

Time Series Weather Forecasting Techniques: Literature Survey

Janhavi Patil¹, Prof. Nirmala Shinde²

¹MTech Computer Engineering Student, KJ Somaiya College of Engineering, Mumbai, Maharashtra, India

²Professor, Dept. of Computer Engineering, KJ Somaiya College of Engineering, Mumbai, Maharashtra, India

Abstract - Forecasting is a part of statistical modelling that is widely used in several fields because of its benefits in decision-making. The purpose of forecasting is to predict the future values of certain variables that range with time using its previous values. Time Forecasting is related to the formation of models and methods that can be used to produce a good forecast. This research is survey paper research that used a systematic mapping study and systematic literature review. Generally, time series forecasting uses linear time series models, specifically the ARIMA model and LSTM that has long been used because it has good forecasting accuracy. The goal of this research is to review time series forecasting methods such as ARIMA, Prophet and LSTM and analyze the working of time series forecasting methods. It also discusses the approaches of different methods used in time series forecasting. Its goal is to increase the amount of awareness regarding time series forecasting and its methods.

Key Words: Time Series models, ARIMA, LSTM, Prophet, Accuracy, Forecasting

1. INTRODUCTION

Time series data forecasting is a part of statistical modelling that is widely used in various departments such as weather stations, Finance, banking, healthcare departments such as covid-19 data analysis because of its benefits in decision-making. Time series forecasting analysis has several objectives, namely, forecasting, modelling, and manage. Forecasting is an element that is important in managing activities because whether or not an effective decision is made depends on several factors that influence, although hidden, when a decision is taken.

The purpose of time series forecasting model is to predict the upcoming values of certain variables that range with time using its previous values. Forecasting is related to the generation of models and methods that can be used to construct a good forecast. In time series data, the doings of past events can be used for forecasting because it is expected that, in the future, the impact of the doings of past events will still occur. The advantages of forecasting can be felt in many fields, including production, marketing, economics and finance. Generally, time series research uses linear time series models, specifically the autoregressive integrated moving average (ARIMA) model, Prophet and LSTM.

2 Methodology

The method used structured mapping study and structured literature review conducted by recognizing and interpreting the clarifying in the literature review in accordance with the topic time series forecasting raised in this paper. The univariate time series made up of a single result over a time period. The multivariate time series made up of more than one result collected over time. Multivariate time series analysis examination is more challenging compared to univariate time series analysis.

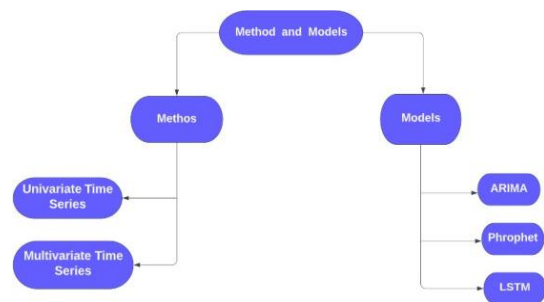


Fig -1: Time Series forecasting Methods and Models

Literature review related to the use and development of time series forecasting models from various studies in various departments then calculated to find timelessness and the latest developments from each method used.

3 Forecasting Model

The research methodology was studied to assess the accuracy of different types of time series models for rainfall, covid data, real-estate, Bit-coin forecasting. Initially, a comprehensive literature survey was carried out to study related research conducted to identify the techniques, datasets and observations of the different methodologies implemented worldwide. Often used time series forecasting models were identified from the literature survey and the models were developed to forecast the rainfall Bit-coin forecasting, covid-19 data, real-estate.

3.1 Forecasting with ARIMA

The first phase in applying ARIMA model is to check whether the time series is stationary or not. Autoregressive Integrated Moving Average works best when data has a fixed design

overtime, meaning that the variance and mean of the data have to remain constant overtime. Thus, when the data has a trend of going upwards or downwards and has a particular pattern (seasonality), then the data is not stationary.[1]

Three parameters constitute the Autoregressive Integrated Moving Average model (AR, I, MA) that have essence on the model accuracy: (p, d, q) indicate the autoregressive, difference order, and moving average window size, respectively. To identify these parameters, we first apply differencing lag-1 for a moving trend or seasonal differencing, then we fit the ARIMA model to the various series. The classic method for determining these parameters is using perceptible examinations of the time series to detect trends as well as looking at the correlation and partial correlation charts.

