WWJMRD 2021; 7(7): 56-60
www.wwjmrd.com
International Journal
Peer Reviewed Journal
Refereed Journal
Indexed Journal
Impact Factor SJIF 2017: 5.182 2018: 5.51, (ISI) 2020-2021: 1.361
E-ISSN: 2454-6615
DOI: 10.17605/OSF.IO/Q8GNK

Fitriana Hayun Anggreini
Information System Management, Postgraduate Program, Universitas Gunadarma Jl. Margonda Raya No. 199, Depok 16424, Indonesia.

Karmilasari
Information System Management, Postgraduate Program, Universitas Gunadarma Jl. Margonda Raya No. 199, Depok 16424, Indonesia.

Forecasting the Number of International Travelers Visiting Bali during the Covid-19 Pandemic using the FBProphet Method

Fitriana Hayun Anggreini, Karmilasari

Abstract
Bali is one of the tourist destinations that is popular with international travelers (tourists). The number of tourist attractions in Bali also establishes a lot of jobs for the people around the tourist attractions. However, challenged with the Covid-19 pandemic, the number of travelers arrivals fell by 99.99% (source: BPS Bali). This study provides a model for forecasting the number of international travelers visiting Bali using the FBProphet method. The dataset will go through the pre-processing stage and processed by FBProphet to match the forecasting model. From this model, it will be used to forecast for 30 days and calculate the MdAPE error rate with an error metric and the result value is 0.94.

Keywords: Forecasting, FBProphet, Time Series, Tourism, Travelers.

Introduction

Bali is a tourist destination that has attractions of its culture, nature, and various culinary experiences. This has made Indonesia known internationally as one of the countries with popular tourist attractions, [1] as shown in Table 1.

Table 1: National Percentage of International Traveler in 2019

No.	Province	Percentage
1	Bali	39.80 %
2	Jakarta	24.13 %
3	Batam	15.05%

Source: BPS Bali Province

Faced with the Covid-19 pandemic in 2020, both domestic and international travelers experienced a very significant reduction. To handle those issues, in March 2020 through the PP Number: 21, 2020 the Indonesian government has imposed Large-Scale Social Restrictions (PSBB). One of the PSBB regulations strictly limits foreigners entering Indonesian territory. Every foreigner who enters Indonesia must have a Health Certificate. The Indonesian government is trying to rehabilitate Bali's economy by implementing health protocols like the Cleanliness, Health, Safety and Environment Sustainability (CHSE) program from the Ministry of Tourism and Creative Economy (Kemenparekraf) in all sectors. This program is based on the Decree of the Indonesian Minister of Health Number: HK01.07/Menkes/382/2020.[2]

By looking at the pandemic situation, forecasting can help the government by describing the number of visits of international travelers to Bali. Forecasting is an approach to forecast the future based on quantitative values, including historical data and knowledge.[3] FBProphet is an alternative machine learning approach for forecasting modelling that eases making time series forecasting models. In addition, FBProphet can process outliers data in a data set and data that undergoes dramatic changes.[4] The purpose of this research is to create a forecasting model made based on the number of international travelers visiting Bali using the FBProphet method. Furthermore, the method can be modified to predict the scenario of a decrease in international travelers' arrival coming to Bali.

Correspondence:
Fitriana Hayun Anggreini
Information System Management, Postgraduate Program, Universitas Gunadarma Jl. Margonda Raya No. 199, Depok 16424, Indonesia.

Materials and Methods
In time series forecasting there is architecture as the guide in forecasting, as can be seen in Fig. 1.

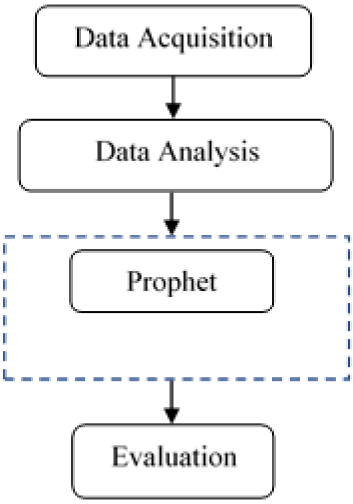

Fig 1: Architectural of Time Series Forecasting.

