Time Series Analysis

Winter Holt's Seasonality Method

Shantanu Pandey
CS20BTECH11046

Koilada Yerni Jyotshna Sri
AI19BTECH11011

1 Introduction

1.1 Time Series Analysis

Time series analysis is a statistical data analysis technique used to analyze data points collected over a period of time which involves studying the patterns and trends in the data to make predictions and forecast future events. Time series analysis has many applications across various fields, including finance, economics, engineering, and science.

The primary objective of time series analysis is to understand the *underlying structure* of the time series data described by the **trend**, **seasonality** and **residual/random** components involved in it. This information is to use this information to make predictions or forecasts about future events. The general notation f a time series data is given by:

$$x_t = \mu_t + S_t + Y_t$$

- μ_t is the trend in the time series which refers to the long-term direction of the data over time which be linear/non-linear, positive/negative
- S_t is the seasonality in the time series which refers to the periodic variations in the data that occur at regular intervals.
- Y_t is the randomness/residual in the time series which refers to the unpredictable fluctuations in the data that are not explained by trend or seasonality

1.2 Applications of Time Series Analysis

There is huge amounts of data in any sector we choose. And of this data we have loads of data like Sales, Weather Forecasting, Trends in Market etc., which are changing with respect to time.

Say a company is beginning to increase it's reach into the market. Hence, they wish to make changes in the infrastructure and logistics of the company like no.of warehouses, change in cost of product, offers/discounts etc., for which they have to analyze the previous data they have (say in the past year) and predict/forecast the increase in sales. These kind of applications use time series analysis and take them seriously since these predictions involve investment of capital and future of the company as well.

There are many models involved into time series analysis like ARIMA (Auto Regressive Integrated Moving Average) models, Fourier analysis etc., In this project, we used two methods: **Holt's Linear Trend Method** and **Winter Holt's Seasonality Method** to analyze a data set.

2 Algorithm & Code

2.1 Data Set

The data set we chose to use here is **Tractor Sales** of a company named *PowerHorse* which is a tractor and farm equipment manufacturing company which was in around 1950s. The data set has **two columns** with **144 rows** of data. The columns indicate the *Number of Tractors Sold in a every month* for 12 year (2003-2014). This data set is chosen because it seems to visually have evident and clear trend and seasonal components it The data set can be found at link under the drop down of Time Series Data to the right.

Out of 144 rows, 124 are used for training and 20 are used to analyse the forecasting which will be made in the later sections.

As we can see in Figure 1, there is a clear trend and seasonal component in the data.

