

*Research Article***Time Series Prediction Based on Facebook Prophet: A Case Study, Temperature Forecasting in Myintkyina****Zar Zar Oo<sup>a</sup>** , **Sabai Phyu<sup>b</sup>** <sup>a,b</sup>Cloud Laboratory, University of Computer Studies, Yangon(UCSY), No.(4) Main Road, Shwe Pyi Thar Township, Yangon, Post code – 11411, Myanmar

## ARTICLE INFO

*Article history:*Received 27 October 2020  
Accepted 1 December 2020*Keywords:*Facebook Prophet Model,  
Time series data,  
Time series prediction.

## ABSTRACT

Temperature forecasting is a progressive and time series analysis process to forecast the state of the temperature for a certain location in coming time. Nowadays, agriculture and manufacturing sectors are mostly dependent on temperature so forecasting is important to be precise because temperature warnings can save life and property. In this work, the Prophet Forecasting Model is used for Myitkyina's annual temperature forecasting using historical (2010 to 2017) time series data. Myitkyina is the capital city of the northernmost state (Kachin) in Myanmar, located 1480 kilometers from Yangon. Prophet is a modular regression model for time series predictions with high accuracy by using simple interpretable parameters that consider the effect of custom seasonality and holidays. In this study, the temperature forecasting model is proposed by using weather dataset provided by an International institution, National Oceanic and Atmospheric Administration (NOAA). This work implements the multi-step univariate time series prediction model and compares the forecasted value against the actual data. Such findings check that the proposed forecasting model provides an efficient and accurate prediction for temperature in Myitkyina.

This is an open access article under the CC BY-SA 4.0 license.  
(<https://creativecommons.org/licenses/by-sa/4.0/>)

**1. Introduction**

A time series is a collection of data points fixed in time. Time series analysis is applied to analyze these time series data by combining different methods to obtain meaningful information. The forecasting of the time series data is a methodology which can help the model predict future values using previously observed historic values [1].

Nowadays, there are important time series problem to be solved such as how much inventory to maintain, how many people will travel by an airplane, how high the temperature will be in next month, how much the price of a tradable financial asset will be near tomorrow. For these problems, every data scientist should know the techniques for time series predicting. Therefore, time based patterns knowledge is very important for any areas. For an organization, Forecasting is an important data

science task required for many activities to be carried out. For weather forecasting, there are many various methods available and many researchers are interested in this because of its impact on living things [2]. Therefore, this paper aims to apply time series analysis for temperature forecasting of a city.

In such analysis of time series forecasting, two main points have been observed in forecasting: (1) Complete automated forecasting methods can be challenging and too inflexible to implement useful assumptions. (2) Due to the requirement of significant experience in data science skill, analysts cannot still forecast in high quality. There exists quite a variety of different ways to forecast future trends, such as, ARCH, ARIMA, artificial neural networks, regressive models,. Among them, the Prophet forecasting model is used in this work to predict the temperature of Myitkyina, Myanmar by dealing the common features of the weather data. For forecasting

time series data, an open-sourced Prophet is a model released by Facebook on 23 February 2017 [3]. In this work, the proposed temperature prediction system is built and it will provide the future temperature value for a city. Thus, it may provide meteorologists in predicting the future temperature value quickly and truthfully.

The remaining part of this paper is organized in the following way. Section 2 addresses earlier temperature prediction systems with multiple learning algorithms previously performed in literature. The most important section 3 outlines the proposed work of using Facebook Prophet to build an effective temperature forecasting method, explains the specifics of the time series based temperature forecasting technique and shows the results for experiments by plotting the performance of the proposed system, then, the paper finishes in Section 4.

## 2. Literature Review

### 2.1. A. Time Series Prediction

Fundamentally, the prediction objective of a time series is to estimate the value at time  $i$ ,  $y_i$  based on its previous historic data  $y_{i-1}, y_{i-2}, \dots$ . If the interested data is  $x = \{ y_{i-k}, y_{i-k+1}, \dots, y_{i-1} \}$ ,  $i = \{ k, \dots, n \}$ , the goal aims at finding a function  $f(x)$  so that  $\hat{y} = f(x)$  is as close to the ground truth  $y_i$  as possible. Some methods are one-step univariate forecasting, multi-step or sequence forecasting and multivariate forecasting, among them, the multi-step univariate is applied in the proposed system. Multivariate forecasting observes various measurements and predicts one or more of them. Time series forecasting is acceptable considering the nature of weather data, since temperature forecasting is temporary and time series operation.

