Comparative Performance Analysis of Machine Learning Algorithms for COVID-19 Cases in India

Apoorva Sharma[1⋆][0009−0001−0390−2910], Maitreyee Dutta[1][0000−0002−2608−6821], and Ravi Prakash[2][0000−0002−1582−9992]

[1] National Institute of Technical Teachers Training & Research (NITTTR), Chandigarh, India
apoorva.cse20@nitttrchd.ac.in, maitreyee@nitttrchd.ac.in
[2] Digital University Kerala (formerly IIITM-K), Thiruvananthapuram, India
1907prakash.ravi@gmail.com

Abstract. A novel corona virus is the cause of the viral infection recognized as COVID-19 (initially named as SARC-CoV-2). Since the pandemic emerged in the Wuhan province of China in November 2019, it has been recognized as a global threat. However, over the next two years, it has been witnessed that the novel corona virus tends to evolve rapidly. In this paper, we leverage our time-series data collected since the initial spread of COVID-19, mainly in India, to better understand the growth of this pandemic in different regions throughout the country. The research is based on cases reported in India in chronological order. In addition to numerous previous works, we have tried to come up with the most appropriate solution to estimate and predict the newly reported COVID-19 cases in the upcoming days, with the least possible error through machine learning. This study also aims to compare multiple machine learning algorithms on various factors and their trade-off for prediction. The experimental results indicate that Orthogonal Matching Pursuit is the best algorithm for this problem. We make our dataset available for further research.

Keywords: COVID-19, Corona, Machine Learning, Regression, Predictive Model.

1 Introduction

The novel corona virus is known to emerge in China in November 2019. It was a viral infarction that rapidly spread from bats to humans, and within a few months, it reached almost every other country worldwide. COVID-19 virus was found to affect the human respiratory system severely. Unfortunately, even after detecting the virus and its behavior in the early days, it took over six months

⋆ The link to access the final code and dataset used for essential data preparation and testing of the model is: https://github.com/apoorva46/COVID-19-Project-2023 to come up with proper vaccination. COVID-19 was declared a global pandemic due to its very sharp growth rate. Moreover, the virus mutated rapidly. It made the situation even more challenging as the virus was also found to evolve over time. Moreover, it was found that different waves of COVID-19 proved fatal for people of different age groups.

Even after a lot of effort regarding its prevention, COVID-19 is still found to be frequently spreading. In such a situation, the estimation of a possible number of new cases seems very difficult. Based on the several approaches involved in predictive data analysis, people have tried to estimate the count of new cases and the rate of its growth [10,23]. This paper brings in a more detailed comparison of different machine learning algorithms. It proposes methodologies to accurately predict the counts of new cases in the following few days based on the analysis. This study describes the disease's state in terms of total confirmed cases till a specific date. Here, the time-series data has been cumulatively stored on a daily basis, along with the counts of death cases and recovered cases every day. In this work, we are predicting the total number of cases in the upcoming days with the help of different supervised ML algorithms and comparing their performance based on the five evaluation metrics MAE, MSE, $RMSE$, $MAPE$ & R^2[20]. Here, we have also analyzed the trends in the number of COVID-19 instances in different states of India.

The organisation of this paper is - Introduction to the problem is discussed in Section 1. It is followed by the Literature Review in Section 2, which covers a detailed description of relevant previous work and the research gaps. Section 3 provides the complete methodology. The outcomes of our work are given in Section 4. Conclusion and Future Scopes of this work is given in Section 5.

2 Literature Review

Our motivation behind this paper is to develop a solution to estimate the possible count of newly reported COVID-19 cases for the near future using some regression-based ML model.

After going through recently published papers, we found that Zhong *et al.*[26] developed a mathematical model to forecast the spread using epidemiological data in March 2020. Over time, it was discovered that it had no effect. In April 2020, Benitez Pena *et al.*[5] used RF and SVM for the analysis of the disease. They leveraged Gurobi to solve the problem.

Chakraborty *et al.*[8] used ML approach for solving two problems, i.e., forecasting short-term future COVID-19 cases and risk assessment based on fatality rate. Here authors explained that hybrid time-series models using ARIMA & wavelet-based forecasting techniques for predicting cases in five different countries could be the best approach for the problem. As the pandemic was dynamically spreading throughout the world, all of these researchers discovered that there were kinds of dissimilarity in the rate of its growth.

