

COVID-19 Outbreak Prediction and Forecasting in Bangladesh using Machine Learning Algorithm

S M Abdullah Al Shuaeb¹, Md. Kamruzaman², Mohammad Al-Amin³

^{1,3}Instructor (Tech), ²Workshop Super (Electronics),

^{1,2}Tangail Polytechnic Institute, Tangail, Bangladesh

³Brahmanbaria Polytechnic Institute, Budhanti, Bangladesh

ABSTRACT

In this time, Novel Corona Virus is an important issue in the world, it also named COVID-19. This virus has been come from Wuhan, China in last December 2019. This virus has created critical circumstances in the whole world especially Bangladesh. The outbreak of COVID-19 is increasing gradually in Bangladesh. To predict and forecasting COVID-19 in Bangladesh we have used machine learning (ML) Linear Regression model. LR model is useful to predict the outbreak of COVID-19 in Bangladesh. It can be helped efficiently to predict some common numerical data like observation day, tested case, affected case, death case, recover cases, and forecast the number of upcoming cases for the next 30 days in Bangladesh. Our paper to study to analyze the epidemic growth of the COVID-19 in Bangladesh. We have applied the mathematical regression model to analyze the prediction and forecast for the effective threat of the COVID-19 in Bangladesh. The main objective of this paper how to predict the virus-affected cases, recover cases, death cases, tested cases, and forecasting the future situation of Bangladesh.

KEYWORD: Machine Learning (ML), Linear Regression (LR), COVID-19, Prediction, Forecast, VIRUS, Bangladesh, etc

How to cite this paper: S M Abdullah Al Shuaeb | Md. Kamruzaman | Mohammad Al-Amin "COVID-19 Outbreak Prediction and Forecasting in Bangladesh using Machine Learning Algorithm" Published in International Journal of Trend in Scientific Research and Development (ijtsrd), ISSN: 2456-6470, Volume-5 | Issue-1, December 2020, pp.829-835, URL: www.ijtsrd.com/papers/ijtsrd38068.pdf



IJTSRD38068

Copyright © 2020 by author(s) and International Journal of Trend in Scientific Research and Development Journal. This is an Open Access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0) (<http://creativecommons.org/licenses/by/4.0>)



1. INTRODUCTION

1.1. Prelude

At present Novel Coronavirus is a hot topic in the world. It creates critical situation in whole world. The outbreak began in last December 2019 in China's Hubei Province. The virus has been out broken to 215 regions of the world, including Bangladesh. The [1]first corona virus was detected in Bangladesh on March 8, 2020. Through Institute of Epidemiology, Disease Control and Research (IEDCR) press conference, they informed that three patients infected with Corona virus were found in Bangladesh on 8 March 2020. Among the first three people infected with the coronavirus are two men and one woman between the ages of 20 and 35. No patients infected with the novel coronavirus were found between March 9 and March 15, 2020, but 5 more patients infected with the novel coronavirus were confirmed on March 16, according to a press conference by IEDCR. A person infected with the COVID-19 died on 18 March in Bangladesh. Due to this epidemic, like other countries of the world, Bangladesh has been forced to go for compulsory lockdown like home quarantine, social distancing, and local or international flight bans, etc. for slowing down the spreading. In Bangladesh COVID-19 infection case is increasing day by day. Then the government declared all school, colleges, universities in Bangladesh closed on 18th March. And a week later, from March 26, the government declared all government and private and semi-government offices remain closed. As of April 14, the number of patients

infected with the COVID-19 had reached 1,012 and it has already spread to most districts of Bangladesh.

In the beginning, the rate of transmission of COVID-19 in Bangladesh was slow. The infection situation continues to rise rapidly in mid-May. The outbreak of the COVID-19 in Bangladesh reached a critical stage in June. The number of new patients daily has decreased over the last two to three months. But at the same time the number of infections and detections is getting lower than before. As a result, it is difficult to say exactly what will happen to the corona virus in Bangladesh in the next two-three months. Outbreaks appear to be exacerbated around the world about 10 months ago. Even after such a long time, no significant change was noticed in the Corona situation. Like the waves of the sea, once it decreases, it increases again.

