ISSN No:-2456-2165

Covid-19 Short-Term Forecasting in Bangladesh Using Supervised Machine Learning

Md Mahbubur Rahman[1]
Computer Science and Engineering Discipline
Khulna University
Khulna – 9206, Bangladesh

Farhana Tazmim Pinki[2]
Computer Science and Engineering Discipline
Khulna University
Khulna – 9206, Bangladesh

Abstract:- COVID-19 is a human-to-human transmissible virus responsible for damage to the human body, and people died all over the world. Bangladesh was affected by COVID-19 on March 8th, 2020. During the pandemic, people and the government struggled to prevent transmission due to an inadequate supply of vaccines and healthcare equipment. Therefore, it is essential to understand the upcoming infected cases for several days. That may help people and the government make pre-decision before the pandemic to save live. In this paper, we proposed a COVID-19 short-term forecasting model using Linear Regression (LR), Least Absolute Shrinkage and Selection Operation (LASSO) Regression, and Support Vector Regression (SVR) to predict the next seven days of COVID-19 infected cases in Bangladesh during the pandemic situation. Here we considered data from 8th May 2021 to 21st July 2021. We analyzed different past data volumes for the model to understand the impact of past data in the model. The result reveals that Support Vector Regression (SVR) performance was better than LR and LASSO in all aspects with high accuracy. The performance also indicated that the high volume of past data helps to increase prediction accuracy.

Keywords:- *COVID-19, Short-term Forecast, New infected cases, Bangladesh, Supervised Machine Learning, LR, LASSO, SVR.*

I. INTRODUCTION

The 21st-century world is facing a new pandemic caused by COVID-19. In December 2019, the first COVID-19 case has initially been identified in Wuhan, China [1]. According to the World Health Organization, COVID-19 symptoms include fever, dry cough, shortness of breath, fatigue, bodily pain, sore throat, etc. [2]. The high human-to-human transmission nature of the disease has been progressively unpredictable and spread rapidly worldwide. As of January 30, 2023, the world has reported over 200 million confirmed cases and over 4 million deaths due to COVID-19 [3]. The pandemic also has impacted global health and economies, causing economic downturns and disruptions to daily life. On March 8, 2020, Bangladesh received its first case report; since then, the number of cases has rapidly escalated. The official statistics of the Institute of Epidemiology, Disease Control and Research (IEDCR) show 2,037,386 infected cases and 29,441 confirmed deaths till 19th January 2023 [4] [5]. In the early stage of the pandemic, Bangladesh’s government adopted strict actions, including lockdowns, stay-at-home recommendations, mass gathering cancellations, school, and non-essential shop closures, banned flights, sealed international borders, and export-import transportation. These actions have significantly impacted the economy, with many businesses struggling and many people losing their jobs. The healthcare system in Bangladesh has also been under significant strain, with hospitals and healthcare facilities facing shortages of personal protective equipment, oxygen, and other essential supplies. The country has struggled to ramp up its testing, and there are concerns about the accuracy of its COVID-19 data. Bangladeshi government carried out COVID-19 mass vaccination campaigns to swiftly contain the disease outbreak and completed 346,605,580 vaccine doses in 3 steps till 15 January 2023 [4]. An inadequate supply of vaccines making challenging to normalize the situation. Also, the virus mutation prevents the vaccine from being fully effective in the human body. Population density is another concern for maintaining the vaccine necessary measures and proper doses of vaccines. In that case, it would be much more effective if the government and healthcare industries could know the upcoming transmission in advance. From the beginning, Bangladesh has faced three pandemic waves of COVID-19. The first wave started on 8th March 2020. The Beta variant's second wave started in mid-March 2021, and the third wave of the Delta variant started on 08 May 2021. Each time a new variant of COVID-19 made a new pandemic situation. It rapidly increases confirmed and death cases [6] [7]. Significant increases in transmission affect human daily life, and healthcare organizations also struggle to prevent transmission. Therefore, measuring the downturn of COVID-19 transmission in that situation is almost impossible. So, it is essential to know and understand how COVID-19 can transmit shortly to make a valuable decision to prevent COVID-19 transmission. In this regard, Short-term Forecast on COVID-19 transmission helps understand and predict upcoming scenarios, manage public health efficiently, and optimize healthcare resources in a pandemic. Also, the healthcare system will get prepared and provide the best support in the pandemic that may help save lives.

II. LITERATURE REVIEW

Presently the application of machine learning is an advanced approach for solving the problem in several fields such as business, industrial, scientific, etc. The Performance of machine learning algorithms in future forecasting using existing data has promising results with high accuracy.

