PROJECT REPORT

Forecast Commuters Inflow For Airline Industry Using Prophet Model

Category: Deep Learning

Guided Project At Smartinternz

Gadari Neha

neha1021g@gmail.com

**CONTENTS**

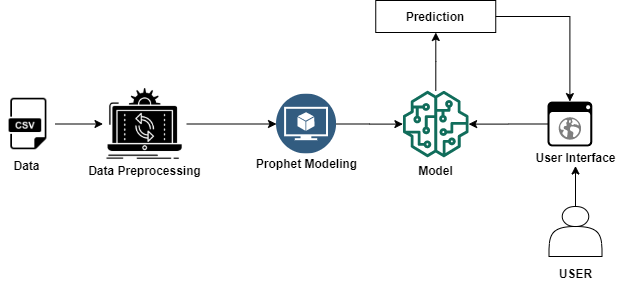
**Page No.**

1. Introduction------------------------------------------------------------1
2. Existing Problem------------------------------------------------------2-4
3. Solution-----------------------------------------------------------------5-6
4. Literature survey-------------------------------------------------------7
5. Experiment and Theoretical Analysis-------------------------------8-11
6. Advantages and Disadvantages------------------------------------12
7. Result-------------------------------------------------------------------13-14
8. Conclusion-------------------------------------------------------------15

**INTRODUCTION**

Air passenger traffic forecast is of great importance for airlines and civil aviation authorities. For airlines, accurate forecasts play an increasingly important role in revenue management. It helps to reduce the airlines’ risk by objectively evaluating the demand of the air transportation business. For civil aviation authorities, air passenger traffic forecast provides a concrete basis for planning decisions in air transport infrastructure. The main objective of this project is to build a prophet time series model that forecasts the passenger traffic for a given date.

**Architecture:**



**1**

**EXISTING PROBLEM**

From time crunch and flight delays to waiting lines and [cumbersome carry-on luggage](https://www.theactivetimes.com/travel/world/airlines-most-likely-lose-your-luggage), airplane travel can stress even the most nomadic traveler. It’s easy to let the hassle get the best of people, but knowing how to manage airline obstacles will help relieve the tension and [get you to your destination with minimal stress](https://www.theactivetimes.com/travel/n/15-tips-make-thanksgiving-travel-breeze).



**Making it through security**



To minimize time spent at the security checkpoint, be prepared and travel light,

**2**

minimizing obstacles to safe, smooth travel. [Make the security checkpoint go by quickly](https://www.theactivetimes.com/travel/gear/worst-things-wear-airport-according-tsa-agents) by emptying pockets ahead of time, removing laptop from bags, and removing shoes and belts to not only make it faster for you, but for those behind you. Also make sure that all liquids are in the appropriately sized containers before heading to the airport.

**TSA and long lines**



While TSA screenings are an important safety measure, [the long lines and extra time](https://www.theactivetimes.com/travel/us/best-and-worst-tsa-checkpoints-us) spent during bag searches and pat-downs can be a hassle.

**Overbooking**



Airlines often compensate [passengers who volunteer](https://www.theactivetimes.com/travel/n/things-flight-attendants-notice-about-you-when-you-board-plane) to give up their seat by paying for all expenses such as hotel and meals, in addition to giving them a flight voucher. If you are in absolutely no rush to get to your destination, it may be something to consider. However, if you’re one of the ones chosen to [give up your seat](https://www.theactivetimes.com/travel/us/airlines-offering-least-and-most-leg-room), but you have to be on that flight for other commitments, explain your situation and politely refuse, all while maintaining an amicable tone.

**3**

**Delays and cancelations**



You can’t always rely on leaving on time. Delays, which can result in missed connections, and cancelations are, unfortunately, part of everyday travel. They often happen without any warning, too.

**4**

**SOLUTION**

**Prophet: forecasting at scale**

Today Facebook is [open sourcing Prophet](https://github.com/facebookincubator/prophet), a forecasting tool available in Python and R. Forecasting is a data science task that is central to many activities within an organization. For instance, large organizations like Facebook must engage in capacity planning to efficiently allocate scarce resources and goal setting in order to measure performance relative to a baseline. Producing high quality forecasts is not an easy problem for either machines or for most analysts. We have observed two main themes in the practice of creating a variety of business forecasts:

* Completely automatic forecasting techniques can be brittle and they are often too inflexible to incorporate useful assumptions or heuristics.
* Analysts who can produce high quality forecasts are quite rare because forecasting is a specialized data science skill requiring substantial experience.

The problems of scale we have observed in practice involve the complexity introduced by the variety of forecasting problems and building trust in a large number of forecasts once they have been produced. Prophet has been a key piece to improving Facebook’s ability to create a large number of trustworthy forecasts used for decision-making and even in product features.

