Geethanjali College of Engineering and Technology

Cheeryal (V), Keesara (M),Medchal District – 501 301 (T.S)

(AUTONOMOUS)



Project on

Univariate Time Series Analysis For Weather Prediction Using Prophet Library With IBM Cloud

**Team Members: (CSE-05)**

T. Manvitha (19R11A05D9)

T. Lakshmi Prasanna (19R11A05E0)

V. Likhitha (19R11A05E4)

V. Nithin (20R15A0516)

N. Kirthana (19R11A05N6)

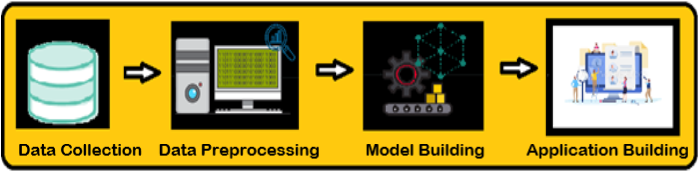
**INTRODUCTION**

* Any data associated with the time that is dependent on time-related matters can be termed as time-series data.
* Time series analysis comprises methods for analyzing time series data in order to extract meaningful statistics like trends, nonstationarity, and seasonality based on a daily, weekly, yearly basis and other characteristics of the data.
* Time series forecasting is the use of a model to predict future values based on previously observed values.
* There are many models present for the predictive analysis of time series like Machine learning ARIMA (Auto-Regressive Integrated Moving Average model), Auto-Regressive model, Exponential Smoothing, LSTM (Long Short Term Memory), etc. These models require the data to be fed and with certain tweaking and fine-tuning they help us to make predictions. .
* But, Facebook Prophet library is a third party library that could perform all the fine-tuning part within and we just need to feed the model.
* The main idea of our project is to create an application for forecasting the weather by using Facebook Prohphet library.

Purpose

* User interacts with the UI (User Interface) to give the date as input
* The date is analyzed by the model which is integrated
* Once model analyses the uploaded date, the prediction is showcased on the UI.

Block diagram



Software Requirements

In order to develop this project we need to install the following software/packages:

### Anaconda Navigator

### Jupyter notebook

### Spyder

### To build Machine learning models you must require the following packages:

### Numpy

### Fbprophet

### To make a responsive python script you must require the following packages:

### Requests

### Flask

### Need to download the required packages:

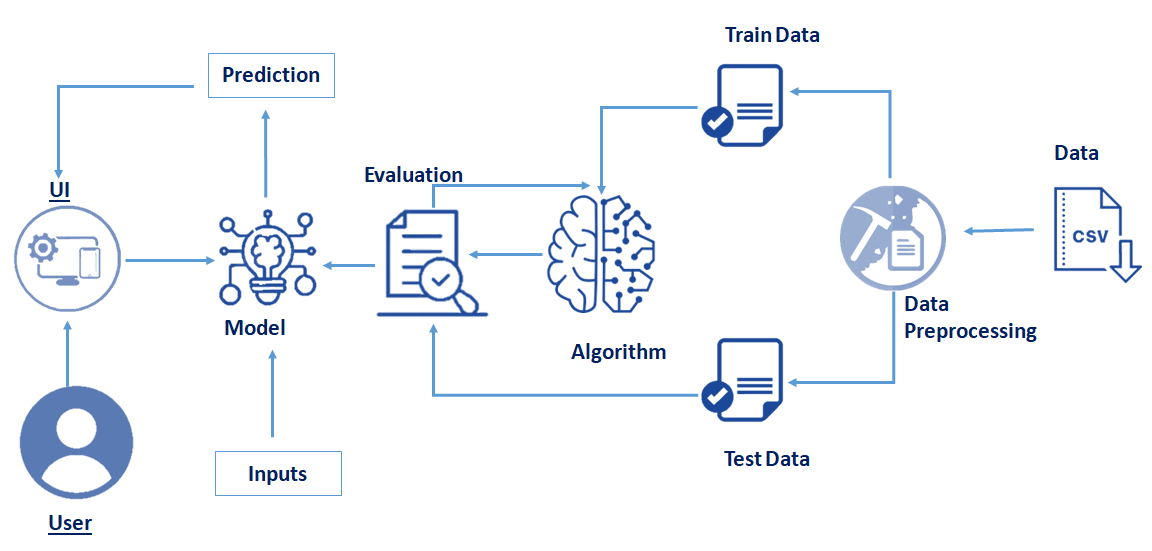
### pip install requests

### pip install Flask

### pip install pystan

### conda install -c  conda-forge fbprophet

**FLOW CHART**



To accomplish the application, we have to complete all the activities and tasks listed below.

**1.Data Collection:**

Collect the data or create the data.

