Predicting the time period of extension of lockdown due to increase in rate of COVID-19 cases in India using machine learning

Parth Wadhwa [a], Aishwarya [a,*], Amrendra Tripathi [b], Prabhishek Singh [a], Manoj Diwakar [c], Neeraj Kumar [d]

[a] *Department of CSE, Amity School of Engineering and Technology, Amity University Uttar Pradesh Noida, India*
[b] *Department of CSE, University of Petroleum and Energy Studies, Dehradun, Uttarakhand, India*
[c] *Department of CSE, Graphic Era (Deemed to be University) Dehradun, Uttarakhand, India*
[d] *Department of IT, Babasaheb Bhimrao Ambedkar University, Central University, Lucknow, India*

ARTICLE INFO

ABSTRACT

The research paper proposes a methodology to predict the extension of lockdown in order to eradicate COVID-19 from India. All the concepts related to Coronavirus, its history, prevention and cure is explained in the research paper. Concept used to predict the number of active cases, deaths and recovery is Linear Regression which is an application of machine learning. Extension of lockdown is predicted on the basis of predicted number of active cases, deaths and recovery all over India. To predict the number of active cases, deaths and recovery, date wise analysis of current data was done and necessary parameters like daily recovery, daily deaths, increase rate of covid-19 cases were included. Graphical representation of each analysis and prediction was done in order to make predicted results more understandable. The combined analysis was performed at the end which included the final result of total cases of coronavirus in India. Combined analysis included the no. of cases from start of COVID-19 to the predicted end of cases all over India.

1. Introduction

1.1. What is Coronavirus?

- The virus is not a living organism, it is a protein molecule covered by protective layer of fat which when observed by the cells changes its genetic code.
- As the virus is a protein molecule but not a living organism that means it can not be killed but it decays on its own. The disintegration time depends on the following factors
  o Humidity
  o Temperature
  o Type of material where it lies
- The virus is very fragile that is why the only thing that protects it is a thin outer layer of fat molecules that is why any soap or detergent is the best way because foam cuts the fat. By dissolving the fat layer, the protein molecule disperses and breaks down on its own.

- The virus molecules remain very stable in external cold or artificial as air conditioners in houses and cars. The virus also needs moisture to stay stable and especially darkness therefore dehumidified, dry, warm and bright environment will degrade it faster.
- The virus cannot go through healthy skin.
- The PH of corona virus varies from 5.5 to 8.5.
- Coronavirus is a zoonotic virus that means it commonly infects animals such as birds bats camels and pigs. In very few cases the coronavirus jump species that means the virus is passed on to humans through direct or indirect contact such event is called zoonotic spill over.

1.2. History of coronavirus

Coronaviruses are family of virus which lies between common cold to MERS (Middle East Respiratory Syndrome) and SARS (Severe Acute Respiratory Syndrome) and causes diseases in mammals and birds. They are one of the largest RNA viruses with the genome (genetic material) size of 26–32 kilobases. The special club shaped

* Corresponding author.
E-mail address: aishwaryasinghal307@gmail.com (Aishwarya).

spikes on the surface reminds of the 'crown' from where it received its name-Coronavirus.

SARS, caused by SARS-CoV. It was believed to be an animal virus mostly found in bats but in 2002 it started spreading to other animals like cats (civet cats) and eventually first human infected by this was found in Guangdong province in southern China in 2002. This disease turns into an epidemic effecting more than 8000 people of 26 countries with more than 750 lives in 2003. Symptoms of SARS-CoV were fever, headache, diarrhoea, and shivering rigors.

MERS, also known as camel flu caused by MERA-CoV. It is one of the 4 family members of coronavirus. MERS-CoV is beta-coronavirus derived from bats. Camels are believed to be infected by bats and hence spread in humans. It was first observed in September 2012 in Jordan then in 2015 in Republic of Korea and then in 2019 worldwide. Till date there no treatment for the disease. The total number of cases are 2519 with 866 deaths. Mortality rate is 1/3 of the total diagnosed cases. Symptoms of MERS are fever, cough, shortness of breath, diarrhoea and abdominal pain.