In April 2020, Vashisht *et al.*[23] explained the growth rate of a novel corona virus in China based on the regression models. Kanagarathinam *et al.*[13] pro-

Table 1. Abbreviations

Abbreviation	Full Forms
ARIMA	AutoRegressive Integrated Moving Average
DT	Decision Tree
ES	Exponential Smoothing
GB	Gradient Boosting
KNN	K-Nearest Neighbors
LR	Linear Regression
MAE	Mean Absolute Error
MAPE	Mean Absolute Percentage Error
ML	Machine Learning
MSE	Mean Square Error
NB	Naive Bayes
NN	Neural Network
PR	Polynomial Regression
RF	Random Forest
RL	Reinforcement Learning
RMSE	Root Mean Square Error
SVM	Support Vector Machine
VAR	Vector AutoRegression

posed the SEIRS model and used data of a month for prediction. Sujath et. al.[21] performed Linear Regression along with a Neural Network based method, by applying Multi-Layer Perceptron and a multivariate solution using Vector AutoRegression on kaggle data having 80 instances of COVID-19.

Rustam *et al.*[19] used Linear Regression, Least Absolute Shrinkage and Selection Operator, Support Vector Machine, and Exponential Smoothing and concluded that ES, LASSO & LR performed better than SVM. Nabi[17] employed Trust Region Reflective algorithm for tentative predictions of the epidemic peak. Goswami *et al.*[11]used the Verhulst Logistic Population Model and also used the Generalized Additive Model of regression to examine the impact of various meteorological parameters on the prediction of COVID-19 instances. In order to estimate the trend of the outbreak, Wang *et al.*[24] incorporated epidemiological data collected before June 16, 2020, into the logistic model.

Burdick *et al.*[6] assessed an ML algorithm's effectiveness for predicting invasive mechanical ventilation in COVID-19 patients within 24 hours of the first contact. Amar *et al.*[4] used logistic growth regression models to analyze COVID-19 data of Egypt. Khanday[14] showed that Logistic Regression and Multinomial NB algorithms produce better results than RF, Stochastic GB, DT, and Boosting. Yadav *et al.*[25] proposed a Novel Support Vector Regression method rather than employing a simple regression line to assess five tasks differently.

Darapaneni *et al.*[9] used RF and found accuracy for the training data as 97.17% and that for testing data as 94.80%. Kumari *et al.*[15] used DT training techniques for splitting data and Autoregression model to predict the possible number in the future. Gupta *et al.*[12] used RF, Linear Model, SVM, DT, and NN for forecasting and found that RF outperformed the others. Mary *et al.*[16]

employed Feature Selection Techniques, SVM, kNN, and NB, and found that SVM outperformed other algorithms. Tiwari *et al.*[22] conducted a comprehensive evaluation and comparison of 37 ML-related studies that covered many ML algorithms. However, all these papers have used limited dataset due to early conduct of their research works. In few papers, vague assumptions have been made. Apart from this, forecast is done for only a week and results is some papers are overestimated.

3 Dataset and Methodology

3.1 Dataset

Our dataset was compiled using multiple sources to ensure its accuracy and completeness. To validate the data and ensure that the most appropriate time-series information was collected, we referred to several sources including Ministry of Health and Family Welfare (MoHFW) [1] and Indian Council of Medical Research (ICMR) [2] of India. The data was collected from as early as January 2020, i.e., since initial days of COVID-19 in the country. We utilized R scripts to fetch and validate the data from the GitHub repository of CSSE, Johns Hopkins University [3] as well. The CSSE database consisted of the data related to not only India but for the entire world. Additionally, we cleaned the data by removing outliers and filling in null values for several states where no cases were reported during the early days.

3.2 Model configuration

We target training and evaluating the performance of multiple ML algorithms by training on the same data. To meet this goal, our research work primarily utilizes the programming language 'Python', and for comparing the performance of different models based on five evaluation metrics, we have employed PyCaret, Table 4 provided in Appendix contains the configuration details used for training the models.

3.3 Methodology

The collected time-series data contained a total count of active, recovered, and death cases in 36 states and union territories of India. The data also consisted of the cases from some unclassified locations those were handled during the data pre-cleaning phase. The date-range of collected data starts from 25th March 2020 till the day when the lockdown was withdrawn, i.e., 28th July 2020. This data was collected every single day and stored in two formats. One formatting was done as a cumulative data for India. The other format of data was purely time series with daily case counts. The final time-series data was taken as the cumulative sum of confirmed cases ζ^i on every i-th day for every j-th state and union territory S_j of India as per Eq.1. Later, the collective sum of counts of every S_j is taken to generate the data for the whole country (Eq.2).

$$\zeta^i = \zeta^i_{S_j} \tag{1}$$

$$\zeta^i_s = \zeta^i_\delta \;\; \forall \;\; \delta = \sum_{36}^{j=1} S_j \tag{2}$$

Working towards our goal, in order to forecast the daily count of unique cases, we created a separate column in the dataset η_i for each day using Eq. 3. The proposed approach performed this computation for each day from 25th March 2020 to 28th July 2020 in India.

$$\eta_i = \zeta^i_S - \zeta^{i-1}_S \tag{3}$$

Firstly, data cleaning was performed that included the following three essential steps:

- **Step-1:** Removing or replacing the redundant values from data
- **Step-2:** Removing the $NULL$ and dealing with blank (no) data points
- **Step-3:** Outliers removal

For further analysis, we have split our entire data in a ratio of 70 : 30 as training and testing data, respectively 85 : 30 rows. In order to make the comparison, the 'PyCaret' library has been used, which was developed by data scientist Moez Ali. This library is capable of building many regression models simultaneously.

