Bangladesh COVID-19 Daily Cases Time Series Analysis using Facebook Prophet Model

[1]Sakib Mahmud, Undergraduate Student (2/2), Department of Economics, Shahjalal University of Science & Technology, Sylhet. Email: sakib13@student.sust.edu

Abstract: The study aims to predict & analyze the daily case data of Bangladesh. In the analysis, secondary data collected & the RStudio version, 1.3.959 was used. The prediction of daily cases was found from 22nd July to 19th September using the Facebook Prophet Model. Though the plot of the daily case shows that the graph is following a downward trend in July, the Facebook prophet model predicted that the daily cases will keep on rising if the trends till July 21st from 9th March continues. But there was found some overestimation & underestimation of the daily cases, the linear model in actual vs predicted cases gave a p-value much lower with a R^2 value of 0.8919. But there is some space if the trend which is seen in July continues than, an updated study using this model would predict the lowering of daily cases.

Keywords: COVID-19; Data Analysis; Prophet Model; Time Series; Bangladesh

Introduction:

The Coronavirus disease or COVID-19 pandemic is caused by a virus named severe acute respiratory syndrome coronavirus 2 or (SARS-CoV-2)(2020a). Till 21st July 2020 all over the world, there are 14,562,550 confirmed cases of COVID-19, including 607,781 deaths. In Bangladesh, the number of confirmed cases is 207,453 with 2,668 deaths(2020b).

The R_0 or the Reproductive Number of COVID-19 was found higher than SARS Coronavirus(Liu et al., 2020), so it spreads faster from people to people. In controlling the pandemic Government measures like early detection, adequate medical supplies, early treatment, providing medical supplies, admitting to the specific hospital & therapeutic strategy can lower the cases of COVID-19(Fang et al., 2020). Another study also suggests that Lockdown & recommended individual hygiene can slow down the outbreak(Paul et al., 2020). In this respect, the government of Bangladesh declared lockdown from 26th March 2020 (Vaidyanathan, 2020) which extended until 30th May 2020. Side by side of declaring nation-wise lockdown the government also kept educational institutions (Shawon,

2020)& international flights closed. Though on 16th June 2020 the international flights were re-opened(Abdullah, 2020). Due to the lockdown initiative to tackle COVID-19 impact in Bangladesh, the income of rural & urban poor has dropped significantly, about 80%. For meeting the consumption need about 50% of the poor population is taking a loan(Mohiuddin, 2020).

In this paper, we tried to summarize & graphically show the daily cases of COVID-19 in Bangladesh, most significantly this paper tried to predict daily cases using Facebook Prophet Model & using the linear model we tried to find the overestimation & underestimation. This analysis might help other researchers to understand the Daily Cases trends of COVID-19 in Bangladesh.

Literature Review:

For the literature review, we tried to review all the latest papers which relate to Bangladesh COVID-19 data analysis. The first case of COVID-19 in Bangladesh was detected on 8 March 2020. An explorative study according to the data from 8 March 2020 to 16 April 2020 found testing for COVID-19 infection is inadequate & not accessible easily. Infections of COVID-19 is occurring to all age groups, & deaths due to covid19 are found to be high. Among the health care workers, Doctors are more susceptible to COVID-19 infection(Hossain et al., 2020). A trend analysis using data until May 09, 2020, suggests that in Bangladesh Death Rate is low in compared to diagnosis, recovery seems to be fast in compared to death, & in compared to two to three days of prior cases the disease is seen to be reducing(Muyeed and Siddiqi, 2020).

The reproductive number or R_0 of COVID-19 seems to vary from country to country, the R_0 & incubation period of COVID-19 plays an important role in controlling the disease(Kamrujjaman et al., 2020). The infectious disease, COVID-19 has got many unclear properties that is why it is tough to get an accurate SEIR(Susceptible, Exposed, Infected, and Recovered) prediction(Hamzah et al., 2020). In the analysis of COVID-19, the time-dependent SIR model found to be more effective & adaptive compared to the traditional SIR model(Chen et al., 2020). A model study using SEIRD (Suspected, Exposed, Infectious, Recovered & dead) March 8 to May 15, 2020, found death rate, recovery rate & transmission rate lower in Bangladesh,

compared to other countries. And R_0 which they found was 2.25(Muka and Sannyal, 2020).

