# Predicting the Case of COVID-19 in Indonesia using Neural Prophet Model

Efraim Yahya Wijaya<sup>1</sup>, Alethea Suryadibrata<sup>2</sup> <sup>1</sup> Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia efraim.wijaya@student.umn.ac.id <sup>2</sup> Department of Informatics, Universitas Multimedia Nusantara, Tangerang, Indonesia alethea@umn.ac.id

> Accepted 10 October 2022 Approved 13 December 2022

*Abstract*— The spread of Coronavirus disease (COVID-19) is constantly changing in Indonesia. It is important to know the trends to help hospitals handle the crisis during outbreaks. Through this research, the prediction method is used to find increasing and decreasing of COVID-19 cases using the Neural Prophet model. Then the model is compared with the Facebook Prophet as comparison model. In this study dataset is used from (covid19.go.id) which was taken on 23 June 2022 with scraping technique. The results of this study indicate that the Neural Prophet model has better value in RMSE, and MAE compared to the Facebook Prophet.

Index Terms—COVID-19; Facebook Prophet; MAE; Neural Prophet; RMSE;

## I. INTRODUCTION

Since the initial entry of coronavirus disease 2019 (COVID-19) in Indonesia, the spread of virus is continuing to increase. The virus can quickly spread in the air through respiratory of an infected person [1]. This disease can cause fever, shortness of breath, even death. To suppress the spread of disease, the government makes a policy such physical distancing and wearing a mask. However, the emerge of new variant mutation such B.1.1.7 (Alpha) and B.1.617.2 (Delta) on 03 May 2021, causing the case to increase up to 56,757 active case on 15 July 2021 (covid19.go.id). This has significant impact on hospitals that provide medical treatment to patient that exposed to the COVID-19. Some patients with severe symptoms require advanced health resources, including respiratory support and intensive care.

Until this day the spread of virus can be controlled by policies dan regularization in every province such as limiting mobility with PPKM, self-isolation for people exposed with the COVID-19 and giving vaccine to the public. These factors create uncertainty to the spread of the COVID-19 virus. Based on COVID-19 information from the government website (covid19.go.id), shows the case of spread increased significantly at the time of long holidays. In addition, the mutation of variant B.1.617.2 (Delta) appeared in May 2021, and a new variant B.1.1,529 (Omicron) appeared in December 2021. Therefore, the forecasting method is needed to predict the number of infected cases in Indonesia.

Auto Regressive Integrated Moving Average (ARIMA) is one of the classical models and commonly used in various forecasting models [2]. Unfortunately, the ARIMA accuracy decreases when performing long range forecast. Moreover, data is not always linear. Therefore, many time series practitioner combining classical model with other model to make high quality forecast [3] [4]. However, the use of the model requires a high domain knowledge and expertise.

Recurrent Neural Network (RNN), Long Term Short Memory (LTSM), and Gated Recurrent Unit (GRU) is one of the deep learning models that can be used for prediction. This model is more robust compared with classical time series. The advantage of this model is that can remember important feature of sequential time series input. This model can produce good forecast in various forecasting models [5] [6]. Even so, this model is overly complex and could reducing the interpretability. Moreover, this model prone to underfit due small amount of data.

Facebook Prophet is one of the most popular models to predict time series data. This model can decompose time series to three main components: trend, seasonality, and holidays [7]. This model use analysisin-the-loop approach which can automatically to modeling time series. Moreover, this model could provide insight and easy to interpret.

Neural Prophet is successor of the Facebook Prophet model [8]. This model has similar characteristic feature as Facebook Prophet but more extensibility like auto-regression (AR), lagged regressor, and feature regressor. This model capable to increase model accuracy with used of neural network. In addition, this model uses Pytorch so that this model can be developed further by time series practitioner.

There have been several studies to predict cases of COVID-19. Wahyudi and Palupi predict the peak of COVID-19 pandemic using Susceptible, Infected, and Removed (SIR) model [9]. Research by Khurana et al. [10] said that the prediction of Neural Prophet is better than other machine learning models. The research by Wildhanrahman et al. [11], said model Facebook Prophet cannot predict the peak of case because the case of COVID-19 is still increasing that time. Then the research by Harahap et al. [12], said Facebook Prophet prediction still need more dataset. Therefore, if dataset updated in range 2 years, the result could be more accurate.