Once the dataset was ready, we began with the regression-based supervised learning approach to build the ML model. Here, we have taken three folds cross-validation in order to increase the efficiency of the models. At the same time, we sorted the performance of multiple regression algorithms based on the following five evaluation metrics MAE, MSE, $RMSE$, $MAPE$ & R^2. We have separately studied cumulative data and data consisting of new cases of all the states.

This allowed us to set up various machine-learning models using the generated data. Here the aim was to predict the total number of new cases of COVID-19 in the country. We tried to explore the maximum possible literature works, pondering and analyzing the work as per the state of the art approaches. Here, we trained many regression algorithms, namely Orthogonal Matching Pursuit, Extra Trees Regressor, Ridge Regression, Light GB Machine, and 15 more algorithms mentioned in Table 2 of Section 4.

4 Results and Observations

We initially tried exploring the data of COVID-19 spread in India both state-wise as well as collectively for the whole country. These graphs typically depict the cumulative number of cases and new cases in all the states of India. The line graphs plotted for this observation had varying trends. After analyzing all graphs, we can observe that results in a few states are dynamically varying, whereas, in some states, the trend is constant. We found that big states like Rajasthan, and Tamil Nadu, as shown in Fig. 1 and Fig. 2, respectively, have fluctuations in new case counts for short duration.

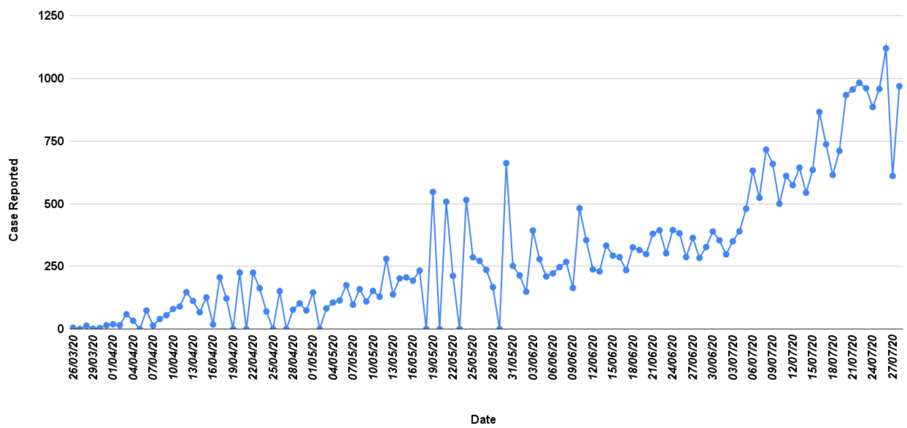

Fig. 1. Fluctuating number of new cases in Rajasthan from 26th March 2020 to 27th July 2020.

However, it can be seen that there is increase in number of new cases reported in a long period of time. One possible reason could be the wide area where it is a challenging task to deploy a complete lockdown in one go. Considering these insights, we found that the large states and those with higher populations decide the overall trend in growth of COVID-19 in the country. On the basis of the hypothesis Fig. 3 shows the reporting of new cases on the daily bases. Similarly in Fig. 1 and Fig.2, it can observed that the count of newly reported COVID-19 cases has increased over the time. It became more clear when we analyzed the cumulative growth of COVID-19 pandemic throughout the country, as shown in Fig. 4.

We have also noticed that few of the states and UTs show some different trends. Delhi is one among these few. We can see from Fig. 5 that cases in Delhi have actually decreased over the time. We can actually trace here a pattern among growth and control in the COVID-19 cases over the time. As a result of this, we can easily predict the total number of instances in the following days. However, without changing its configurations, it is unlikely that a model can accurately predict the upcoming cases for states like Tamil Nadu as precisely as it can for Delhi. Meanwhile, the decrease in number of new cases emphasizes how the usage of masks and creating awareness, and imposing lockdowns helped in the management of this lethal spread.