### 2.2. Related Work

In these days, many previous studies have been performed on predicting temperature. Some of those are described below. Shaminder Singh, Pankaj Bhambri and Jasmeen Gill [4] implemented a time series dependent temperature prediction model by combining back propagation with the genetic algorithm that takes various population sizes. A sliding window of size 5 is used to get the moving average from the full dataset. After that, the dependent parameters are acquired and fed into the system as an input for the network training. An initial population of chromosomes is also generated randomly in network training, and the weights are extracted from each chromosome.

Dr. S. Santhosh Baboo and I.Kadar Shereef [5] described Back Propagation Neural network based algorithm for predicting the temperature. That proposed model can capture the complex relationships between many dependent factors that offer to assured temperature. It is confirmed with the real time dataset and compared with actually working of meteorological department.

Kuldeep Goswami and Arnab N. Patowary [6] developed a Seasonal Autoregressive Integrated Moving Average (SARIMA) model to forecast temperature on monthly and seasonal time scale. The analysis used long term temperature data of Dibrugarh, Assam for the period of fifty years (1966-2015). In the analysis, a seasonal ARIMA model for monthly minimum temperature data and a seasonal ARIMA model for monthly maximum temperature data are developed.

In order to analyze the monthly records of absolute surface temperature, a relevant environmental parameter, a generalized, structural, time series modeling system was developed using a deterministic combined stochastic (DSC) method by YE Liming and YANG Guixia [7]. Although their system development was focused on characterizing the variation patterns of a global dataset, the methodology could be applied to any absolute monthly temperature record. Y.Radhika and M.Shashi apply Support Vector Machines (SVMs) to predict the next day's maximum temperature for a given location based on time series data [8].

## 3. Prediction Model

The proposed temperature prediction system is to help meteorologist to better estimate weather in future for a specific location. Even though so many different kinds of time-series exist, Facebook prophet is used for this work. This is a Facebook library which works excellent because more metrics increasing the accuracy of the models are included.

### 3.1. Multi-step Forecasting Model

The main purpose of the real forecasting problems is to predict an interested data value ahead in time and some data values in a certain time forecast horizon  $k$ . The forecast horizon  $k$  is the region of time in the future and expected values for that should be prepared. This method of forecasting is called multi-step forecasting, which can be applied using two different techniques, unless the forecast horizon is one:

- the direct strategy: the model is trained explicitly to predict some several steps ahead
- the iterative method: iterates predictions up to the desired horizon by doing repeated one-step predictions

In this work, the direct strategy is used for the multi-step univariate forecasting model.

### 3.2. Multi-step Forecasting Model

As a classical time series, the next following set of data members are expected to draw only on a certain number of their immediate predecessors. Univariate forecasting problem is the forecasting problem which is comprised of one single series. When the historical data for some time series is presented as  $x_0, x_1, x_2, \dots, x_{n-1}, x_n$ .

As there is some functional dependency between historical and future time series data points, the required forecasted value  $x'_{n+1}, x'_{n+2}, \dots, x'_{n+k-1}, x'_{n+k}$  for the k forecast horizon is a function of the previous n data points. Therefore, this dependency can be described in the following:

$$x'_{n+1}, x'_{n+2}, \dots, x'_{n+k-1}, x'_{n+k} = f(x_0, x_1, x_2, \dots, x_{n-1}, x_n)$$

$f$  might be a machine learning method. And, as a prediction model, Facebook Prophet Model is used for this work.

**3.3. Facebook Prophet Model**

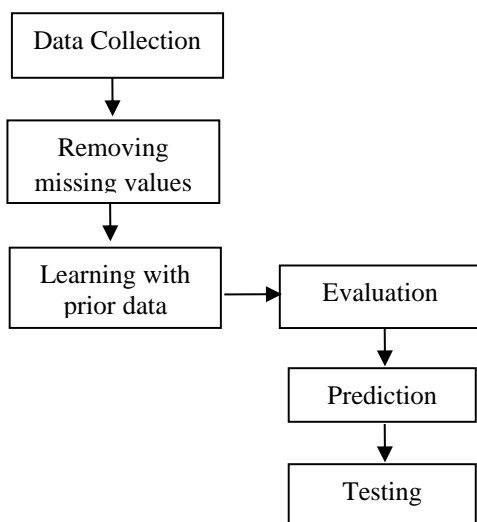
Prophet has key components: trends  $g(t)$ , seasonality  $s(t)$ , holidays  $h(t)$ . They are composed as

$$y(t)=g(t)+s(t)+h(t)+\varepsilon(t) , \text{ where } \varepsilon(t) \text{ is error term.}$$

