Thursday Learning Hour – June 2021

Modern Timeseries Forecasting

Prabakaran Chandran

10 June 2021

“Prediction is truly very difficult, especially if it’s about the unknown future”

– Nils Bohr, Nobel laureate

Agenda:

- Introduction to Time series forecasting

- Classic Time series forecasting
  - Auto regression , Moving average Family of Algorithms
  - Exponential smoothing Family of Algorithms
  - Conventional ML based algorithms

- Modern Time series forecasting
  - Facebook's Prophet
  - Amazon's Deep AR
  - LinkedIn's Silver kite

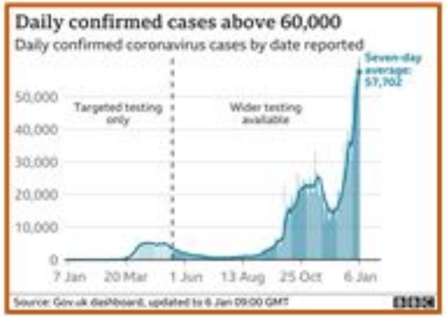
Introduction to Time series forecasting:

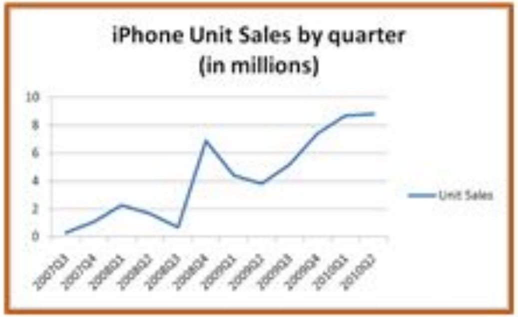
Time Series pertains to the sequence of observations collected in constant time intervals be it daily, monthly, quarterly or yearly

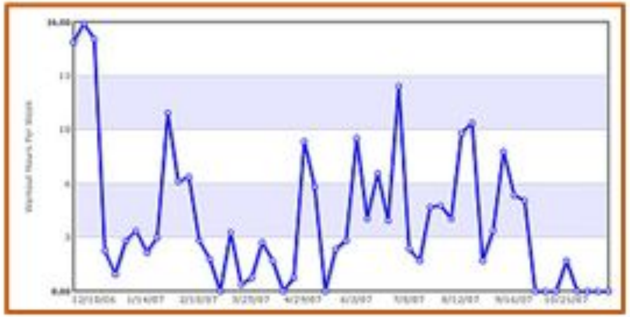
Time Series Analysis involves developing models(statistical / machine learnt / deep learnt) used to describe the observed time series and understand the "why" behind its dataset. This involves creating assumptions and interpretations about a given data.

Time Series Forecasting makes use of the best fitting model essential to predicting the future observation based on complex processing current and previous data.

Based on the granularity , we have many time series that are part of our life – covid cases, price of commodities(Doge , Ada , Bitcoin , Shiba) , Revenue , Sales , Fitness , Number of TLHs

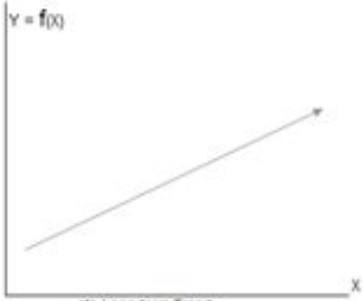




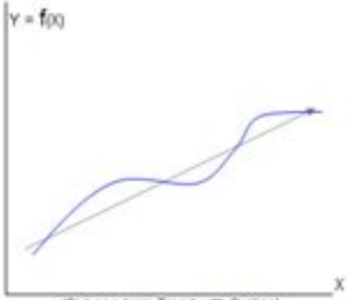


Components of Time series

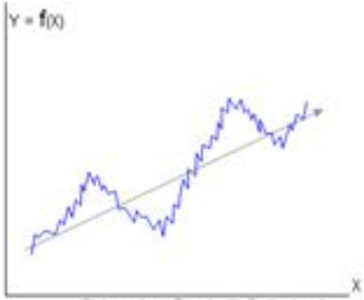
A time series can be decomposed into three components: the trend (long term direction), the seasonal (systematic, calendar related movements) and the irregular (unsystematic, short term fluctuations). $y_t = T_t + S_t + C_t + R_t.$ $y_t = T_t \times S_t \times C_t \times R_t$ $y_t = T_t + S_t + C_t R_t.$