After carefully examining all the graphs, it is clear that the deployment of a number of stringent regulations and public safety measures has been successful

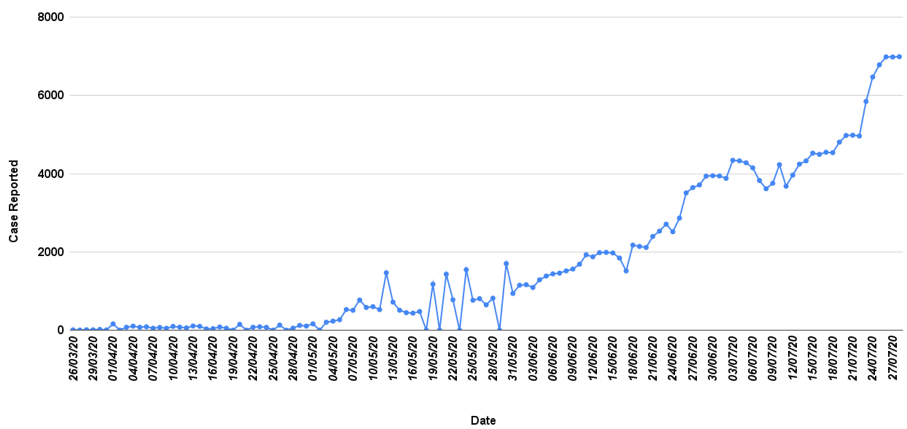

Fig. 2. Sharp increase in number of new cases in Tamil Nadu from 26th March 2020 to 27th July 2020.

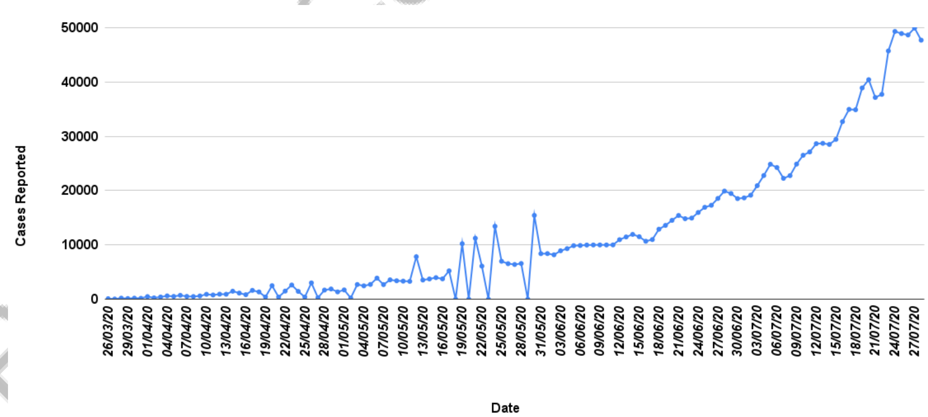

Fig. 3. Fluctuation in number of new cases in India during lockdown from 26th March to 27th July 2020.

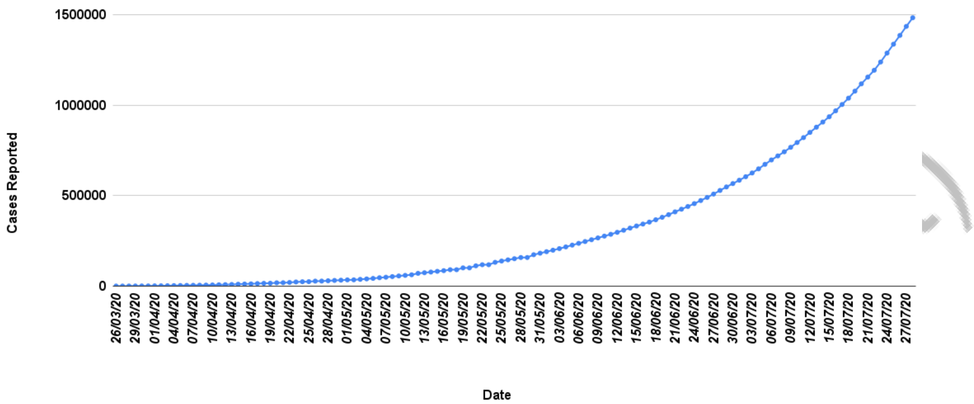

Fig. 4. Constant increase in number of total active cases in India from 26th March to 27th July 2020.

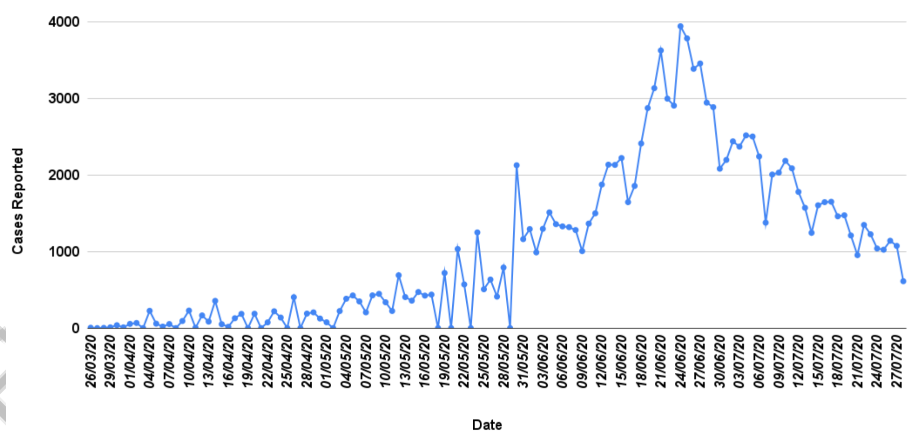

Fig. 5. Trend of COVID-19 daily spread in Delhi between 26th March 2020 and 27th July 2020.

