International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

A New AGE Forecasting Model PROPHET

[1]Rishabh Das, [2]Apeksha Raut, [3]Prof.Shubhangi Mahadik, [4]Ms. Gunjan Behl

[1]Bharati Vidyapeeth's Institute of Management & Information Technology, Mumbai University.

[2]Bharati Vidyapeeth's Institute of Management & Information Technology, Mumbai University.

[3]BVIMIT Associate Professor, Navi Mumbai

[4]BVIMIT Associate Professor, Navi Mumbai

Abstract

Product analysis is the most crucial part for all manufacturing business. It delivers the sales record of their present manufactured product and also helps the business owners to predict there product performance in the future. For this study, a PROPHET model is used with Time series forecasting. This paper will unfold the need of such model rather than using a simple regression model to predict the order demand. This study evaluates the Prophet model to forecast an sales and demand for the Product over a time period. Sales and Demand in Product is the integral part for planning all processes in supply chain, and therefore deciding on Product demand is crucial for supply chain. This research work has proposed the FBProphet tool for the Sales and Demand prediction. The findings demonstrate that the model can be used to predict future demand in the product manufacturing sector. These findings will aid business owners in developing a solid decision-making framework.

Keyword: PROPHET, Facebook, Seasonality, Holiday

I. INTRODUCTION

In an unceasingly evolving manufacturing business, adapting to the changes in supply and demand is of prime significance. Currently, organizations are more gravitating towards developing the most effectual demand supply chain model. The manufacturing market has been involved with customers demanding and discriminating against the supplier of what products they want and when they need them.

Inventory management has a difficult time forecasting demand. Inventory management levels depend on the demands coming from the customers. The incorrect evaluation of demand can cause a huge loss to the manufacturing companies, which proves that the process is not correct.

However, large investments are required in inventories. You cannot define when there is no demand and have continued demand for stock. There is a time when the product has several periods of no demand. This will cause difficulties for traditional time-demand forecasting methods to predict the demand for the future.

For this, time series models are used, which consider time as the main property for predictions. In the time series analysis, the datapoints are sequentially measured at particular time periods. This method tries to understand the content of the data points or make a prediction about the future values of those particular data points.

Many methods can be used for forecasting, but there are challenges of surety and availability of the product which are needed at all time. Therefore, it needs demand prediction for product orders and inventory to assure the product in hand. The main reason for this study is for manufacturers to manufacture the product according to the demand and make the decision so there will be no chances for any losses.

II. LITERATURE REVIEW

A. Forecasting

Forecasts are critical to a company's existence, as changes in today's companies are changing well-established organizational structures quickly and dramatically. All company markets, in this situation, require accurate and practical right future readings. Forecasting involves using variables to anticipate the future, which can include demand, supply, or price. The majority of the time, these variables are driven by demand. Future values are assumed in forecasting by considering these variables.

Forecasting demand while producing is the most problematic aspect of inventory management. These are frequently taken into account during operation planning, assembly processes, capacity planning, and the acquisition of used products. The demand forecast is used as a foundation for supply planning in the supply chain. There are procedures that are referred to as "pulls." customer demand procedures, as well as push processes for evaluating or hoping for customer demand. The most crucial decision a corporation must make is which forecasting approach to use. When choosing a method, these businesses must consider the pull and push processes. There are four types of forecasting models that are commonly used which can be applied to qualitative, time series, causal, and simulation applications. The time series forecasting methodology is used to predict order demand in these cases.

B. PROPHET

Prophet is a method for forecasting time series data based on an additive model with yearly, weekly, and daily seasonality, as well as holiday effects. It works best with time series with strong seasonal effects and data from multiple seasons. Prophet is resistant to missing data and trend shifts, and it typically handles outliers well. The Core Data Science team at Facebook released Prophet as open source software. It's available for free on CRAN and PyPI.

Because the Prophet forecasts time series "at scale," memory usage and computation complexity are not major concerns when making a forecast. It can fit time-series data with non-linear trends and holiday effects. It works well with data that has daily, weekly, monthly, and/or yearly seasonality, as well as when we have several seasons of historical data to make future forecasts. It has time-series forecasting APIs in R and Python. It employs the Stan platform to generate forecasts quickly and with easily interpretable parameters.