The first wave of corona virus in Bangladesh is not over yet, the number of daily detections has not dropped below 10 percent and the community transmission did not stop. As a result, before the end of the first wave of Corona virus, Bangladesh may start a second wave like other countries in the world. Due to the dry environment in winter, people are at higher risk of spreading the pandemic virus through sneezing and coughing. In addition, as the number of respiratory patients in our country increases in winter,

children and the elderly have more problems. In winter, the elderly and children will be risk than before. As a result, children and adults need to be more careful. Until the corona vaccine is available, all people should maintain social or physical distance. Adherence to hygiene rules is the only way to prevent infection, and if it fails, many people will die as health and medical systems collapse.

In Machine learning (ML) can help us to solve problems caused by the COVID-19 epidemic. This is not just a novelty however it will affect the information and innovation of the people who use it. Machine learning (ML) works by predicting designing in chronicled training information. We can practice from a situation and apply these to COVID-19 situations, marking on our progressive information to invent the best discussion on what may action what may predict.

1.2. Motivation

In this part, we will introduce the COVID-19 problem. There are very few papers COVID-19 related tasks that have been published in various international journals. COVID-19 pandemic is very challenging of our daily lives. Machine learning (ML) regression algorithm help to predict the corona virus cases in Bangladesh. In [2] ML can be utilized huge data to predict the breakout of the disease and they used remote cloud model to prediction the corona virus. In paper, we have discussed about the number of COVID-19 cases, the affected cases, tested cases, recovered cases death cases, and the fall of the number of cases in near future situation in Bangladesh and the date when various countries may expect the pandemic to finish.

1.3. Objective of the Study

- To ensure early identification of cases among the people.
- To control the rapid action of COVID-19.
- To create awareness among general people and identify infection.
- To predict the result of affected, death, recover with COVID-19.
- To forecast the future situation of COVID-19 in Bangladesh.

2. Literature Review

In[3] this paper, sheds light on COVID-19 preparedness of the different factor in Bangladesh. In Bangladesh the healthcare system by examining the spatial distribution of isolation beds across districts and divisions. To forecasting the number of ICU units that may be required in the short term. It is analyzing the availability of frontline healthcare workers to combat the pandemic. In[4] this paper, discuss about the spread of COVID-19 pandemic using machine learning algorithm. It also come up with corona virus cases for global epidemic prediction and simple statistical models. It is used two types machine learning algorithm like ANN (Artificial Neural Network) and MLP (Multi-Layer Perceptron) to predict global COVID-19 cases and future situation of the world. In[5] Machine learning models used to predict the corona virus. It also talks about future situation and possible global threats, clinical procedures and treatments. In[6] Linear regression models have used to predict the COVID-19 epidemic in India and it also predict the near future situation of India to use the LR models. In[7] this paper they discuss about the COVID-19 cases. This paper, highlights most important research questions

concerning virus transmission, asymptomatic and virus shedding, diagnosis, treatment, vaccine development, origin of virus and viral pathogenesis, etc.

3. Methodology and Dataset

3.1. Introduction

In[8] Machine learning is a sub set of Artificial Intelligence. It enables a machine to automatically learn from data improve the performance from past experiences and predict things without being explicitly program. We have study ML algorithms will help us to improve our models. These algorithms have the ability to learn from past mistakes and minimize the error over time.

3.2. Linear Regression Model

In[9] Regression analysis is a statistical method. This model displays the relationship between a dependent (Target) and independent (Predictor) variables with one or more independent variables. It helps us to guess how the value of the dependent variable is changing corresponding to an independent variable are held fixed. It also predicts continuous or real values. LR is one of the simple ML algorithms. It is statistical method that is used for predictive analysis. LR model depicts the relationship between a dependent(Y) and one or more independent(X) variables that is called linear regression. In this work, we have used linear regression model for prediction of COVID-19 cases. It is a simple model which is used to find the relation between a dependent and an independent variable. It uses the value of intercept and slope to predict the output variable.

LR model provides a slop straight line representing the relationship between the variables. Consider the below figure.