because the number of cases has decreased whenever travel restrictions and a lockdown have been put in place. Also, we found that it is very difficult to build one identical model to forecast the results for all different states. Therefore, we tried to collectively analyze the data for the whole country and forecast the results accordingly. It can also be seen in Fig. 4 that there is less fluctuation in data as compared to each individual state or UTs.

After going through the existing literature, we found that the Random forest [5,9,12,14,22], Linear Regression [19,21,22,23,25], SVM [5,12,16,19,22,23,25], Decision Trees [9,14,22], and a few additional algorithms are frequently used for prediction. Other models, namely Hybrid ARIMA Wavelet [8], SIR [26], SEIRS [13], and SEI_DI_HQHRD [17] have also been used for some other types of predictive analysis for COVID-19. Fig.6 shows the complete distribution of all of the algorithms considered in total $N = 20$ papers.

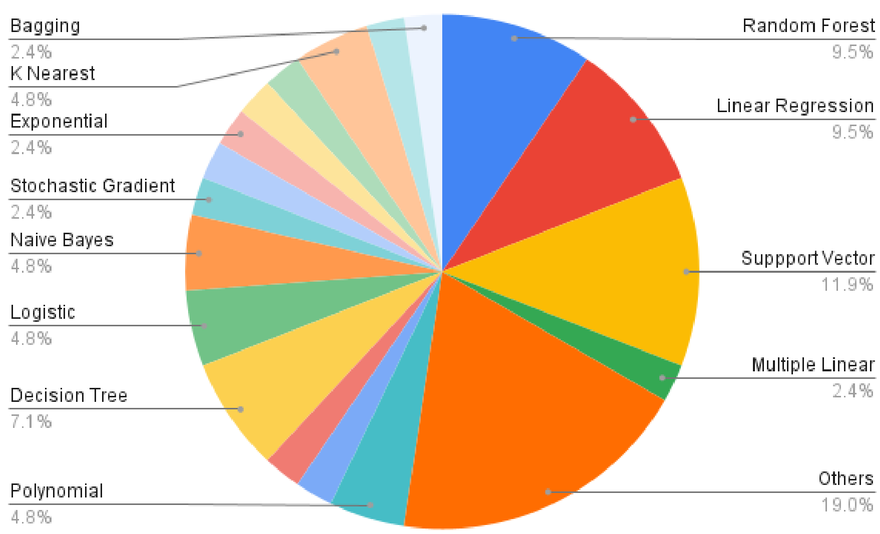

Fig. 6. Distribution of various ML algorithms that have been used in existing papers

In our analysis, we have considered five performance metrics namely MSE, R^2, MAPE, MAE, and RMSE [7] for each model. In addition to that, the time taken in milliseconds (*msec.*) by each of these algorithms, except the LR, is given in Fig. 7. It also shows that OMP is the second-best model in terms of time consumption with just 16.7 *msec.* On this scale, we found that LR is the worst prediction algorithm which takes as much as 1103.3 *msec.* as shown in Fig. 8, which is much more in comparison to any other algorithms.

After implementing our model, we found that Orthogonal Matching Pursuit[18] outperforms all other algorithms and fits best for the data we prepared in order