Several Researchers have checked and performed future forecasts using different machine-learning methods in Bangladesh. From traditional methods to deep learning, research is always an open interest. [8] Developed a cloud-based model for forecasting COVID-19 infected cases in Bangladesh. Several machine learning models were employed to forecast the next seven days scenario by training the previous sample data. They considered Bangladesh's COVID-19 infection and fatalities data from the beginning to 10th June 2020. Total data was split into 35 subsets, and each subset was trained and predicted individually. All models were evaluated by RMSE, MAE, and R^2 values. When it came to both infection and death cases, the prophet model produced the best results. Another study by [9] demonstrates how machine learning models can be used to predict COVID-19 virus transmission in the upcoming days. Linear regression, support vector machines, lasso regression, and exponential smoothing methods were used to forecast the next ten days of COVID-19 infected cases. The result proves that the ES performs best among all models. LR and LASSO perform better for newly confirmed, death, and recovery cases. SVM exhibits low performance in all prediction scenarios for the given dataset. Another study by [10] evaluated the effectiveness of convolutional as well as recurrent neural networks in predicting COVID-19. Long short-term memory (LSTM), gated recurrent unit (GRU), convolutional neural network (CNN), and multivariate convolutional neural network (MCNN) were used to perform probable future outbreaks in Brazil, Russia, and UK. The Study highlighted that CNN performed better than other algorithms when making forecasts using a few features and historical data. [7] looked into Bangladesh's COVID-19 third-wave data and forecasted the COVID-19 situation. The data from Mar 01, 2021, to July 31, 2021, was used to generate the Auto-Regressive Integrated Moving Average (ARIMA) prediction model. The finding shows that the next three months from 1st August are frightening in Bangladesh. [1] Compare the prediction accuracy between Auto-Regressive Integrated Moving Average (ARIMA) and eXtream Gradient Boosting (XGBoost) in Bangladesh. 633 days of the daily conformed dataset and death cases were collected from the Directorate General of Health Service (DGHS) and IEDCR. The final finding shows that the ARIMA model performs better than the XGBoost model in predicting confirmed and death cases in Bangladesh. Most of the work has been done based on a specific range of datasets where they did not contribute to the explanation of the past data volume for higher accuracy in forecasting. We aim to investigate the past data volume effect on model accuracy using different data scenarios on COVID-19 transmission and create a short-term forecasting model using linear regression (LR), Least Absolute Shrinkage and Selection Operation (LASSO) regression, and support vector regression (SVR). Many academics have researched using a short range of the past dataset for forecasting COVID-19 up to 7-30 days [6] [7]. “Forecasts are more accurate for shorter than longer time horizons. The shorter the time horizon of the forecast, the lower the degree of uncertainty. Data do not change very much in the short run” [11]. [12] also informed that “According to NOAA (National Oceanic and Atmospheric Administration U.S. Department of Commerce), a five-day forecast can accurately predict the weather 90 percent of the time. That drops to 80 percent for a seven-day forecast and then down to 50 percent for a 10-day forecast” for weather data. In our case, we attempt to forecast up to 7 days for higher prediction accuracy for COVID-19 future forecast.

III. RESEARCH METHODOLOGY

This section will describe a detailed overview of the proposed methodology: data collection, preprocessing, and future forecasting of COVID-19 transmission in Bangladesh.

A. *Dataset Description*

Nowadays, the online source is the most authentic and acceptable source for COVID-19 data. Many Academic researchers use research data from online sources [9] [13]. COVID-19 data was obtained from the GitHub repository of Johns Hopkins University's Center for System Science and Engineering for this work [14]. The dataset folder name (csse_covid_19_time_series) includes the COVID-19 daily time series data report, and the data updated frequency is once a day. The dataset has six attributes. However, Bangladeshi confirmed case data had been used for our work.

B. *Data Preprocessing*

Fig. 1 shows the total COVID-19 infected cases scenario from 2020-01-22 to 2022-01-18, where the highest alarming situation happened on 2021-07-28. The graph shows that infected cases increased very fast from 2021-05-08 to 2021-07-28. Therefore, in this study, we used COVID-19 infected cases data from 2021-05-08 to 2021-07-21 to understand future infected cases in a pandemic. We extracted daily Bangladeshi confirmed cases data along with the continuous date value from 8th May 2021 to 21st July 2021. 75 days of confirmed cases data were selected to predict COVID-19 transmission. The data has split into three different datasets for the forecasting models. The datasets contain past data volumes of 45, 60, and 75 days. The COVID-19 dataset is supervised data. Moreover, we checked our data's inconsistencies and completed all data processing work using the Python programming language and the Jupiter notebook platform.

C. *Methodology*

We are proposing three regression methods Linear Regression (LR), Least Absolute Shrinkage and Selection Operation (LASSO) Regression, and Support Vector Regression (SVR) methods to create the COVID-19 prediction model. For the COVID-19 prediction analysis, the observations dataset has been recorded over daily time frequency. The time step feature has been constructed for the forecasting models input feature using day numbers derived directly from the dataset date index column. Moreover, confirmed cases are the target feature. Prepared data has two sets: 70% are training data, and 30% are testing data for all regression models. *Fig. 2* shows the proposed workflow.