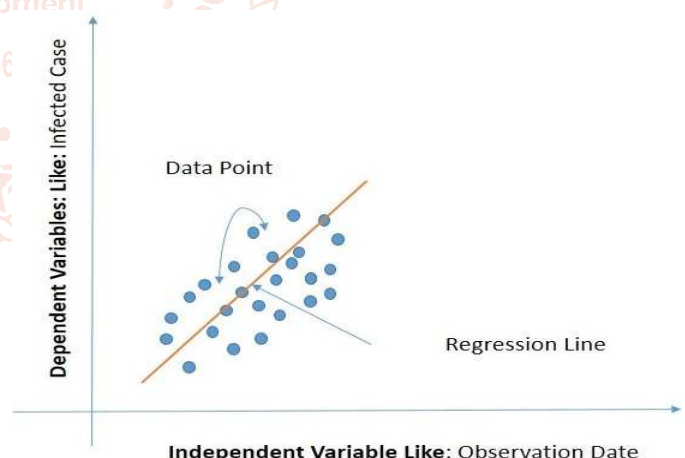


Fig.3.1: Linear Regression

Mathematically, we can represent a linear regression as: $Y = a_0 + a_1x + \epsilon$. Here, y = Dependent Variable (Target Variable), x = Independent Variables (Predictor variables), a_0 = intercept of the line, a_1 = Linear regression coefficient, ϵ = random error. The values for x and y variables are trainging datasets for LR model representation.

3.3. Dataset for this study

The goal of this section is to provide the study area, dataset source, dataset used, etc. We also study the technique of preprocessing the dataset. Finally, we have prepared various datasets to get desire goals.

3.3.1. Data Source

During a pandemic, humanitarian crisis may arise in a developing country like Bangladesh we had taken expert suggestions, consultations from a different group of people. Google form can be used to conduct the survey. An online database system can have applied by expert group of Bangladesh. These may be shared through email, Facebook, Messenger, LinkedIn, WhatsApp, etc. We have used data from online papers and different Medias. We have used data which were obtained from different PCR(Polymerase Chain Reaction) lab. There are many TV channels telecast news of the COVID-19 every day which may be one of the data source. There are different online journals or newspaper which helps us to collect accurate data. These data were collected day by day.

3.3.2. Data used

There are many types of data which can be used in this regards. But we have used numerical data. The total case of affected people, amount of tests, number of death, number of recovers, observation date is known as numerical data. We can store these types of data in a database. Data can be store in access file, excel file, CSV file. The data can be obtaining from those file according to their formats.

4. Result Analysis

In this section, we have discussed about COVID-19 situation in Bangladesh. In this work, Results from Linear regression analysis is used five types parameters in this regard like these arguments are tested, confirmed, recover, and death cases as input and predicted the result.

3.3.3. Study Area

Novel Corona virus is increasing day by day in Bangladesh. It is increasing rapidly in Dhaka City. COVID-19 affects different peoples and different ways. Most common symptoms are fever, dry cough, and tiredness. The less common symptoms are aches and pains, sore throat, diarrhea, headache, loss of test or smell, a rash on skin or discoloration of fingers, or toes, etc. Forty-five areas in the capitals, eleven areas in the ports city of Chittagong, the entire Gazipur District, and few areas of Narayanganj District, and Narsingdi District have been identified as red zones. On 1 June 2020 the Government decided to divide the entire country into the three zones- red, yellow, and green. These were named based on affected people with COVID-19. Almost in every hospital in Bangladesh are takes steep to test this virus. Different PCR lab are set in the different private and government hospitals. By the following different symptoms, the test case is done. Different public University are also engaged to test the symptoms of the people. Bangladesh government has taken steps to reduce this problem.