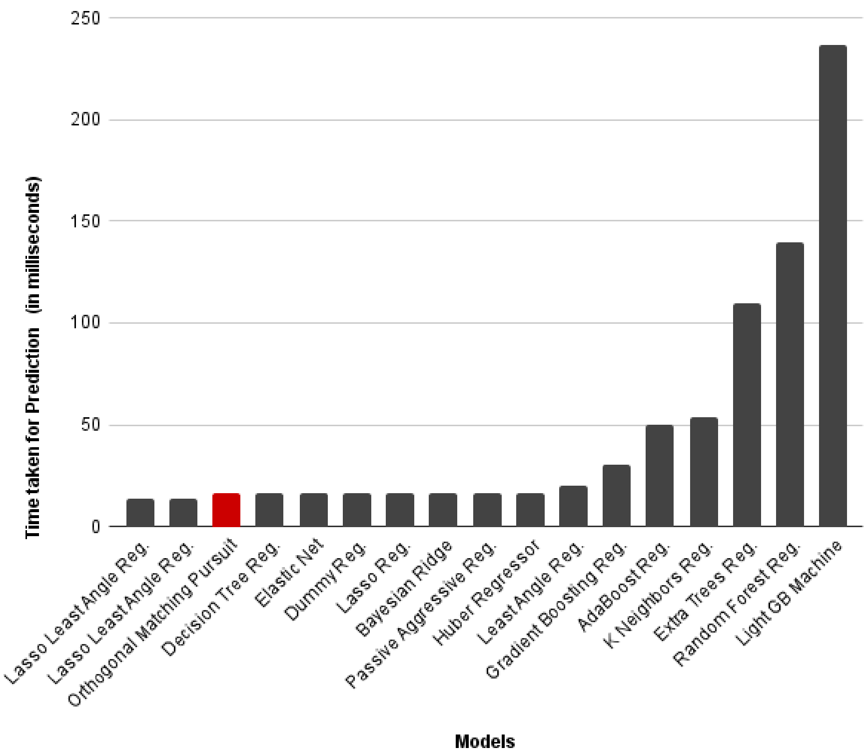

Fig. 7. Time taken by different algorithms for predicting the COVID-19 cases in India

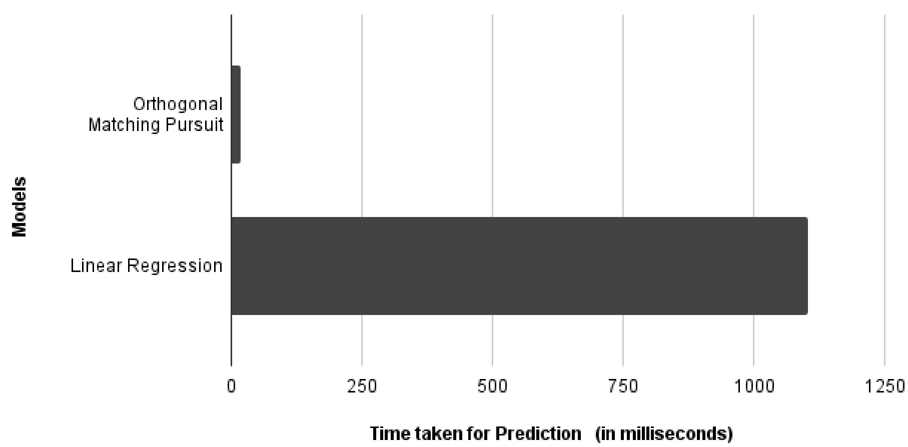

Fig. 8. Comparison of OMP with LR for time consumption

to forecast the estimated growth of COVID-19 cases in India. To validate our results, we have evaluated and compared $N = 18$ ML algorithms in Table 2. The details related to the model selected as per our research are described in Section 3.2.

Table 2. Results of all evaluation parameters for 18 different Machine Learning Algorithms.

Model	**MAE**	**MSE**	**RMSE**	$\mathbf{R^2}$	**MAPE**
Orthogonal Matching Pursuit	1.0776 e3	4.1988 e6	1.5092 e3	-1.1100 e-1	6.7410 e-1
kNN Regression	1.3975 e3	6.2503 e6	1.7938 e3	-3.2040 e-1	4.4610 e-1
Decision Tree Regression	1.4608 e3	6.6125 e6	1.8587 e3	-4.2380 e-1	4.6390 e-1
Extra Trees Regression	1.4677 e3	6.6382 e6	1.8609 e3	-4.2320 e-1	4.6880 e-1
Ridge Regression	1.4689 e3	5.8834 e6	1.9554 e3	-1.6667	9.9870 e-1
Random Forest Regression	1.4710 e3	6.6667 e6	1.8551 e3	-3.9440 e-1	4.6670 e-1
Gradient Boasting Regression	1.4976 e3	6.9410 e6	1.8910 e3	-4.3200 e-1	4.7890 e-1
AdaBoost Regression	1.5014 e3	6.8975 e6	1.8819 e3	-4.1460 e-1	4.7400 e-1
Elastic Net	1.6851 e3	7.0733 e6	2.1302 e3	-1.9207	1.0471
Linear Regression	1.7562 e3	9.6891 e6	2.2832 e3	-1.0814	8.2000 e-1
Light Gradient Boasting Machine	1.8583 e3	9.7754 e6	2.2605 e3	-9.4800 e-1	4.7970 e-1
Lasso Least Angle Regression	1.9754 e3	1.1123 e7	2.3863 e3	-1.0622	5.1250 e-1
Dummy Regression	1.9754 e3	1.1123 e7	2.3863 e3	-1.0622	5.1250 e-1
LASSO Regression	1.9970 e3	1.1112 e7	2.5932 e3	-2.3992	1.1260
Bayesian Ridge	2.4297 e3	1.9990 e7	3.3005 e3	-3.0900	1.1927
Passive Aggressive Regression	3.4993 e3	5.4454 e7	4.8409 e3	-5.0192	7.2030 e-1
Huber Regression	5.2911 e3	1.0623 e8	7.1759 e3	-1.3597 e1	1.9984
Least Angle Regression	3.0938 e9	6.9931 e19	4.8282 e9	-694.03 e11	2.3890 e6