There are a variety of functional benefits of this approach. Due to weekly and annual seasonality, the seasonal component  $s(t)$  provides a flexible model of periodic changes. The  $h(t)$  portion reflects predictable annual abnormal days including those on irregular schedules. The error term,  $\varepsilon(t)$  reflects information not expressed in the model. It is typically modeled as normally distributed noise [2].

**3.4. Methodology**

The proposed model collects time series weather data and selects the weather parameters to be predicted by extracting the relation between the different weather parameters. Missing data for some columns are replaced as a zero value. From the data, training data set which contains inputs and outputs and test data set with only inputs are created. To learn the historic data of the city weather conditions and to predict the future temperature value, the Prophet model is used.



**Figure 1.** Methodology for the Proposed Model

Prophet model has a great handling power for non-experienced data scientist to forecast in their practices. This model support data frame with only two columns

which are ds for datetime and y for values and must be numeric and this is the forecasting to get. After preparing the raw data for the model, the learning model is trained with the historical data. Due to the complexity of values, it is need to make it simpler using log function and exponential function to reverse it to its normal state. So, prophet style of ds and y features nomenclature must be obeyed. The core part of the model is creating a new data frame to save new predicted values and predicting the target values. After that, this work is validated by the actual data.

**Table 1.** The collected average temperature

| DATE      | TAVG | TMAX | TMIN |
|-----------|------|------|------|
| 1/1/2010  | 69   |      | 50   |
| 1/2/2010  | 67   |      |      |
| 1/3/2010  | 60   | 81   | 46   |
| 1/4/2010  | 62   | 81   | 45   |
| 1/6/2010  | 65   |      | 49   |
| 1/7/2010  | 66   | 85   | 48   |
| 1/8/2010  | 63   |      | 48   |
| 1/9/2010  | 65   |      | 47   |
| 1/10/2010 | 70   | 83   |      |

**3.4.1. Data Collection and Preprocessing**

The initial stages of the prediction model are data collection and preprocessing. Since only valid data generates accurate performance, the main stage is preprocessing the data. Daily average temperature in degree Fahrenheit ranging from January, 2010 till December, 2017 is accessed from the above stated weather stations. One problem that needed to be resolved with the time series datasets was the missing data. Therefore, weather station with the least amount of missing data was selected for this work. For this work, the 8 years average temperature data of Myintkyina, Myanmar from NOAA is used. The collected weather data is noisy data with few missing values and it is essential to handle these data. In the preprocessing stage of the data, every lost value is changed with 0. After that, data is ready for learning as listed in TABLE 1 which shows some data of the weather station at Latitude 25.3946° N, 97.3841° E. The collected average temperature and time stamps are used for the prediction model.

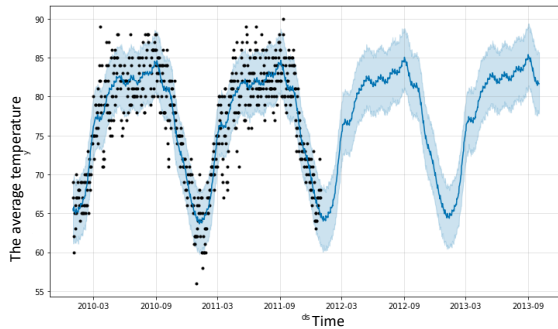
**3.4.2. Experimental Results**

The forecasting model is trained using the daily weather data of the years 2010-2017 and it forecasts the average temperature of one thousand days in next three years. Figure 2 indicates the predicted temperature value. To describe the prediction accuracy of the model, two plots are compared in Figure 3. These graphs are compared the real values and predicted values. In left plot, 2010 and 2011 for next 2 years is predicted. By comparing the predicted data (left plot) and real data for 2010 to 2013 (right plot), it can be clearly seen that the proposed temperature prediction model has the high