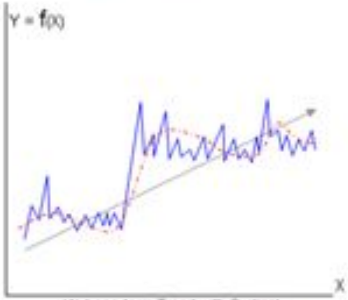
(1) Long-term Trend



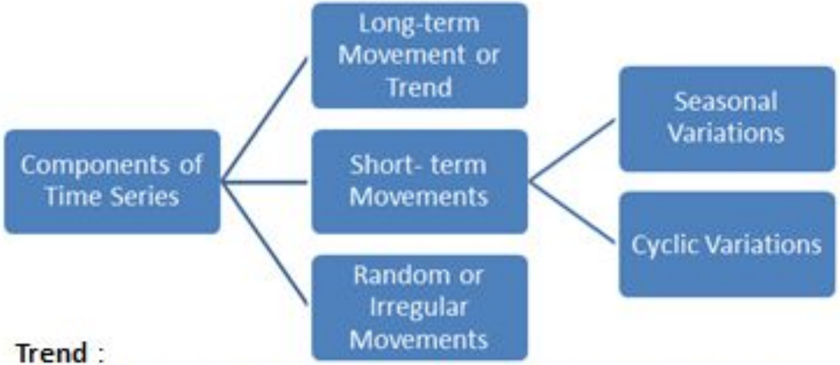
(2) Long-term Trend with Cyclical Variations/Movements



(3) Long-term Trend with Cyclical and Seasonal Variations/Movements



(4) Long-term Trend with Cyclical, Seasonal, and Random Variations/Movements



Trend :

→ the general tendency of the data to increase or decrease during a long period of time

Seasonality :

→ the rhythmic forces which operate in a regular and periodic manner over a span of less than a year.

Cycle:

→ the variations which operate over a span of more than one year are the cyclic variations. It is called as Business cycle

Irregularities:

→fluctuations are unforeseen, uncontrollable, unpredictable, and are erratic

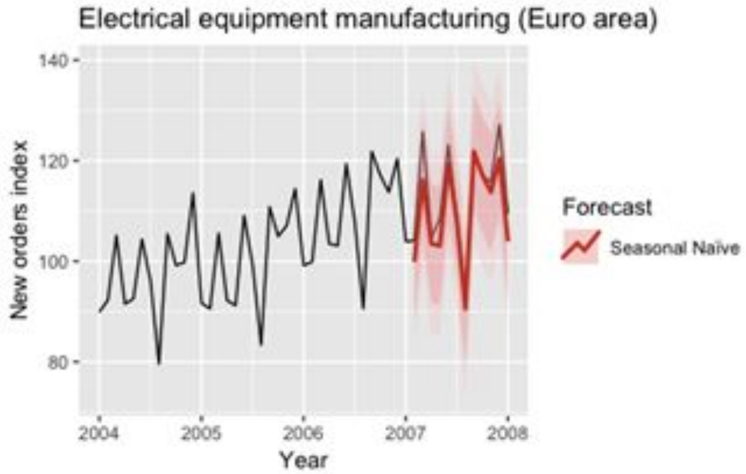
Naïve time series forecasting methods

Naïve model :

- the forecasts for every horizon correspond to the **last observed value.**
- $\hat{Y}(t+h|t) = Y(t)$

S Naïve (Seasonal Naïve)

- the forecasts for the following *T* time steps are equal to the previous *T* time steps.
- the forecasts for the next year is equal to the last year's observations.
- $\hat{Y}(t+h|t) = Y(t+h-T)$



Exponential Smoothing – Time series forecasting methods

Exponential smoothing was first suggested in the statistical literature without reference to previous work by Robert Goodell Brown in 1956 and then expanded by Charles C. Holt in 1957.

Exponential smoothing is a broadly accurate principle for smoothing time series data using the exponential window function.