Some forecasting research works show significant findings. Using SEI_DI_UQHRD; Susceptible-Exposed-Symptomatic Infectious-Asymptomatic Infectious-Quarantined-Hospitalized-Recovered-Dead deterministic compartmental model, it was found that infectious cases could reach the peak around the beginning of June 2020 with cases around 2,078-2,840 in Bangladesh. And R_0 value was also calculated & was found between 3.12-3.94 using the data till May 11, 2020(Nabi, 2020). Between May 28 & May 30, 2020, Bangladesh has crossed the inflection point for COVID-19. Their forecasting using LSTM (Long Short Term Memory) networks & Logistic Curve method predicts that at the end of the epidemic in Bangladesh there will be around 187-193 thousand confirmed cases & around 3,600-4,000 death toll. The study also adds that if the authority takes strict measures than by the end of January 2nd, 2021 the COVID-19 cases in Bangladesh will be under control(Hridoy et al., 2020). A modified SIR(Susceptible-Infected-Recovered) predicted that the peak of COVID-19 will reach at the end of June 2020(Arifutzzaman et al., 2020). If 60% lockdown in Bangladesh is maintained SIR (Susceptible, Infected, Recovered) model predicts that final COVID-19 cases on 92nd day will be 3782,558 & after 193 days the cases will steadily decrease to zero(Rahman et al., 2020). By analyzing all the related literatures we found that there is no work with time series analysis using Facebook Prophet Model.

Methodology:

We collected the new cases or daily cases data of Bangladesh from Our World in Data (https://ourworldindata.org/coronavirus/country/bangladesh?country=~BGD) from 9th March to 21st July 2020. For making the analysis we used 4 libraries of R- 'dplyr', 'prophet', 'lubridate', 'ggplot2'. RStudio version 1.3.959 was used for the overall analysis. We used Facebook Prophet Model (Taylor and Letham, 2018) for predicting Bangladesh Daily Cases of COVID-19 up to 19 September 2020. This model is an additive regressive model or $Y(t)= G(t)+S(t)+H(t)+E_t$ where, G(t)= Trend Factor; H(t)= Holiday Component; S(t)= Seasonality Component; & E_t= Error Term.

R Codes used for the data analysis:

```
library(dplyr)
library(prophet)
library(lubridate)
library(ggplot2)
Daily_Cases <- read_excel("Daily Cases.xlsx")
attach(Daily_Cases)
names(Daily_Cases)
summary(`Daily Cases`)
tDaily_Cases$Dates <- ymd(Daily_Cases$Dates)
str(Daily_Cases)
qplot(Dates,`Daily Cases`,data = Daily_Cases, main = 'COVID-19 Daily Cases in Bangladesh')
ds <- Daily_Cases$Dates
y <- Daily_Cases$`Daily Cases`
df <- data.frame(ds, y)
m <- prophet(df)
future <- make_future_dataframe(m, periods = 60)
forecast <- predict(m, future)
plot(m, forecast)
dyplot.prophet(m,forecast,main='COVID19 Daily Cases Prediction')
prophet_plot_components(m, forecast)
pred <- forecast$yhat[1:135]
actual <- m$history$y
plot(actual, pred)
abline(lm(pred~actual),col='red')
summary(lm(pred~actual))
```

Results & Discussion:

We summarized the Daily Cases-

Table: Summarization of Daily Cases

Daily Cases	Min	Maximum	Date Coming Maximum Cases	Average (Mean)	Median
Total of 135 Days: 207453	0	4019	2020-07-03	1537	1041

In a total of 135 days from 9th March 2020 the total number of cases in Bangladesh was found to be 207,453 with average (Mean= 1,537 & Median= 1,041), & maximum cases of 4,019 which was seen in 3rd July 2020.

