p–ISSN: 2723 - 6609 e-ISSN: 2745-5254
Vol. 5, No. 01 January 2024 http://jist.publikasiindonesia.id/

# DASHBOARD USING R SHINY FOR ARIMA AND PROPHET COMPARISON

**Christopher Caleb**
Universitas Tarumanagara Jakarta, Indonesia
Email: christopher.825200077@stu.untar.ac.id

*Correspondence

**Keywords:** Solusi Dokumen, Forecasting, ARIMA, Prophet.

**ABSTRACT**

As technology progresses, more and more business processes are digitised. This research was made to determine how the document industry is affected and make an educated guess on revenue forecasts based on previous data. Data will be taken daily for five years, from January 2017 to December 2022. Revenue data will be forecasted using two methods: ARIMA and Prophet. Data will be plotted on a graph in a user-authenticated dashboard made with R Shiny. The results of each forecast will be error-tested with MAE and RMSE. The results show that Prophet consistently produces a smaller number in both tests, which shows that Prophet is the more accurate method.

## Introduction

A significant boost in office digitalisation, mainly documentation, occurred with the COVID-19 virus pandemic, which WHO declared as a public health emergency in January 2020 the impact of COVID-19 on the workforce at various levels (Ella & Andari, 2022). Industry 4.0 ideas, solutions, and digital transformation have become a panacea for many affected by the pandemic. By moving towards digitisation, companies are reducing the need for paper and providing access to information to all employees without the costs associated with printing and storing physical paper documents (Jamaludin et al., 2022).

This research is a case study of a company that provides document solutions. Document solutions provide transformation products and services, from hardware-based service providers to solution-based services that cover all aspects of the document lifecycle, from input (creating, scanning, merging, editing, capturing), digital document management (sharing, indexing, storing, archiving, distributing), to document output (printing, faxing, scanning, copying, emailing, web-viewing) (Aksenta et al., 2023).

In all areas of business, planning and decision-making are necessary. For effective implementation, it is necessary to carry out proper analysis. One of the fruits of a mature analysis is forecasting with a solid foundation (Sahrudin, 2016). Forecasting covers many fields, including business and industry, government, economics, environmental, medicine, social sciences, politics, and finance. Predictions can be made for the next few years or even just a few minutes. Some things are more accessible to predict than others. The accuracy of predictions depends on several factors, including:

1. how deep the understanding is related to the factors involved;
2. how much data there is;
3. and whether the prediction targets will be affected by making predictions.

Most prediction cases involve the use of time-series data. A time series is a chronological sequence of observations of variables (MANOGARI, 2019). Many business forecasting applications use daily, weekly, monthly, and yearly data but are not limited to those time vulnerabilities (Aisah, Zaqiah, & Supiana, 2021). Predictions for future events are critical information in various planning and decision-making processes.

This study aims to compare the performance of several forecasting models for business forecasting applications. The first model is ARIMA, or autoregressive integrated moving average, which combines autoregressive (AR) and moving average (MA) processes and builds a composite model of the observed time series (Lince, 2022). The second method is Prophet, a model created by Facebook. Prophet is optimised for business forecasting observed by Facebook, such as time, daily, weekly, past data observations, data per year, outliers, trend changes, missing observations, and non-linear trends.

The study aimed to produce a simple dashboard that displays graphs of recorded data and data prediction of the results of two methods: ARIMA and Prophet Facebook. The dashboard will be designed using the R language with Shiny components to create dashboards**.**

Business Intelligence (BI) can be defined as a set of techniques and tools used to acquire and transform raw data into meaningful and valuable information for business analysis purposes (Yahya, 2023). The difference between BI and BA (business analytics), according to Thomas Davenport, professor of information technology and management at Babson College, is that BI is divided into querying, reporting, online analytical processing (OLAP), and BA focuses on statistics, prediction, and optimisation rather than reporting. Business Intelligence and Analytics is a term that combines both concepts (Maziyyah, 2022). Some analytical types include**:**

1. Decision analytics (Decision Analytics): assists in decision-making by humans with visual analytics of user models to reflect thinking
2. Descriptive Analytics: gain insights from past data with reports, clustering, etc.
3. Predictive Analytics: using predictive models using statistical and machine learning techniques
4. Prescriptive Analytics: recommend decisions with optimisation, simulation, etc.