Exponential Smoothing Formula

$s_t = \alpha x_t + (1 - \alpha)s_{t-1} = s_{t-1} + \alpha(x_t - s_{t-1})$

Here,

s_t = smoothed statistic, it is the simple weighted average of current observation x_t

s_{t-1} = previous smoothed statistic

α = smoothing factor of data; $0 < \alpha < 1$

t = time period

If the value of the smoothing factor is larger, then the level of smoothing will reduce.

Value of α close to 1 has less of a smoothing effect and give greater weight to recent changes in the data, while the value of α closer to zero has a greater smoothing effect and are less responsive to recent changes.

Exponential Smoothing – Time series forecasting methods

3 different types of Exponential smoothing

Single exponential smoothing

If the data has no trend and no seasonal pattern, then this method of forecasting the time series is essentially used

$$s_t = \alpha x_t + (1-\alpha)s_{t-1}$$

$$= s_{t-1} + \alpha(x_t - s_{t-1})$$

Forecast y_t = s_t

Double exponential smoothing

linear trend and no seasonal pattern

$$s_t = \alpha x_t + (1-\alpha)(s_{t-1} + b_{t-1})$$
$$b_t = \beta(s_t - s_{t-1}) + (1-\beta)b_{t-1}$$

St → level
bt → trends
Forecast y_t = s_t + b_t

Triple exponential smoothing

the time series when the data has both linear trend and seasonal pattern.

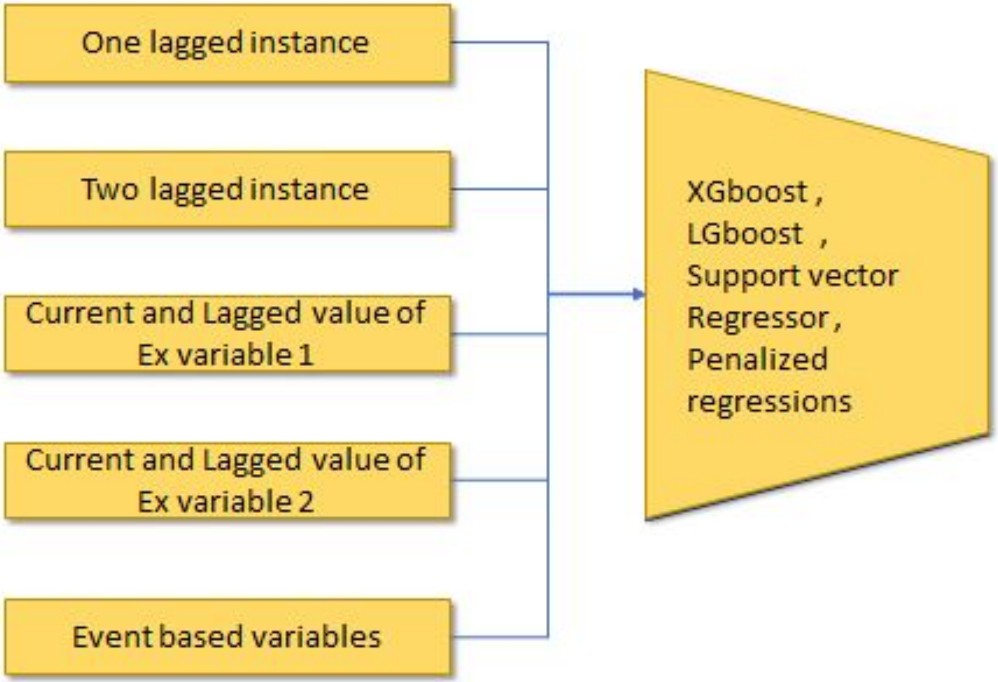
$$s_t = \alpha x_t + (1-\alpha)(s_{t-1} + b_{t-1})$$
$$b_t = \beta(s_t - s_{t-1}) + (1-\beta)b_{t-1}$$
$$C\_t = \gamma(y_t - b_{t-1} - s_{t-1}) + (1-\gamma)C_{t-m}$$