The forecasting equation is constructed as follows. First, y_t denotes the d th difference of Y , which describes:

$$\text{If } d=0: y_t = Y_t \tag{1}$$

$$\text{If } d=1: y_t = Y_t - Y_{t-1} \tag{2}$$

$$\text{If } d=2: y_t = (Y_t - Y_{t-1}) - (Y_{t-1} - Y_{t-2}) = Y_t - 2Y_{t-1} + Y_{t-2} \tag{3}$$

The general forecasting model equation is, In terms of y is as follows:

$$\hat{y}_t = \mu + \phi_1 y_{t-1} + \dots + \phi_p y_{t-p} - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \tag{4}$$

Frequently the parameters are denoted there by AR(1), AR(2) and MA(1), MA(2), etc. To identify the appropriate Autoregressive Integrated Moving Average model for Y , we start by identifying the order of differencing needing to stationarise the series and cancel the gross features of seasonality [5]

3.2 Forecasting with Prophet

PROPHET is a software which is open-source that is available in R and Python for forecasting time series model data. PROPHET is published by Facebook's team of Core Data Science. It is based on a contribution of model where trends are nonlinear are fit with yearly and weekly holidays and plus seasonality. PROPHET is well built to missing data, capturing the shifts in the trend and large outliers. In summation, it gets a suitable estimate of the mixed data without spending manual work [7]. PROPHET is effective for business forecast that are observed on Facebook. For example, time, weekly, daily observations of history, within a year, missing observation, trend changes, large outliers and trends that are non-linear growth curves [7]

Prophet parameters consist of changepoints, capacities, seasonality, holiday and smoothing parameters that can be interpretably applied to improve the model performance.[8]

Prophet can be considered a nonlinear regression model, of the form:

$$y_t = g(t) + s(t) + h(t) + \epsilon_t, \quad y_t = g(t) + s(t) + h(t) + \epsilon_t, \tag{1}$$

describes a piecewise-linear trend (or "growth term"), $s(t)$ describes the various seasonal patterns, $h(t)$ captures the holiday effects, and ϵ_t is a white noise error term.

3.3 Forecasting with LSTM

Extension of recurrent neural network are Long Short-Term Memory (LSTM), which essentially deepen their memory. Along these lines it is significant to acquire from vital encounters that have long situations slacks in the middle. The units of Long-Short Term Memory networks are used as building units for the layers of a recurrent neural network, which is then often known as a Long Short-Term Memory network. Long Short-Term Memory allow recurrent neural networks to recall their inputs over a long period of time. This is because recurrent neural network contains their data in memory that is much similar to the memory of a computer in the light of the fact LSTM can erase, read and write data from its memory. [9]

The cell situation in LSTM helps the knowledge to flow through the units without being modified by allowing only a few linear exchanges. Each unit has an input, output and a forget gate which can count or eliminate the information to the cell state. The forget gate chooses which information from the previous cell state should be obliterated for which it uses a sigmoid function. The input gate rules the information flow to the current cell state using a point-wise multiplication operation of 'tanh' and 'sigmoid' respectively. Finally, the output gate determines which information should be passed on to the next unrevealed state

4 Literature Survey

A list of research areas in time series forecasting using ARIMA, Prophet and LSTM. The summary of methods used in literature is given in following table.