SARS-CoV-2, virus strain causing Corona Virus Disease 2019 (COVID-19) name given by WHO on 11 February 2020. It is also called novel coronavirus because it's the new strain of coronavirus different from both MERS and SARS, but closely related to SARS. COVID-19 was discovered in December 2019 in Wuhan, China. It was detected when a large number, greater than normal people were infected symptoms of pneumonia illness similar to flu. It has a very close genetic similarity to bat coronavirus which suggest that this virus is bat-borne virus. The first case known was found on 17 November 2019 after this, it spread to whole China and now more than 150 other countries. 30 January 2020, WHO declared SARS-CoV as Public Health Emergency of International Concern (PHEIC) and later on 11 March 2020 WHO declared pandemic.

There were more than 82,000 infected in China, from where this pandemic started. However, China handled the situation rapidly which helped them to control the death rate. Although the proportion of infected cases and the progress of diagnosed is not clear. It was estimated that75, 815 people were infected in Wuhan, China by 25 January 2020 and within a month, 95% deaths of the infected cases were recorded in Wuhan. Most affected countries include USA with total case of more than 5,00,000 and 20,000 + deaths, Spain with more than1,60,000 and 16,000 + deaths, Italy with 1,52,000 + cases and 19,000 + deaths.

To study the detection of coronavirus or any new virus, we need to know the working of the immune system of our body. Virus are microorganisms which enter into our body to live and reproduce. It is just an unwanted thing to the body, by seeing this immune system realises that it's not a part of the body and attacks it, after killing it, the immune system remembers the virus. Viruses change themselves to survive in the environment and the new virus with different structures are formed.

People with weak immune system are most likely to be affected. Older people, children and people with different medical condition like diabetes, sugar and other heart disease are at high risk. Effects of virus in children is not known, as the number of infected children is very less. The virus can be deadly but only in few cases, among senior citizens with previous medical history.

1.3. Transmission of the virus

While sneezing and coughing, an infected person transmits the virus from one person to another. These respiratory droplets have the range of 6 ft. and any person within this range can get infected. Individuals can also catch infection by touching surfaces contaminated or infected with the virus and then touching their face (e.g., eyes, nose, mouth). Later in a study it was observed that this virus may remain viable on different objects like

Steel or plastic	Up to 3 days
Cardboard	Not more than a day
Copper	Not more than 4 h
Soap	Inactive

1.4. Symptoms of COVID-19

These are not the symptoms of the virus, but the actions of the immune system to get rid of the virus. These actions required a lot of energy and hence make the body restless and tired. Since COVID-19 is a respiratory disease, it majorly effect nose, throat and lungs. So, these are the common symptoms-

1. Fever
2. Cough
3. Sore throat
4. Fatigue
5. Shortness of breath

Symptoms of Covid-19 are quite similar with common cold and flu. So, if someone have these symptoms may have COVID-19 and one should not panic and take the required measures.

2. Literature survey

Many different researches and works had been performed on Covid-19, its analysis and prediction. Some of the works and analysis are described below.

Lixiang Li et al. in 2020 in paper [1] proposed an approach a study to analysis the transmission and process of Covid-19. Advantage of this paper is its accuracy; error curve is quite small. The author used gaussian distribution theory to build the prediction model. This research strictly covers Hubei and non-Hubei provision if China, and other countries link South Korea, Iran and Italy.

Anuradha Tomar and Neeraj Gupta in 2020 in the paper [2] used various methods of deep learning like recurrent neural networks with long short-term memory for predicting the increase in number of cases in next 30 days. Accuracy is good and the error percentage is also less. The parameters used to predict the new cases were total active cases and recovery rate. The authors also compared Covid-19 with other diseases like Ebola, bird flu, SARS etc.

Susanta Mahato, Swades Pas and Krishna Gopal Ghosh in 2020 in the paper [3] study the impact of lockdown on environment, specially air pollution in Delhi, India. The author's study shows that the air quality index decreased gradually as the duration of lockdown increased. The study was displayed using various graphs and heatmaps.