We use a Google Collaboratory to work with FBProphet by calling some libraries in Python, namely Pyplot, Numpy and Pandas. Pyplot is an additional library that functions to visualize data in graphic form. Numerical Python (Numpy) is a library which is utilized to process arrays more efficient and quickly. Next is Python for Data Analysis (Pandas) to import the library that can be written by typing Python code as shown in Fig. 2.[5]

```
from fbprophet import Prophet
from matplotlib import pyplot
import numpy as np
import pandas as pd
```

Fig 2: Python Code to Import Libraries

Data Acquisition
In this study, the dataset used is a dataset of the number of International travelers visiting Bali during the period of June 2019 – May 2021 (source: BPS Bali, Central Bureau of Statistics) because June is the start of summer break for the international travelers, especially American and European.

Data Analysis
1. Data analysis is carried out by ensuring the column that will be used for forecasting because the FBProphet Algorithm depends on two main parameters namely ds column (Datestamp) – it must be in YYYY-MM-DD format. For this reason, the attributes that will be used in forecasting are the "date" attribute as the ds column and "total" as the y column as shown in Fig. 3.

```
bali_df['ds'] = bali_df.date
bali_df['y'] = np.log(bali_df.total)
bali_df[['ds','y']]
```

Fig 3: The Python code to analyze history data.

2. The "ds" column needs to be expanded with a certain number of dates by creating a column with predicted data (make_future_dataframe) and then specify the number of days in the period parameter as shown in Fig. 4.

```
future_data = bali_model.make_future_dataframe(periods=180, freq = 'd')
future_data.tail(10)
```

Fig 4: Python code to make a future data frame.

3. Activate the prediction method that generates each row with predicted date and value which matches by "yhat" by writing Python code in Fig. 5.

```
bali_pdf = bali_model.predict(future_data)
bali_pdf[['ds','yhat','yhat_lower','yhat_upper']]
```

Fig 5: Python code to make predictive data.

4. Visualize the forecasting results using the Plot function by writing Python code in Fig. 6.

```
bali_model.plot(bali_pdf).show()
```

Fig 6: Python code to make plot function.

5. Cross-validation is assigned to the model with initial and horizon parameters by writing the python coding in Fig. 7.

```
from fbprophet.diagnostics import cross_validation
bali_cv = cross_validation(bali_model, initial='180 days',
period='15 days', horizon = '30 days')
bali_cv
```

Fig 7: Python Code to Cross-Validation.

6. Employ performance_metrics to obtain various evaluation metrics for the horizon by comparing the actual value and estimated value (yhat).

Facebook Prophet (FBProphet)
The model on FBProphet relies on additive regression which identifies trends, seasons, holidays and jointly consolidates using Equation 1.

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

Trend
The trend is an increase or decrease in data modelled by making a piecewise linear curve, which is a connected line segment as shown in Fig. 8. A piecewise linear function is a function that has a domain that can be decomposed into parts where the model is not affected by spikes or missing data [6].

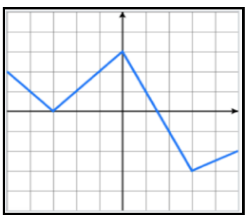

Fig 8: A Continuous Piecewise Linear Function Graph

The piece linear model is the default approach to forecasting and can be computed using Equation (2). The piece linear model is a flexible model for linear and non-linear functions as well as robust for outliers data model. [7]

$$g(t)=(k+a(t)^{T}\delta)t+(m+a(t)\gamma \qquad (2)$$

Seasonality

Seasonality is a recurring pattern studied by FBProphet described in Equation 3.

$$s(t)=\sum_{n=1}^{N}\left(a_n \cos\left(\frac{2\pi nt}{P}\right)+b_n \sin\left(\frac{2\pi nt}{P}\right)\right) \qquad (3)$$

Holiday or Event

FBProphet allows analysis to provide a list of past and future events habits. Each day in the custom is separated by parameters using additional parameters to model Holiday or Event effects as shown in Table 2.