## Research Methods

### Data Preprocessing and Application Planning

The data used in the test was taken from Document Solution's company database for 81 months every day from January 2017 to December 2022. The data retrieved has four attributes: month, COGS, MOP, and Revenue. Each attribute is individual, and this study focuses on univariate predictions, specifically on revenue figures.

The dashboard application will be designed using R Shiny, with R version 4.3.2. R is a programming language commonly used in data analysis, and in this study, it was developed into a dashboard**.**

## Database Planning

Figure 1 shows the star schema of the data warehouse used for dashboard design.

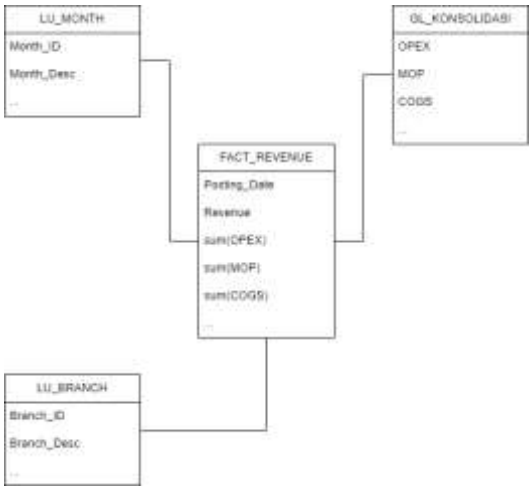


**Figure 1 Star Schema Dashboard**

## Use Case Diagram

Use case diagrams are created to plan what each actor can do in the use of the application or system to be created. Figure 2 shows the administrator use case diagram, and Figure 3 shows the user use case.

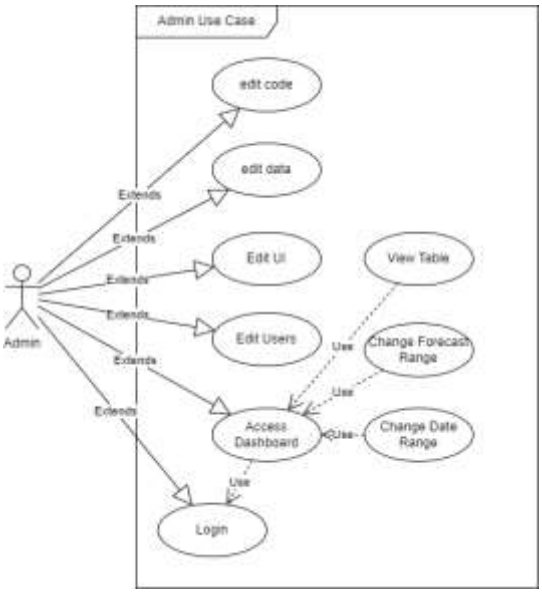


**Figure 2 Administrator Use Case Diagram**

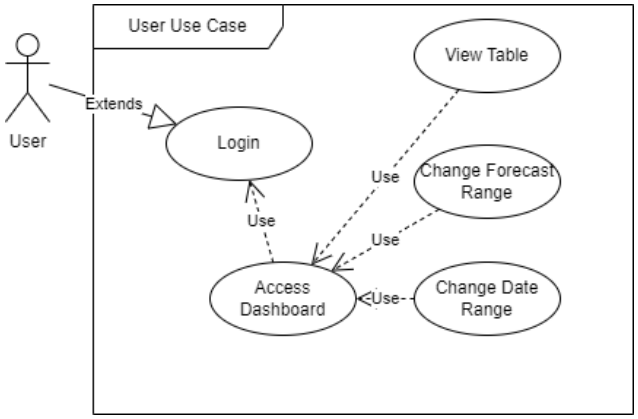


**Figure 3 Administrator Use Case Diagram**

## User Interface

The user must obtain a username and password from the administrator, and the user cannot change the data. Figure 4 shows the first dashboard page, which shows the KPI. Each value can be adjusted by the date. The amount of data will also have an impact on the test results. The sidebar shows several tabs that show the pages present in the application. There are four pages: the KPI page, the Forecast page shown in Figure 5, the Comparison page in Figure 6, and the data table page. Below the tab, there is a section that accepts user input for the number of times to be predicted by each method. By default, it is set to 30 days. Below the data graph, there are MAE and RMSE calculations, which will change based on the distance of the first and last data and the number of days predicted**.**