Sunita Tiwari, Sushil Kumar, Kalpna Guleria in 2020 in the paper [4] proposed a method to predict the outbreak of novel corona virus. The prediction was based on the outbreak of same virus in China and the data on which this method of machine learning was applied was a Kaggle dataset which have confirmed cases, recovery rate and death rate. Comparison between India and China is displayed using graphs.

Atanu Bhattacharjee, Mukesh Kumar and Kamalesh Kumar Patel in 2020 in paper [5] stated a method to predict recovery rate and cases load rate on the basis of active cases and death rate. The study suggests that the recovery rate will be more than the active cases rate after 20 May 2020. All the history and efforts by the government are also well written. For more accurate results machine learning can be used.

Ardabili et al. in 2020 in paper [6] proposed prediction model for Covid-19. The authors used machine learning approach and soft computing models. Both methods were compared and the most effective results are shown by multi- layered perceptron (MLP) and adaptive network based fuzzy inference system (ANFIS). This method is applied to the data of countries.

Li et al. in 2020 in paper [7] took step to filter information which is related to Covid-19 so that only relevant information is shared and mislead from unauthorised information can be stopped. In this paper, the dataset used is Weibo dataset. NLP based approach is used to classify the information. Doing manual labelling the accuracy is less and there are many limitations because of less data.

Kang et al. in 2020 in paper [8] proposed an application for detection coronavirus disease using machine learning algorithms to reduce the large need to doctors and nurses. CT scan (Chest computed tomography) has been used for the process of diagnosis. Neural network is used for viewing the CT images from different angles. The results shown in this paper that there is large difference types of pneumonia.

Dong et al. in 2020 in paper [9] proposed an idea to improve the diagnosis techniques. To help doctors, CT (computed tomography) plays a vital role in diagnosis and treatment. The technique of imaging data using artificial intelligence is used. The study suggests that the combine use of AI and CT scans can result to support rapid diagnosis as the changing in lung lesions can be observed more easily.

Oh et al. in 2020 in paper [10] proposed the idea to analyse the images of X-rays. The used is batched Convolution neural network, the data set was small else deep learning can give more effective results.

Yule Kim in 2020 in paper [11] discussed the impact of covid-19 on consumers and the uncertainty about consumers returning to the particular shops. The paper showed significant results on how the growth of e-commerce sector has increased during the pandemic. The author also suggests that the managers should start developing digital sales to boom their state.

Zheng, el at. In 2020 in the paper [12] proposed a predicting model to predict the development and spread of the pandemic. In the paper, author used a hybrid artificial intelligence (AI) for the prediction and other latest technology like NLP and LSTM network to predict the effects of prevention measures. This method was applied to various different cities of China like Wuhan, Beijing, Shanghai.

3. Proposed methodology

3.1. Proposed methodology description

Step1. Data related to COVID-19 was collected from website [13]. Data collected included parameters like total number of cases, deaths and recovery all over India. Rate of increase of cases, deaths and recovery on daily basis was also collected.
Step2. Analysis of data was done to keep the relevant data out of the whole dataset and delete the irrelevant material from the dataset.
Step3. Study was done on the data and machine learning algorithm was performed to predict number of cases, deaths and recovery in coming month all over India.
Step4. Prediction of lockdown period was made assuming that if number of cases stop increasing in month of June how much more lockdown period should be extended to stop the spread of COVID-19 cases.
Step5. Prediction of lockdown was done on the basis of predicted active cases and predicted deaths and recoveries using machine learning algorithms.

Step6. Graphical representation of each analysis was done to make it more understandable and clearer. Graphical representation of prediction of COVID-19 cases and extension of lockdown period was done to generate the final result of the prediction that how much time should lockdown be extended in order to make India COVID-19 free.

3.2. Proposed methodology flowchart

Fig. 1.

4. Experimental analysis and prediction

4.1. Present condition of COVID-19 cases

The graph shown above is the graphical representation of total number of active cases, deaths and recovered cases all over India till 10th may 2020. As shown above the graph of all the three parameters (total cases, deaths and recovery) is increasing day by day (See Fig. 2).