Forecast y_t = s_t + b_t + c_t

s_t = smoothed statistic, it is the simple weighted average of current observation x_t
s_{t-1} = previous smoothed statistic
α = smoothing factor of data; $0 < \alpha < 1$
t = time period
b_t = best estimate of a trend at time t
β = trend smoothing factor; $0 < \beta < 1$
c_t = sequence of seasonal correction factor at time t
γ = seasonal change smoothing factor; $0 < \gamma < 1$

Forecasting as a Conventional Regression problem :-

Other than Classical Statistical models , we often use ML based algorithms for time series forecasting .Here we consider this forecasting problem as a regression problem



This can be used for Univariate time series forecasting and also used for multivariate time series forecasting

Multivariate time series forecasting is a common problem in Time series forecasting

This structure shows the importance of the feature engineering

Modern Time Series Forecasting : - Facebook's Prophet

Prophet was developed internally at Facebook by Sean J. Taylor and Ben Letham in order to build a flexible forecasting tool and robust forecasting tool which is easy to use

Prophet was designed to optimally handle business forecasting tasks, which typically feature any of these attributes:

- Time series data captured at the hourly, daily, or weekly level with ideally at least a full year of historical data

- Strong seasonality effects occurring daily, weekly, and/or yearly

- Holidays and other special one-time events that don't necessarily follow the seasonality patterns but occur irregularly

- Missing data and outliers

- Significant trend changes that may occur with the launch of new features or products, for example

- Trends that asymptotically approach an upper or lower bound

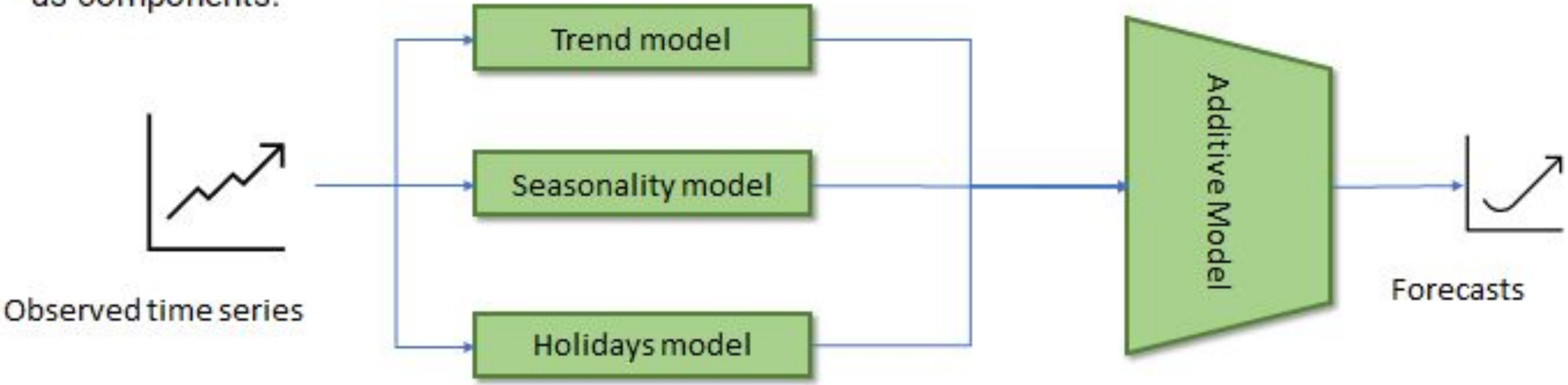
Modern Time Series Forecasting : - Facebook's Prophet

Forecast : $Y_t = growth\ term\ g(t) + seasonal\ term\ s(t) + holiday\ term\ h(t) + residuals$

Looks like a Generalized Additive Models

- **g(t)**: piecewise linear or logistic growth curve for modelling non-periodic changes in time series
- **s(t)**: periodic changes (e.g. weekly/yearly seasonality)
- **h(t)**: effects of holidays (user provided) with irregular schedules
- **ε_t**: error term accounts for any unusual changes not accommodated by the model

- Using time as a regressor, Prophet is trying to fit several linear and non linear functions of time as components.



Modern Time Series Forecasting : - Facebook's Prophet

Trend Model :

1. two trend models that cover many Facebook applications: a saturating growth model, and a piecewise linear model.

saturating growth model

-- Trend growth will reach the saturation point at C – capacity.