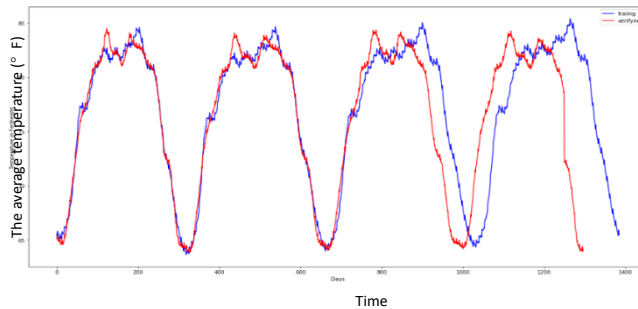
accuracy. Both the predicted value and the actual temperature value are compared on a graph in Figure 4 to see the variance value and the accuracy can be seen significantly in the last two years period. As the measurement of the prediction accuracy of the model, the Root Mean Square Error (RMSE) has been at 5.7573 for 2012 and 2013 years.

**4. Conclusion**

In this work, univariate time series prediction is made to forecast the temperature by learning the historic data which is inputted to the prediction model.



**Figure 3.** For 2012 and 2013, forecasted value vs. real value



**Figure 4.** Forecasted value vs. real values graph of the average temperature.

The result with the appropriate RMSE has shown strong benefits for weather forecasting. Moreover, the prediction results revealed that the model is satisfactorily fitted to the historical data. Through the developing of the proposed system, it is proved; Prophet Model is able to yield good results for temperature prediction and can be used as an alternative to conventional meteorological methods.

**Acknowledgment**

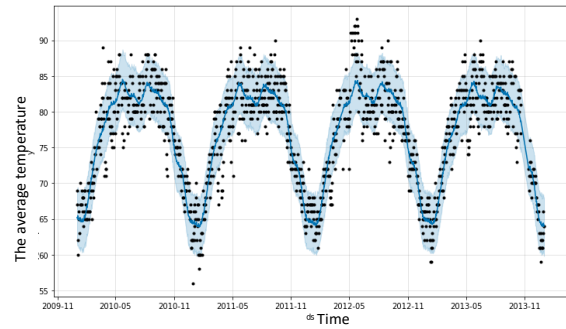
The writers recognize the research facilities of the University of Computer Studies, Yangon, Myanmar.

**Author's Note**

Abstract version of this paper was presented at 9th International Conference on Advanced Technologies (ICAT'20), 10-12 August 2020, Istanbul, Turkey with the



**Figure 2.** The predicted average temperature value (2018 to 2020). In the plotted graph, X-axis is time axis and Y-axis is the average temperature in °F.



title of “Time Series Prediction Based on Facebook Prophet A Case Study Temperature Forecasting in Myintkyina”.

**References**

- [1] R. Adhikari and R. K. Agrawal, “An Introductory Study on Time Series Modeling and Forecasting”, M. Tech. thesis, Jawaharlal Nehru University, New Delhi, India, 2013.
- [2] Sean J. Taylor and Benjamin Letham, “Forecasting at Scale”, September 2017.
- [3] <https://www.kaggle.com/armamut/predicting-transactions-fb-prophet-tutorial>.
- [4] Shaminder Singh, Pankaj Bhambri and Jasmeen Gill, “Time Series based Temperature Prediction using Back Propagation with Genetic Algorithm Technique”, IJCSI International Journal of Computer Science Issues, Vol. 8, Issue 5, No 3, September 2011, pp. 28-32.
- [5] Dr. S. Santhosh Baboo and I.Kadar Shereef, “An Efficient Weather Forecasting System using Artificial Neural Network”, International Journal of Environmental Science and Development, Vol. 1, No. 4, October 2010, pp. 321-326.
- [6] Kuldeep Goswami and Arnab N. Patowary, “Monthly

Temperature Prediction Based On ARIMA Model: A Case Study In Dibrugarh Station Of Assam, India”, International Journal of Advanced Research in Computer Science Volume 8, No. 8, September-October 2017, pp.292-298.

- [7] Y. Liming, Y. Guixia and E. V. Ranst, “Time-Series Modeling and Prediction of Global Monthly Absolute Temperature for Environmental Decision Making”, Advances in Atmospheric Sciences, Volume. 30, No. 2, 2013, pp.382–396.
- [8] Y.Radhika and M.Shashi, Atmospheric Temperature Prediction using Support Vector Machines, International Journal of Computer Theory and Engineering, Vol. 1, No. 1, April 2009, pp. 55-58.