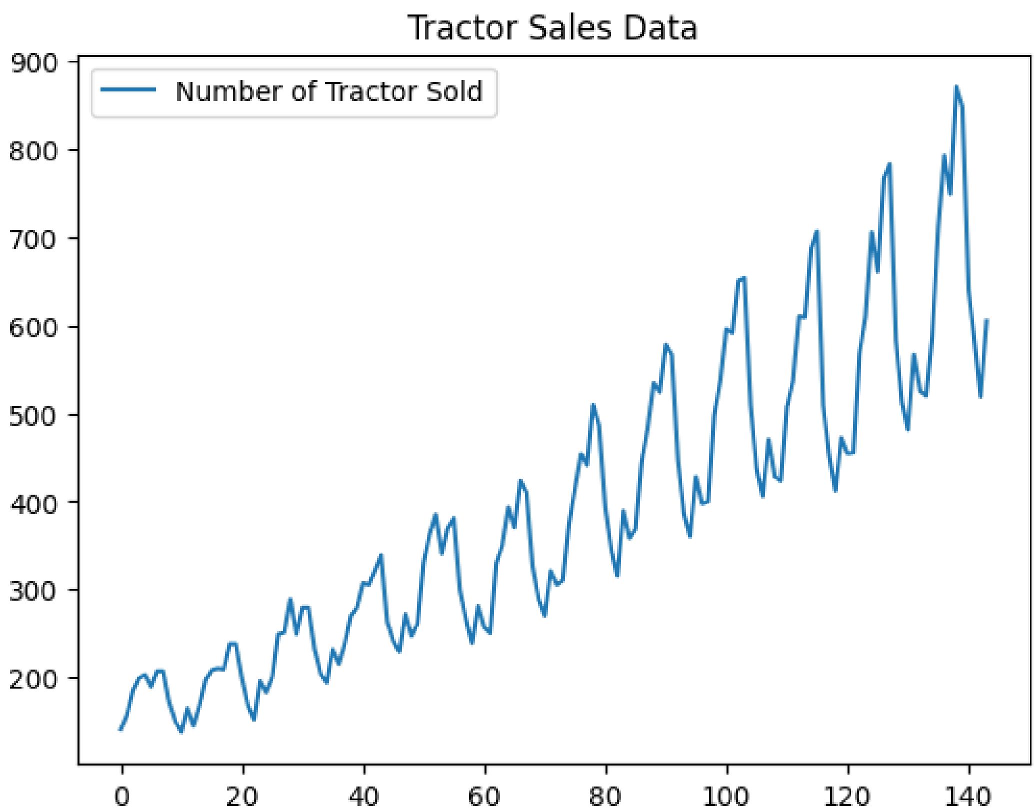

Figure 1: Plot of Tractor Sales Data set

2.2 Holt's Linear Trend Method

Holt's method is a popular technique for time series forecasting that can be used to model data with a trend. The method is an extension of simple exponential smoothing that incorporates a trend component into the model. Holt's method involves estimating two smoothing parameters α (for the level of the series) and γ (for the trend of the series).

Holt's linear trend method can be represented mathematically as:

Equation for Level :

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

Equation for Trend :

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}$$

Forecast at time $t+h$:

$$\hat{x_t}(h) = L_t + hT_t$$

where:

- x_t is the data point in the time series at time t.
- L_t is the level of the time series at time t.
- T_t is the trend of the time series at time t.
- α and γ are the smoothing parameters,
- h is the number of periods ahead for which the forecast is being made.

α and γ control the weights given to the actual value x_t, the level L_{t-1}, and the trend T_{t-1} in the smoothing process.

Holt's Linear Method

```
[ ] def holt_linear(data, alpha, gamma):
      level = [0]*len(data)
      level[0] = data[0]
      trend = [0]*len(data)

      for i in range(1, len(data)):
        level[i] = alpha*data[i] + (1-alpha)*(level[i-1]+trend[i-1])
        trend[i] = gamma*(level[i]-level[i-1]) + (1-gamma)*trend[i-1]

      error = np.sum((data - list(np.array(level)+np.array(trend)))**2)

      return (alpha, gamma, level, trend, error)
```

```
[ ] min_error = 1000000

    for a in tqdm(np.linspace(0, 1, 100)[1:-1]):
      for g in np.linspace(0, 1, 100)[1:-1]:
        (al, ga, l, t, e) = holt_linear(train_data, a, g)
        if e < min_error:
          (alpha, gamma, level, trend, min_error) = (al, ga, l, t, e)

    100%|██████████| 98/98 [00:09<00:00, 10.72it/s]
```

```
[ ] linear_forecasts = []

    for i in range(1, len(test_data)+1):
      linear_forecasts.append(np.exp(level[-1] + i*trend[-1]))
```

To choose α and γ, a simple grid search is used over a linear space of values from 1 to 100. Below are the predictions obtained by this method.

Smoothing Parameter	Value
α	0.989898
γ	0.010101

The values of α and γ show that the models does much more smoothing for the trend component.