Tabel 2: Seasonality and Holiday Parameters.

No.	Parameter	Description
1	Yearly_seasonality	Fit yearly seasonality
2	Weekly_seasonality	Fit weekly seasonality
3	Daily_seasonality	Fit daily seasonality
4	Holidays	Feed data frame containing holiday name and date
5	Seasonality_prior_scale	Parameter for changing the strength of seasonality model
6	Holiday_prior_scale	Parameter for changing the strength of holiday model

Results

Pre-process data on the number of international travelers visiting Bali will be presented here:

The Results of Features Selection.

Fig. 9 Shows data from the dataset starting from June 1, 2019, to May 31, 2021. The number continues to decrease until it no longer detects new international tourist arrivals and a total of 0 arrivals on May 31, 2021.

	date	airport	harbor	country	province	total
0	1-Jun-19	35462	5	Indonesia	Bali	35467
1	2-Jun-19	28342	3	Indonesia	Bali	28345
2	3-Jun-19	28133	4	Indonesia	Bali	28137
3	4-Jun-19	37432	2	Indonesia	Bali	37434
4	5-Jun-19	17384	1	Indonesia	Bali	17385
...	...	...	...	...	...	...
726	27-May-21	0	0	Indonesia	Bali	0
727	28-May-21	0	0	Indonesia	Bali	0
728	29-May-21	0	0	Indonesia	Bali	0
729	30-May-21	0	0	Indonesia	Bali	0
730	31-May-21	0	0	Indonesia	Bali	0

Fig 9: Actual Data of International Traveler Visiting Bali.

Dateframe History Analyze

Based on Fig 10, it is known that the data set has 712 data. The dataframe that will be used is date and total visitor, because the input required by FBProphet is always a dataframe with two columns “ds” and “y”.

	ds	y
0	1-Jun-19	10.476358
1	2-Jun-19	10.252206
2	3-Jun-19	10.244841
3	4-Jun-19	10.530335
4	5-Jun-19	9.763363
...	...	...
697	28-Apr-21	0.693147
709	10-May-21	0.000000
710	11-May-21	0.693147
711	12-May-21	1.386294
712	13-May-21	0.000000

Fig 10: The Dataframe of Foreign Tourist Visits to Bali.

The Results of Historical Dataframe Analyze

The forecasting model created will then be used to generate new visit values. Then to create a forecast value, a new data frame is required. This data frame consists of one column "ds". Fig. 11.

	ds	y
0	1-Jun-19	10.476358
1	2-Jun-19	10.252206
2	3-Jun-19	10.244841
3	4-Jun-19	10.530335
4	5-Jun-19	9.763363
...	...	...
697	28-Apr-21	0.693147
709	10-May-21	0.000000
710	11-May-21	0.693147
711	12-May-21	1.386294
712	13-May-21	0.000000

Fig 11: Dataframe of Foreign Tourist Visits to Bali.

The Result of Forecasting New Number of Foreign Traveler Visits

FBProphet assigns approximate values for each date in the “ds” column into a new column named “yhat”.[8] The data frame of forecasting results can be seen in Fig. 12 which displays the “yhat” column along with the uncertainty interval.

	ds
649	2021-10-31
650	2021-11-01
651	2021-11-02
652	2021-11-03
653	2021-11-04
654	2021-11-05
655	2021-11-06
656	2021-11-07
657	2021-11-08
658	2021-11-09

Fig 12: Forecasting results Dateframe

	ds	yhat	yhat_lower	yhat_upper
0	2019-06-01	10.354089	8.943930	11.756860
1	2019-06-02	10.097059	8.738189	11.385332
2	2019-06-03	9.471720	8.050324	10.824482
3	2019-06-04	9.606041	8.135262	10.848464
4	2019-06-05	9.648152	8.439264	10.969824
...	...	...	...	...
654	2021-11-05	1.322928	-1.397756	4.109562
655	2021-11-06	1.684269	-1.086473	4.282007
656	2021-11-07	1.429678	-1.323810	4.215835
657	2021-11-08	0.806779	-1.853791	3.486638
658	2021-11-09	0.943539	-1.856856	3.694105

Fig 13: Forecasting Date Dataframe.