The comparison among similar approaches has been done against the proposed approach in Table 3. Values of evaluation parameters that are not found as per the existing literature for analysis are denoted by cross mark ($\times$). Kumari *et al.* [15] worked on the spread of COVID-19 in various geographic areas of India and proposed a model for forecasting the quantity of confirmed, recovered, and fatal cases. They forecasted the potential number of instances of new COVID cases using multiple linear regression and autoregression. Rustam *et al.* [19] used LR, LASSO, SVM, ES for the prediction of number of newly infected, deaths, and recoveries in the next ten days and found ES outperforms others with values given in Table 3. Chakraborty *et al.*[8] proposed ARIMA, wavelet-based, and hybrid ARIMA wavelet-based model to forecast the number of daily confirmed cases for five countries. However, we have considered the performance of their proposed model for India only. Vashisht *et al.* [23] have estimated the possible rate of active cases in China for the upcoming week and compared SVM, kNN, LR, Polynomial Regression models. They found that PR performs better compared to other algorithm. The evaluation parameters obtained by them for the best model are mentioned in Table 3. It can be clearly seen in the last row of Table 3 that the results obtained by our study are better considering the amount of dataset which has been used for analysis.

Table 3. Evaluation parameters obtained by other Authors

Authors	**Model**	**MAE**	**MSE**	**RMSE**	$\mathbf{R^2}$
Kumari R. *et al.*[15]	Multiple LR	×	×	3085.43	0.999
Rustam F. *et al.*[19]	Exponential Smoothing	406.08	66.22e4	813.77	0.98
Chakraborty T. *et al.* [8]	ARIMA	16.07	50.83	×	×
Vashisht G. *et al.*[23]	Polynomial Regression	×	×	0.582	0.999
Proposed Model	OMP [18]	1.08e3	4.20e6	1.51e3	-1.11e-1

5 Conclusion and Future Scope

Through this research, we have collected the COVID-19 data for India from various data repositories and processed the same to obtain the most suitable dataset. Later, we have tested as many as eighteen different machine learning algorithms and compared their performance to come up with a best possible solution in order to forecast COVID-19 instances in India. The findings of this paper indicate that the likelihood of India reporting new daily COVID-19 cases is heavily influenced by the trends in its densely populated and larger states. However, it also emphasizes the importance of considering states with dissimilar trends to avoid any potential biases in the final results.

This study opens a wide area of research opportunities and societal benefits like deciding effective management strategies for fatal diseases like COVID-19. The results can be applied to various predictive classifications also, allowing for timely warnings and implementation of appropriate safety measures. Particularly in the light of the current economic downturn, this information is invaluable for the country's financial planning and highlights areas that require immediate attention. Additionally, this work provides more processed data that may be used for determining the allocation of resources towards constructing new hospitals/isolation centers, acquiring COVID-19 test kits, medical equipment, and improving care and treatment.

Our research work can be further enhanced by adding more attributes and comparison based on the neural network driven models that have not yet been tested.

Acknowledgment

We would like to express our gratitude to Dr. Jagriti Saini, Siddheshwari Dutt Mishra, and Mohammad Ahsan Siddiqui from the Department of Computer Science & Engineering, NITTTR, Chandigarh as well as Deepak Jaglan from Central University of Haryana, for their technical support at various stages of this research work.

Appendix

Table 4. Configuration for PyCaret to evaluate the performance of multiple models.

Parameter/Description	**Value**
Original Data	(85, 5)
Missing Values	FALSE
Numeric Features	3
Transformed Train Set	(85, 24)
Transformed Test Set	(40, 24)
Shuffle Train-Test	TRUE
Fold Generator	TimeSeriesSplit
Fold Number	3
Use GPU	FALSE
Experiment Name	reg-default-name
USI	4a9b
Imputation Type	simple
Numeric Imputer	mean
Iterative Imputation Numeric Model	None
Categorical Imputer	constant
Unknown Categoricals Handling	least_frequent
Transformation Method	None
Transform Target	TRUE
Transform Target Method	box-cox

