**Stock Prophet Deployment – Rubric Questions**

1. **How does the Prophet Algorithm differ from an LSTM?**

Prophet was developed by Facebook as an algorithm for the in-house prediction of time-series values for different business applications. It is an additive model consisting of four components:

$Y\_{t}$ = g(t) + s(t) + h(t) + $\epsilon(t)$

**g**(t): It represents the *trend* and the objective is to capture the general trend of the series. For example, the number of advertisements views on Facebook is likely to increase over time as more people join the network.

s(t): It is the Seasonality component. The number of advertisement views might also depend on the season. For example, in the Northern hemisphere during the summer months, people are likely to spend more time outdoors and less time in from of their computers. Such seasonal fluctuations can be very different for different business time series. The second component is thus a function that models seasonal trends.

h(t): The Holidays component. We use the information for holidays which have a clear impact on most business time series. Note that holidays vary between years, countries, etc. and therefore the information needs to be explicitly provided to the model.

The error term εt stands for random fluctuations that cannot be explained by the model. As usual, it is assumed that εt follows a normal distribution N (0, σ2) with zero mean and unknown variance σ that has to be derived from the data.

LSTM stands for Long short-term memory. LSTM cells are used in recurrent neural networks that learn to predict the future from sequences of variable lengths. Note that recurrent neural networks work with any kind of sequential data and, unlike ARIMA and Prophet, are not restricted to time series.

The main idea behind LSTM cells is to learn the important parts of the sequence seen so far and forget the less important ones. This is achieved by the so-called gates, i.e., functions that have different learning objectives such as:

1. a compact representation of the time series seen so far
2. how to combine new input with the past representation of the series
3. what to forget about the series
4. what to output as a prediction for the next time step.

Why does an LSTM have poor performance against ARIMA and Profit for Time Series?

My personal experience with LSTM is that it LSTM model is too advanced for a rather small datasets and is prone to overfitting. Despite adding regularization terms such as dropout, its still continued to overfit.

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

Exponential smoothing is a time series forecasting method for univariate data that can be extended to support data with a systematic trend or seasonal component.

Time series methods like the Box-Jenkins ARIMA family of methods develop a model where the prediction is a weighted linear sum of recent past observations or lags.

Exponential smoothing forecasting methods are similar in that a prediction is a weighted sum of past observations, but the model explicitly uses an exponentially decreasing weight for past observations.

Collectively, the methods are sometimes referred to as ETS models, referring to the explicit modeling of Error, Trend and Seasonality.

There are three main types of exponential smoothing time series forecasting methods.

* Single exponential smoothing, SES for short, also called Simple Exponential Smoothing, is a time series forecasting method for univariate data without a trend or seasonality. It requires a single parameter, called *alpha* (*a*), also called the smoothing factor or smoothing coefficient. 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.
* Double exponential smoothing: This is an extension to Single exponential smoothing that explicitly adds support for trends in the univariate time series.

Double Exponential Smoothing with an additive trend is classically referred to as Holt’s linear trend model, named for the developer of the method Charles Holt.

* Additive Trend: Double Exponential Smoothing with a linear trend.
* Multiplicative Trend: Double Exponential Smoothing with an exponential trend.
* Triple exponential smoothing: This method is sometime called Holt-Winters Exponential Smoothing, named for two contributors to the method, Charles Holt and Peter Winters. As with the trend, seasonality may be modeled as either additive or multiplicative process for a linear or exponential change in seasonality.
* Additive Seasonality: Triple Exponential Smoothing with a linear seasonality.
* Multiplicative Seasonality: Triple Exponential Smoothing with an exponential seasonality.

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

Seasonality in time series data means periodic fluctuations. It is often considered when the graph of the time series resembles a sinusoidal shape, which means that the graph looks like a sine function or shows repetitions after every fixed interval of time. This repetition interval is known as period.

Stationarity is one of the most important characteristics of time series data. A time series is said to be stationary if it has constant mean, variance and the covariance is independent of time. In ideal situations we would prefer a stationary series, but in real world that’s not the case. There are different types of stationary time series as follows:

* Stationary process: A process that generates a stationary series of observations.
* Stationary model: A model that describes a stationary series of observations.
* Stationary trend: A time series that does not show a trend.
* Seasonal stationary: A time series that does not display seasonality.
* Strictly stationary: A mathematical definition of a stationary process, specifically that the joint distribution of observations is invariant to time shift.

Statistical tests:

The most famous one is the Augmented Dickey-Fuller test (ADF). ADF uses an autoregressive model and optimizes an information criterion across multiple different lag values. The simple idea behind this is looking at p-value. If the p-value is <= 0.05 then we reject the null hypothesis as the data does not have a unit root and is stationary.

Autocorrelation Function (ACF) plots:

Autocorrelation is the correlation of a signal with a delayed copy or lag of itself as a function of the delay. When plotting the value of the ACF for increasing lags (a plot called a orrelogram), the values trend to degrade to zero quickly for stationary time series, while for non-stationary data the degradation will happen more slowly.

* Stationarity means that the statistical properties of a time series (or rather the process generating it) do not change over time.
* Stationarity is important because many useful analytical tools and statistical tests and models rely on it.

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

Seasonal patterns have constant length, while cyclic patterns have variable length.