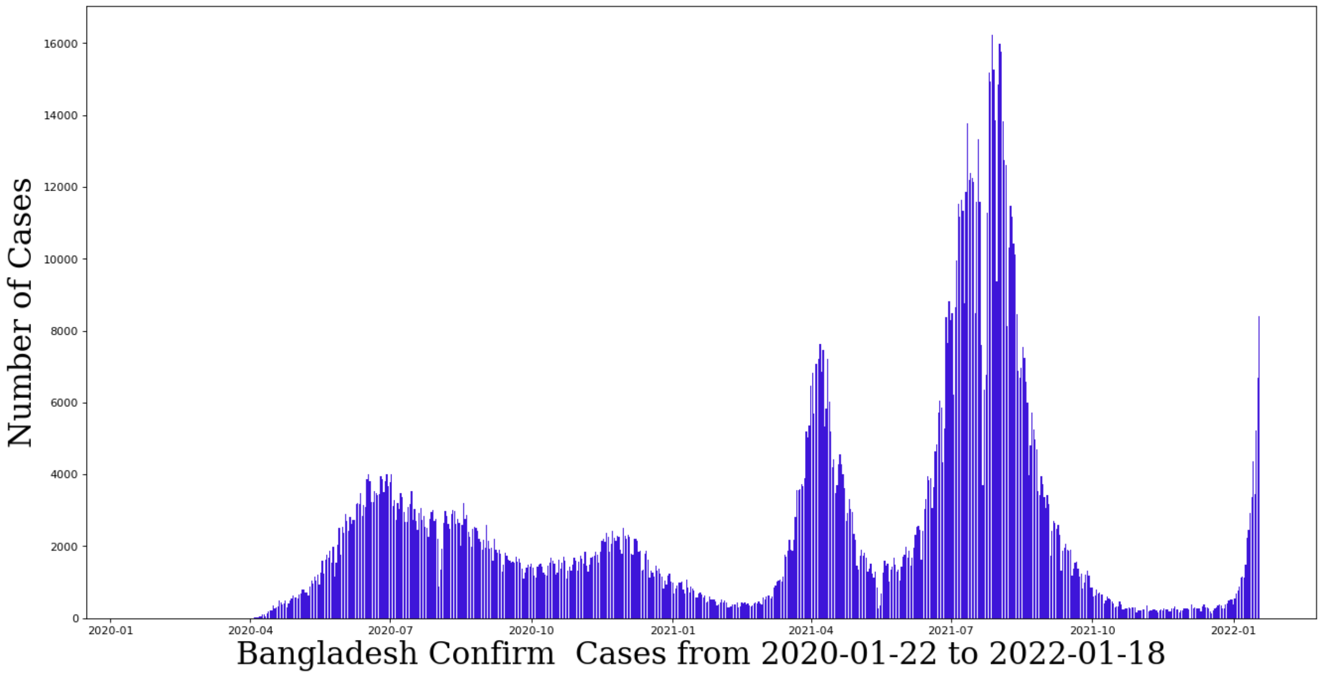

Fig. 1. Bangladesh daily new infected cases data from 2020-01-22 to 2022-01-18

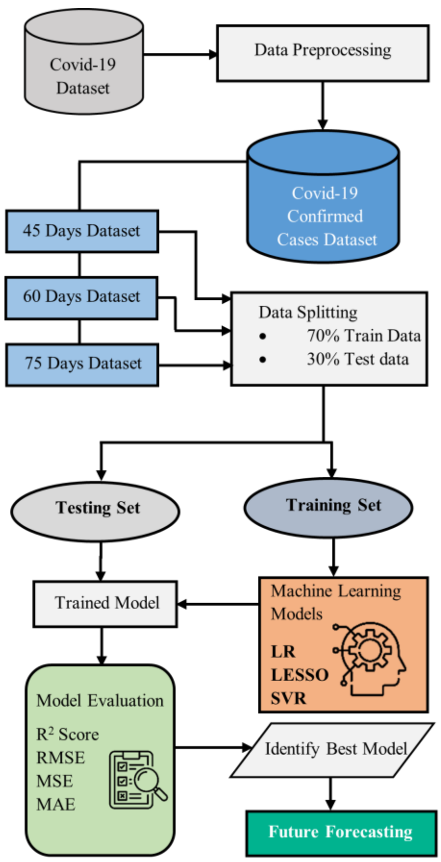

Fig. 2. Proposed workflow.

- Linear Regression:

Linear regression is a popular prediction analysis technique in machine learning. This technique calculates how strong the relationship between two variables is. The aim is to make a best-fit linear line for the observed data with minimal error. Observations depend on two variables, x and y where x is the independent variable, and y is the dependent variable. Below is the linear regression equation showing how the independent variable y is related to x.

$$y = \beta_0 + \beta_1 x + \varepsilon \quad \textbf{(1)}$$

$$E(y) = \beta_0 + \beta_1 x \quad \textbf{(2)}$$

Here, y is the dependent variable, x is the independent variable, β_0 represents the y-intercept, β_1 is the regression coefficient, and the estimate's error term is called epsilon ε. The total error rate should be minimized to get the best-fit line. In that case, the difference between the actual and predicted data point values should be minimal. The minimization equation can be defined as:

$$c = \frac{1}{n}\sum_{i=1}^{n}(pred_i - y_i)^2 \quad \textbf{(3)}$$

Here, c is called a cost function, $pred_i$ is the predicted value, y_i is the actual value, and n is the number of the total data point.