References

1. Ministry of Health and Family Welfare website (1947)
2. Indian Council of Medical Research website (1949)
3. CSSE - Johns Hopkins University website (2019)
4. Amar, L.A., Taha, A.A., Mohamed, M.Y.: Prediction of the final size for covid-19 epidemic using machine learning: A case study of egypt. Infectious Disease Modelling 5, 622–634 (2020)
5. Benıtez-Pena, S., Carrizosa, E., Guerrero, V., Dolores, M.: Short-term predictions of the evolution of covid-19 in andalusia. an ensemble method. Preprint (2020)
6. Burdick, H., Lam, C., Mataraso, S., Siefkas, A., Braden, G., Dellinger, R.P., McCoy, A., Vincent, J.L., Green-Saxena, A., Barnes, G., et al.: Prediction of respiratory decompensation in covid-19 patients using machine learning: The ready trial. Computers in biology and medicine 124, 103949 (2020)
7. Chai, T., Draxler, R.R.: Root mean square error (rmse) or mean absolute error (mae). Geoscientific model development discussions 7(1), 1525–1534 (2014)
8. Chakraborty, T., Ghosh, I.: Real-time forecasts and risk assessment of novel coronavirus (covid-19) cases: A data-driven analysis. Chaos, Solitons & Fractals 135, 109850 (2020)
9. Darapaneni, N., Singh, A., Paduri, A., Ranjith, A., Kumar, A., Dixit, D., Khan, S.: A machine learning approach to predicting covid-19 cases amongst suspected cases and their category of admission. In: 2020 IEEE 15th International Conference on Industrial and Information Systems (ICIIS), pp. 375–380. IEEE (2020)
10. Devarajan, J.P., Manimuthu, A., Sreedharan, V.R.: Healthcare operations and black swan event for covid-19 pandemic: A predictive analytics. IEEE Transactions on Engineering Management 0, 1–15 (2021)
11. Goswami, K., Bharali, S., Hazarika, J.: Projections for covid-19 pandemic in india and effect of temperature and humidity. Diabetes & Metabolic Syndrome: Clinical Research & Reviews 14(5), 801–805 (2020)
12. Gupta, V.K., Gupta, A., Kumar, D., Sardana, A.: Prediction of covid-19 confirmed, death, and cured cases in india using random forest model. Big Data Mining and Analytics 4(2), 116–123 (2021)
13. Kanagarathinam, K., Sekar, K.: Estimation of the reproduction number and early prediction of the covid-19 outbreak in india using a statistical computing approach. Epidemiology and health 42, 1–5 (2020)
14. Khanday, A.M.U.D., Rabani, S.T., Khan, Q.R., Rouf, N., Mohi Ud Din, M.: Machine learning based approaches for detecting covid-19 using clinical text data. International Journal of Information Technology 12, 731–739 (2020)
15. Kumari, R., Kumar, S., Poonia, R.C., Singh, V., Raja, L., Bhatnagar, V., Agarwal, P.: Analysis and predictions of spread, recovery, and death caused by covid-19 in india. Big Data Mining and Analytics 4(2), 65–75 (2021)
16. Mary, L.W., Raj, S.A.A.: Machine learning algorithms for predicting sars-cov-2 (covid-19) – a comparative analysis. In: 2021 2nd International Conference on Smart Electronics and Communication (ICOSEC), vol. 0, pp. 1607–1611 (2021)
17. Nabi, K.N.: Forecasting covid-19 pandemic: A data-driven analysis. Chaos, Solitons & Fractals 139, 110046 (2020)
18. Pati, Y.C., Rezaiifar, R., Krishnaprasad, P.S.: Orthogonal matching pursuit: Recursive function approximation with applications to wavelet decomposition. In: Proceedings of 27th Asilomar conference on signals, systems and computers, pp. 40–44. IEEE (1993)

19. Rustam, F., Reshi, A.A., Mehmood, A., Ullah, S., On, B.W., Aslam, W., Choi, G.S.: Covid-19 future forecasting using supervised machine learning models. IEEE Access 8, 101489–101499 (2020)
20. Schneider, P., Xhafa, F.: Anomaly Detection and Complex Event Processing Over IoT Data Streams: With Application to EHealth and Patient Data Monitoring. Academic Press (2022)
21. Sujath, R., Chatterjee, J.M., Hassanien, A.E.: A machine learning forecasting model for covid-19 pandemic in india. Stochastic Environmental Research and Risk Assessment 34, 959–972 (2020)
22. Tiwari, S., Chanak, P., Singh, S.K.: A review of the machine learning algorithms for covid-19 case analysis. IEEE Transactions on Artificial Intelligence 4(1), 44–59 (2023)
23. Vashisht, G., Prakash, R.: Predicting the rate of growth of the novel corona virus 2020. International Journal on Emerging Technologies 11(3), 19–25 (2020)
24. Wang, P., Zheng, X., Li, J., Zhu, B.: Prediction of epidemic trends in covid-19 with logistic model and machine learning technics. Chaos, Solitons & Fractals 139, 110058 (2020)
25. Yadav, M., Perumal, M., Srinivas, M.: Analysis on novel coronavirus (covid-19) using machine learning methods. Chaos, Solitons & Fractals 139, 110050 (2020)
26. Zhong, L., Mu, L., Li, J., Wang, J., Yin, Z., Liu, D.: Early prediction of the 2019 novel coronavirus outbreak in the mainland china based on simple mathematical model. IEEE Access 8, 51761–51769 (2020)