Table -1: Literature Survey

| Reference | Method | Observation | Dataset |
|--|-----------------------------|---|---|
| Satrio, Christophorus Beneditto Aditya, et al.2021 [1] | ARIMA and PROPHET | PROPHET has good accuracy in predicting the confirmed cases with 91% precision, while ARIMA did not even pass through half precision. Both models also have negative R2 values. | The dataset of Covid- 19 consists of 27618 rows and 8 columns. |
| Alghamdi, Taghreed, et al.2019 [2] | ARIMA | ARIMA model achieves excellent performance with a confidence level at 95%. | Dataset contains 2175 observations that been obtained in different 3 months and 13 attributes with different data types |
| Raymond, Y. C. 1997 [3] | ARIMA | Results strongly show that there exist cyclical trends in the office and industrial property prices in Hong Kong. | Real-estate prices |
| Geetha, A., and G. M. Nasira 2016 [4] | ARIMA, Statistical measures | Results obtained through this model are well acceptable with the prediction accuracy range of 80%. | Rainfall of a coastal region, five-year dataset (2009-2013) consisting of temperature dew point |
| Yenidoğan, Işil, et al. 2018 [6] | Prophet | While the PROPHET model makes predictions quite close to reality, that is up to 94.5% precision, the ARIMA model provided only 68% precision. | The dataset selected for this study starts from May 2016 and ends in March 2018, Bitcoin value |

| Reference | Method | Observation | Dataset |
|---|---------------------------|--|---|
| Salehin, Imrus, et al. 2020 [7] | SARIMA , Facebook Prophet | The result indicates that Facebook Prophet,with the lowest Mean Squared Error (MSE) and Root Mean Squared Error (RMSE), is the fittest model to predict the monthly rainfall in Central Jakarta | The datasets used in the study consist of rainfall time series datasets from January 1st,2008, to August 31, 2020 |
| Salehin, Imrus, et al. 2020 [8] | LSTM, Neural Network | They got 76% accuracy in LSTMmodel | Rainfall Dataset in Dhaka city from 2000 to 2014 |
| Fente, Dires Negash, and Dheeraj Kumar Singh 2018 [9] | LSTM | In this model the data is trained using LSTM algorithm. From experimental result, it is observed that Long-Short Term Memory neural network gives substantial results with high accuracy among the other weather forecasting techniques. | The historical weather dataset is taken from national climate datacentre (NCDC) from November 2007 to October 2017.The dataset contain many weather attributes like temperature , humidity, dew point, pressure |

5 Conclusion

This paper has discussed different Time Series Weather Prediction techniques like ARIMA, Prophet, LSTM used to develop data-driven weather forecasting systems. While there are many promising DL applications in other fields and numerous attempts were made to replace the widely used Numerical Weather Prediction (NWP), there are specific properties of weather data that require the development of new approaches beyond the classical. New concepts derived from computer vision, speech recognition, or DL solutions for many of these issues are still under development.

REFERENCES

- [1] Satrio, Christophorus Beneditto Aditya, et al. "Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET." *Procedia Computer Science* 179 (2021): 524-532.
- [2] Alghamdi, Taghreed, et al. "Forecasting traffic congestion using ARIMA modeling." 2019 15th International Wireless Communications & Mobile Computing Conference (IWCMC). IEEE, 2019.
- [3] Raymond, Y. C. "An application of the ARIMA model to real-estate prices in Hong Kong." *Journal of Property Finance* (1997).
- [4] Geetha, A., and G. M. Nasira. "Time-series modelling and forecasting: Modelling of rainfall prediction using ARIMA model." *International Journal of Society Systems Science* 8.4 (2016): 361-372.
- [5] Duke "<https://people.duke.edu/~rnau/411arim.html>"
- [6] Yenidoğan, Işil, et al. "Bitcoin forecasting using ARIMA and PROPHET." 2018 3rd international conference on computer science and engineering (UBMK). IEEE, 2018.
- [7] Sulasikin, Andi, et al. "Monthly Rainfall Prediction Using the Facebook Prophet Model for Flood Mitigation in Central Jakarta." 2021 International Conference on ICT for Smart Society (ICISS). IEEE, 2021.
- [8] Salehin, Imrus, et al. "An artificial intelligence-based rainfall prediction using LSTM and neural network." 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-ECE). IEEE, 2020.
- [9] Fente, Dires Negash, and Dheeraj Kumar Singh. "Weather forecasting using artificial neural network." 2018 second international conference on inventive communication and computational technologies (ICICCT). IEEE, 2018.