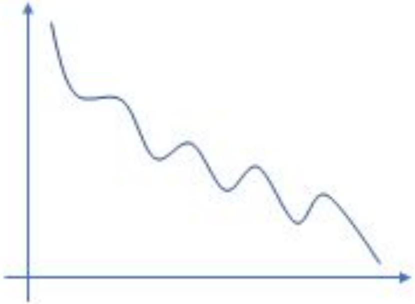
-- Typically, saturating growth is modeled as Logistic growth

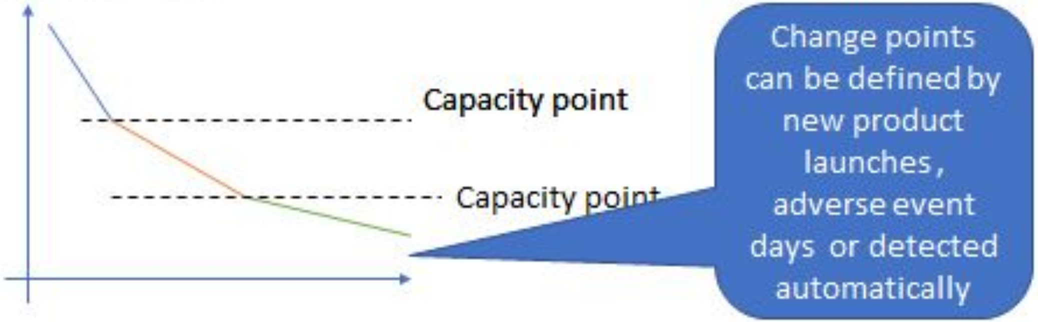
$$g(t) = \frac{C}{1 + \exp(-k(t - m))},$$

with C the carrying capacity, k the growth rate, and m an offset parameter.

The piecewise logistic growth model is then

$$g(t) = \frac{C(t)}{1 + \exp(-(k + \mathbf{a}(t)^{\intercal}\boldsymbol{\delta})(t - (m + \mathbf{a}(t)^{\intercal}\boldsymbol{\gamma})))}.$$





At change points , Growth rate can be changed – based on Gamma values (adjustment ration)

Modern Time Series Forecasting : - Facebook's Prophet

Seasonality Model :

Fourier series to provide a flexible model of seasonal components

$$s(t) = \sum_{n=1}^{N} \left(a_n \cos\left(\frac{2\pi nt}{P}\right) + b_n \sin\left(\frac{2\pi nt}{P}\right) \right)$$

a_n → Fourier series , P → Number of period effects , n → weekly , daily , yearly seasonality

Holidays and Events:

- Holiday effects are added as simple dummy variables.
- It is often important to include effects for a window of days around a particular holiday, such as the weekend of Thanksgiving.
- To account for that we include additional parameters for the days surrounding the holiday, essentially treating each of the days in the window around the holiday as a holiday itself

Holiday	Country	Year	Date
Thanksgiving	US	2015	26 Nov 2015
Thanksgiving	US	2016	24 Nov 2016
Thanksgiving	US	2017	23 Nov 2017
Thanksgiving	US	2018	22 Nov 2018
Christmas	*	2015	25 Dec 2015
Christmas	*	2016	25 Dec 2016
Christmas	*	2017	25 Dec 2017
Christmas	*	2018	25 Dec 2018

Beyond everything Prophet stands as a unique tool,
-- automatic evaluation of forecasts in which packages tries to calculate the accuracy and the large forecast error
-- Robust to outliers and missing values

Modern Time Series Forecasting : - LinkedIn's Silver kite (a new-born baby)

The Silverkite algorithm works well on time series with (potentially time-varying) trends and seasonality, repeated events/holidays, and/or short-range effects.

applied it to a wide variety of metrics in different time frequencies (hourly, daily, weekly, etc.), as well as various forecast horizons, e.g., 1 day ahead (short-term) or 1 year ahead (long-term).