In this research will be using Neural Prophet for predicting the case of COVID-19 in Indonesia by days. The case is from every positive confirmed case for all province in Indonesia. This model was preferred because it easy to interpret and can increase the accuracy of prediction with extensibility of neural network. This model will be compared with Facebook Prophet as a comparison model. For the metrics, it will use scale dependent error such as Root Mean Square Error (RMSE) and Mean Absolute Error (MAE). The dataset which was taken from government website (covid19.go.id). This is public data which was updated every day from Kementerian Kesehatan (Kemenkes) to inform the surge increase of COVID-19 in Indonesia. The data is taken from 23 June 2022 with scraping technique. The model is expected to be used as anticipatory action against the surge increase of COVID-19 cases in Indonesia.

#### II. LITERATURE REVIEW

## A. Coronavirus disease (COVID-19)

Coronavirus disease (COVID-19) are infectious diseases by SARS-CoV-2 viruses [1]. This virus first time appear in Wuhan, China in 2019. This virus can spread quickly through respiration of infected person. According to WHO [1], people who infected by this virus will experience mild respiratory illness and recover without requiring medical attention. Generally, people who has underlying medical condition such cardiovascular disease, diabetes, chronic respiratory disease, or cancer are more likely to develop serious illness. Unfortunately, every people who infected by this virus will seriously ill even death. This virus first entry on 02 March 2020. Afterwards the virus mutated and become B.1.617.2 (Delta) in May 2021, and B.1.1.529 (Omicron) in December 2021. Until this day, virus spread continue. Therefore, we need forecast model to predict the future.

#### B. Time Series Forecasting

Forecasting is used to get information and used to as a reference to take decision in long term strategic planning. According to Hyndman et al. [13], forecast model is very dependent on availability of data. If the data is not available and not relevant, then the qualitative method preferred. Otherwise, if data is available and relevant, then quantitative method is preferred. Time series is one method to forecast quantitative method. As Hyndman et al. [13], said time series have criteria such numerical information about the past and having similar pattern in the future. Time series have composition such trend, seasonality, and cyclic

- 1. Trends occur when there is sudden increase or decrease in long term sequence of time series.
- 2. Seasonality occurs when time series affected by seasonal factor such a week, month, or even year. Seasonality occurs if data appear to be rise or fall and not in fixed frequency.
- 3. Cyclic normally occur by the fluctuation of economic, and always associated with business cycle. Cyclic happen with unknown factor.

# Facebook Prophet

Facebook Prophet is a model that develop by Facebook Ai Researcher for solve a problem in forecasting. The first problem in time series forecasting is choosing a method for solving time series problems. Second, expertise required in designing a time series model. Therefore, Facebook Ai Researcher developing a framework that resulting a high-quality forecast. The Facebook Prophet uses analyst-in-the-loop which can allow time series practitioner to automate the model. In other word, if there is an error in the model, Facebook Prophet will give feedback to practitioner for model inspection [7]. In Taylor and Letham research, said that Facebook prophet uses decomposable time series or dividing time series to be three main model: trend, seasonality, and holidays. This model is similar with Generalized Additive Models (GAM), which the output from function sum is the result of prediction. Here is the following formulation of Facebook Prophet (1).

$$y(t) = g(t) + s(t) + h(t) + \varepsilon(t)$$
(1)

С.

Where, g(t) is a trend function, which model nonperiodic changes in value of the time series. s(t) is seasonality function, which has periodic change in (days, weekly, and yearly). h(t) is a holidays function, that effect time series in one day or more days.  $\varepsilon(t)$  is a error term function, where represent any idiosyncratic changes which was not accommodate by the model.

#### D. Auto-regressive Neural Network (AR-Net)

Auto-regressive Neural Network (AR-Net) is a model that developed by Triebe et al. [14], this model aims to

In Fig. 1, Triebe et al. [14], said that AR-Net model can mimic the process of traditional AR with neural network. AR-Net is designed so that the parameters of the first layer is the same as AR-coefficient. AR-Net could be extended to many hidden layers for increasing model accuracy. Where AR-Net architecture with lag y(t-1), ..., y(t-p) is input and  $y_t$  is target, every weight in line is  $w_t, ..., w_p$ , and  $H_1, ..., H_k$ .