- Lasso Regression:

"Lasso" stands for Least Absolute Shrinkage and Selection Operation. When linear regression cannot give the required best-fit line or the predicted outcome is overfitted, lasso regression comes here to reduce overfitting and improve the prediction accuracy. The shrinking of the extreme coefficient values toward the central value makes the model stable for overfitting reduction. The L1 regularization technique adds a penalty term equal to the absolute value of the coefficient's magnitude. The model automatically penalizes the coefficient value equal to or close to zero. That means the feature is unable to improve the best-fit line. The coefficient magnitude can be set to zero or equal to zero. The above process makes the model sparse with the

ISSN No:-2456-2165

selected coefficient. The goal of the model is to reduce the following:

$$\sum_{i=1}^{n}(y_i - \sum_{j} x_{ij}\beta_j)^2 + \lambda \sum_{j=1}^{p}|\beta_j| \quad \textbf{(4)}$$

or Similarly

The sum of square residuals + λ |slope|

Here, λ is a tuning parameter for L1 Regularization where $0 \leq \lambda \leq \infty$ and λ |slope| is the penalty term. We aim to find a best-fit line with low bias and high variance.

- Support vector regression:

Support vector regression is well known for regression-based model prediction and analysis. This method is used for both linear and non-linear data. Most linear regression tries to minimize the sum of the residual square to get the best-fit line where SVR creates a hyperplane and two parallel marginal planes to find out a soft marginal space for the given dataset. The main aim is maximizing the marginal plane distance, minimizing the error rate, and creating a hyperplane for the best-fit line. SVR can set the acceptable error rate using ϵ for the model and find the best-fit line. If any data point falls outside of ϵ, SVR can still calculate the deviation from the margin by ξ. The final objective function and constraints are as follows:

Minimize:

$$MIN\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}|\xi_i| \quad \textbf{(5)}$$

Constraints:

$$|y_i - w_i x_i| \leq \varepsilon + |\xi_i| \quad \textbf{(6)}$$

Here, $\|w\|$ is a weight vector, C is the marginal distance from the hyperplane where $-\varepsilon \geq c \leq \varepsilon$, ξ is the distance between the marginal plane to an outside data point, y_i is the actual value, and $w_i x_i$ predicted value. *Fig 3* shows the Illustrative Example of the support vector machine. Nonlinear input data transform into n-dimension using a nonlinear function called kernel function. The basic understanding is that the input space is transferred from low dimension to higher dimensional space. This way, data will be linearly separable in the new space using a mapping technique. The kernel technique helps us to find a hyperplane with high accuracy and low error. We will check SVR polynomial kernel methods with different tuning parameters and finally pick the best model for the given dataset.

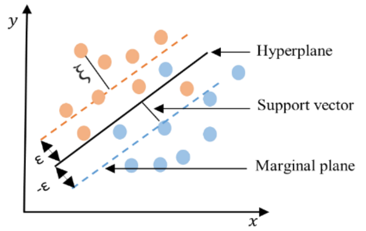

Fig. 3. Support vector Machine

D. Evaluation Parameters:

The study will evaluate all the learning models using the most widely used evaluation metrics. Model's performance will be measured by R-squared (R^2), Mean Square Error (MSE), Mean Absolute Error (MAE), and Root Mean Square Error (RMSE).

- R-squared

R square is a statistical measure that determines the goodness of fit in a regression model by comparing the residual sum of squares (SSres) with the total sum of squares (SStot). R^2 ideal value is 1. The regression model value closer to 1 indicates a better fit of the model.

$$R^2 = 1 - \frac{SS_{RES}}{SS_{TOT}} \quad \textbf{(7)}$$

$$= 1 - \frac{\sum_i(y_i - \hat{y}_i)^2}{\sum_i(y_i - \bar{y}_i)^2}$$

Here, y_i is actual data point, $\hat{y}_i$ predicted data point, and $\bar{y}_i$ is the predicted mean value.

- Mean Square Error

Mean Square error determines the average distance between the data point and regression line by squaring the distance value. Square is important to replace the negative value sign and gain more weight to a larger difference. A Smaller value close to 0 of MSE indicates a better fit of the model, whereas a high value means a high error rate.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2 \quad \textbf{(8)}$$

n is the total number of data points, y_i is the actual value and $\hat{y}_i$ is the predicted value of the data point.