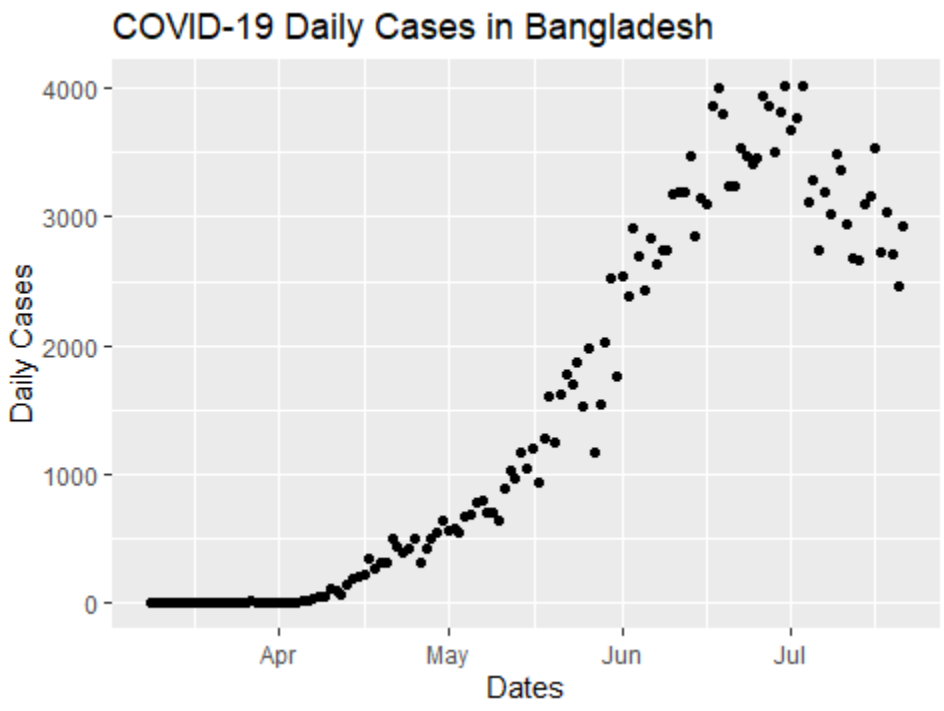

Figure: Daily Cases Plot

In this plot of Daily Cases in Y-axis & Dates in X-axis, it is seen that in June the graph is taking Concave shape so it was increasing in a decreasing rate, but in July (That is after 3rd July 2020) the graph is sloping downward & the same scenario is seen continued throughout July.

Now we are going to run the prophet model using our time series data, without considering daily & yearly seasonality as we don't enough data to measure daily & yearly seasonality.

We are predicting 60 days from 21st July that is up to 19 September 2020. These predictions are valid assuming the current condition continues.

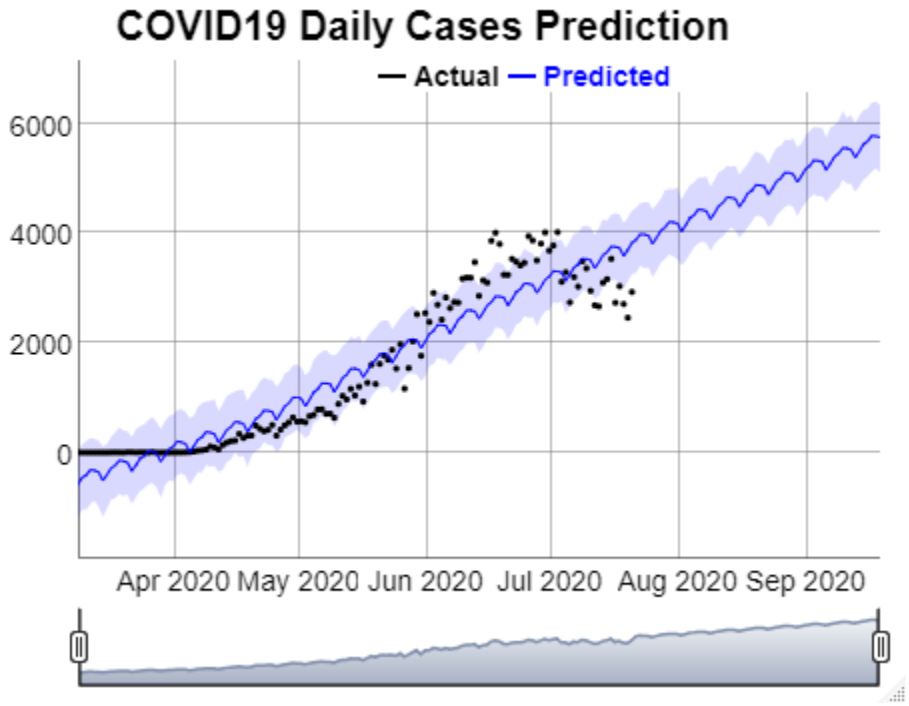

Figure: Time Series Prediction Using Prophet Model

This model predicts that at the end of July 2020 we will see about 4,201 cases, in the middle of August 4,628 cases, at the end of August 5,042 cases & in the last predicted day that is on 19 September 2020 we will see 5,750 cases.