Descriptive Statistics

	N	Minimum	Maximum	Sum	Mean	Std. Deviation
ObservationDate	261	08-MAR-20	23-NOV-20	41730506 2...	16-JUL-20	75 11:43:18....
DailyTested	261	4	18498	2667064	10218.64	5349.876
DailyInfected	261	0	4019	449905	1723.77	1102.692
DailyDead	261	0	64	6409	24.56	15.387
DailyRecovered	261	0	17533	364606	1396.96	1494.739
Valid N (listwise)	261					

Fig.4.1: Descriptive Analysis

In this analysis, we see that COVID-19 data from 8th March 2020 to 23rd November 2020 is discussed here. We have also seen that highest tested case is one day 18498 and lowest tested case is 0, and highest infected, highest dead case, and highest recovered case are 4019,64, 17533 respectively in Bangladesh.

Now we have displayed the COVID-19 tested case in Bangladesh from 8th March 2020 to 23rd November 2020.

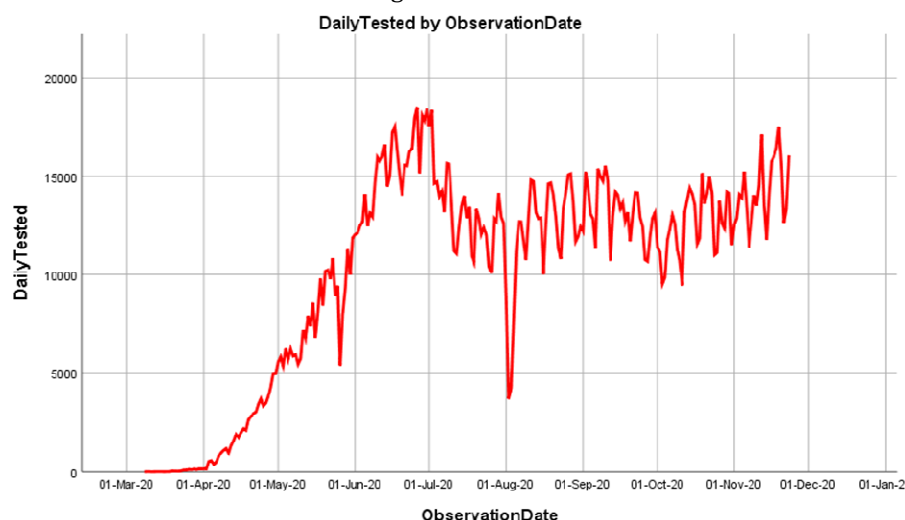


Fig.4.2 Daily Tested Cases by Observation Date

From the above Fig.4.2 it is understood that last March, April and May were comparatively less compassionate test in Bangladesh but June to July comparatively more compassionate test every day. In Bangladesh has so far conducted the highest number of sample tests in 18498 one day. At present, experts in the country and abroad are urging to increase the number of COVID-19 patient identification tests to control the COVID-19 situation, but in Bangladesh it has decreased. COVID-19 tests have dropped 26 percent in the past month. A review of the data shows that the number of coronavirus cases in Bangladesh has decreased by 1 lakh 23 thousand 750 in the fifth month as compared to the fourth month. Experts say an estimated 20,000 tests a day were needed at this time to understand the true picture of infection in Bangladesh. From the beginning, Bangladesh lagged far behind in terms of population ratio. As of mid-July, Bangladesh was ranked 145 out of 215 countries and territories. From the graph above we can see that in August-September-October the corona test is much less and again in November COVID-19 test is increasing comparatively every day. On the other hand, experts say there is no alternative to increasing the rate of sample testing to deal with the second stage of infection. This means that the sooner an infected person is identified by increasing the number of tests, the easier it will be to treat his or her family and isolate others in the family and take other preventive measures.

