Department of Computer Science and Engineering (Data Science)

Subject: Time Series Analysis

Experiment 4

Name: Divyesh Khunt **Sapid:60009210116** **Batch:D12**

(Smoothing Methods)

Aim: Implement various Smoothing Methods (Simple exponential, double exponential and Triple exponential) on a given dataset.

Theory:

Smoothing:

Smoothing is a statistical method we can use to create an approximation function to remove irregularities in data and attempt to capture significant patterns.

The smoothing technique is a family of time-series forecasting algorithms, which utilizes the weighted averages of a previous observation to predict or forecast a new value. The main idea of this technique is to overweight recent values in a time series. It is utilized for short-term forecasting models. This technique is more efficient when time-series data is moving slowly over time. It harmonizes errors, trends, and seasonal components into computing smoothening parameters. These components are pooled either additively or multiplicatively.

Simple Exponential Smoothing:

Simple exponential smoothing (SES) is one of the minimal models of the exponential smoothing algorithm. Such techniques are employed for a univariate observation (data), which has no clear trend or seasonal patterns. It involves a parameter called alpha, which is the smoothing parameter. The core idea is to employ a weighted moving average, including exponentially decreasing weights, which designates higher weight to the most recent observation. The concept behind SES is forecasting future values by using a weighted average of all the previous values in the series. This method can be used to predict series that do not have trends or seasonality.

Assume that a series has the following:

- Level (L_t)
- No trends
- No seasonality
- Noise (somewhat constant)

SES is the forecast estimate level at a most recent point in time.

$$F_{t+k} = L_t$$

Let's estimate and update the level (L_t). Here is the level updating equation:

$$L_t = \alpha Y_t + (1-\alpha) L_{t-1}$$

We are taking the level at time L_t and updating the previous level, L_{t-1}, by integrating information from most recent data point, Y_t. We can see that it is a weighted average, where α and $1-\alpha$ are the weights.

The smoothing constant equals α. The $0 \leq \alpha \leq 1$ alpha value lies between 0 and 1.

The algorithm learns a new level from the new data that it sees. One of the possible ways of initializing the whole system is to assign L1 equal to the first record in the series.

$$F_1 = L_1 = Y_1$$

Why is it called exponential smoothing?
Here is the level update equation:

$$\mathbf{L_t = \alpha Y_t + (1\text{-}\alpha)\, L_{t-1}}$$

Substitute L_t with its formula, as shown here:

$$L_t = \alpha Y_t + (1\text{-}\alpha)\,[(\alpha Y_{t-1} + (1\text{-}\alpha)L_{t-2}]$$
$$= \alpha Y_t + \alpha(1\text{-}\alpha)\, Y_{t-1} + (1\text{-}\alpha)^2\, L_{t-2}$$
$$= \ldots = \alpha Y_t + \alpha(1\text{-}\alpha)\, Y_{t-1} + \alpha\,(1\text{-}\alpha)^2\, Y_{t-2} + \ldots$$

We can see that we are ending up with an average of all the values in a series, but they have weights that are decaying exponentially into the past. Hence, this is called exponential smoothing when $\alpha = 1$ and past values do not influence the forecast. It's called under-smoothing when $\alpha = 0$ and past values have equal weight in the average. In this case, we do not give more weight to recent information, which is called over-smoothing.
Here is the effect of alpha:

α	α (1- α)	α (1- $\alpha)^2$	α (1- $\alpha)^3$
0.9	0.089	0.0089	0.00089
0.5	0.25	0.125	0.0625
0.1	0.09	0.081	0.0729

In all cases, decay as we go into the past with alpha is fast, whereas with a smaller alpha decay it is slower.

Double Exponential Smoothing:

This is a more reliable method for handling data that consumes trends without seasonality than compared to other methods. This method adds a time trend equation in the formulation. Two different weights, or smoothing parameters, are used to update these two components at a time. Holt's exponential smoothing is also sometimes called double exponential smoothing. The main idea here is to use SES and advance it to capture the trend component. Assume that a series has the following:

- Level
- Trends
- No seasonality
- Noise

Forecast = estimate level + trend at the most recent time point.

$$\mathbf{F_{t+k} = l_t + kT_t}$$

l is the most recent level. Tt is the most recent time point. K is how many steps into the future we are trying to forecast.