4.2. Prediction of cases

The graph shown above is the graphical representation of predicted active cases, deaths and recovery in upcoming month of COVID-19 all over India. Linear regression that is an algorithm of
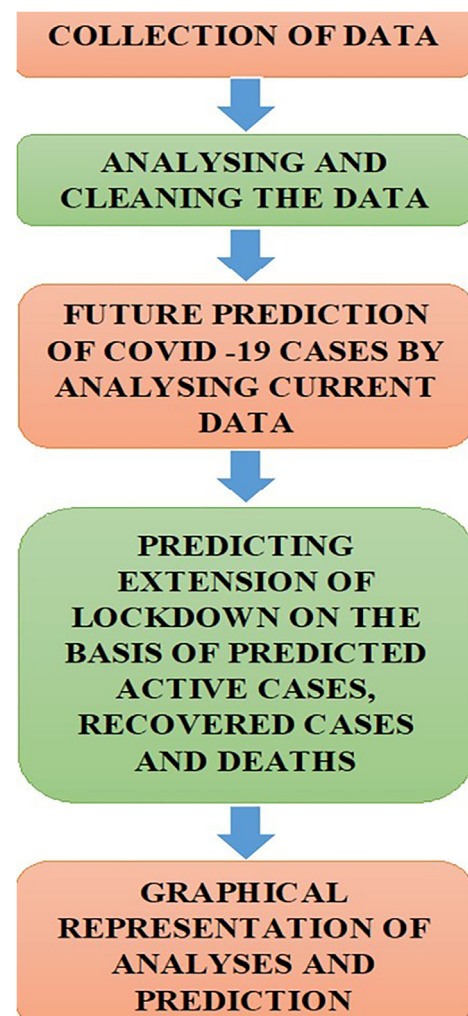

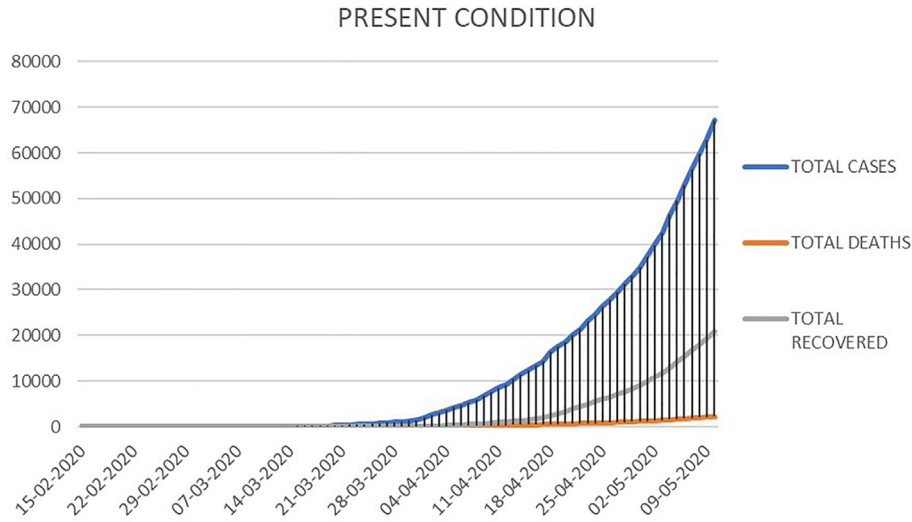**Fig. 1.** Flowchart of proposed methodology.

Fig. 2. Analysis of present condition of COVID-19 all over India.

machine learning is used to predict these three parameters shown above, this is the reason for predicted graph being increased linearly as shown above in (See Fig. 3).

4.3. Prediction of lockdown

The graph shown above is the graphical representation of analysis done to prediction the extension of lockdown period. Analysis is done on the basis of predicted total number of cases, deaths and recovery of COVID-19 all over India. As linear regression is used to predict the time period of lockdown therefore graph is decreasing linearly that is number of cases are decreasing linearly. According to all the analysis made lockdown period should be extended to September 9 so that there are no cases of COVID-19 in India Fig. 4.