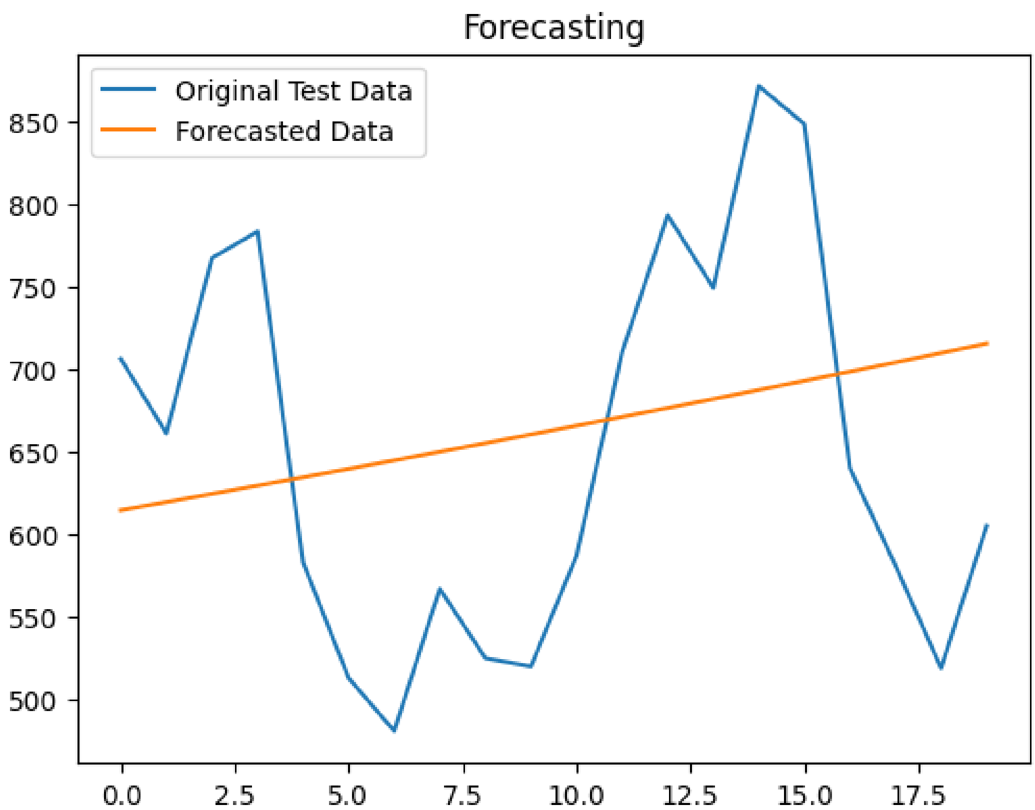

Figure 2: Forecasting of Holt's Linear Trend method for 20 months

Since Holt's Linear Trend method only analyzes the trend component, the forecasting (in Figure 2) are in a straight line (and increasing).

2.3 Holt Winter's Seasonal Method

Winter's Holt method is an extension of Holt's method that can be used to model time series data with a seasonal component. The method incorporates an additional smoothing parameter than Holt's Linear Trend method for the seasonal component of the data.

Winter Holt's Seasonal trend method can be represented mathematically as:

Equation for Level :

$$L_t = \alpha(x_t - S_{t-d}) + (1-\alpha)(L_{t-1} + T_{t-1})$$

Equation for Trend :

$$T_t = \gamma(L_t - L_{t-1}) + (1-\gamma)T_{t-1}$$

Equation for Seasonality :

$$S_t = \delta(x_t - L_t) + (1-\delta)S_{t-d}$$

Forecast at time $t+h$:

$$\hat{x_t}(h) = L_t + hT_t + S_{t+h-d}$$

where:

- x_t is the data point in the time series at time t.
- L_t is the level of the time series at time t.

- T_t is the trend of the time series at time t.
- S_t is the seasonal component of the time series at time t.
- α, γ and δ are the smoothing parameters,
- d is the seasonal fluctuation (generally considered as 12
- h is the number of periods ahead for which the forecast is being made.

Winter Holt's Method

```
[ ] def winter_holt(data, alpha, gamma, delta):
      level = [0]*len(data)
      level[0] = data[0]
      trend = [0]*len(data)
      seasonal = [0]*len(data)
      d = 12

      for i in range(1, len(data)):
        if i-d >= 0:
          level[i] = alpha*(data[i]-seasonal[i-d]) + (1-alpha)*(level[i-1]+trend[i-1])
          trend[i] = gamma*(level[i]-level[i-1]) + (1-gamma)*trend[i-1]
          seasonal[i] = delta*(data[i] - level[i]) + (1-delta)*seasonal[i-d]
        else:
          level[i] = alpha*data[i] + (1-alpha)*(level[i-1]+trend[i-1])
          trend[i] = gamma*(level[i]-level[i-1]) + (1-gamma)*trend[i-1]
          seasonal[i] = delta*(data[i] - level[i])

      error = np.sum((data - list((np.array(level)+np.array(trend))*np.array(seasonal)))**2)

      return (alpha, gamma, delta, level, trend, seasonal, error)
```

```
[ ] min_error = 100000000

    for a in tqdm(np.linspace(0, 1, 100)[1:-1]):
      for g in np.linspace(0, 1, 100)[1:-1]:
        for d in np.linspace(0, 1, 100)[1:-1]:
          (al, ga, delt, l, t, s, e) = winter_holt(train_data, a, g, d)
          if e < min_error:
            (alpha, gamma, delta, level, trend, seasonal, min_error) = (al, ga, delt, l, t, s, e)

    100%|██████████| 98/98 [22:07<00:00, 13.55s/it]
```

```
[ ] winter_forecasts = []

    for i in range(1, 13):
      winter_forecasts.append(np.exp(level[-1] + i*trend[-1] + seasonal[len(seasonal)-1+i-12]))
```

To choose α, γ and δ, a simple grid search is used over a linear space of values from 1 to 100. Below are the predictions obtained by this method.

Smoothing Parameter	Value
α	0.010101
γ	0.010101
δ	0.989898

The values of α, γ and δ show that the models does much more smoothing for the seasonal component.