- Root Mean Square Error

RMSE is another error metric for numerical prediction. This process calculates the standard deviation of the residuals. Standard deviation σ measures the spread of data points around the mean, and the residual prediction error is measured by calculating the distance between the predicted value and the actual value. RMSE's outcome is the average magnitude of inaccuracy. A lower error value close to 0 indicates a better fit of the model. The formula of RMSE follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y}_i)^2} \quad \textbf{(9)}$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2} \quad \textbf{(10)}$$

n is the total number of data point, y_i is the actual value, $\hat{y}_i$ is the predicted value of the data point, x_i is the actual data point, μ is the mean of the actual data point.

- Mean Absolute Error

Mean absolute error determines the mean difference between the actual and predicted values of a model's total variance. MAE score initiated from 0 to infinity, where a low score indicates the improved performance of the anticipated model. The formula is as follows:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y}_i| \quad \textbf{(11)}$$

n is the total number of data points y_i is the actual value, $\hat{y}_i$ is the predicted value.

E. Hyperparameter Tuning

Sometimes, the default settings of the regression model did not show the best result. Therefore, parameter tuning is essential to improve the results. In our cases, several parameters were tuned to get the best result using LR, LASSO, and SVR. Table 1 shows the hyperparameter tuning for LR, LASSO, and SVR in all different data scenarios.

Table 1. The Hyperparameters of best-performing regression models.

Models	**Hyperparameter Tuning**		
	45 Days Dataset	60 Days Dataset	75 Days Dataset
LR	Random State = 13	Random State = 58	Random State = 3
LASSO	Random State = 13	Random State = 58	Random State = 3
SVR	Random State = 3, kernel="poly", C=0.3, gamma=0.3, degree=3	Random State =7, kernel="poly", C=0.3, gamma=0.3, degree=3	Random State =5, kernel="poly", C=0.3, gamma=0.3, degree=3

IV. RESULTS AND DISCUSSION

This work uses machine learning techniques to create a short-term forecasting model for daily new infections caused by the COVID-19 virus. The dataset in this study included the frequency of infections that occur each day in Bangladesh. From Fig 1, data shows a significant increase in new variants, which is alarming for the country. This study attempts to develop a forecasting model that can predict the upcoming 7 days of newly infected cases by analyzing the previous infected cases data. Three different datasets (45 days, 60 days, and 75 days dataset) have been used for the model. The study also measured the impact of past data volume on future forecasting. Three machine learning models, LR, LASSO, and SVR, were used to forecast the upcoming 7 days of newly infected cases.

A. Forecasting using 45 days Dataset

The study predicted newly infected cases by analyzing the past 45 days infected cases dataset. According to the model's performance and results, SVR performs better than LR and LASSO models. LR and LASSO performances were almost the same, with equal R^2 scores. In comparison, SVR gives the best result for the given dataset. The outcome is displayed in Table 2.

Table 2. Model's performance on future forecasting for infected cases.

Model	R^2 Score	MSE	MAE	RMSE
LR	0.80	227815.61	363.27	477.30
LASSO	0.80	227854.33	363.31	477.34
SVR	0.93	79352.30	254.05	281.69

Figures 4, 5 and 6 show the performance of LR, LASSO and SVR models respectively. Figures showing the predicted infected cases have an increasing trend for the upcoming 7 days which is very alarming.

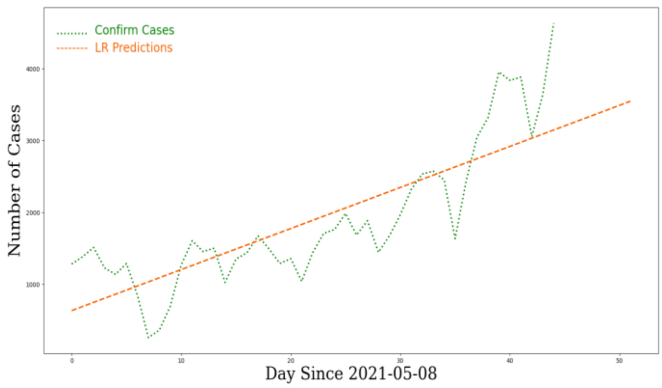

Fig. 4. Infected cases prediction by LR for the upcoming 7 days.

ISSN No:-2456-2165

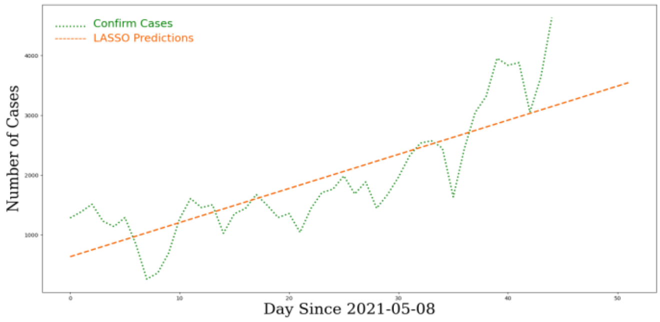

Fig. 5. Infected cases prediction by LASSO for the upcoming 7 days.