Some key benefits:

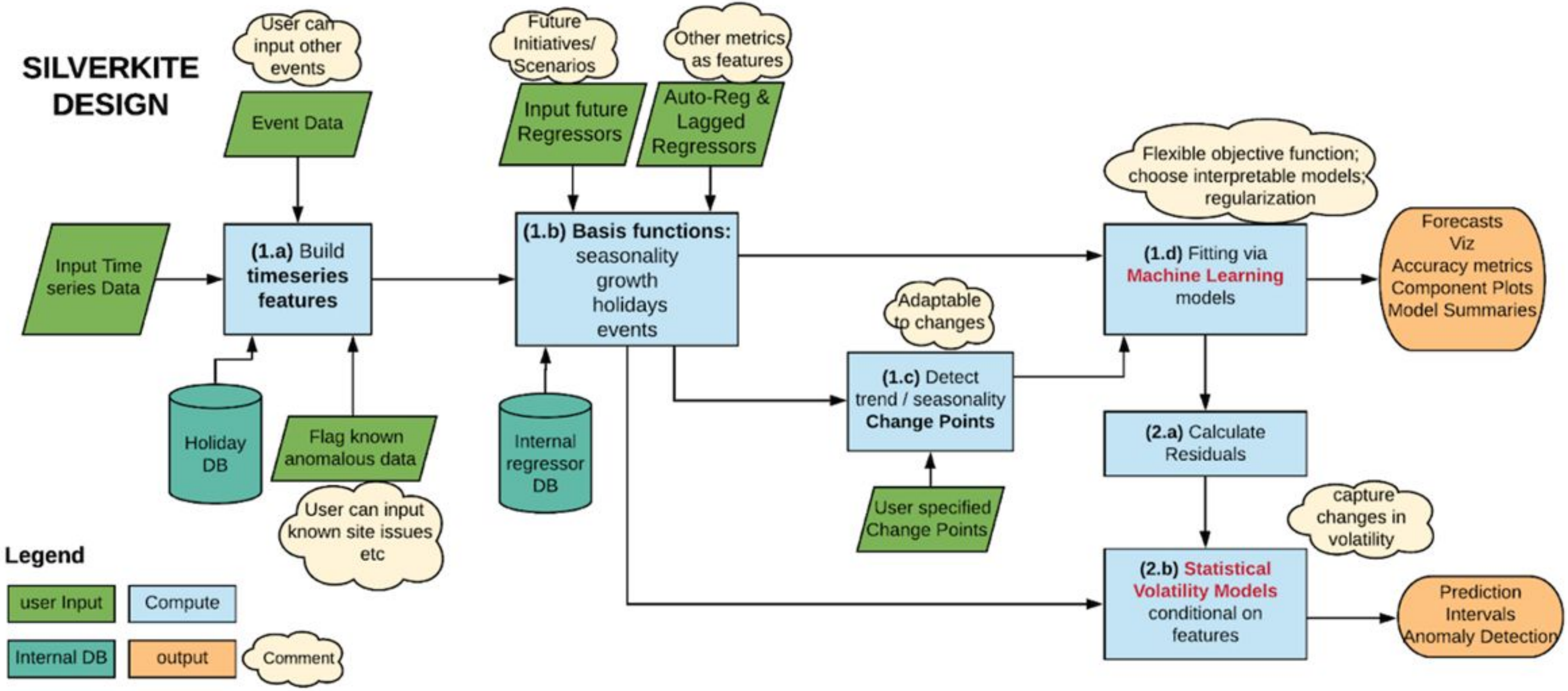
- **Flexible:** provides time series regressors for trend, seasonality, holidays, changepoints, and autoregression; users select the ones they need and fit the machine learning model of their choice.
- **Intuitive:** provides exploratory plots, templates for tuning, and explainable forecasts with clear assumptions.
- **Fast:** allows for quick prototyping and deployment at scale.

At LinkedIn, we use time series forecasts for resource planning, performance management, optimization, and ecosystem insight generation.
For example:

1. To provision sufficient infrastructure to handle peak traffic.
2. To set business metric targets and track progress for operational success.
3. To optimize budget decisions by forecasting growth of various markets.
4. To understand which countries are recovering faster or slower after a shock like the COVID-19 pandemic.