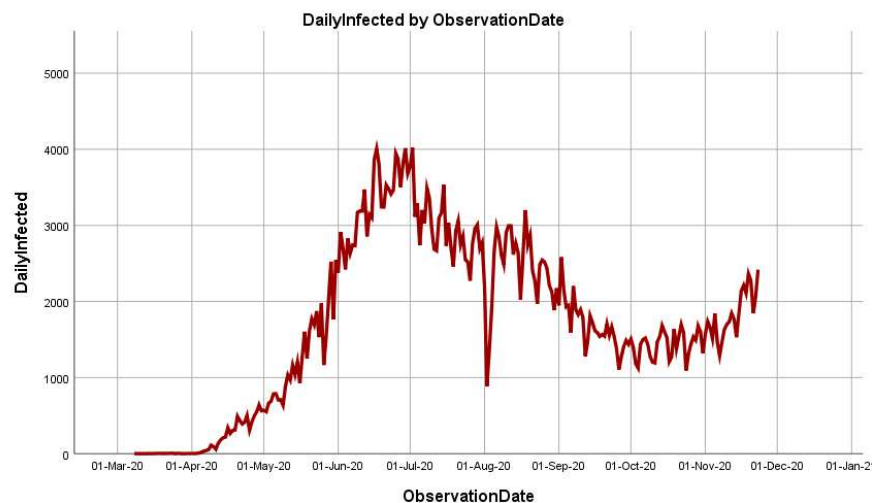


Fig.4.3 Daily infected case by Observation Date

The first case of COVID-19 disease in Bangladesh was reported on March 8, 2020, after originating from Wuhan, Hubei Province, China. The epidemic had been declining in Bangladesh since mid-May and intensified in June-July. In August-September, the number of corona virus cases in Bangladesh was much less than in June-July. The number of patients infected with COVID-19 was found to be 4019 maximum patients in one day. In many European countries, the rate of infection is declining and rising again, which is called the second wave. When the infection rate drops to 5% and stays the same for at least three weeks, according to the WHO (World Health Organization) guidelines, it can be called the first wave outbreak of the Corona virus. So far, an average of 1723.77 patients are being infected in Bangladesh every day. From the Fig.4.3 above we can see that in the third week of November in Bangladesh, the number of patient deaths has suddenly increased again. Earlier, for more than two months, the number of patients diagnosed with the virus and the death rate has been steadily declining. However, compared to the sample test, the rate of infection detection has never dropped below 10 percent in Bangladesh. Since the withdrawal of the lockdown in Bangladesh, the social distance has been largely indifferent to the issue of wearing masks and hand washing by disinfectants. Since the withdrawal of the lockdown in Bangladesh, the social distance has been largely indifferent to the issue of wearing masks and hand washing by disinfectants. The incidence of coronavirus and other winter viruses may increase at an alarming rate. At the beginning of the winter, the crowds in the tourist centers of the common people of Bangladesh are increasing abnormally where there is no health scourge. Where the Bangladesh government has not been able to decide to reopen the educational institutions by putting the lives of the students at risk, such neglect of the common man's epidemic could lead to a second wave of winter woes.