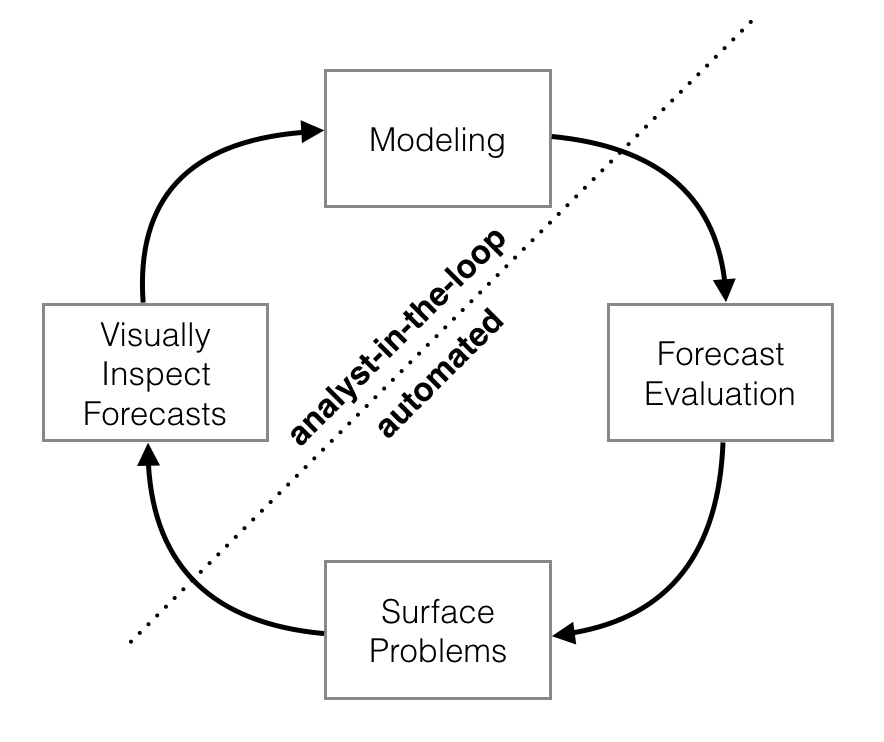
Not all forecasting problems can be solved by the same procedure. Prophet is optimized for the business forecast tasks we have encountered at Facebook, which typically have any of the following characteristics:

* hourly, daily, or weekly observations with at least a few months (preferably a year) of history
* strong multiple “human-scale” seasonalities: day of week and time of year
* important holidays that occur at irregular intervals that are known in advance (e.g. the Super Bowl)
* a reasonable number of missing observations or large outliers
* historical trend changes, for instance due to product launches or logging changes
* trends that are non-linear growth curves, where a trend hits a natural limit or saturates

We have found Prophet’s default settings to produce forecasts that are often accurate as

**5**

those produced by skilled forecasters, with much less effort. With Prophet, you are not stuck with the results of a completely automatic procedure if the forecast is not satisfactory — an analyst with no training in time series methods can improve or tweak forecasts using a variety of easily-interpretable parameters. We have found that by combining automatic forecasting with analyst-in-the-loop forecasts for special cases, it is possible to cover a wide variety of business use-cases. The following diagram illustrates the forecasting process we have found to work at scale:



**6**

**LITERATURE SURVEY**

While the literature search was used primarily to identify the main drivers of airline upgauging (and downgauging), the project team prepared an online survey for airports to evaluate the usual impacts and challenges that airports experience because of the rapid variation of traffic and changes in aircraft size.

To create a list of airports that is representative of all types and sizes of facilities and operations, the team performed an analysis of aviation activity at all U.S. airports and added to the screening process various selection criteria relevant to the synthesis topic:

• 10-year historical enplanements (2005–2016 data from the FAA Terminal Area Forecast (database), to identify any significant variation in air traffic activity, either representing major growth (potential candidate for recent airline upgauging), or a drastic decrease (possibly due to a loss of service)

• NPIAS category and hub size (non-hub, small hub, medium hub, large hub), to provide a representative sample for operations and sizes

• FAA regions (nine airport divisions), to provide an adequate representation of various geographical locations

• FAA Part 139 Certification

Airport choices were finalized in consultation with the panel of experts guiding this project. The collection of data and the results presented can be categorized as descriptive statistics summarizing the responses of a relatively small and select convenience sampling. The analysis of survey responses led to numerous insights about the issues and challenges associated with airline upgauging at various types of airports. The project team approached 20 airports. Eighteen completed the survey, while the other two did not respond before the deadline, so a 90% response rate was recorded.