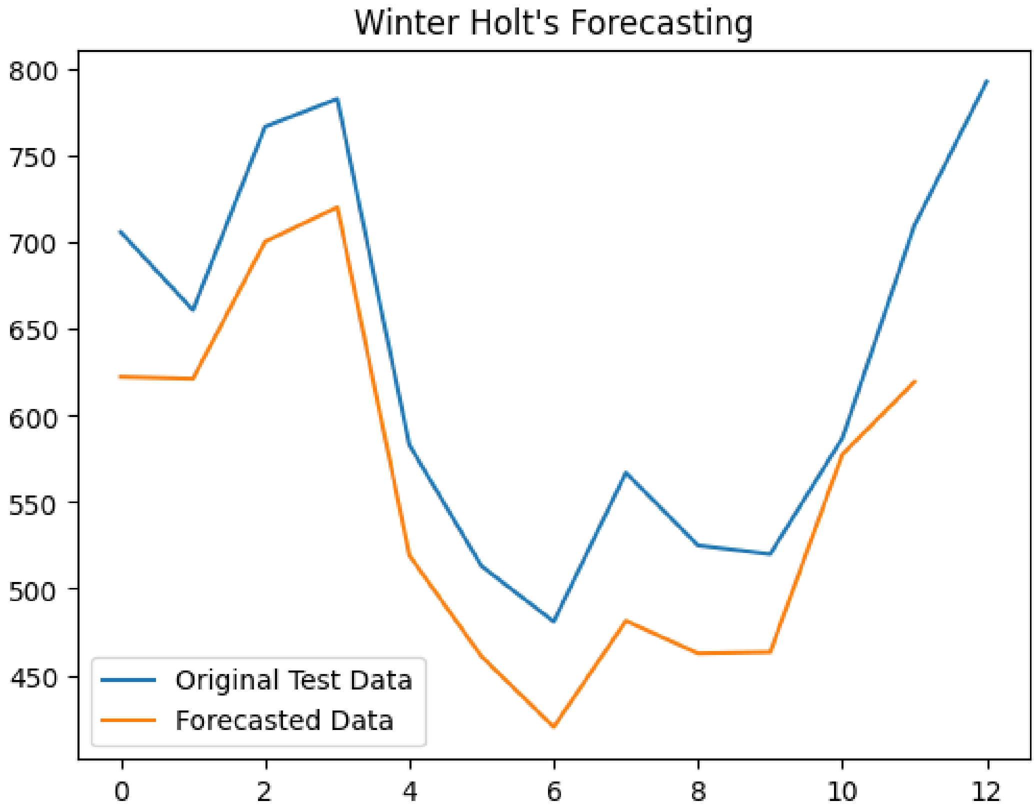

Figure 3: Forecasting of Winter Holt's Seasonality method for 20 months

Definitely, Winter Holt's Seasonality method does better prediction for the chosen data set. Below are the plots of original data and the forecasts side by side.

2.4 Final Plots

Below is the plot showing the forecasting.

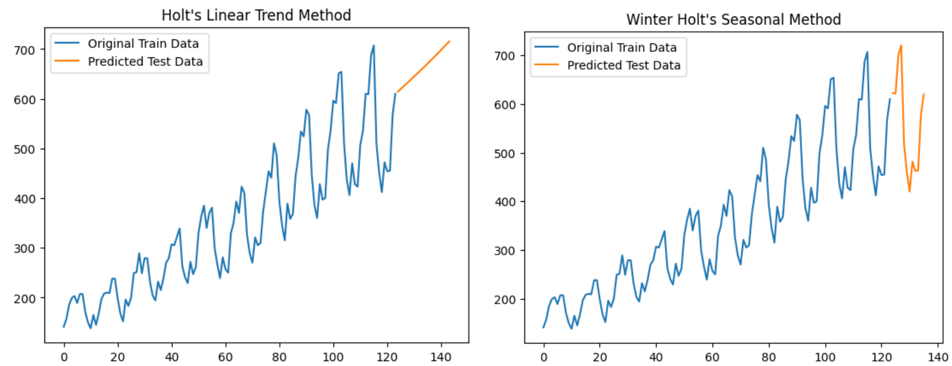

Figure 4: Forecasting of both the methods

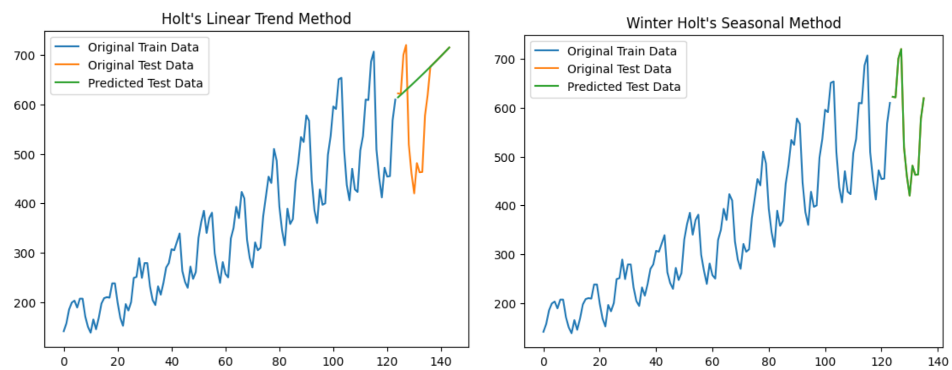

Figure 5: Forecasting of both the methods with test data plotting

3 Code

The code implemented to obtain the results and plots can be found at link