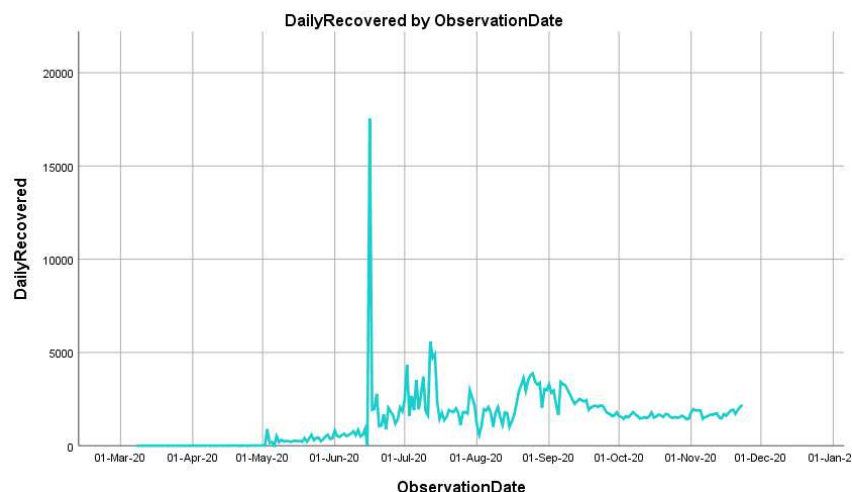


Fig.4.4 Daily recovered Case by Observation Date

In medical guidelines state that a person infected with the coronavirus can recover in 14 days if they have mild symptoms. Many times it can take up to 20 days to get rid of the virus. If there is a physical complication, it takes more time to recover. In the beginning if the infected patient tested negative for two consecutive days, he would be called a healthy patient out of compassion. As a result, the number of healthy patients from Bangladesh COVID-19 has been very slow in the beginning. Then the government moves away from this strategy and if corona virus is negative in one test, he seems to be a healthy patient from COVID-19. As a result, the number of healthy patients in Bangladesh is gradually increasing day by day. In above Fig.4.4 Suddenly one day it was seen that 17500 out of 33 patients recovered from corona virus together which is the highest number of healthy patients so far in one day. So far, Bangladesh is recovering on an average of 1382.44 per day from novel corona virus.

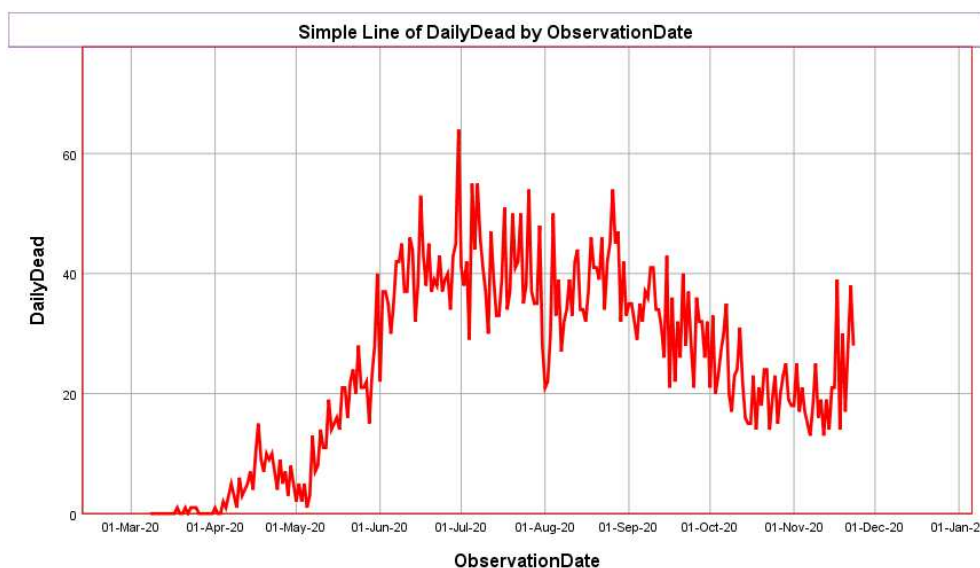


Fig.4.5 Daily Dead case by Observation

Through IEDCR's press conference, we know that the first covid-19 infected patient in Bangladesh was caught with corona virus on March 8, 2020. The news of the death of the first corona patient in Bangladesh was confirmed on 18th March 2020. Towards the end of May and the beginning of June-July-August, and September, the number of deaths due to Corona virus in Bangladesh continued to rise significantly. In mid-July, a maximum of 64 people died in a single day due to the mercy of Bangladesh, which is the highest number of deaths in a single day so far. We then notice from the Fig.4.5 above that the death toll from the corona virus was somewhat stable from the last week of September 2020 to the whole of October 2020 and the first week of November 2020, with an average of about 18 to 19 people dying from the virus every day but from the second week of November. So far, an average of 24.56 people has died of the corona virus every day in Bangladesh. Experts are already saying that the number of deaths due to corona virus in winter will gradually increase in Bangladesh. So in order to get rid of this virus, we should be aware and not make any concessions about it.