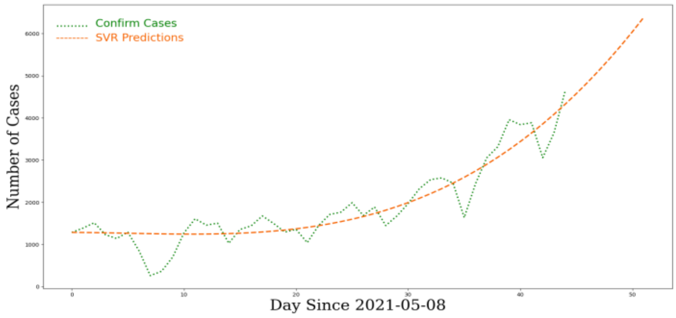

Fig. 6. Infected cases prediction by SVR for the upcoming 7 days.

B. Forecasting using 60 days Dataset

The Prediction model based on the past 60 days data has better performance than 45 days dataset. All models LR, LASSO and SVR performance increased and results show that the SVR model performs better than all other models. LR and LASSO performance were almost identical with equal R^2 scores. Table 3 displays the outcomes.

Table 3. Model's performance on future forecasting for infected cases.

Model	R2 Score	MSE	MAE	RMSE
LR	0.81	1320050.68	936.60	1148.93
LASSO	0.81	1320058.97	936.59	1148.93
SVR	0.97	170617.92	311.09	413.05

Figures 7, 8 and 9 show the performance of LR, LASSO and SVR models in graphs, respectively. Prediction indicates an increasing trend of infected cases for the upcoming 7 days. In comparison to all models, SVR performance is better in this situation.

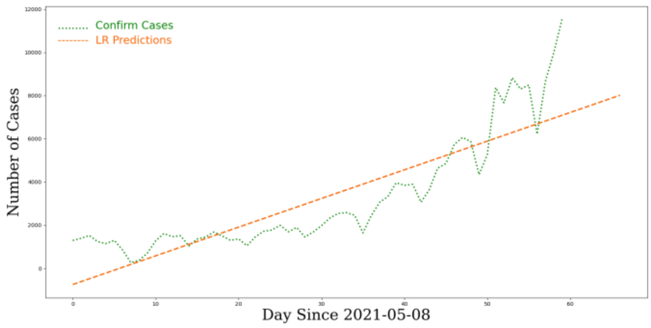

Fig. 7. Infected cases prediction by LR for the upcoming 7 days.

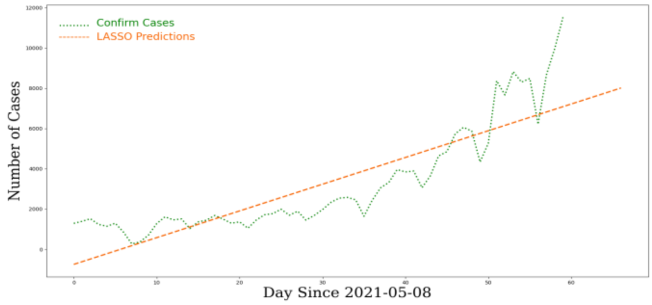

Fig. 8. Infected cases prediction by LASSO for the upcoming 7 days.

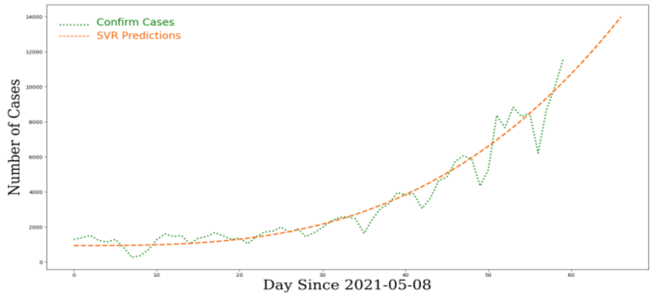

Fig. 9. Infected cases prediction by SVR for the upcoming 7 days.

C. Forecasting using 75 days Dataset

The Performance of all models using 75 days dataset was quite promising and model performance improved. The result shows that the SVR performs better than the others model. LR and LASSO performance also improved and both have almost identical performance with equal R^2 scores. Table 4 displays the findings.

Table 4. Models performance on future forecasting for infected cases.

Model	R^2 Score	MSE	MAE	RMSE
LR	0.88	1954244.90	1217.47	1397.94
LASSO	0.88	1954258.42	1217.47	1397.94
SVR	0.96	643586.37	593.40	802.23

Figures 10, 11 and 12 show the performance of LR, LASSO and SVR models respectively, where the model's prediction is quite promising. In the 75 days dataset, the SVR model gives better prediction accuracy. LR and LASSO performance also improved for the given dataset, where both LR and LASSO models have the same results with equal R^2 scores.