The core idea is based around the structural decomposition:

$$X_t = T_t + S_t + H_t + \epsilon_t$$

where

- Tt : trend component
- St : seasonal component (weekly, yearly)
- Ht : deterministic irregular component (holidays)
- ϵt: noise

The core mathematical idea behind Prophet is that the Kolmogorov-Arnold representation theorem, which states that multivariate function might be represented as sums and compositions of univariate functions:

$$f(x_1, \ldots, x_n) = \sum_{q=0}^{2n} \Phi_q \left(\sum_{p=1}^{n} \phi_{q,p}(x_p) \right)$$

The theorem has no constructive proof suitable for modeling ⟹⟹ simplification is necessary:

$$f(x_1, \ldots, x_n) = \Phi \left(\sum_{p=1}^{n} \phi_p(x_p) \right)$$

where ΦΦ could be a smooth monotonic function. This equation gives a general representation of GAM models and a well-recognized variant of this approach is that the class of Generalized Linear Models :

$$\Phi^{-1}\left[\mathrm{E}(Y)\right] = \beta_0 + f_1(x_1) + f_2(x_2) + \cdots$$

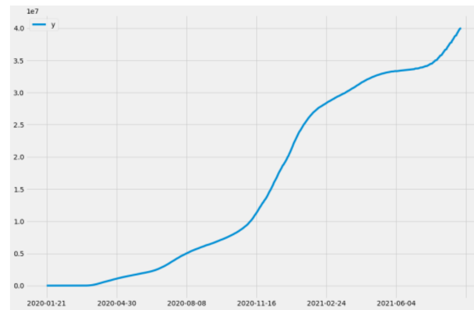
The smooth functions within the context of Prophet are the trend, seasonal and holiday components - we are able to isolate each individual function and evaluate its effect in prediction, which makes such models easier to interpret. we estimate through backfitting algorithm →→ convergence.

So how does that employment in practice? We take a GAM-style decomposition as our starting point:

$$X_t = T(t) + S(t) + H(t)$$

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

Unpacking the equation:
• time is that the only regressor
• easy accomodation of latest components
• multiple seasonal patterns
• forecasting
• no need for normal spacing
• fast fitting with backfitting algorithm
• probabilistic aspects
• works in additional general cases, but "designed" for daily data

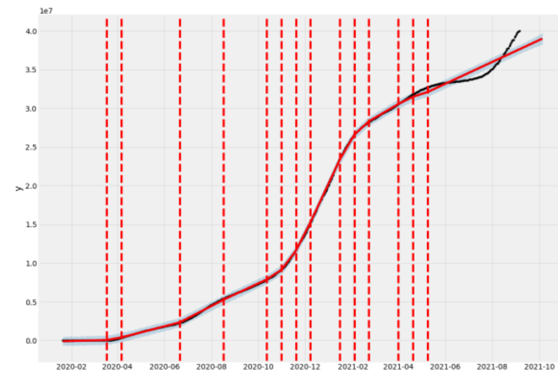


Piecewise Linear Model with a constant rate of growth

C. Dataset

We have use different types of dataset to show different features of the Prophet Model.

- PJM Hourly Energy Consumption Data. Over 10 years of hourly energy consumption data from PJM in Megawatts.
- New York Times data on Covid cases
- who-cases-dataset-and-wdi-country-population
- air_passengers dataset
- Norway bicycles Dataset



automatic detection of changepoints

D. Method of Execution

There are two potential trend models that the Prophet library uses.