**7**

**EXPERIMENTAL AND THEORETICAL ANALYSIS**

* We will be building a Web application where
* The user selects the date from User Interface(UI)
* The passenger traffic for the selected date is analysed by the model
* The count of passengers for the selected date is displayed on UI
* To accomplish this, complete all the milestones & activities listed below.

Data Collection.

* Create or Collect the dataset.

Data Pre-processing.

* Importing of Libraries.
* Importing of Dataset & Visualisation.

Model Building.

* Fitting the prophet library.
* Cross validation of the model.
* Evaluation of the model.
* Save the model.

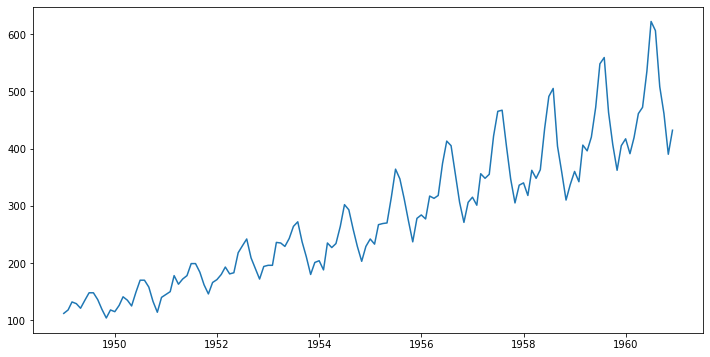
Application Development.

**Time Series Analysis with Prophet: Air Passenger Data**

Using Prophet to conduct time series analysis:

Prophet is a forecasting model by Facebook that forecasts time series using special adjustments for factors such as seasonality, holiday periods, and changepoints.

**8**



With the data now prepared,and ready to use the Prophet library to produce forecasts of our time series.

As we can see, the time series shows quite a stationary pattern (one where there is a constant mean, variance and autocorrelation.

We will not formally test for this condition here, but it is also evident that there appears to be significant seasonality present in the dataset — i.e. significant shifts in the time series trend that occur at certain time intervals.

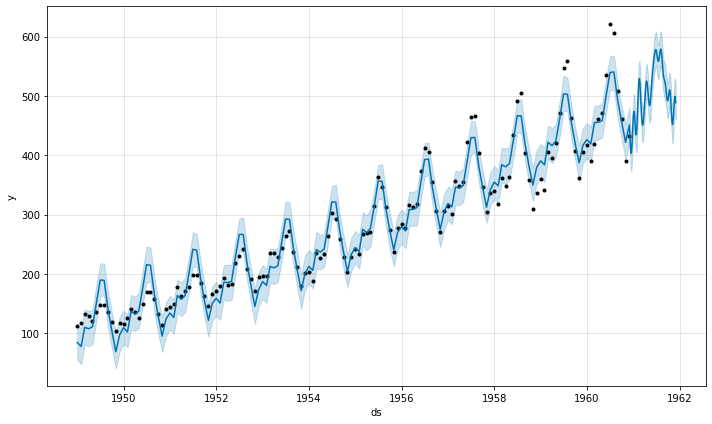
**Model Building**

The Prophet library to predict future values of our time series. The authors of Prophet have abstracted away many of the inherent complexities of time series forecasting and made it more intuitive for analysts and developers alike to work with time-series data.

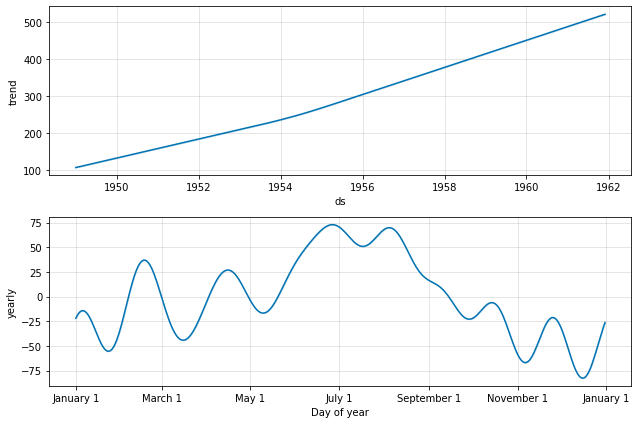
**Obtaining the Forecasts**

We use the predict method to make future predictions. This will generate a dataframe with an yhat column that will contain the predictions.  
 If we check the head for our forecast dataframe we’ll notice that it has very many columns. However, we are mainly interested in ds, yhat, yhat\_lower and yhat\_upper. yhat is our predicted forecast, yhat\_lower is the lower bound for our predictions and yhat\_upper is the upper bound for our predictions.

**9**



Here are the components of the forecast:



**10**

However, given that we are working with a monthly dataset — we will not use Prophet to explicitly model seasonality in this instance.