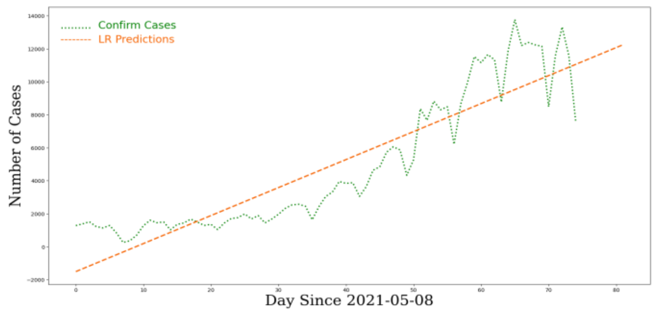

Fig. 10. Infected cases prediction by LR for the upcoming 7 days.

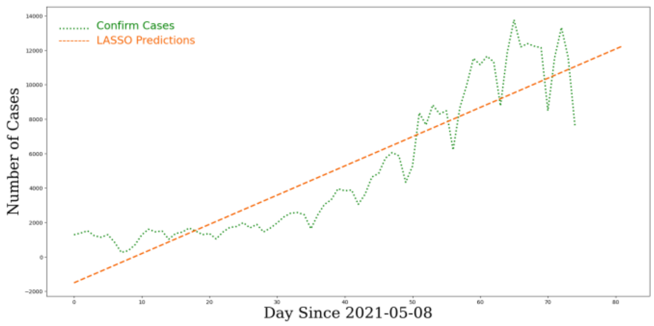

Fig. 11. Infected cases prediction by LASSO for the upcoming 7 days.

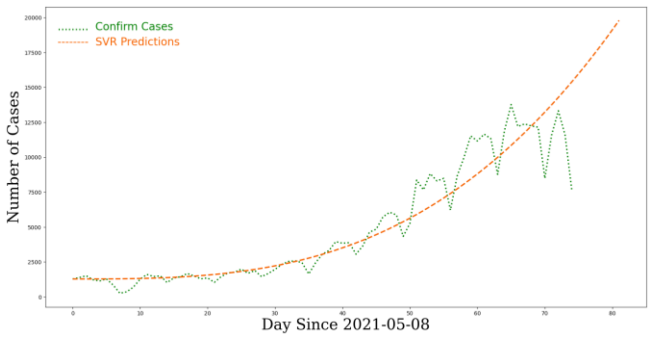

Fig. 12. Infected cases prediction by SVR for the upcoming 7 days.

The overall summary of the experimental results is shown in Table 5.

Table 1. Experimental Results for all regression models in different data scenarios.

Models	**Evaluation Parameters**	**Accuracy**		
		45 days	60 days	75 days
LR	**R^2 Score**	0.80	0.81	0.88
	MSE	227815.61	1320050.68	1954244.90
	MAE	363.27	936.60	1217.47
	RMSE	477.30	1148.93	1397.94
LASSO	**R^2 Score**	0.80	0.81	0.88
	MSE	227854.33	1320058.97	1954258.42
	MAE	363.31	936.59	1217.47
	RMSE	477.34	1148.93	1397.94
SVR	**R^2 Score**	0.93	0.97	0.96
	MSE	79352.30	170617.92	643586.37
	MAE	254.05	311.09	593.40
	RMSE	281.69	413.05	802.23

From the table, we can see that the performance of the models has increased for high-volume datasets.

V. CONCLUSION

COVID-19 creates health and economic crises all over the world. Every time new variants of COVID-19 make a pandemic wave, many people are affected worldwide. As a result, it causes damage to the human body, and people die. In this study, we propose a short-term future forecasting model for newly infected case prediction for the upcoming 7 days. LR, LASSO, and SVR regression models were used for future forecasting. Findings show that the SVR model performance was better than LR and LASSO models in all three data scenarios. Results also indicate the increases in past data volume, increasing the result accuracy for all models. Overall, we conclude that the short-term forecasting model using SVR gives better results for upcoming days of infected cases, which may help to understand the upcoming situation. This forecasting model also can be helpful for people and the government to make decisions and prepare for the upcoming situation. The healthcare system could prepare and provide the best support in the pandemic situation that may help save lives. In the future, it is intended to adopt deep learning approaches to forecast the COVID-19 transmission for small datasets. Several relational features like population density, people's age range, human health condition etc. can be considered to better understand the future case scenario of COVID-19.

REFERENCES

[1]. Md. Siddikur Rahman, Arman Hossain Chowdhury, Miftahuzzannat Amrin. 2022. "Accuracy comparison of ARIMA and XGBoost forecasting models in predicting the incidence of COVID-19 in Bangladesh." PLOS GLOBAL PUBLIC HEALTH. doi:https://doi.org/10.1371/journal.pgph.0000495.

[2]. Noara AlHusseini, Muhammad Sajid, Afaf Altayeb, Shahd Alyousof, Haifa Alsheikh, Abdulrahman Alqahtani, Afrah Alsomali. 2021. "Depression and Obsessive-Compulsive Disorders Amid the COVID-19 Pandemic in Saudi Arabia." The Cureus Journal of Medical Science. doi:https://doi.org/10.7759/cureus.12978.

