

SEAS 6414 - Python Applications in Data Analytics

Homework 8

Due Date: March 9, 2024 (10:00am EST)

Instructions: To complete the following task using Python, please download an Integrated Development Environment (IDE) of your choice. Ensure that your solution includes both the written code (input) and its corresponding output. Once completed, upload your solution in PDF format or any other format you prefer. **The question is worth 100 points.**

Introduction

This homework focuses on analyzing and forecasting merchant transaction data using various time series techniques. The primary objective is to apply different forecasting models and evaluate their performance in predicting future transaction volumes.

1 Data Preparation

- Analyze merchant transaction data from January 1, 2033, to January 1, 2035.
- Resample the data to obtain daily transaction sums (**in Dollars**).

2 Forecasting Methods

2.1 Holt-Winters (Triple Exponential Smoothing)

Apply Holt-Winters method to capture seasonality, trend, and level in the daily transaction data.

2.2 Time Series Cross-Validation

Use time series cross-validation for evaluating the performance of the models.

2.3 Feature Engineering with Lags

Create new features based on lagged values of the time series for linear regression and random forest models.

2.4 Linear Regression Model

Implement a linear regression model using the lagged features.

2.5 Random Forest Model

Similarly, use the lagged features in a random forest model for forecasting.

2.6 Facebook Prophet

Implement the Facebook Prophet model for forecasting.

3 Model Validation

3.1 Mean Absolute Percentage Error (MAPE)

Calculate and compare the MAPE for each model to assess their accuracy.

3.2 Confidence Intervals and Anomaly Detection

Plot the confidence intervals for each model's forecasts and identify any anomalies.

4 Report Writing

- Summarize the findings from each model, including their performance metrics and insights.
- Discuss the implications of the results and suggest the most suitable model for forecasting merchant transactions.