1. **How does the Prophet Algorithm differ from an LSTM?  
   Why does an LSTM have poor performance against ARIMA and Profit for Time Series?**

|  |  |
| --- | --- |
| **Prophet Algorithm** | **LSTM** |
| Prophet is a procedure for forecasting time series data based on an additive model where non-linear trends are used. It works best with time series data that has strong seasonal effects. | The LSTM is capable of capturing the patterns of both long term seasonality’s such as a yearly pattern and short term seasonality’s such as weekly patterns. |
| Prophet's algorithm requires less hyperparameter tuning as it is specifically designed to detect patterns in business time series data. | LSTM model needs careful hyperparameter tunning. |
| Prophet relies on assumption that data has Date field. | LSTM does not make assumption on data for existence of Date field. |
|  | The LSTM requires more computation than other recurrent neural networks. The main reason is that it has more parameters, which are used for demand forecasts. |
|  | LSTM addresses the vanishing gradient problem that makes network training difficult for a long sequence of words or integers. |

|  |  |
| --- | --- |
| **ARIMA** | **LSTM** |
| **ARIMA** - Auto Regressive Integrated Moving Average is actually a class of models that ‘explains’ a given time series based on its own past values, that is, its own lags and the lagged forecast errors, so that equation can be used to forecast future values.  An ARIMA model is characterized by 3 terms: p, d, q where,  p - the order of the AR term, Auto Regression  q - the order of the MA term, moving average  d - the number of differencing required to make the time series stationary. | **LSTM** unit is composed of **a cell, an input gate, an output gate and a forget gate**. The cell remembers values over arbitrary time intervals and the three gates regulate the flow of information into and out of the cell.  **Complex layered architecture.** |
| Only p,q,d values need to be calculated | Needs multiple hyper parameter tunning. |
| Can be trained for small datasets. | Needs large amounts of data to train |
| Lesser computational resources | More computational resources, |
| Faster and better predictions | Takes longer time to train and predict |

1. **What is exponential smoothing and why is it used in Time Series Forecasting?**

Exponential Smoothing is a technique used for forecasting where the forecast is made through the exponentially weighted average of prior observations.

The largest weight is provided to present observations, less weighted are projected to immediately preceding observations, more less weighted to the observation earlier to that, and so on such that weighted values follow/ reflect exponential decay in terms of influence of past data.

The exponential smoothing of time series data allocates the exponentially decaying weights from newest to oldest observations, ie. analyzing data from a specific period of time via providing more importance to recent data and less importance to former data. This method produces “smoothed data”, the data that has a noise removed, and allows trends and patterns to be more clearly visible.

The essential aim of exponential smoothing is to make original series smooth and make use of smoothed data for forecasting futures values of the variable of interest.

|  |  |  |
| --- | --- | --- |
| **Single Exponential Soothing** | **Double Exponential Soothing** | **Triple Exponential Soothing** |
| Simple Exponential Smoothing, is a time series forecasting method for univariate data without a trend or seasonality. | Double Exponential Smoothing is an extension to Exponential Smoothing that explicitly adds support for trends in the univariate time series. | Triple Exponential Smoothing is an extension of Exponential Smoothing that explicitly adds support for seasonality to the univariate time series. |
| It requires a single parameter, called alpha (a), also called the smoothing factor or smoothing coefficient. | In addition to the alpha parameter for controlling smoothing factor for the level, an additional smoothing factor is added to control the decay of the influence of the change in trend called beta (b). | In addition to the alpha and beta smoothing factors, a new parameter is added called gamma (g) that controls the influence on the seasonal component. |
| This parameter controls the rate at which the influence of the observations at prior time steps decay exponentially. Alpha is often set to a value between 0 and 1. Large values mean that the model pays attention mainly to the most recent past observations, whereas smaller values mean more of the history is taken into account when making a prediction. | The method supports trends that change in different ways: an additive and a multiplicative, depending on whether the trend is linear or exponential respectively. | As with the trend, the seasonality may be modeled as either an additive or multiplicative process for a linear or exponential change in the seasonality. |
| Hyperparameters:   * **Alpha**: Smoothing factor for the level. | Hyperparameters:   * **Alpha**: Smoothing factor for the level. * **Beta**: Smoothing factor for the trend. * **Trend Type**: Additive or multiplicative. * **Dampen Type**: Additive or multiplicative. * **Phi**: Damping coefficient. | Hyperparameters:   * **Alpha**: Smoothing factor for the level. * **Beta**: Smoothing factor for the trend. * **Gamma**: Smoothing factor for the seasonality. * **Trend Type**: Additive or multiplicative. * **Dampen Type**: Additive or multiplicative. * **Phi**: Damping coefficient. * **Seasonality Type**: Additive or multiplicative. * **Period**: Time steps in seasonal period |