**2.Data Preprocessing:**

Data Pre-processing includes the following main tasks

* Importing the required libraries

The first step is usually importing the libraries that will be needed in the program.

The required libraries to be imported to Python script are:

1. Numpy
2. Pandas
3. Matlplotlib
4. Fbprophet

* Importing the dataset

Load the csv data file into pandas using the **read\_csv() function.**

* Analyse the data

head() method is used to return top n (5 by default) rows of a DataFrame or series.

tail() method is used to return bottom n (5 by default) rows of a DataFrame or series.

* Resampling the dataset

Converting the datatype**:**

When a csv file is imported and a Data Frame is made, the Date time objects in the file are read as a string object rather than Date Time object and Hence it’s very tough to perform operations like Time difference on a string rather a Date Time object. Pandas to\_datetime() method helps to convert string Date time into Python Date time object.

Setting the index:

The set\_index() function is used to set the DataFrame index using existing columns.

Set the DataFrame index (row labels) using one or more existing columns or arrays of the correct length. The index can replace the existing index or expand on it. datetime\_utc has data and time.As we are concentrating on only date setting the datetime\_utc column as index and resampling it.

* Pre processing the data

For time series analysis,we are interested about two columns i.e., date(datetime\_utc) and temperature(\_tempm).So let us concentrate only on them.

* Taking care of Missing Data

1. After loading the dataset, it is important to check the complete information of such as null values in a column or a row

2.Check for the null values. if i present then the following tasks can be done,

a.Imputing data using Imputation method in sklearn

b.FillingNaN values with mean, median, and mode using fillna() method.

c.Delete the records

* Prophet Library naming convention

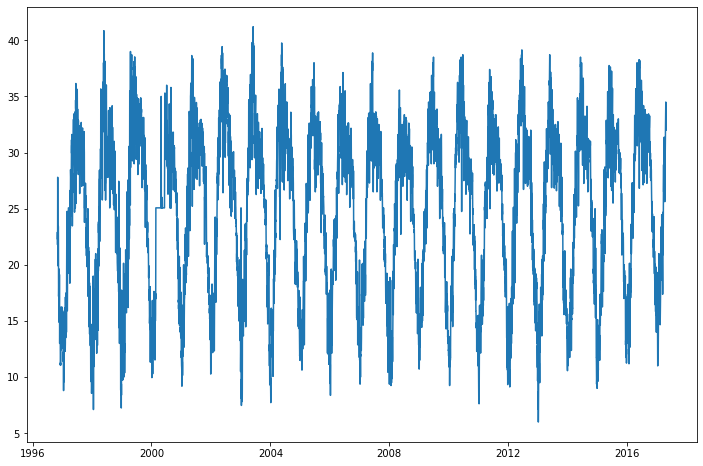
Pandas Series.reset\_index() function generate a new DataFrame or Series with the index reset.

For this univariate analysis Prophet expects the dataset to have two columns named as ds and y(proper name convention). ds is the date column while y is the column that we are forecasting. so let us rename them.

* Data visualization

A time-series dataset does not make sense to us until we plot it.Plotting a time series helps us actually see if there is a trend, a seasonal cycle, outliers, and more. It gives us a feel for the data.

We can plot the data easily by calling the plot() function on the DataFrame.



**3. Model Building:**

**Model Fitting:**

The API implementation for Prophet and Scikit-learn are very similar as we’ll see below. To use Prophet for forecasting, first, a Prophet() object is defined and configured, then it is fit on the dataset by calling the fit() function and passing the data.

The Prophet() object takes arguments to configure the type of model you want, such as the type of growth, the type of seasonality, and more. By default, the model will work hard to figure out almost everything automatically.

The fit() function takes a DataFrame of time series data. The DataFrame must have a specific format. The first column must have the name ‘ds‘ and contain the date-times. The second column must have the name ‘y‘ and contain the observations.

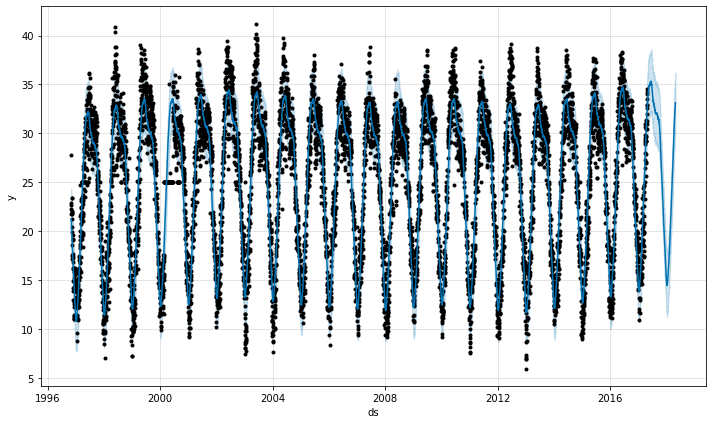
# **Making Future Predictions :**

# The next step is to prepare our model to make future predictions. This is achieved using the Prophet.make\_future\_dataframe method and passing the number of days we’d like to predict in the future. We use the periods attribute to specify this. This also include the historical dates.