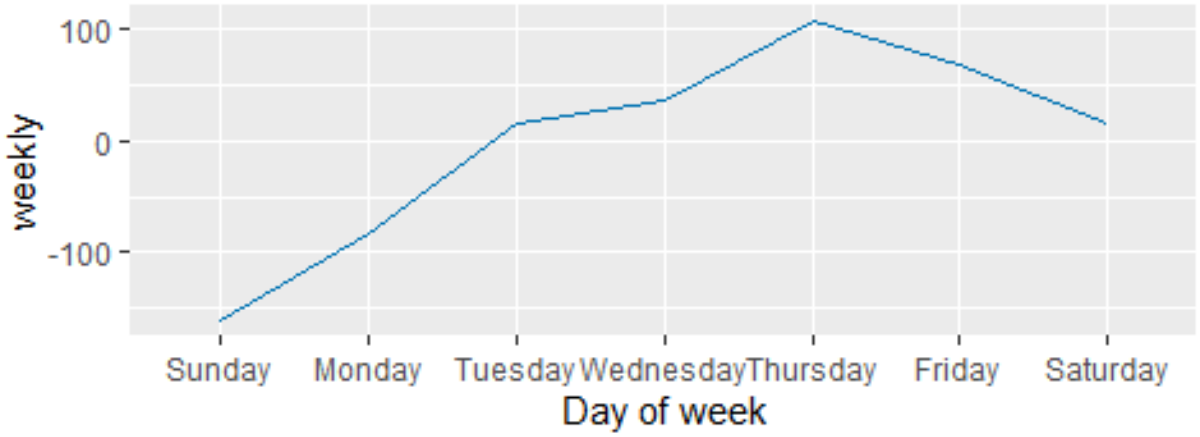

Figure: Number of Daily Cases trends observed Weekly

This plot shows the number of daily cases is often found more reported on Thursdays & Fridays in compared to other days of the week. But these do not mean people have more risk of COVID-19 on Thursday & Friday.

Now to analyze our model we plot actual & predicted data. Then we use the linear model in this plot.

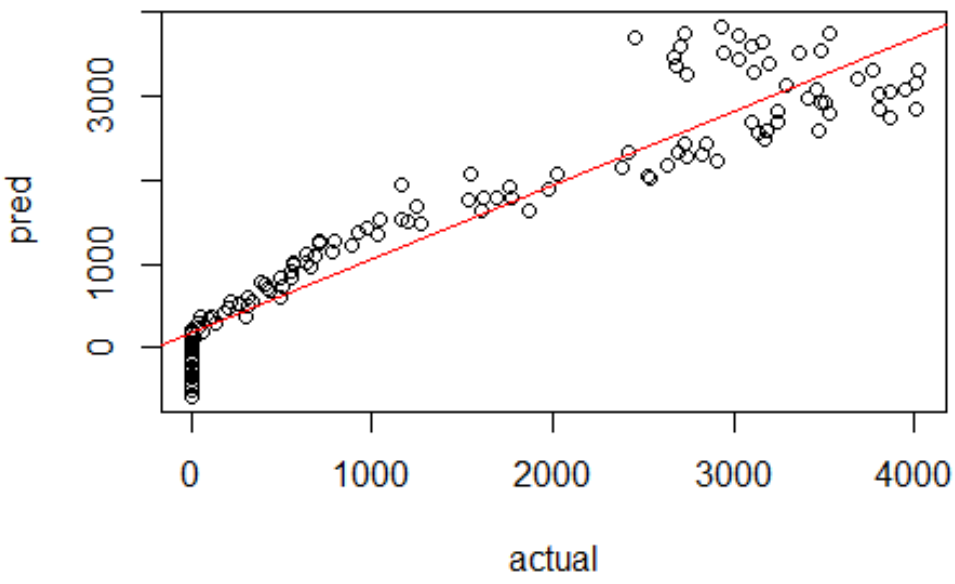

Figure: Predicted Cases & Actual Cases Plot with Linear Line

In actual vs predicted values we can see both underestimation & overestimation.

Summary of this Linear Model:

```
lm(formula = pred ~ actual)

Residuals:
    Min      1Q  Median      3Q     Max
-864.83 -342.61   10.28  263.90 1346.74

Coefficients:
             Estimate Std. Error t value   Pr(>|t|)
(Intercept) 183.47450   55.43354    3.31   0.0012 **
actual        0.88090    0.02659   33.13   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 435.3 on 133 degrees of freedom
Multiple R-squared:  0.8919,   Adjusted R-squared:  0.8911
F-statistic:  1098 on 1 and 133 DF, p-value: < 2.2e-16
```

So in this model were are getting a high R^2 value of .8919 with much lower p-value; which indicates that this model relation is statistically significant.

Conclusion:

The COVID-19 cases of Bangladesh seem to be decreasing day by day after the maximum number of cases seen on 3rd July 2020. The predictions using facebook prophet model suggested the opposite of what seen in daily cases in July month, it says that if the trends till 21st July continues than COVID-19 cases will keep on increasing & in the last predicted day (19 September 2020) we will see the daily cases reach to 5,750. In the case of weekly analysis, the study suggests that most cases are reported on Thursday & Friday in Bangladesh. When we ran a linear model in predicted vs. actual cases from 9th March to 21st July 2020, we found a statistically significant relation, & very smaller difference observed & predicted values though there was found some overestimation & underestimation. This analysis prediction is true based on current trends till 21st July 2020, but as we say the cases are dropping day by day an updated study after 1 or 2 months might predict different results.