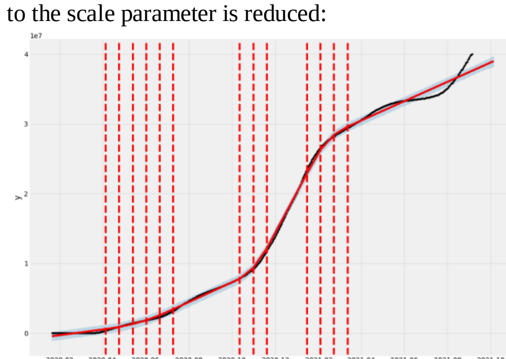
The linear trend

The first trend model is a straightforward piecewise linear model with a fixed rate of growth. Because a piecewise linear function can roughly mimic a large class of shapes, it is best suited for issues without saturation growth.

$$T(t) = \left[k + a(t)^T \delta\right] t + \left[m + a(t)^T \gamma\right]$$

Prophet by default calculates 25 model changepoints, or more than 80% of the dataset (those parameters can be adjusted). Changepoints are defined as deviations from the trajectory and can be manually entered or estimated; the latter is preferable if the analyst has access to domain expertise.

We may regulate this behaviour by imposing more regularisation. It looks that the default settings are too lenient when it comes to allocating changepoints. In order to achieve this, the value of the changepoint before to the scale parameter is reduced:



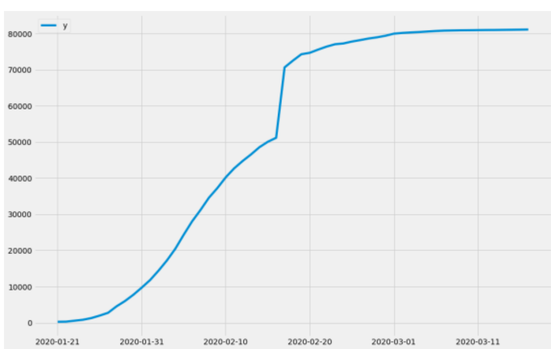
Nonlinear growth

Nonlinear, Saturating Growth is the first. The logistic growth model serves as a representation of it:

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

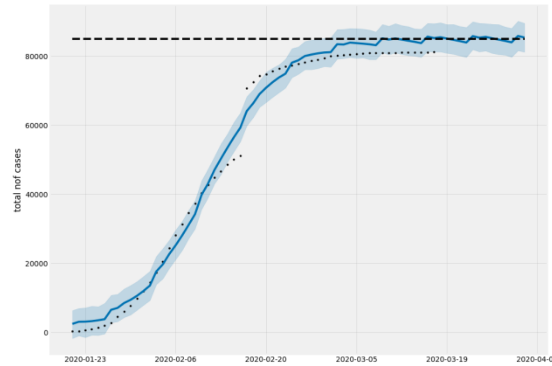
where C denotes carrying capacity (maximum amount) and k denotes growth rate ("steepness" of the trend curve). C and k may be stable throughout time or fluctuate. This logistic equation can simulate non-linear growth with saturation, which happens when the growth rate of a value declines as it increases. Variable tuning is possible with Prophet both automatically and manually. The library may choose the best points of trend shift by fitting the given historical data.

To continue the topic that has been the focus of everyone's attention since 2020, we will use data on COVID-19 instances in China.



Nonlinear, Saturating Growth

The curve appears to be flattening slightly; in Prophet, we can incorporate such knowledge into the model by setting a "cap" (an upper limit on the forecast value):



Seasonality

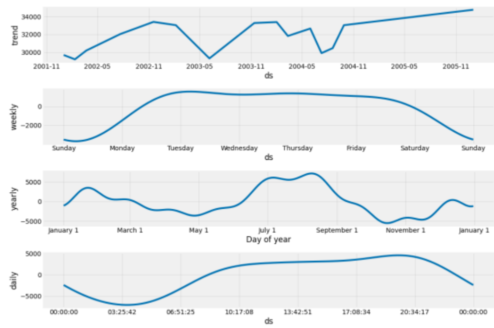
When dealing with data in practical applications, it is frequently important to take into account numerous seasonal patterns that are present at the same time. Energy consumption data is a classic example; there are trends in the morning and evening, workdays and weekends, and throughout the year (annual). One of the issues Prophet was created to address was the time-consuming nature of modelling them explicitly (you need to add extra equations in exponential smoothing or insert dummies in ARIMA). The basic idea is based on the Fourier expansion:

$$S(t) = \sum_{i=1}^{N} \left(a_n cos\left(\frac{2\pi it}{P}\right) + b_n sin\left(\frac{2\pi it}{P}\right)\right)$$

Deconstructing this formula:

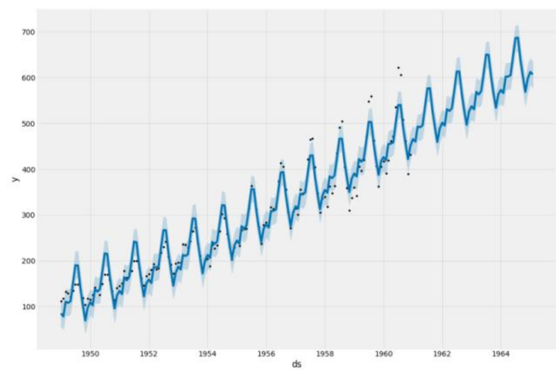
- The sin and cos functions are orthogonal.
- Every function can therefore be written as a combination, as in the equation above. Keep in mind that time is the only regressor in the GAM configuration, hence a time series can be seen as a function of time S. (t).
- By cutting off the expansion for a specific N, we can remove high frequency oscillations through low-pass filtering

Because Prophet is designed for univariate data, we will just use one series. Notice that we are using hourly data, which we will gradually aggregate to lower frequencies to demonstrate Prophet's out-of-the-box functionality.

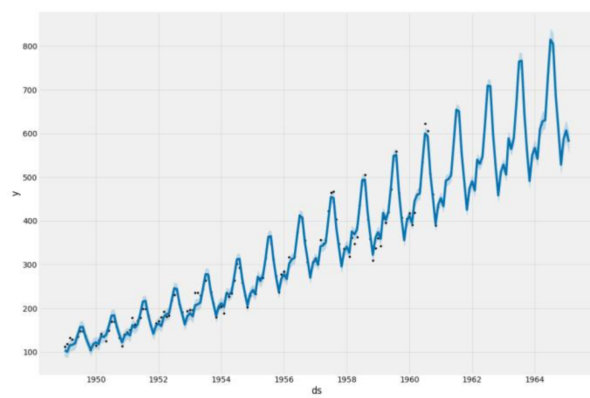


Seasonality specification

Aside from explicitly picking which frequencies to mimic, we have more options for configuring our Prophet model. The first is seasonality mode, which might be additive or multiplicative:

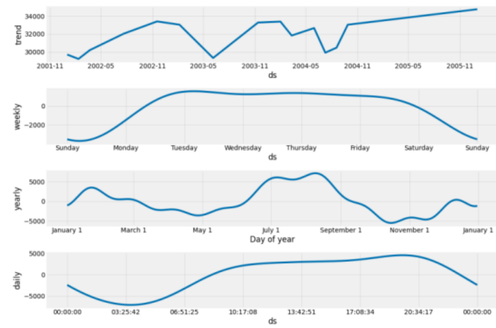


Additive mode



Multiplicative Mode

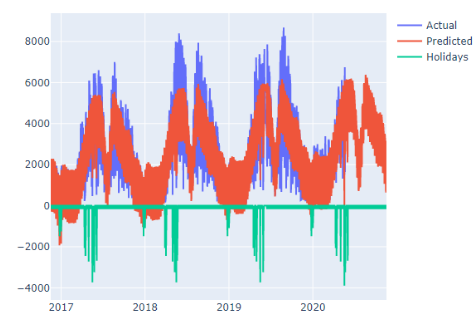
Last but not least, we can—and should—use interval forecasting, which combines our point estimates with uncertainty. Because the option mcmc samples is set to 0 by default, full Bayesian sampling is required to obtain the interval around seasonality; uncertainty around trend can be determined using the Maximum A Posteriori (MAP) estimate.



Special days

So we've dealt with trend and seasonality, but it doesn't mean everything else belongs in the random error component. There are data points that are not necessarily random, but can nonetheless have an impact on the model's performance:

- statutory holidays (Christmas, Easter, New Year, Black Friday)
- special occasions (World Cup)
- important events, such as the commencement of a conflict or terrorist attacks - when assessing the model's historical performance, it is frequently advantageous to exclude those



When it comes to representing unusual days, Prophet uses a bespoke list of occurrences. We assume that the effects are

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 06 Issue: 07 | July - 2022 Impact Factor: 7.185 ISSN: 2582-3930

unrelated; therefore, if a sporting event falls on a public holiday, the effects will be recorded separately.