5. Controlling the spread of COVID-19

5.1. Measures to be taken

- Stay home when sick
- Ultra Violet light on any object that may contain it breaks down the virus protein. For example, to use mask is perfect.
- Disposing used tissue immediately.

- Washing hands very often at least for 20 s because rubbing the hands for 20 s or more makes a lot of foam which is the only way to cut the fat layer of the virus.
- use of alcohol-based hand rub (sanitizer) with minimum 60% alcohol content because alcohol or any mixture with alcohol content dissolves any fat, especially the external lipid layer of the fat.
- Heat melts fat, this is why it is so good to use water above 25 degrees Celsius for washing hands, clothes and everything. Adding to this hot water makes more foam and that makes is even more useful.
- Social distancing
- If anyone possesses any kind of COVID-19 symptoms, then first call your doctor and then take that person to the hospital according to the advice given by the doctor.
- Never shake used or unused clothing sheets or cloth.
- Wash your hands before and after touching of mucosa, food, locks, knobs, switches, remote control etc.
- Keep your nails short so that the virus does not hides there.
- As PH of corona virus is from 5.5 to 8.5 so therefore, we need to take more of alkaline food above the level of pH of corona virus. Some of them are: -
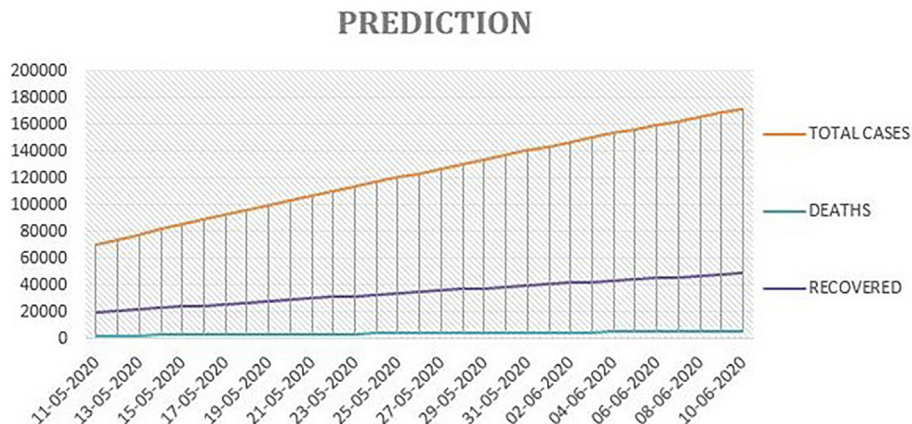- o Lemon-9.9ph
- o Avocado-15.6ph

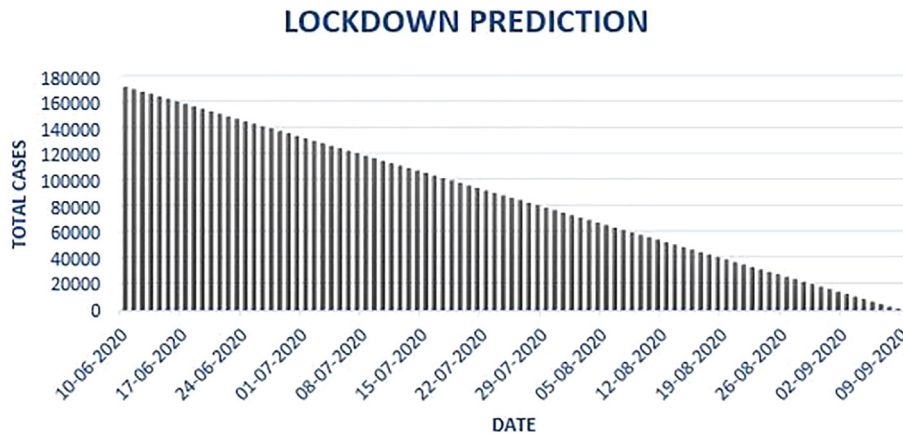**Fig. 3.** Analysis of predicted cases of COVID-19 all over India.