Modern Time Series Forecasting : - LinkedIn's Silver kite (a new-born baby)

Algorithm design



Modern Time Series Forecasting : - LinkedIn's Silver kite (a new-born baby)

Two phases of Forecasting:

•Phase (1): the conditional mean model.– to fil the base line predictions
•Phase (2): the volatility/error model.– to fit the residuals

Phase (1) can be broken down into these steps:

•(1.a) Extract raw features from timestamps, events data, and history of the series (e.g., hour, day of week);

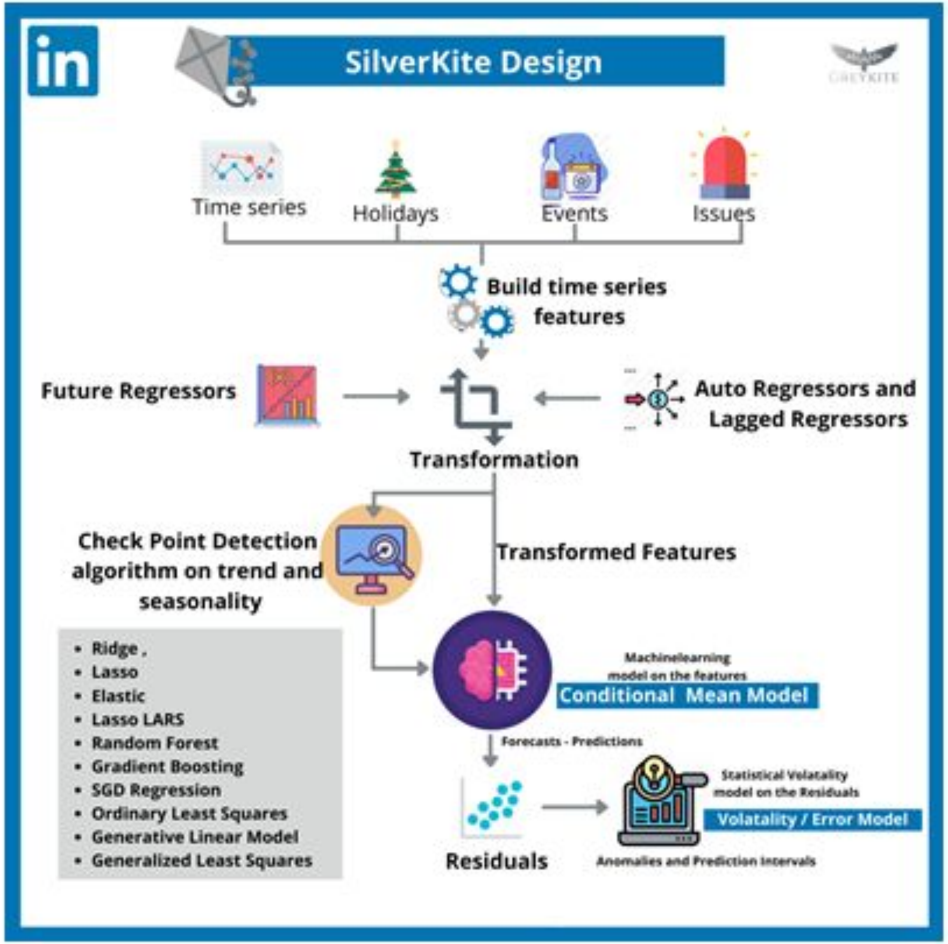
•(1.b) Transform the features to appropriate basis functions (e.g., Fourier series terms for various time scales);

•(1.c) Apply a changepoint detection algorithm to the data to discover changes in the trend and seasonality over time;

•(1.d) Apply an appropriate machine learning algorithm to fit the features from (1.b) and (1.c) (depending on the objective).

•In Phase (2), a simple conditional variance model can be fitted to the residuals, which allows for the volatility to be a function of specified factors, e.g., day of the week.

Modern Time Series Forecasting : - LinkedIn's Silver kite (a new-born baby)



Changepoints are detected in trend and seasonality terms

1. Piece wise curve fitting is used to calculate the trend seasonality along with the penalization

1. Seasonality effect is modeled with Fourier series terms in the model, and this makes it easy to include seasonality changepoints

Once Feature Engineering is done , any kind of regression model can be utilized to fit the conditional mean model.

Later , the residuals will be utilized to build the volatility model which predicts the upper and lower band of uncertainty

Thanks for attending!

Appendix