## E. Neural Prophet

Neural Prophet is not much different from Facebook Prophet. This model was developed by Triebe et al. [8], to simplify the forecasting models and maintaining the purpose of Facebook Prophet such as interpretation and configuration. The Neural Prophet provides more extensibility features such as automatic differencing with Pytorch as backend. Neural Prophet is made fully modular, so that people could scale by adding more components.

According to Triebe et al. [8], the main concept of Neural Prophet model is modular composability, where every module will contribute to additive component which will be used in forecast. Every module has input and their process models. However, all modules must generate h output, where h is number of steps which will be forecasted in the future simultaneously [8]. Neural Prophet formulation found in Equation 2.

$$\check{y} = T(t) + S(t) + E(t) + F(t) + A(t) + L(t) \quad (2)$$

solve the scalability problem in Classical Autoregression (AR) model. The problem occurs when dealing with long-range dependencies, Classical AR could be slow when fitting data with large scale. According to Triebe et al. [14], model sequence-to-sequence such as Recurrent Neural Network (RNN) could solve the problem. However, Triebe said that RNN is too complex for typical time series problem and reduce interpretability. Therefore, Triebe et al. created the AR-Net model by combining Classical AR and feed-forward neural network.

Where, T(t) is trend function in time t. S(t) is seasonality in time t. E(t) is event function in time t. F(t) is regression function in time t for future-known exogenous variable. A(t) is autoregression function in time t based on previous observation. L(t) is regression function in time t for lagged observation of exogenous variable. All the components can be configurable and combined to make a model. If all component not configurable then only trend and seasonality are used. [8]

# F. Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) is scale dependent error, where the error value is dependent on scale and cannot be compared with different scale. For calculating the RMSE is square root of mean value of predicted and actual.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \check{y}_i)^2}{n}}$$
(3)

Where,

n : Amount of data.

$$\check{y}_i$$
: Prediction value.

y: Actual value.

G. Mean Absolute Error (MAE)

Mean Absolute Error (MAE) is scale depended on error, where the error value dependent with scale and cannot be compared with different scale. For calculating the MAE is mean of absolute value from data predicted and actual.

$$MAE = \frac{\sum_{i=1}^{n} |y_i - \check{y}_i|}{n} \tag{4}$$

Where,

n : Amount of data.

 $\check{y}_i$ : Prediction value.

y: Actual value.

## III. METHODOLOGY

## A. Data Retrieval

The data was taken from the government website (covid19.go.id) with scraping technique. The observation data are only daily confirmed cases from all provinces in Indonesia. The daily case is including holidays and weekend. In this research, the range of data used is from 02 March 2020 to 23 June 2022 or 843 days since the initial entry of COVID-19 in Indonesia.

## B. Model Flowchart

Fig. 2, explain the approach in this research. First is data retrieval from the website government. Then changing the format data. After that, adding holidays into the models. Next, if we do experiment 1 then we split dataset into two, namely train data and test data. Different with experiment 2 we use all the datasets. After that, we fit the model with dataset. Then plotting all model result. Lastly, the models will be compared with RMSE and MAE.



#### C. Implementation Method

In this research, we aim to implement Neural Prophet model. This model is easy to interpret like Facebook Prophet and it can increase model accuracy by extensibility of autoregression (AR) and neural network. First, we input the model parameter. Then, split the data. Adding Event. After that, creating a model. Next, plotting the prediction result. Lastly, we compared the model with Facebook Prophet as comparison. RMSE and MAE will be used as a metrics. Table 2 explain the model parameter that will be used.

TABLE I. MODEL PARAMETER

Models	Parameters				
	Trend	Seasonality	Event	AR	
Neural Prophet	Piece- wise linear	Weekly & yearly	Holidays & Event	AR-NN	
Facebook Prophet	Piece- wise linear	Weekly & yearly	Holidays & Event	-	

# 1. Experimentation

In this research there are two experiments. The first experiment, we split the data into two data, such as train data and testing data. The second experiment we use all the datasets. The result of the experiment will be resulting prediction value and model component.