LOCKDOWN PREDICTION

Fig. 4. Analysis of predicted time period of lockdown.

○ Garlic-13.2ph
○ Mango-8.7ph
○ Pineapple-12.7ph
○ Orange-9.2ph

5.2. Treatment of COVID-19

No vaccine is available till date, but symptoms of covid-19 can be cured. Early and proper care from doctors and nurses' results in curing this disease. Various clinical trials are being conducted to assess potential therapeutics.

No bactericide serves. The virus is not a living organism like bacteria. We cannot kill what is not alive with antibiotics [14].

First of all, it is decided whether the person having symptoms will be tested for COVID-19 or not. Person with following are at high risk of COVID-19 positive

- Person from travel history from different countries (countries which already have COVID-19 patients)
- Close contact with person diagnosed positive
- Living locality (if the area is highly infected, possibility of test positive will be high)

These points and symptoms are kept in mind for the treatment of a patient.

- Any mix with 1-part bleach and 5 parts water directly dissolves the protein molecule and breaks it down from the inside.
- Oxygenated water helps long after soap, alcohol and chlorine, because peroxide dissolves the virus protein, but is is useful and pure and also hurts the skin.

6. Result

The graph shown above is the graphical representation of combined analysis done on COVID-19. This graph shows when did COVID-19 begin in India and according to the analysis and prediction done when will COVID-19 eradicate in India. As shown in Fig. 5 the graph is first increasing that is no. of cases are increasing but if the lockdown is followed properly keeping all the protective measures in mind, the cases will start decreasing and all the cases will be resolved till 9 september2020. The cases are decreasing linearly because linear regression is used for the prediction.
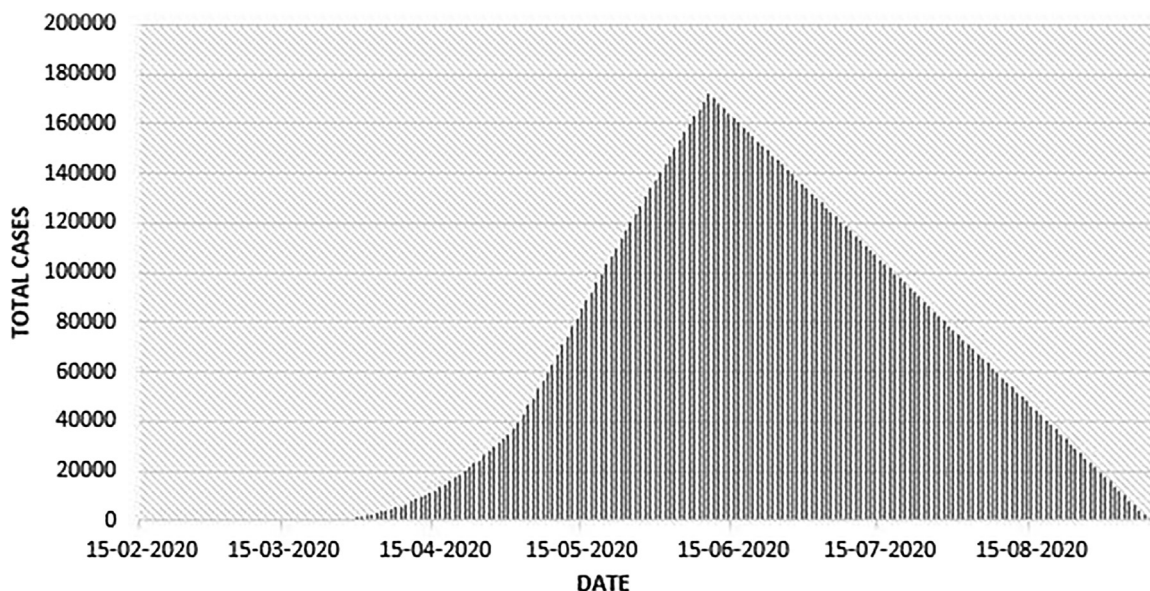

Fig. 5. Combined analysis of total no. of cases of COVID-19.