Forecasting COVID-19 in Bangladesh

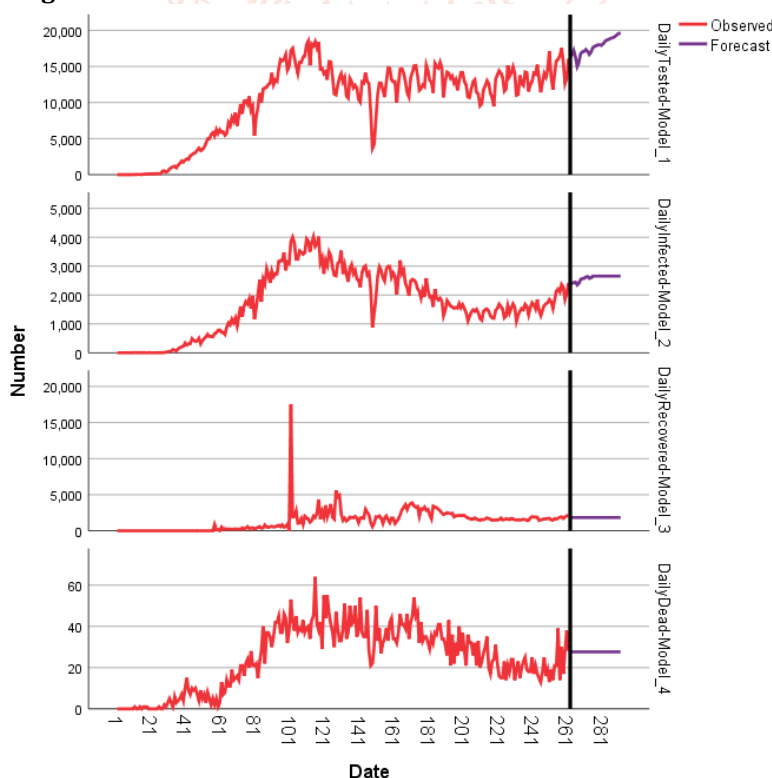


Fig.4.6: Observation and forecast Graph

ML forecasting is everywhere. At present people have been forecasting various topic like weather patterns, economic events, political events, sports outcomes, COVID-19 and more. Because we try to predict so many different events, there are a wide variety of ways in which forecasts can be developed. In the Fig.4.5 represent the forecasting the spread of COVID-19 in Bangladesh. Our aims how to predict the virus-affected cases, recovered cases, death cases, tested cases, and forecasting the future situation of Bangladesh and also forecast the number of upcoming cases for the next 30 days in Bangladesh. We notice from the above Fig.4.5 the machine learning forecasting graph shows cumulative tested case in Bangladesh will increase in next 30 days. In the next 30 days, corona sample test will be held at home average 16500 every day. In addition, the number of patients infected with the corona virus will increase in the next 30 days.

	Model											
	DailyTested-Model_1			DailyInfected-Model_2			DailyRecovered-Model_3			DailyDead-Model_4		
	Forecast	UCL	LCL	Forecast	UCL	LCL	Forecast	UCL	LCL	Forecast	UCL	LCL
262	16273	19365	13385	2411	2903	1920	1840	4182	-503	28	40	15
263	16616	20747	12835	2417	2994	1839	1840	4204	-524	28	40	15
264	17223	21968	12920	2422	3075	1770	1840	4225	-545	28	41	15
265	16325	21211	11932	2447	3129	1764	1840	4245	-566	28	41	14
266	14993	19878	10639	2349	3060	1638	1840	4266	-587	28	42	14
267	15565	20585	11087	2407	3145	1669	1840	4286	-607	28	42	13
268	16769	22133	11978	2541	3306	1776	1840	4307	-627	28	42	13
269	17039	22886	11857	2573	3396	1749	1840	4327	-647	28	43	13
270	17074	23379	11537	2573	3452	1695	1840	4347	-667	28	43	12
271	17355	23964	11574	2624	3554	1694	1840	4366	-687	28	43	12
272	17085	23856	11193	2642	3621	1663	1840	4386	-707	28	44	12
273	16656	23522	10714	2577	3602	1551	1840	4405	-726	28	44	11
274	16971	24035	10864	2592	3662	1522	1840	4425	-745	28	44	11
275	17601	24966	11240	2655	3768	1542	1840	4444	-765	28	45	11
276	17797	25475	11194	2655	3839	1471	1840	4463	-784	28	45	10
277	17843	25798	11034	2655	3905	1404	1840	4482	-803	28	45	10
278	18001	26194	11010	2655	3969	1341	1840	4501	-821	28	46	10
279	17993	26367	10873	2655	4029	1280	1840	4519	-840	28	46	9
280	17912	26443	10684	2655	4087	1222	1840	4538	-858	28	46	9
281	18128	26870	10737	2655	4143	1166	1840	4556	-877	28	47	9
282	18483	27482	10885	2655	4197	1113	1840	4574	-895	28	47	8
283	18650	27896	10867	2655	4249	1061	1840	4593	-913	28	47	8
284	18729	28198	10783	2655	4299	1011	1840	4611	-931	28	47	8
285	18862	28544	10760	2655	4348	962	1840	4628	-949	28	48	8
286	18951	28824	10710	2655	4395	914	1840	4646	-967	28	48	7
287	19010	29065	10638	2655	4441	868	1840	4664	-985	28	48	7
288	19184	29446	10656	2655	4486	823	1840	4681	-1002	28	48	7
289	19417	29907	10716	2655	4530	779	1840	4699	-1020	28	49	7
290	19571	30279	10709	2655	4573	736	1840	4716	-1037	28	49	6
291	19679	30591	10669	2655	4615	694	1840	4734	-1054	28	49	6

