.144257
3	SARIMA	3.141460	3.947274	15.580972	21.728248
4	ARIMAX	3.163564	3.946457	15.574522	23.689286
5	SARIMAX	3.146013	3.945624	15.567950	23.237196

Why we chose SARIMAX?

- Highlight SARIMAX's ability to incorporate both seasonality and exogenous variables.
- SARIMAX had one of the lowest RMSE and MSE compared to other models like ARIMAX or Dynamic Multivariate Regression.
- Flexibility and robustness in handling seasonal patterns and external predictors like clicks and impressions.

Forecasting using SARIMAX



Key Insights:

- **Trend Consistency:** The forecasted values align closely with historical trends, suggesting that future 'Quantity' values will continue to follow similar patterns, with some seasonal fluctuations.
- **Impact of Exogenous Variables:** The forecast reflects the influence of external factors such as 'Clicks' and 'Impressions,' indicating that these variables play a crucial role in determining the 'Quantity' outcome.
- **Forecast Stability:** Despite some variability, the forecasts remain stable, reinforcing the model's ability to predict future values accurately, even with changing external conditions.
- **Seasonal Patterns:** The model captures recurring seasonal trends, providing insights into how 'Quantity' behaves during specific times of the year, which can help in planning for demand spikes or dips.

Conclusion

Time series analysis helped us understand the patterns and trends in the data, providing valuable insights for future forecasting. By using models like SARIMAX, we were able to account for seasonality, trends, and external factors, improving the accuracy of our predictions. Forecasting plays a critical role in decision-making, as it helps anticipate future values and prepare for changes in demand or behavior. This analysis can be applied to optimize strategies, resource allocation, and planning, ensuring more informed and effective business decisions.

Thank You