Based on approximate date frame in Fig 13, there will be 3 visits (exponent of 0.9453539) with uncertainty levels from 0 visits (exponent -1.856856) to 40 people (exponent 3.694105).

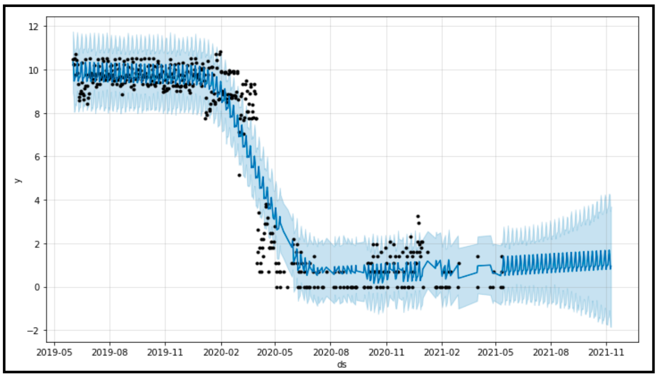

Fig 14: The New Visits Chart.

The x-axis represents the date range in "ds" column and the y-axis represents the value of the number of visits in the log. The Black Dot represents the historical "y" value in the most recent visit historical data dataframe. The dark blue line represents the result of the "yhat" forecast that has been made along with the level of uncertainty represented by the light blue area around the "yhat" line. Visualization of the comparison of actual data with forecasting results can be seen in Fig. 14 and Fig. 15.

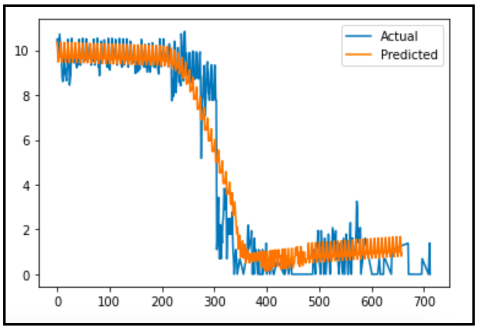

Fig 15: The Graph Comparison of actual data with predicted data

Discussions

From generated forecasting model, data will go through a cross-validation process to calculate the error rate as shown in Fig. 16. The results of the cross-validation will be employed by FBProphet to calculate the model error rate using Equation 4.

$$p_t = \left|\frac{y_t - f_t}{y_t}\right| = \left|\frac{e_t}{y_t}\right| \quad (4)$$

	ds	yhat	yhat_lower	yhat_upper	y	cutoff
0	2019-12-06	9.731899	9.356720	10.118278	9.873801	2019-12-05
1	2019-12-07	10.221211	9.850741	10.576817	10.198058	2019-12-05
2	2019-12-08	9.810786	9.435897	10.147621	9.936390	2019-12-05
3	2019-12-09	9.285845	8.903400	9.640736	9.440261	2019-12-05
4	2019-12-10	9.232686	8.872518	9.618716	9.348536	2019-12-05
...	...	...	...	...	...	...
558	2021-04-28	0.863741	-0.559546	2.193688	0.693147	2021-04-13
559	2021-05-10	0.741638	-0.696660	2.082856	0.000000	2021-04-13
560	2021-05-11	0.881560	-0.519378	2.230159	0.693147	2021-04-13
561	2021-05-12	0.906622	-0.569231	2.300551	1.386294	2021-04-13
562	2021-05-13	0.967279	-0.377000	2.326211	0.000000	2021-04-13

Fig 16: The Dataframe of Cross-Validation.