2. Data split

Data split is used to determine the accuracy of the models. Table 2 explain the number of distribution dataset. In the first experiment, we split 776 datasets into 80:20, or 80% train data and 20% test data. The train data is 620 days, calculated from 02 March 2020 to 12 November 2021. While the test data is 156 days, calculated from 12 November 2021 to 17 April 2022. In the second experiment, we feed all data to the models with additional 67 days, so the total would be 843 days, calculated from 02 March 2020 to 23 June 2022.

#### 3. Adding Event

In event we add the date of COVID-19 variant of concern first time found in Indonesia. In this case Alpha, Beta, Delta, Omicron, with additional Son of Omicron (BA1) and (BA.2). In holidays we add every holidays in Indonesia.

. Creating Model

Modeling is done by adding all the model parameters into the model. For Neural Prophet, the model will be trained with mini batch SGD, the model will be measured with L1Loss or MAE. We use L1Loss because it robust against outliers. For Facebook Prophet, model will fit directly with datasets.

Hyper-parameter will be used to tune the models. By doing so we could increase model accuracy. For Neural Prophet we can adjust learning rate, batch size, epoch, number of hidden layers, and number of nodes. For Facebook Prophet we can adjust changepoint prior scale, seasonality prior scale, and holidays prior scale. 5. Model Prediction

In first experiment, we use all models to predict 156 days ahead or all the data test (insample). In second experiment, we use all models to predict all the datasets. Lastly, in every experiment we predict 30 days ahead (outsample).

## D. Evaluation Methods

After experiment, the models can be compared with actual data. This is done because no observed value in future, so that the data prediction will be compared with actual data. We evaluate the prediction and actual value using RMSE and MAE. We used this metrics, because its commonly used in regression models. For comparison, the smallest result is the best model.

#### IV. RESULT AND DISCUSSION

## A. Computational Resources

We use these computational resources to create a model. This specification is related to hardware and software. Here is the specification:

- Hardware: Macbook Air 2017: Intel i5, RAM 8 GB, Storage 512 GB.
- Software: Google Collab with Python 3.7 installed, and Python library (such as: Pandas, Matplotlib, NumPy, and SNS).

## B. First Experiment

In first experiment we divide dataset to 80% train data, and 20% test data from 776 days. Then every model we add parameter. After that, we train the model with 620 data to predict 156 days ahead. The next step is plotting the prediction results. Then we evaluate the result with actual data (data set). Lastly, every model will predict 30 days ahead, start from 17 April 2022 to 17 May 2022. Here is the result of each model:

## 1. Neural Prophet

First, the Neural Prophet model will train the data. Then the model will train with 500 epoch data. Fig. 3, showing the error function with log-L1Loss, the log loss is enabled by default in Neural Prophet. In the graph show that the model is learning in 70 iteration and diverge after. So that, this model did not learn the pattern of dataset. This happen because if dataset train and dataset test are imbalance then the plot loss will have a gap. If the dataset has sufficient data, then the result will be good.



Fig. 3. Loss Function in Neural Prophet

Fig. 4, showing that the prediction of Neural Prophet is closely too accurate. This model cannot directly predict 156 days into the future with train data, because of the limited lag data in AR component. Therefore, this model will predict the test data (in-sample). In the figure show that Neural Prophet model can predict the increase in second wave of COVID-19 in data test. Unfortunately, the result sloping to negative values. This happen because the model does not have restriction to predict.



Fig. 4. Prediction of Neural Prophet



Second, the model will decompose the main model to trend, seasonality (weekly, and yearly), AR, and event. In Fig. 5, the trend shows detrending case from January to May 2022. In seasonality yearly, show the surge increase in January and July. In seasonality weekly, show the peak of the case is mostly on Thursday. In AR 1-th step ahead, show that the prediction in AR with range of data test. In additive event, show that the event which affect the trend forecast.

Lastly, in Fig. 6, shows the prediction of 30 days ahead (out-sample) using test data. The prediction shows that the case decreasing to negative value for the next 30 days ahead based on 10 predictions of 30 *nth step forecast* parameters.



Fig. 6. Prediction 30 days ahead Neural Prophet

## 2. Facebook Prophet

First, the model will fit the data with 620 days of train data. This model does not have training process, so we directly using train data to predict into 156 days into the future. Because of this, the model will add the train data plus 156 days prediction into the trend component. In Fig. 7, showing that the Facebook Prophet is able to fit in train data but failed to predict the test data. This happen because model is overfit and resulting negative prediction.