There are two reasons for this:

* Detection of seasonality would be more accurate if we were using daily data — but we are not in this case.
* Making an assumption of yearly seasonality may not be particularly accurate in this case. Inspecting the dataset shows that while certain seasonal shifts occur every year, others occur every 6 to 8 months. Therefore, explicitly defining a seasonality parameter in the model may do more harm than good in this instance.

**11**

**ADVANTAGES AND DISADVANTAGES**

**Advantages:**

We have frequently used Prophet as a replacement for the forecast package in many settings because of two main advantages:

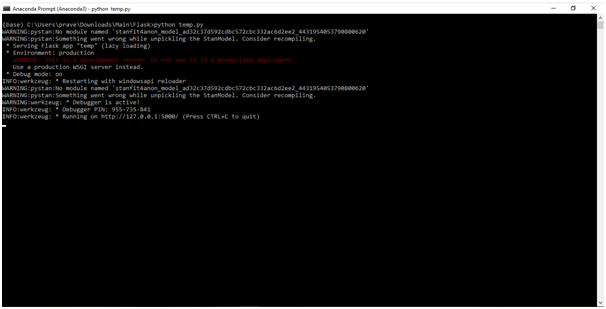
* Prophet makes it much more straightforward to create a reasonable, accurate forecast. The forecast package includes many different forecasting techniques (ARIMA, exponential smoothing, etc), each with their own strengths, weaknesses, and tuning parameters. We have found that choosing the wrong model or parameters can often yield poor results, and it is unlikely that even experienced analysts can choose the correct model and parameters efficiently given this array of choices.
* Prophet forecasts are customizable in ways that are intuitive to non-experts. There are smoothing parameters for seasonality that allow you to adjust how closely to fit historical cycles, as well as smoothing parameters for trends that allow you to adjust how aggressively to follow historical trend changes. For growth curves, you can manually specify “[capacities](https://en.wikipedia.org/wiki/Logistic_function#Time-varying_carrying_capacity)” or the upper limit of the growth curve, allowing you to inject your own prior information about how your forecast will grow (or decline). Finally, you can specify irregular holidays to model like the dates of the Super Bowl, Thanksgiving and Black Friday.

**Disadvantages:**

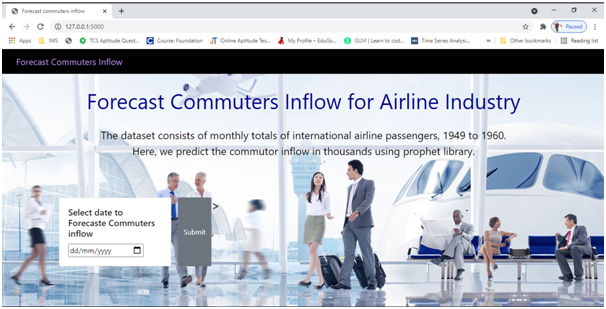
* Works well with only stationary data.
* Prior knowledge is required.
* Can’t handle data with seasonal components and irregularities.
* Becomes complex and fails to predict accurately.
* Real time data is often nonlinear.

**12**

**Result**

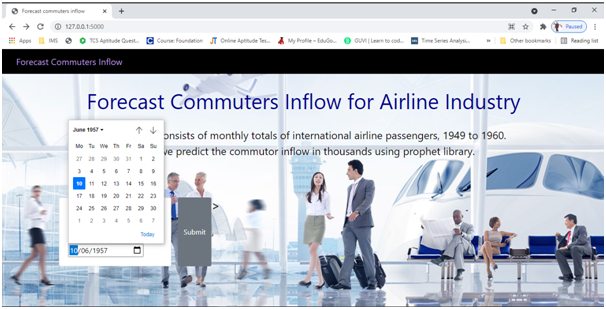


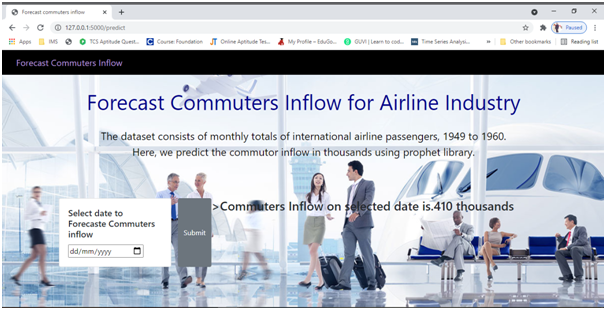
**UI will look like**



**13**

**Select the date you would like to predict and click on submit.**  
**The output prediction will be like:**





**14**

**CONCLUSION**

**By this project:**

* Will be able to know the fundamental concepts of time series forecasting.
* Working with Prophet library
* Flask Application Development.
* How Prophet can be used to make time series forecasts
* How to analyse trends and seasonal fluctuations using Prophet

**15**