References:

2020a. *Naming the coronavirus disease (COVID-19) and the virus that causes it* [Online]. World Health Organization. Available: https://www.who.int/emergencies/diseases/novel-coronavirus-2019/technical-guidance/naming-the-coronavirus-disease-(covid-2019)-and-the-virus-that-causes-it [Accessed 7/22/2020.

2020b. *WHO Coronavirus Disease (COVID-19) Dashboard* [Online]. World Health Organizatin. Available: https://covid19.who.int/ [Accessed 2020/7/22.

ABDULLAH, M. 2020. Bangladesh to resume international flights on June 16. *Dhaka Tribune*, June 11th, 2020.

ARIFUTZZAMAN, A., FARGANA, A. & RAKHIMOV, A. 2020. Spreading Analysis of COVID-19 Epidemic in Bangladesh by Dynamical Mathematical Modelling. *medRxiv*.

CHEN, Y.-C., LU, P.-E. & CHANG, C.-S. 2020. A Time-dependent SIR model for COVID-19. *arXiv preprint arXiv:2003.00122*.

FANG, Y., NIE, Y. & PENNY, M. 2020. Transmission dynamics of the COVID-19 outbreak and effectiveness of government interventions: A data-driven analysis. *Journal of medical virology,* 92**,** 645-659.

HAMZAH, F. B., LAU, C., NAZRI, H., LIGOT, D., LEE, G. & TAN, C. 2020. CoronaTracker: worldwide COVID-19 outbreak data analysis and prediction. *Bull World Health Organ,* 1**,** 32.

HOSSAIN, I., KHAN, M. H., RAHMAN, M. S., MULLICK, A. R. & AKTARUZZAMAN, M. 2020. The epidemiological characteristics of an outbreak of 2019 novel coronavirus diseases (COVID-19) in Bangladesh: A descriptive study. *Journal of Medical Science and Clinical Research,* 8.

HRIDOY, A.-E. E., NAIM, M., EMON, N. U., TIPO, I. H., ALAM, S., AL MAMUN, A. & ISLAM, M. S. 2020. Forecasting COVID-19 Dynamics and Endpoint in Bangladesh: A Data-driven Approach. *medRxiv*.

KAMRUJJAMAN, M., JUBYREA, J. & ISLAM, M. S. 2020. Data analysis and mathematical model: control measures and prediction to prevent COVID-19 outbreak. *Arabian Journal of Medical Sciences,* 3**,** 5-9.

LIU, Y., GAYLE, A. A., WILDER-SMITH, A. & ROCKLÖV, J. 2020. The reproductive number of COVID-19 is higher compared to SARS coronavirus. *Journal of travel medicine*.

MOHIUDDIN, A. K. 2020. A Pandemic Review of Covid-19 Situation In Bangladesh. *The American Journal of Medical Sciences and Pharmaceutical Research,* 2**,** 38-50.

MUKA, A. & SANNYAL, M. 2020. Transmission Dynamics of COVID-19 in Bangladesh-A Compartmental Modeling Approach.

MUYEED, A. & SIDDIQI, M. N. A. 2020. Prevalence and Severity of COVID-19 Disease in Bangladesh: A Trend Analysis. *Journal of Health & Biological Sciences,* 8**,** 1-8.

NABI, K. N. 2020. Forecasting COVID-19 pandemic: A data-driven analysis. *Chaos, Solitons & Fractals***,** 110046.

PAUL, A., CHATTERJEE, S. & BAIRAGI, N. 2020. Prediction on Covid-19 epidemic for different countries: Focusing on South Asia under various precautionary measures. *medRxiv*.

RAHMAN, M. M., AHMED, A., HOSSAIN, K. M., HAQUE, T. & HOSSAIN, M. A. 2020. Impact of control strategies on COVID-19 pandemic and the SIR model based forecasting in Bangladesh. *medRxiv*.

SHAWON, A. A. 2020. https://www.dhakatribune.com/bangladesh/2020/05/27/state-minister-update-of-lockdown-coming-soon. *Dhaka Tribune*, May 27th, 2020.

TAYLOR, S. J. & LETHAM, B. 2018. Forecasting at scale. *The American Statistician,* 72**,** 37-45.

VAIDYANATHAN, R. 2020. Bangladesh fears a coronavirus crisis as case numbers rise. *BBC News*, 16 June 2020.