Fig. 7. Prediction of Facebook Prophet

Second, the model will decompose the main model into trend, event, seasonality weekly, and seasonality yearly. In Fig. 8, show the trend is declining in the next 156 days after train data. In event, shows the date of holidays that affect the trend. In seasonality weekly, show that the peak of the case is on Friday. In seasonality yearly, show the peak of the case is in July to September.



Fig. 8. Model Decompose of Facebook Prophet.

Lastly, in Fig. 9, shows that the next 30 days ahead using data test (out-sample). The prediction shows negative result in the next 30 days ahead.



Fig. 9. Prediction 30 days ahead Facebook Prophet

## C. Second Experiment

In second experiment we will use all the dataset to fit both models. Then every model we add parameter and hyper-parameter. Next, plotting the prediction results. After that, we try to predict 30 days ahead (outsample). Then, we evaluate the result of the prediction with all the data. Here is the result of each model:

## 1. Neural Prophet

First, the model will train all the dataset from 02 March 2020 to 23 June 2022 or 843 days after the first case of COVID-19 in Indonesia. Then the model will be fine-tuned with hyper-parameters. In Fig. 10, shows that the Neural Prophet model can predict all datasets accurately. This model can accurately predict the first, second, and third wave of COVID-19 in Indonesia.



Fig. 10. Prediction of Neural Prophet

Second, the model will decompose the main model into trend, seasonality yearly, seasonality weekly, AR, and Event. In Fig. 11, shows the almost similar component in experiment 1 (Fig. 5), but with all the datasets. Thus, this model can capture more insight.



Lastly, in Fig. 12, shows that the next 30 days ahead (out-sample) prediction. The prediction shows surge increase in the next 30 days ahead. This prediction based on AR nth step ahead which n is 30 steps from forecast origin. In this prediction only 10 step origin forecasts shown. The yhat1 is the recent origin and yhat10 is previous forecast.



Fig. 12. Prediction 30 Days Ahead Neural Prophet

#### 2. Facebook Prophet

First, the model will be fitted to all datasets from 02 March 2020 to 23 June 2022 or 843 days. Then the model will be fine-tuned with hyper-parameters. In Fig. 13, shows the Facebook Prophet can accurately predict all the dataset.



Fig. 13. Prediction of Facebook Prophet

Second, the model will decompose the main model into trend, holidays, seasonality weekly, and seasonality yearly. In Fig. 14, shows the smoother component after adding more data and fine-tuned hyper parameter. This shows that Facebook Prophet could compete with Neural Prophet with more data and proper fine-tune. Unfortunately, Facebook Prophet fail to capture yearly seasonality data.



Fig. 14. Model Decompose of Neural Prophet

Lastly, in Fig. 11, shows that the next 30 days ahead (out-sample). The prediction shows that COVID-19 case will be increased in the next 30 days.



Fig. 15. Prediction 30 Days Ahead Facebook Prophet

# D. Model Evaluation

The result of the model will be split into two parts, namely the first experiment and the second experiment. For first experiment, the model will be compared with data test. For the second experiment, the model will be compared with all datasets. The testing data is done by calculating the actual value with the predicted value, then calculating the error with MAE and RMSE. The result of the experiment is seen in Table 2. From the prediction results, the Neural Prophet model has the smallest MAE and RMSE compared with Facebook Prophet. But with second experiment by adding more data, the Facebook Prophet can compete with Neural Prophet.

TABLE II. PREDICTION RESULTS

Metrics	First Experiment		Second Experiment	
	Facebook Prophet	Neural Prophet	Facebook Prophet	Neural Prophet
RMSE	28,798.97	5,728.69	2,122.16	2,040.5
MAE	19,702.26	4,007.67	1,239.67	941.5

## V. CONCLUSION

Implementation of the Neural Prophet model to predict COVID-19 cases in Indonesia has been completed. The experiment show that the evaluation of Neural Prophet has better RMSE, and MAE compared to Facebook Prophet. The first experiment, the Neural Prophet has the result values of RMSE 5,728.69 and MAE 4,007.66. Then for the Facebook Prophet has the following result values of RMSE 28,798.97 for RMSE and 19,702.26 for MAE. The second experiment, the Neural Prophet model has the result of RMSE 2,040.5 and MAE 941.73. While the Facebook Prophet model has the result of RMSE 2,122.16 and MAE 1,239.67. In the future, the forecast method could be implemented with UI so that people can easily see the prediction. Also in future, the model could automatically add the data or holidays. In addition, if the number of datasets sufficient, the deep learning models could be used in future research.