**Figure 4 KPI Dashboard page**

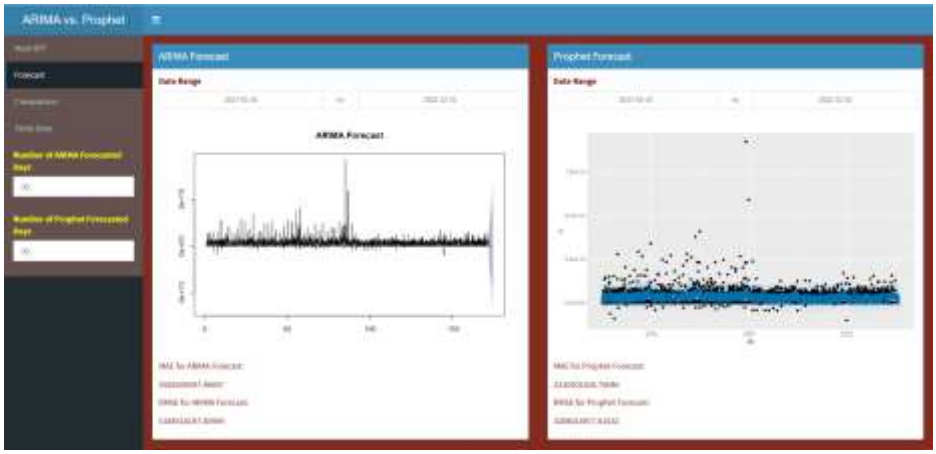


**Figure 5 Forecast Dashboard page**


**Figure 6 Comparison Dashboard page**

## Results and Discussion

Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) are basic metrics used for error testing and assessing the accuracy of predictive models. The MAE measures the average absolute difference between the predicted value and the actual value, thus directly indicating the overall error of the model (Firdaus, Putra, Arifandi, Anam, & Lathifah, 2023). It offers a simple, easy-to-interpret measure of accuracy, where smaller MAE values indicate better predictive performance. On the other hand, the RMSE extends this concept by taking the square root of the mean squared difference between the predicted value and the actual value. The RMSE penalises more significant errors, making them sensitive to outliers and emphasising the importance of minimising significant errors (PURWANTI, Rochim, & Warsito, 2022). MAE and RMSE serve as valuable tools for model evaluation, allowing data scientists and analysts to measure and compare the performance of various models, helping them make informed decisions about the suitability of predictive models for specific tasks.

After conducting a thorough analysis and comparison of forecasting models, it can be concluded that the Prophet's forecasting model outperformed ARIMA in terms of accuracy, as evidenced by the lower Root Mean Square Error (RMSE) and Mean

Absolute Error (MAE) metrics. Prophet, a time series forecasting tool developed by Facebook, has proven accurate in predicting time series data (Junus, Tarno, & Kartikasari, 2023). Flexibility in handling different types of data and the ability to capture seasonality and holidays make it an excellent choice for time series forecasting tasks. In contrast, ARIMA, although a popular choice in time series analysis, may struggle with complex, nonlinear data patterns. Prophet's superior performance, as demonstrated by lower RMSE and MAE values, confirms its efficacy as a forecasting tool in a wide range of applications. Table 3.1 shows the RMSE and MAE for prediction data 30 days after the last date.

**Table 1**
**Comparison of MAE and RMSE ARIMA and Prophet**

| Algoritma | ARIMA | Prophet |
|---|---|---|
| Data Duration | 01 January 2017-31 December 2022 | 01 January 2017-31 December 2022 |
| Day Prediction | Next 30 days | Next 30 days |
| RMSE | 5148516197.40999 | 4208014977.61532 |
| MAE | 3582009597.46667 | 2210501626.75646 |

| Algoritma | ARIMA | Prophet |
|---|---|---|
| Data Duration | 01 January 2020-31 December 2022 | 01 January 2020-31 December 2022 |
| Day Prediction | Next 100 days | Next 100 days |
| RMSE | 5134521157.16281 | 2436799575.52544 |
| MAE | 3787018641.84 | 1575306588.29659 |

## Conclusion

In conclusion, R Shiny can create simple dashboard designs. Using other available modules or components, R Shiny can create dashboards with attractive designs and ease of use. The results of RMSE and MAE calculations show that Prophet has the upper hand in forecasting data. The overall conclusion is as follows:

1. R shiny can be used to create dashboards, showing calculations and graphs of results
2. ARIMA and Prophet can be used to predict time series data.