FBProphet can measure several errors, one of which is the Median Absolute Percentage Error (MdAPE). MdAPE is determined by ordering the Absolute Percentage Error (APE) from smallest to largest and using the middle value (or the average of the two middle values if N is an even number) as the median as shown in Equation 5.

$$MdAPE = median\,(p_1,\ p_2,\ \ldots, p_N) \quad (5)$$

MdAPE is calculating the percentage error between the forecast data and the final result and MdAPE is more resistant to outliers than MAPE and sMAPE, as can be seen in Fig. 17.

	horizon	mse	rmse	mae	mdape	coverage
0	1 days	3.437315	1.854000	1.402265	1.382621	0.347826
1	2 days	5.282608	2.298392	1.634950	0.913577	0.409091
2	3 days	5.592620	2.364872	1.676254	0.649676	0.277778
3	4 days	4.691804	2.166057	1.420791	0.563231	0.500000
4	5 days	3.850948	1.962383	1.445400	0.317746	0.444444
5	6 days	5.833547	2.415274	1.684877	0.949874	0.476190

Fig 17: The Error Rate Calculation Results.

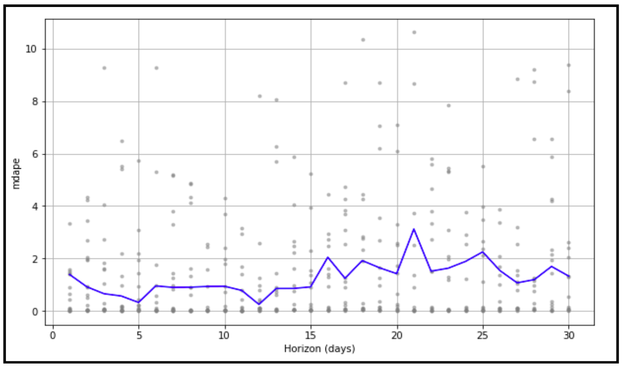

Fig 18: The Metric Graph of MdAPE.

Fig. 18 shows the absolute percent error for each prediction in the cross-validation data frame. It can be seen for the next 30 days, the forecasting results are good because the results of percentage error are close to 0%.

Conclusions

By using the FBProphet method resulted in forecasting new visits from June 1, 2021 to November 9, 2021. The Accuracy is calculated by the MdAPE with results of the percentage error are close to 0%. With this forecast, the FBProphet forecasting model is expected to be used as a fast and good forecasting alternative as a long-term forecasting tool. The results of the forecast and visualization from the analysis can be used as consideration for both tourism businesses and the Indonesian government in determining a strategy for the recovery of the tourism sector in Bali from the impact of the Covid-19 pandemic.

References

1. Friscintia PBA, Alamsyah A. Indonesian Tourism Demand Forecasting Using Time Series Approach to Support Decision Making Process 1Putu Bella Ayastri Friscintia, 2Andry Alamsyah 1Telkom University bellafriscintia@gmail.com. 2019;23(2):12.
2. Maemunah I. Implementation of Cleanliness, Health and Environmental Sustainability Guidelines in Restaurants around the Southern Java Crossing Route in Ciamis Regency. Bp Int Res Crit Inst BIRCI-J Humanit Soc Sci. 2021 May 21; 4(2):2635–42.
3. Anandatirtha JS. Analysis and Forecasting COVID-19 in India. Int J Res Appl Sci Eng Technol. 2020 Aug 31; 8(8):70–6.
4. Petrevska B. Predicting tourism demand by A.R.I.M.A. models. Econ Res-Ekon Istraživanja. 2017 Jan; 30(1):939–50.
5. Adawadkar K. Python Programming-Applications and Future. Int J Adv Eng Res Dev. 2017 Apr; 4.
6. Subashini A, K S, Saranya S, Harsha U. Forecasting Website Traffic Using Prophet Time Series Model. Int Res J Multidiscip Technovation. 2019 Jan 25; 56–63.
7. Taylor SJ, Letham B. Forecasting at scale [Internet]. PeerJ Preprints; 2017 Sep [cited 2021 Mar 20]. Available from: https://peerj.com/preprints/3190v2
8. Battineni G, Chintalapudi N, Amenta F. Forecasting of COVID-19 epidemic size in four high hitting nations (the USA, Brazil, India and Russia) by Fb-Prophet machine learning model. Appl Comput Inform [Internet]. 2020 Dec 10 [cited 2021 Mar 22].