1. **What is stationarity? What is seasonality? Why Is Stationarity Important in Time Series Forecasting?**

**Stationarity:** Data is said to be stationery when the statistical properties do not change over time.(ex: mean, variance do not change over time).  When forecasting or predicting the future, most time series models assume that each point is independent of one another. The best indication of this is when the dataset of past instances is stationary.

Graphical user interface, chart, line chart

Description automatically generated

**Seasonality:** Recurring pattern at a fixed and known frequency based on a time of the year, month, week, or day. Seasonality is **a characteristic of a time series in which the data experiences regular and predictable changes that recur every calendar year**. Any predictable fluctuation or pattern that recurs or repeats over a one-year period is said to be seasonal.

Say the data is not stationary as shown below. The mean and variance keep changing over time. Then the time series predictions cannot be made because the time series models assume that the data is stationery to make predictions.

Chart, line chart

Description automatically generated

1. **How is seasonality different from cyclicality? Fill in the blanks:  
   \_\_Seasonality\_ is predictable, whereas \_\_Cyclicality\_ is not.**

|  |  |
| --- | --- |
| **Seasonality** | **Cyclicality** |
| A seasonal pattern occurs when a time series is affected by seasonal factors such as the time of the year or the day of the week. | A cycle occurs when the data exhibit rises and falls that are not of a fixed frequency. |
| Seasonality is always of a fixed and known frequency. | These fluctuations are usually due to economic conditions and are often related to the “business cycle”. The duration of these fluctuations is usually at least 2 years. |
| Seasonality is Predictable | Cyclicality is not predictable. |
|  | the average length of cycles is longer than the length of a seasonal pattern |
|  | magnitudes of cycles tend to be more variable than the magnitudes of seasonal patterns |

[**https://neptune.ai/blog/arima-vs-prophet-vs-lstm#:~:text=Prophet's%20advantage%20is%20that%20it,are%20only%20a%20special%20case**](https://neptune.ai/blog/arima-vs-prophet-vs-lstm#:~:text=Prophet's%20advantage%20is%20that%20it,are%20only%20a%20special%20case)**.**

[**https://www.capitalone.com/tech/machine-learning/understanding-arima-models/**](https://www.capitalone.com/tech/machine-learning/understanding-arima-models/)

[**https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/**](https://www.machinelearningplus.com/time-series/arima-model-time-series-forecasting-python/)

[**https://medium.com/mlearning-ai/the-classification-of-text-messages-using-lstm-bi-lstm-and-gru-f79b207f90ad**](https://medium.com/mlearning-ai/the-classification-of-text-messages-using-lstm-bi-lstm-and-gru-f79b207f90ad)

[**https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/**](https://machinelearningmastery.com/exponential-smoothing-for-time-series-forecasting-in-python/)

[**https://www.analyticssteps.com/blogs/tutorial-exponential-smoothing-and-its-types**](https://www.analyticssteps.com/blogs/tutorial-exponential-smoothing-and-its-types)

[**https://towardsdatascience.com/why-does-stationarity-matter-in-time-series-analysis-e2fb7be74454#:~:text=Stationarity%20is%20an%20important%20concept,is%20independent%20of%20one%20another**](https://towardsdatascience.com/why-does-stationarity-matter-in-time-series-analysis-e2fb7be74454#:~:text=Stationarity%20is%20an%20important%20concept,is%20independent%20of%20one%20another)**.**

[**https://otexts.com/fpp2/tspatterns.html**](https://otexts.com/fpp2/tspatterns.html)