#### REFERENCES

- WHO, "Coronavirus disease (COVID-19): how is it transmitted?," www.who.int, 13 December 2020. [Online]. Available: https://www.who.int/news-room/questions-andanswers/item/coronavirus-disease-covid-19-how-is-ittransmitted. [Accessed 7 February 2022].
- [2] Y. Ardila, "Hybrid ARIMA and multilayer perceptron for short term forecasting of electricity consumption in East Java," Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, 2016.
- [3] M. Khashei and Z. Hajirahimi, "A comparative study of series ARIMA/MLP hybrid models for stock price forecasting," *Communications in Statistics - Simulation and Computation*, vol. 48, no. 9, pp. 2625-2640, 2018.
- [4] C. B. A. Satrioa, W. Darmawan, B. T. U. Nadia and N. Hanafia, "Time series analysis and forecasting of coronavirus disease in Indonesia using ARIMA model and PROPHET," *Procedia Computer Science*, vol. 179, p. 524–532, 2021.

- [5] K. ArunKumar, D. V. Kalaga, C. M. S. Kumar, M. Kawaji and T. M. Brenza, "Forecasting of COVID-19 using deep layer recurrent neural networks (RNNs) with gated recurrent units (GRUs) and long short-term memory (LSTM) cells," *Chaos, Solitons & Fractals*, vol. 146, p. 110861, 2021.
- [6] S. Hansun and A. Suryadibrata, "Gold price prediction in COVID-19 era," *International Journal of Computational Intelligence in Control*, vol. 13, no. 2, pp. 1-4, 2021.
- [7] S. J. Taylor and B. Letham, "Forecasting at Scale," *The American Statistician*, vol. 72, no. 1, pp. 37-45, 2017.
- [8] O. Triebe, H. Hewamalage, P. Pilyugina, N. Laptev, C. Bergmeir and R. Rajagopal, "NeuralProphet: explainable forecasting at scale," arXiv:2111.15397, 2021.
- [9] B. A. Wahyudi, and I. Palupi, "Prediksi puncak pandemi Covid-19 di Indonesia dengan model SIR," *Jurnal Teknologi dan Sistem Komputer*, vol. 9, no. 1, pp. 49-55, Jan. 2021.
- [10] S. Khurana, G. Sharma, N. Miglani, A. Singh, A. Alharbi, W. Alosaimi, H. Alyami and N. Goyal, "An intelligent fine-tuned forecasting technique for COVID-19 prediction using neuralprophet model," *Computers, Materials & Continua*, vol. 71, no. 1, pp. 629-649, 2022.
- [11] F. I. Wildhanrahman, C. L. Abdulbaaqiy and M. F. Dirhami, "Prediction of the peak of the COVID-19 outbreak in Indonesia using a logistic model and forecasting with FB prophet," *Unri Conference Series: Community Engagement*, vol. 2, pp. 352-358, 2020.
- [12] M. Harahap, A. Z. Andika, A. M. Husein and A. Dharma, "Analisis tren dan perkiraan pandemi COVID-19 di Indonesia menggunakan peramalan metode prophet :sebelum dan sesudah aturan new normal," *Jurnal Teknologi Informasi Dan Ilmu Komputer*, vol. 9, no. 1, pp. 51-60, 2019.
- [13] R. J. Hyndman and G. Athanasopoulos, Forecasting: principles and practice, Melbourne, Australia: OTexts, 2021.
- [14] O. Triebe, N. Laptev and R. Rajagopal, "AR-Net: a simple auto-regressive neural network for time series," arXiv:1911.12436, 2019.
- [15] S. P. COVID-19, "COVID-19 Response Acceleration Task Force," Response Acceleration Task Force, [Online]. Available: https://covid19.go.id. [Accessed 20 02 2022].