**Obtaining the Forecasts :**

We use the predict method to make future predictions. This will generate a dataframe with a yhat column that will contain the predictions.

# **Plotting the Forecasts :**

****

# **Cross Validation:**The cross\_validation method allows us to compare the predicted values with the actual values in Prophet.

from fbprophet.diagnostics import cross\_validation

df\_cv = cross\_validation(model, initial='730 days', period='180 days', horizon = '365 days')

df\_cv.head()

# **Visualizing Performance Metrics**: The performance Metrics can be visualized using the plot\_cross\_validation\_metric utility. Let’s visualize the RMSE below.

# **Saving the model file :**

import pickle

pickle.dump( model, open( "weather\_prediction.pickle", "wb" ) )

**4.Application Building:**

**Frontend UI Design :**

Here we use html ndcss code to develop frontend UI and the code will be available in APPENDIX part of doc.

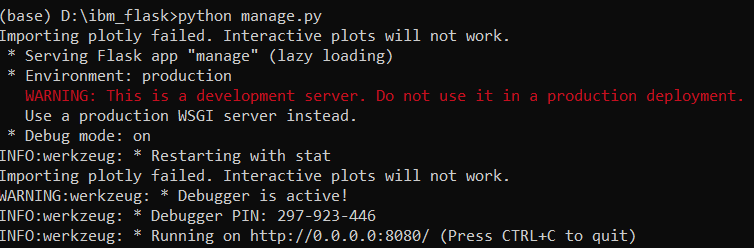
**Flask:**

Here we use the flask to integrate python script and html code . and here we load our model in the code.

**5.Run The App :**

Run The app in local browser

* Open anaconda prompt from the start menu
* Log on to the created virtual environment by using the command - “conda activate environment\_name”
* Navigate to the folder where your python script is.
* Now type “python app.py” command
* Navigate to the localhost where you can view your web page



### Now open your browser and navigate lo localhost:8080 where we can view our webpage

### RESULT :

### The output of the project :

### 

### Here I give some input date and clicked the submit then it showed the predicted temperature

### 

**Applications**

**.** The Time Series Analysis is applied for various purposes, such as:

* Stock Market Analysis

* Economic Forecasting

* Inventory studies

* Budgetary Analysis

* Census Analysis

* Yield Projection

* Sales Forecasting

**Conclusion**

Univariate Time Series Analysis is a complex and challenging science that depends on the efficient interplay of weather observation, data analysis by meteorologists and computers, and rapid communication systems. But it has enormous uses which makes it an essential operation.

**Future scope**

Time series analysis helps in analyzing the past, which comes in handy to forecast the future. The method is extensively employed in a financial and business forecast based on the historical pattern of data points collected over time and comparing it with the current trends.This is the biggest advantage used by organizations for decision making and policy planning by several organizations

**Bibliography**

Abraham, B. (1982). Temporal aggregation and time series, International Statistical Review,50, 285-291.

Abraham, B., and Box, G. E. P. (1979). Bayesian analysis of some outlier problems in time series,Biometrika, 66, 229-236.

Abraham, B., and Chuang, A. (1989). Outlier detection and time series modeling, Techitometrics,31, 241-248,

Abraham, B., and Ledolter, J. (1983). Statistical Methods for Forecasting, John Wiley, New York.

Abraham, B., and Ledolter, J. (1984). A note on inverse autocorrelations, Biometrika, 71, 609-614.

Abraham, B., and Wei, W. W, S. (1984). Inferences about the parameters of a time series model with changing variance, Metrika, 31, 183-194.

Agiaklogion, C., and Newbold, P. (1992), Empirical evidence on Dickey—Fuller-type tests, Journal of Time Series Analysis, 13, 471-483.

Absanullah, M., and Wei, W. W. 8. (1984a). Effects of temporal aggregation on forecasting in an

ARMA(I, 1) process. ASA Proceedings of Business and Economic Statistics Section, 297-302.

Ahsanullah, M., and Wei, W. W. S. (1984b). The effects of time aggregation on the AR(1) process,Computational Statistics Quarterly, 1, No. 4, 343-352.

**Appendix:**

**Github link:**

[**https://github.com/smartinternz02/SI-GuidedProject-2419-1622113188**](https://github.com/smartinternz02/SI-GuidedProject-2419-1622113188)

**THANK YOU**