STM JOURNALS

International Journal of
Information Security Engineering

Volume 1, Issue 1, 2023
January—June
DOI (Journal): 10.37591/IJISE

https://journals.stmjournals.com/ijise

Review IJISE

Time Series Forecasting Based on PyAF and fbProphet

Shailesh Yadav[1,*], Vikas Yadav[2]

Abstract

Time series forecasting is the technique of predicting future events using previous data. Time series data includes information that is collected and recorded at regular intervals, such as daily stock prices, monthly sales figures, or hourly temperature readings. The purpose of time series forecasting is to use previous data to create accurate forecasts about the future values of a given variable. This can be beneficial for a range of applications, including financial forecasting, weather forecasting, and energy demand forecasting. Time series forecasting is an important tool in many fields, including business, economics, finance, and engineering. It can help one make informed judgements about things like production planning, inventory management, and investment strategies. It is also employed in a variety of domains, including meteorology, epidemiology, and ecology. Its applications are far-reaching and contribute to making well-informed decisions in areas like production planning, inventory management, and investment strategies. Moreover, this technique finds extensive use in diverse fields like meteorology, epidemiology, and ecology, where it aids in predicting weather patterns, analyzing disease trends, and understanding ecological systems. By utilizing time series forecasting, professionals in different domains can harness the power of historical data to gain insights and make informed decisions.

Keywords: Structured, big data, sortable, data processing, unstructured, semi-structured.

INTRODUCTION

This research paper explores the use of the PyAF (Python Automatic Forecasting) and fbProphet (Facebook Prophet) libraries for time series forecasting. The paper presents an introduction to the concepts of time series forecasting, the specific characteristics and features of the PyAF and fbProphet libraries, and a comparison of their performance on various time series datasets. The introduction section of the paper probably provides an overview of the importance of time series forecasting in various fields, such as business, economics, and engineering. It also provides a brief overview of the different techniques and algorithms that are commonly used for time series forecasting, such as moving averages, exponential smoothing, and ARIMA (autoregressive integrated moving average) [1].

This research paper delves into the utilization of two libraries, PyAF and fbProphet, for time series forecasting. This paper aims to introduce the fundamental concepts of time series forecasting, highlight the distinctive characteristics and features of the PyAF and fbProphet libraries, and compare their performance using diverse time series datasets. Additionally, the introduction section of the paper provides a comprehensive overview of the significance of time series forecasting in various fields, including business, economics, and engineering [2].

***Author for Correspondence**
Shailesh Yadav
E-mail: shaileshniraj56764@gmail.com

[1,2]Research Scholar, MCA Thakur Institute of Management Studies, Career Development & Research (TIMSCDR) Mumbai, India

Received Date: May 12, 2023
Accepted Date: July 04, 2023
Published Date: July 07, 2023

Citation: Shailesh Yadav, Vikas Yadav. Time Series Forecasting Based on PyAF and fbProphet. International Journal of Information Security Engineering. 2023; 1(1): 32–36p.

The fbProphet library, developed by Facebook's Core Data Science team, has gained popularity for its ease of use and excellent forecasting capabilities. It is specifically designed for time series analysis

with an emphasis on capturing seasonality, trend changes, and holiday effects. The library provides a straightforward interface for model fitting, parameter estimation, and prediction generation. Additionally, it incorporates Bayesian modeling techniques to account for uncertainty in the forecasting process [3].

The PyAF library is a powerful open-source tool that offers a wide range of functionalities for time series forecasting. It is designed to automate the process of building accurate and robust forecasting models by incorporating advanced statistical algorithms and machine learning techniques. PyAF provides various features, such as automatic selection of model components, handling of multiple seasonalities, and support for incorporating exogenous variables. These features contribute to the library's flexibility and effectiveness in handling diverse time series datasets. The main focus of the paper is on the PyAF and fbProphet libraries. The authors present an overview of the libraries, their features, and the specific techniques they use for time series forecasting. They also provide a comparison of the performance of the two libraries on various time series datasets, evaluating the accuracy and efficiency of the forecasts generated by each library [3].

The research paper presents a detailed comparison of the performance of PyAF and fbProphet on various time series datasets. This comparison might include metrics such as forecasting accuracy, computational efficiency, and the ability to handle different types of time series patterns. By evaluating these aspects, the authors provide insights into the strengths and limitations of each library and offer recommendations for selecting the appropriate tool based on the specific requirements of the forecasting task [4].

In conclusion, this research paper explores the application of the PyAF and fbProphet libraries in time series forecasting. It introduces the concepts of time series forecasting, highlights the importance of this technique in different fields, and provides an overview of the features of PyAF and fbProphet libraries. Furthermore, the paper includes a comprehensive comparison of the performance of these libraries on diverse time series datasets. By examining their capabilities, researchers and practitioners can gain valuable insights into selecting the most suitable library for their forecasting needs [5].