For each model, forecasts start after the last non-missing in the range of the requested estimation period, and end at the last period for which non-missing values of all the predictors are available or at the end date of the requested forecast period, whichever is earlier.

Fig. 4.7 Forecast tested, confirmed, recovered, and death cases in Bangladesh.

We have also seen that, In Bangladesh on an average 28 people die every day in Bangladesh due to COVID-19. From the forecasting table we can see that Bangladesh will recover an average of 1840 people from coronavirus every day.

Conclusion: In this work assessed the capability of machine learning model LR for predicting the COVID-19 outbreak in Bangladesh. The models display hopeful out come in terms of predicting the forecast without the doubt that epidemiological models need. ML models, as a substitute to epidemiological models, represent potential in predicting COVID-19. No well-tested drug for the compassion virus has been discovered so far. So apparently we have no choice but to maintain the only lockdown and social distance and public awareness. Experts now say, it is important to plan and plan ahead if the infection situation deteriorates rapidly. In the last six and seven months, many people in Bangladesh have lost their jobs and suffered financially. So in order to get rid

of the corona virus, we need to pay attention to public awareness and tighten the rules of hygiene.

References

- [1] M. H. Rahaman Khan and A. Hossain, "COVID-19 Outbreak Situations in Bangladesh: An Empirical Analysis," pp. 1–15, 2020.
- [2] S. Tuli, S. Tuli, R. Tuli, and S. S. Gill, "Predicting the growth and trend of COVID-19 pandemic using machine learning and cloud computing," *Internet of Things*, vol. 11, p. 100222, 2020.
- [3] T. Howlader and H. R. Khan, "Battling the COVID-19

- Pandemic Is Bangladesh Prepared," *Medrxiv*, pp. 1–20, 2020.
- [4] S. F. Ardabili *et al.*, "COVID-19 Outbreak Prediction with Machine Learning," *SSRN Electron. J.*, 2020.
- [5] K. C. Santosh, "AI-Driven Tools for Coronavirus Outbreak: Need of Active Learning and Cross-Population Train/Test Models on Multitudinal/Multimodal Data," *J. Med. Syst.*, vol. 44, no. 5, pp. 1–5, 2020.
- [6] R. Sujath, J. M. Chatterjee, and A. E. Hassanien, "A machine learning forecasting model for COVID-19 pandemic in India," *Stoch. Environ. Res. Risk Assess.*, vol. 34, no. 7, pp. 959–972, 2020.
- [7] K. S. Yuen, Z. W. Ye, S. Y. Fung, C. P. Chan, and D. Y. Jin, "SARS-CoV-2 and COVID-19: The most important research questions," *Cell Biosci.*, vol. 10, no. 1, pp. 1–5, 2020.
- [8] Y. Baştanlar and M. Özuysal, "Introduction to machine learning," *Methods Mol. Biol.*, vol. 1107, pp. 105–128, 2014.
- [9] N. Altman and M. Krzywinski, "Simple linear regression," *Nat. Methods*, vol. 12, no. 11, pp. 999–1000, 2015.