[3]. WHO. 2023. Coronavirus disease (COVID-19) pandemic. World Health Organization. Accessed March 15, 2023. https://www.who.int/emergencies/diseases/novel-coronavirus-2019.

[4]. IEDCR. 2023. Institute of Epidemiology, Disease Control and Research. Accessed march 15, 2023. https://old.dghs.gov.bd/index.php/bd/component/content/article?id=5391.

[5]. DGHS. 2023. COVID-19 Dynamic Dashboard for Bangladesh. Accessed March 15, 2023. http://dashboard.dghs.gov.bd/webportal/pages/covid19.php.

[6]. Mossamet Kamrun Nesa, Md. Rashed Babu, Mohammad Tareq Mamun Khan. 2022. "Forecasting COVID-19 situation in Bangladesh." Biosafety and Health vol.4: pp.6-10. doi:https://doi.org/10.1016/j.bsheal.2021.12.003.

[7]. Hafsa Binte Kibria, Oishi Jyoti, Abdul Matin. 2022. "Forecasting the spread of the third wave of COVID-19 pandemic using time series analysis in Bangladesh." Informatics in Medicine Unlocked vol.28 (100815). doi:https://doi.org/10.1016/j.imu.2021.100815.

[8]. Md. Shahriare Satu, Koushik Chandra Howlader, Mufti Mahmud, M. Shamim Kaiser, Sheikh Mohammad Shariful Islam, Julian M. W. Quinn, Salem A. Alyami, Mohammad Ali Moni. 2021. "Short-Term Prediction of COVID-19 Cases Using Machine Learning Models." Applied Sciences (MDPI) Vol. 11(9) (4266). doi:https://doi.org/10.3390/app11094266.

[9]. Furqan Rustam, Aijaz Ahmad Reshi, Arif Mehmood, Saleem Ullah, Byung-Won On, Waqar Aslam and Gyu Sang Choi. 2020. "COVID-19 Future Forecasting Using Supervised Machine Learning Models." in IEEE Access (in IEEE Access) vol. 8: pp. 101489-101499. https://ieeexplore.ieee.org/document/9099302.

[10]. Abdelkader Dairi, Fouzi Harrou, Abdelhafid Zeroual, Mohamad Mazen Hittawe, Ying Sun. 2021. "Comparative study of machine learning methods for COVID-19 transmission forecasting." Journal of Biomedical Informatics vol. 114. doi:https://doi.org/10.1016/j.jbi.2021.103791.

[11]. Nada R. Sanders, R. Dan Reid. 2012. "PRINCIPLES OF FORECASTING." In Operations Management: An Integrated Approach, 5th Edition, by R. Dan Reid and Nada R. Sanders. Wiley. https://www.oreilly.com/library/view/operations-management-an/9781118122679/.

[12]. Moore, Audra. 2021. Forecast Accuracy : Long-Range Forecasting is Hard. KMTV 3 News Now. Accessed March 19, 2023. https://www.3newsnow.com/weather/weather-blog/forecast-accuracy.

[13]. Sanzida Solayman, Sk. Azmiara Aumi, Chand Sultana Mery, Muktadir Mubassir, Riasat Khan. 2023. "Automatic COVID-19 prediction using explainable machine learning techniques." International Journal of Cognitive Computing in Engineering (Elsevier) Vol. 4: PP. 36-46. doi:https://doi.org/10.1016/j.ijcce.2023.01.003.

[14]. GitHub, Inc. 2023. time_series_covid19_confirmed_global.csv. https://github.com/CSSEGISandData/COVID-19/tree/master/csse_covid_19_data/csse_covid_19_time_series.

[15]. Abu Kaisar Mohammad Masuma, Sharun Akter Khushbua, Mumenunnessa Keyaa, Sheikh Abujara, Syed Akhter Hossaina. 2020. "COVID-19 in Bangladesh: A Deeper Outlook into The Forecast with Prediction of Upcoming Per Day Cases Using Time Series." 9th International Young Scientist Conference on Computational Science (YSC 2020). Procedia Computer Science. 291–300.

[16]. Iqra Sardar, Muhammad Azeem Akbar, Victor Leiva, Ahmed Alsanad, Pradeep Mishra. 2023. "Machine learning and automatic ARIMA/Prophet models-based forecasting of COVID-19: methodology, evaluation, and case study in SAARC countries." Stochastic Environmental Research and Risk Assessment vol.37: 345–359. doi:https://doi.org/10.1007/s00477-022-02307-x.

[17]. Khondoker Nazmoon Nabi, Md Toki Tahmid, Abdur Rafi, Muhammad Ehsanul Kader, Md. Asif Haider. 2021. "Forecasting COVID-19 cases: A comparative analysis between recurrent and convolutional neural networks." Results in Physics Vol. 24 (104137). doi:https://doi.org/10.1016/j.rinp.2021.104137.