Updated equation for level: $\mathbf{L_t = \alpha Y_t + (1\text{-}\alpha)\,(L_{t-1} + T_{t-1})}$

Lt-1 adjusts the previous level. Tt-1 adds the trend. (Lt-1 + Tt-1) adjusts the previous level by adding a trend.
Here is the updated equation for trends: **$T_t = \beta (L_t - L_{t-1}) + (1-\beta) T_{t-1}$**
Take the trend in the previous time period, Tt-1, and update it by showing the difference in most recent level estimates: Lt-Lt-1. This setup allows the trend to vary over time adaptively.
β is controlling the speed of adjusting the trend. If the trend changes quickly in the series, we want it to learn faster, so we have this: **$\alpha > 0 < 1$ $\beta > 0 < 1$**

Triple Exponential Smoothing:
This method can be applied when the data consumes trends and seasonality over time. It includes all smoothing component equations such as trends and seasonality. Seasonality comprises two different types, such as additive and multiplicative, which is a similar operation in mathematics. The winter method uses the idea of the Holt method and adds seasonality. Winter was a student of Holt and extended his approach by adding an additional equation to update seasonality. This triple exponential smoothing is also known as the Holt-Winters method. Let's assume that a series has the following:

- Level
- Trend
- Seasonality
- Noise

The forecasting equation can be adjusted to handle additive or multiplicative trends and seasonality. This is the additive seasonality model:
Forecast = estimate level + trend + seasonality at most recent time point.

$$F_{t+k} = l_t + kT_t + S_{t+k-M}$$

This is the multiplicative seasonality model with the additive trend: Forecast = estimate (level × trend) × seasonality at most recent time point.

$$F_{t+k} = (l_t + kT_t)\, S_{t+k-M}$$

Assume our series contains the level, trend, and seasonality with M seasons with noise. In Holt-Winter exponential smoothing, we have three smoothing constants.
Level: **$L_t = \alpha (y_t/S_{t-M}) + (1-\alpha)(L_{t-1} + T_{t-1})$**
S is a seasonal component. When we yt/St-M, we are the deseasonalizing value of y. In the level equation, we are updating the previous level by Lt-1, adding Tt-1, and then combining and then combining the deseasonalizing value of yt.
Trend: **$T_t = \beta (L_t - L_{t-1}) + (1-\beta) T_{t-1}$ (Additive trend)**
We can update the previous trend by considering the latest difference between levels.
Seasonality: **$S_t = \gamma (Y_t/L_t) + (1+\gamma) S_{t-M}$ (multiplicative seasonality)**
Yt is divided by level component Lt. This gives the detrended value of Y. So, the seasonality is updated by combining the most recent seasonal component St-M with the detrended value of Yt.

Department of Computer Science and Engineering (Data Science)

When to Use Holt-Winters Single, Double and Triple Smoothing Models

Algorithm	Level?	Trend?	Seasonal?	Cyclic?
Holt-Winters Single Exponential Smoothing	Yes	No	No	No
Holt-Winters Double Exponential Smoothing	Yes	Yes	No	No
Holt-Winters Triple Exponential Smoothing	Yes	Yes	Yes/No*	Yes/No*
*At least one of seasonality/cyclical component should be present				

Lab Assignments to complete:
Perform the following tasks using the datasets mentioned. Download the datasets from the link given:

Link:
https://drive.google.com/drive/folders/107ecl29wJUZdnKl--PqvRfkUe8MaQHse?usp=sharing

Colab Links:
Simple Exponential Smoothing:
https://colab.research.google.com/drive/1q9Oay3iUqxTsEjsZt1mKJ6ZEkKOpfb62#scrollTo=gujoc7gaFc2P
Double Exponential Smoothing:
https://colab.research.google.com/drive/1Z_DPy9qQQcPIatU5zA-mzGrflM6qDTH-
Triple Exponential Smoothing:
https://colab.research.google.com/drive/1eu5zM8F3_J_HFTEbIi36cAmo4bZDy99B

Dataset 1: Facebook Stock Market Performance

1. Implement Simple, Double and Triple Exponential Smoothing and analyse the best method to smoothen the dataset.

CODE:
https://colab.research.google.com/drive/1RkHuntRGaR1psg5BBCCJy4fGoYSfoxv?usp=sharing