7. Conclusion

The research paper proposes a methodology to predict the extension of lockdown in order to eradicate COVID-19 from India. The proposed methodology works on the norm of machine learning more precisely linear regression. Extension of lockdown is predicted on the basis of predicted number of active cases, deaths and recovery all over India. To predict the number of active cases, deaths and recovery, date wise analysis of current data was done and necessary parameters like daily recovery, daily deaths, increase rate of covid-19 cases were included. Graphical representation of each analysis and prediction was done in order to make predicted results more understandable. The combined analysis was performed at the end which included the final result of total cases of coronavirus in India. Combined analysis included the no. of cases from start of COVID-19 to the predicted end of cases all over India. Cases predicted till 10th June came out to be 1,71,943 and according to the analysis and prediction done if lockdown will be followed properly and there will be no external factors that will affect the rate of increase of COVID-19 cases then it will take around 3 months to eradicate all the COVID-19 cases on the basis of recovery rate and death rate from India. According to the prediction and analysis done lockdown should be carried on till 9th September 2020 in order to eradicate COVID-19 from India. There are some assumptions that were included to perform the analysis and prediction-

- There is were low rise in rate of increase of cases after June 10th 2020 that is cases are near to the predicted cases
- No external factor effects rate of increase of COVID-19 cases.
- Lockdown is predicted taking help of predicted death rate and recovery rate.

All the results were predicted and analysed keeping these assumptions in mind.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] L. Li et al., Propagation analysis and prediction of the COVID-19, Infectious Disease Modelling 5 (2020) (2020) 282–292.
[2] Anuradha Tomar et al., "Prediction for the spread of COVID-19 in India and effectiveness of preventive measures", in Elsevier Science of The Total Environment, Volume 728, August 2020, 138762.
[3] Susanta Mahato et al., "Effect of lockdown amid COVID-19 pandemic on air quality of the megacity Delhi, India", in Elsevier Science of The Total Environment, Volume 730, August 2020, 139086.
[4] Sunita Tiwari et al., "Outbreak Trends of Coronavirus Disease–2019 in India: A Prediction", in Disaster Medicine and Public Health Preparedness,1-6, April 2020.
[5] Atanu Bhattacharjee et al,"When COVID-19 Will Decline in India? Prediction by Combination of Recovery and Case Load Rate", in Preprints,May 2020, 2020050044.
[6] Sina F. Ardabili, et al," COVID-19 Outbreak Prediction with Machine Learning.", in 10.20944/preprints, March 2020.
[7] L. Li et al., Characterizing the propagation of situational information in social media during COVID-19 epidemic: a case study on weibo, IEEE Trans. Comput. Soc. Syst. 7 (2) (2020) 556–562.
[8] H. Kang et al., "Diagnosis of Coronavirus Disease 2019 (COVID-19) with Structured Latent Multi-View Representation Learning," in IEEE Transactions on Medical Imaging, doi: 10.1109/TMI.2020.2992546.
[9] D. Dong et al., "The role of imaging in the detection and management of COVID-19: a review," in IEEE Reviews in Biomedical Engineering, doi: 10.1109/ RBME.2020.2990959.
[10] Y. Oh, S. Park and J. C. Ye, "Deep Learning COVID-19 Features on CXR using Limited Training Data Sets," in IEEE Transactions on Medical Imaging, doi: 10.1109/TMI.2020.2993291.
[11] R. Y. Kim, "The Impact of COVID-19 on Consumers: Preparing for Digital Sales," in IEEE Engineering Management Review, doi: 10.1109/EMR.2020.2990115.
[12] N. Zheng et al., "Predicting COVID-19 in China Using Hybrid AI Model," in IEEE Transactions on Cybernetics, doi: 10.1109/TCYB.2020.2990162.
[13] https://www.who.int/.
[14] Wadhwa, Parth, and Prabhishek Singh. "A New Sentiment Analysis based Application for Analyzing Reviews of Web Series and Movies of Different Genres." 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2020.