Approach for Forecasting Expenditures Based on Past Data

Introduction:

The task involves building a machine learning model to forecast expenditures based on past data using the ARIMA (AutoRegressive Integrated Moving Average) time series model. This report outlines my approach to solving this problem and highlights the challenges I faced during the project.

Approach:

1. Understanding the Problem:

- I started by thoroughly understanding the project requirements, which included forecasting expenditures based on historical data.
- I Can't get Dataset of expenditures for any company because it's provided inside the company only so I generated synthetic dataset to gain sample insights and forecasting future spent of the company.

2. Data Preprocessing:

- I loaded the expenditure data from the provided CSV file into a Pandas DataFrame in Python.
- I checked for missing values, outliers, and inconsistencies in the data.
- I ensured that the data was suitable to build ARIMA Machine Learning Model.

3. Data Exploration and Visualization:

- I visualized the data to identify trends, seasonality, and any underlying patterns.
- I used line plots, box plots, and decompositions to understand the data's characteristics.

4. **Time Series Decomposition:**

- I decomposed the time series data into its components: trend, seasonality, and residual error.
- This step helped in understanding the underlying patterns in the data.

5. **Model Selection (ARIMA):**

- I used the ACF (Autocorrelation Function) and PACF (Partial Autocorrelation Function) plots to determine the orders (p, d, q) for the ARIMA model.
- I fitted an ARIMA model to the stationary data using the chosen orders.

6. **Model Training and Evaluation:**

- I trained the ARIMA model on the training data and evaluated its performance using appropriate metrics such as Mean Absolute Error (MAE) and Mean Squared Error (MSE).
- I used the training data to validate the model's accuracy and adjust hyperparameters if necessary.

7. **Forecasting:**

- Finally, I used the trained ARIMA model to forecast future expenditures based on past data for the next 4 months.

Challenges:

1. **Understanding Time Series Concepts:**

- Initially, grasping ARIMA Model and Generating Synthetic dataset, and understanding concept of forecasting are hard.

2. **Model Evaluation:**

- For Selecting and interpreting evaluation metrics for time series models are hard.

3. **Visualization:**

- To Create meaningful visualizations to understand and communicate time series patterns effectively is essential.

- Thorough error handling was required to ensure a pleasant user experience and prevent inaccurate input.

4. **Future Forecast Interpretation:**

- Interpreting and communicating the implications of the forecasts for future business decisions require domain knowledge so I need to learn about the domain more.

- To forecast more accurate data of future spent, large amount of dataset is required.

Conclusion:

Despite the challenges, I build a time series forecasting model using ARIMA to predict future spent of a company and I would like to emphasize that while the current results may not be as accurate as desired, I am genuinely committed to improving my skills and learning more to contribute effectively to the company. This project has been a valuable learning experience, Thank you for the opportunity to work on this project, and I look forward to further growth and contributions within the company.

GitHub Link: https://github.com/marayyy/Krayo.git