Introduction

TBATS is a time series forecasting method that stands for Trigonometric seasonality, Box-Cox transformation, ARIMA errors, Trend, Seasonal components. A time series will have 3 components: the trend, seasonality, and the remainder component. The TBATS model is used for time series data with complex seasonalities, meaning that the data has multiple seasonal patterns. For example, daily data may have a weekly, monthly, and annual pattern. The TBATS model can also deal with non-integer seasonalities.

TBATS was developed by De Livera, Hyndman, & Snyder (2011) and considers multiple seasonalities in a time series. Each seasonality is modeled within a Fourier series, which is a model that represents a period function as a sum of sine and cosine waves. Each term in the Fourier series represents a seasonality of the time series. TBATS is an exponential smoothing model, which is a forecasting method where forecasts are weighted averages of past observations, with the weights decaying exponentially as the observations get older. Thus, more recent observations have more of an impact on forecasts. In the case of the TBATS model, the sin and cosine functions in the Fourier series are weighted based on their recency, accordingly with the exponential smoothing model.

When implemented, TBATS will build out multiple models, such as models with Box-Cox transformation, trends, trend damping, and ARMA errors. The Box-Cox transformation is a transformation that normalizes the observations in a time series. Trend is the overall direction of the time series, and trend dampening is a method to increase or decrease the growth or decay rate of the time series trend. Residuals of a time series can be modeled by ARMA models, and so the TBATS model can use this ARMA process to correct for errors. After these models are built, TBATS will use the Akaike information criterion to estimate the prediction errors and relative qualities of the models to select the model with the best parameters.

Methodology

For the Kaggle Sales Forecasting Data, there are daily observations for every day in 2017. Since this is daily data, it made sense to use the TBATS model because there were multiple seasonalities (weekly, monthly, quarterly, etc.). The complete dataset has 20 independent variables that are used to predict sales. The dataset includes 54 individual stores; the dataset was subsetted for each individual store, creating 54 individual subsets. For each subset, the TBATS model was trained using the training dataset provided by Kaggle. The model was then applied for each subset to predict the sales for a store for the 15 days directly following the end-date of the training dataset. The predicted sales for the next 15 days for each store were then averaged for each day, creating one predicted dataset that is the average of the predicted sales for each store for the next 15 days.

References

Hyndman, Rob J., and George Athanasopoulos. Forecasting: principles and practice. OTexts, 2018.

De Livera, Alysha M., Rob J. Hyndman, and Ralph D. Snyder. "Forecasting time series with complex seasonal patterns using exponential smoothing." *Journal of the American statistical association* 106.496 (2011): 1513-1527.

Naim, Iram, Tripti Mahara, and Ashraf Rahman Idrisi. "Effective short-term forecasting for daily time series with complex seasonal patterns." Procedia computer science 132 (2018): 1832-1841.

Skorupa, Grzegorz. “Forecasting Time Series with Multiple Seasonalities Using TBATS in Python.” *Medium*, Intive Developers, 14 Jan. 2019, https://medium.com/intive-developers/forecasting-time-series-with-multiple-seasonalities-using-tbats-in-python-398a00ac0e8a.

Nadeem. “Time Series Forecasting Using TBATS Model.” *Medium*, Analytics Vidhya, 30 Nov. 2021, https://medium.com/analytics-vidhya/time-series-forecasting-using-tbats-model-ce8c429442a9#:~:text=TBATS%20is%20a%20forecasting%20method,errors%2C%20Trend%20and%20Seasonal%20components.