3. Prophet is superior to ARIMA; consistent results can be proven by changing the date distance and the number of days for forecasting. Prophet consistently delivers lower RMSE and MAE results, with varying long-term and predicted forward-day numbers.

## Bibliography

Aisah, Heti, Zaqiah, Qiqi Yulianti, & Supiana, A. (2021). Implementasi Kebijakan Asesmen Kemampuan Minimum (AKM): Analisis Implementasi Kebijakan AKM. *Jurnal Pendidikan Islam Al-Affan*, *1*(2), 128–135.

Aksenta, Almasari, Irmawati, Irmawati, Ridwan, Achmad, Hayati, Nur, Sepriano, Sepriano, Herlinah, Herlinah, Silalah, Ayupitha Tiara, Pipin, Sio Jurnalis, Abdurrohim, Iim, & Boari, Yoseb. (2023). *Literasi Digital: Pengetahuan & Transformasi Terkini Teknologi Digital Era Industri 4.0 dan Sociaty 5.0*. PT. Sonpedia Publishing Indonesia.

Ella, Susy, & Andari, Rosita Novi. (2022). Manajemen Kontinuitas Bisnis Badan Usaha Milik Desa di tengah COVID-19: Studi Kasus BUMDes Gua Bahu Desa Wisata Kertayasa Kabupaten Pangandaran. *Jurnal Kepariwisataan: Destinasi, Hospitalitas Dan Perjalanan*, *6*(2), 253–272. https://doi.org/10.34013/jk.v6i2.824

Firdaus, Tengku Alvin, Putra, Rinno Hendika, Arifandi, Fajar, Anam, M. Khairul, & Lathifah, Lathifah. (2023). Implementasi Sistem Rencana Pembelajaran Semester Berbasis Web Untuk Mempermudah Proses Pembelajaran. *Jurnal Teknoinfo*, *17*(1), 156–169.

Jamaludin, M., Sulistianto, S. W., MI, A., MM, M., Marthalia, Debby, Wikansari, Rinandita, Fachrurazi, H., Hiswanti, S., Indrayani, Etin, & Khasanah, S. Pd. (2022). *Transformasi Digital Dalam Dunia Bisnis*. Cendikia Mulia Mandiri.

Junus, Chea Zahrah Vaganza, Tarno, Tarno, & Kartikasari, Puspita. (2023). Klasifikasi Menggunakan Metode Support Vector Machine Dan Random Forest Untuk Deteksi Awal Risiko Diabetes Melitus. *Jurnal Gaussian*, *11*(3), 386–396.

Lince, Leny. (2022). Implementasi kurikulum merdeka untuk meningkatkan motivasi belajar pada sekolah menengah kejuruan pusat keunggulan. *Prosiding Seminar Nasional Fakultas Tarbiyah Dan Ilmu Keguruan IAIM Sinjai*, *1*, 38–49.

Manogari, Nathanael. (2019). *Analisis Deret Waktu Menggunakan Metode State Space Dan Aplikasinya Pada Harga Saham*.

Maziyyah, Atika Afifah. (2022). *Penggunaan Self-service Business Intelligence Dalam Pengambilan Keputusan Di Umkm Pada Masa Pandemi Covid-19: Studi Kasus Toko Sayur Keluarga Yogyakarta*. Universitas Islam Indonesia.

PURWANTI, Hani, Rochim, Adian Fatchur, & Warsito, Budi. (2022). *Prediksi Faktor-Faktor Penentu Adopsi Manajemen Hubungan Pelanggan Sosial Berdasarkan Pls-Sem Dan Jaringan Syaraf Tiruan Perambatan Balik*. School of Postgraduate Studies.

Sahrudin, Asep. (2016). Implementasi model pembelajaran means-ends analysis untuk meningkatkan kemampuan pemecahan masalah matematika mahasiswa. *JUDIKA*

*(JURNAL PENDIDIKAN UNSIKA)*, *4*(1). https://doi.org/10.35706/judika.v4i1.233

Yahya, M. Rizal. (2023). Pengaruh Business Intelligence Dan Cloud Computing Terhadap Keamanan Sistem Informasi (Studi Pada BUMN di Provinsi Aceh). *JURNAL ILMIAH MAHASISWA EKONOMI ISLAM*, *5*(2), 212–243.