Holidays

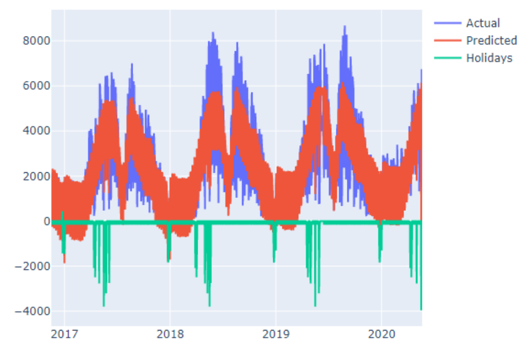
The dataset we will use to test the Prophet functionality around special days is daily count of norway-bicycles commuters in Oslo. We start with the fast and easy way of adding holidays: using the built-in list of country holidays:

```
0                  Søndag
1           Første påskedag
2          Arbeidernes dag
3          Første pinsedag
4           Første juledag
5        Første nyttårsdag
6             Grunnlovsdag
7            Andre juledag
8            Skjærtorsdag
9              Langfredag
10         Andre påskedag
11  Kristi himmelfartsdag
12          Andre pinsedag
dtype: object
```

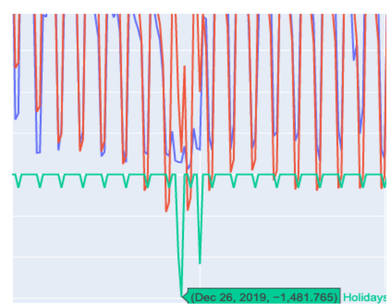
A short search on Google Translate reveals that Christmas is not mentioned, despite the fact that it is a significant public holiday in Europe. We'll add it later; for now, let's see how well the model performs right out of the box.

We enhance by adding to the list of holidays; we do so by generating a new dataframe called "Christmas," which is then sent to Prophet as an input. The 'holiday' entry is largely for the sake of interpretation; the following sections are relevant:

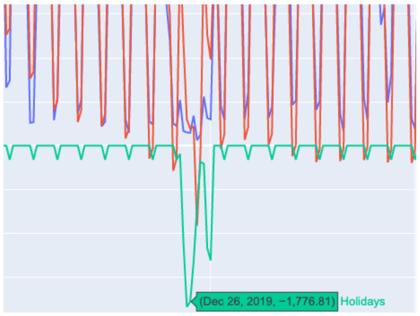
- DS-So we know when the holiday of interest happens.
- The lower window and upper window: those two parameters allow us to incorporate the effect before and after the date, respectively. In our example below, the lower window equals -1 in our example below, indicating that we expect a decline in the number of commuters a day before Christmas, whilst the upper window equals 7, indicating that, with many people taking time off between Christmas and New Year, bicycle traffic is anticipated to reduce for around a week.



Comparing the two graph:



Before adding Christmas



After adding Christmas

E. Conclusion

From the above data we see many features of the Prophet Model. Time Series Forecasting: Prophet, a library from Facebook, is a strong library with numerous built-in features for dealing with the issues of time series forecasting. Users with little programming experience can develop complicated time series forecasting models. The main benefit of Facebook's Prophet is that it can be used by anyone with no prior experience or in-depth knowledge of time series modeling. Prophet enables parameter tweaking and specific seasonality components, which can help improve forecasts.

F. References

1. https://www.kaggle.com/felixzhao/productdemandforecasting
2. Papacharalampous, G. A., and Tyralis, H. (2018). Evaluation of random forests and Prophet for daily streamflow forecasting. Advances in Geoscience 45: 201-208. https://doi:10.5194/adgeo-45- 201-2018
3. APPLICATION OF FACEBOOK'S PROPHET ALGORITHM FOR SUCCESSFUL SALES FORECASTING BASED ON REAL-WORLD DATA by Emir Žunić1,2, Kemal Korjenić1 , Kerim Hodžić2,1 and Dženana Đonko2
4. Forecasting at Scale Sean J. Taylor∗ † Benjamin Letham†