The conclusion section of the paper summarizes the main findings of the study and provides recommendations for future research. The authorsalso suggest which library is more suitable for specific use cases or applications. Overall, the research paper aims to provide a comprehensive analysis of the performance of PyAF and fbProphet libraries for time series forecasting and provide a guidance for practitioners and researchers who are working in the field of time series forecasting.

LITERATURE REVIEW

Before building a model, it is important to understand the characteristics of the time series data, such as the presence of trends, seasonality, and outliers. This can be done by visualizing the data and using statistical tests. The model used will be determined by the properties of the time series data [6].

For example, if the data has a clear seasonal component, a model such as seasonal ARIMA (SARIMA) would be a good choice. Many time series forecasting methods assume that the data is stationary, which means that the series' statistical features do not change over time. If the data is non-stationary, it may be necessary to differentiate the series or use a more complex model such as an ARIMA model. When evaluating the performance of a time series model, it is critical to utilize proper assessment measures such as mean absolute error (MAE), mean absolute percentage error (MAPE), and root mean squared error (RMSE). Time series data can change over time, so it's important to monitor the model’s performance and retrain it as necessary to account for changes in the data. When working with a specific use case, it is important to consider the specific requirements and constraints of the problem and use an approach that is tailored to it. For example, if one is working with a problem with a high amount of missing data, it is important to use an approach that is robust to missing data. When

working with a specific use case, it is important to consider the specific requirements and constraints of the problem and use an approach that is tailored to it. There are various existing models/algorithms which work on time-series forecasting but they are not useful in every scenario/dataset, one of them is AR (autoregressive) [7].

AR models are a type of time series forecasting model in which future values of a time series are assumed to be a linear function of its past values. They are particularly useful for modeling time series data with a strong dependence on past values, such as stock prices, weather data, and electricity consumption.

An AR model is a time series model that forecasts future values based on previous values of the same variable. It assumes that the value at a given time point is a linear combination of its previous values, with an added error term. The word "autoregressive" refers to the model's ability to regress the variable on itself.

AR models are widely used in various fields, including finance, economics, and meteorology, for time series analysis and prediction. They provide a versatile and interpretable framework for capturing data's temporal patterns and dependencies. AR models, on the other hand, assume stationarity (i.e., the statistical features of the series do not change over time), which may limit their applicability to non-stationary series. Nonetheless, AR models serve as valuable tools for understanding and forecasting time-dependent data [8].

An AR model works on the premise that the present value of a time series is a function of its previous values as well as some random noise. More formally, an AR(p) model is defined as:

$$y_t = c + a_1 * y_{(t-1)} + a_2 * y_{(t-2)} + ... + a_p * y_{(t-p)} + e_t$$

where y_t is the current value of the time series, y_(t-1), y_(t-2), ..., y_(t-p) are the past p values of the time series, c is a constant, a_1, a_2, ..., a_p are the coefficients of the model, and e_t is the random noise.

There are some models/algorithms which can be used in every scenario/dataset like PyAF and fbProphet.

PyAF provides a simple and consistent interface to a wide range of forecasting methods, which makes it easy to experiment with different models and compare their performance. Additionally, it also provides several advanced features such as feature extraction, which can help improve the performance of the model. PyAF is a powerful and flexible library for time series forecasting that can be very effective in certain use cases. However, it is important to evaluate its performance on a case-by-case basis, considering the specific data and use case.

Facebook Prophet (fbProphet) is an open-source time series forecasting library developed by Facebook. It is built on top of the PyStan library and is designed to make it easy to create high-quality forecasts for business and other time series data. The performance of Prophet will depend on several factors, including the specific data and use case, the quality of the input data, and the parameters used to configure the model. The library requires very little time-series-specific knowledge and can provide high-quality forecasts with minimal user input.

The core concept of Prophet is to model time series as a combination of three main components: trend, seasonality, and holidays. The trend component captures the overall direction of the time series data, allowing for both linear and non-linear patterns. Prophet automatically detects and adapts to changes in trend over time, making it suitable for datasets with fluctuating patterns [9].

The seasonality component accounts for periodic patterns, such as daily, weekly, monthly, or yearly fluctuations. Prophet utilizes a Fourier series to model these repetitive patterns, providing flexibility in capturing different types of seasonality. It can handle multiple seasonality's simultaneously and automatically identifies the most significant ones based on the data.

The holiday component in Prophet is used to model the impact of holidays or important events on the time series. It allows users to specify custom holiday dates and associated effects. Prophet automatically detects recurring holidays and includes them in the model, helping to capture their influence on the forecast.

Prophet incorporates a Bayesian approach to handle uncertainty in the forecasting process. It models the error term as a combination of trend and seasonality residuals, allowing for more accurate estimation of forecast uncertainty. Prophet generates uncertainty intervals around the point forecasts, providing valuable insights into the range of possible outcomes.

Prophet's user-friendly design and easy workflow are two of its standout qualities. It provides a straightforward API (application programming interface) for data input and offers various options for customization, including the ability to specify seasonality and holiday effects. The library is implemented in Python and supports both interactive use and integration into production systems.

Prophet's scalability is another advantage, as it can efficiently handle large-scale time series data. It utilizes a distributed computing framework that allows for parallelized model fitting and forecasting, enabling fast and efficient processing even with extensive datasets.

CHARACTERISTICS OF BIG DATA

Choosing the right model: Both PyAF and Prophet provide a wide range of forecasting methods, and choosing the right model can be difficult if you are not familiar with the specific characteristics of the dataset and the problem you are trying to solve. *Data preprocessing:* Time series data can be messy, and it is important to perform the necessary preprocessing steps such as handling missing values, outliers, and transforming the data to make it suitable for modeling. *Model tuning:* Both libraries require the user to set various parameters and options to configure the models. Finding the perfect collection of characteristics that will provide the best performance might be difficult. Nonstationary data handling: PyAF and Prophet assume that the data is stationary, which means that the statistical features of the series do not change over time. If the data is non-stationary, it may be necessary to differentiate the series or use a more complex model such as an ARIMA model. *Handling external factors:* Time series data is often influenced by external factors such as holidays, events, or economic indicators. It is important to take these factors into account when building a model and make sure the model is able to account for them. *Handling high-frequency data:* Both PyAF and Prophet are designed for handling low-frequency data, such as daily or weekly data. Handling high frequency data such as minute or second data can be challenging and may require specialized libraries or approaches. Implementing PyAF and Prophet for time series forecasting can be challenging, but by understanding the specific characteristics of the dataset and problem, choosing the right model, preprocessing the data, and carefully tuning the model parameters, one can achieve good performance. It is also important to keep in mind the specific requirements of one's use case and to be aware of the limitations of the libraries.an organization to offload infrequently accessed data.

CONCLUSION

In summary, both PyAF and Prophet are powerful libraries for time series forecasting, and the choice between them will depend on the specific characteristics of the dataset and the problem one is trying to solve. If one is working with a problem that has a strong seasonal and/or trend component and one does not have specific requirements, Prophet might be an easier and faster option to work with. Otherwise, PyAF could be a good choice for its flexibility and wide range of forecasting methods.

REFERENCES

1. Taylor SJ, Letham B. Business Time Series Forecasting at Scale. PeerJ Preprints 5: E3190v2, 35 (8), 48–90.
2. Murphy JE, Wo JY, Ryan DP, Clark JW, Jiang W, Yeap BY, Drapek LC, Ly L, Baglini CV, Blaszkowsky LS, Ferrone CR. Total neoadjuvant therapy with FOLFIRINOX in combination with losartan followed by chemoradiotherapy for locally advanced pancreatic cancer: a phase 2 clinical trial. JAMA oncology. 2019 Jul 1;5(7):1020–7.
3. Rafferty G. Forecasting Time Series Data with Facebook Prophet: Build, improve, and optimize time series forecasting models using the advanced forecasting tool. Packt Publishing Ltd; 2021 Mar 12.
4. Moffitt C. Forecasting Website Traffic Using Facebook's Prophet Library - Practical Business Python. Pbpython.com. 2017. Available from: https://pbpython.com/prophet-overview.html
5. Prusti D, Tripathy AK, Sahu R, Rath SK. Bitcoin Price Prediction by Applying Machine Learning Approaches. In Advances in Distributed Computing and Machine Learning: Proceedings of ICADCML 2023 2023 Jun 28 (pp. 305-319). Singapore: Springer Nature Singapore.
6. Hyndman RJ, Koehler AB, Ord JK, Snyder RD. Forecasting with Exponential Smoothing: The State Space Approach. Berlin, Germany: Springer; 2008.
7. Lim B, Zohren S. Time-series forecasting with deep learning: a survey. Philosophical Transactions of the Royal Society A. 2021 Apr 5;379(2194):20200209.
8. Divina F, Garcia Torres M, Gomez Vela FA, Vazquez Noguera JL. A comparative study of time series forecasting methods for short term electric energy consumption prediction in smart buildings. Energies. 2019 May 20;12(10):1934.
9. Jin X. Modern machine learning in time series forecasting. University of California, Santa Barbara; 2022.

