

Transfer-Recursive-Ensemble Learning for Multi-Day COVID-19 Prediction in India using Recurrent Neural Networks

Debasrita Chakraborty¹, Debayan Goswami², Susmita Ghosh^{2,*}, Ashish Ghosh¹, Jonathan H. Chan³, and Lipo Wang⁴

¹Technology Innovation Hub (TIH), Indian Statistical Institute, Kolkata, India

²Department of Computer Science and Engineering, Jadavpur University, Kolkata, India

³Innovative Cognitive Computing (IC2) Research Center, King Mongkut's University of Technology Thonburi, Thailand

⁴School of Electrical and Electronic Engineering, Nanyang Technological University, Singapore

*susmitaghoshju@gmail.com

ABSTRACT

The COVID-19 pandemic has put a huge challenge on the Indian health infrastructure. With a larger number of people getting affected during the second wave, hospitals were overburdened, running out of supplies and oxygen. Hence, predicting new COVID-19 cases, new deaths, and total active cases multiple days in advance can aid better utilization of scarce medical resources and prudent pandemic-related decision-making. The proposed method uses gated recurrent unit networks as the main predicting model. A study is conducted by building four models pre-trained on COVID-19 data from four different countries (United States of America, Brazil, Spain, and Bangladesh) and fine-tuned on India's data. Since the four countries chosen have experienced different types of infection curves, the pre-training provides a transfer learning to the models incorporating diverse situations into account. Each of the four models then gives 7-day ahead predictions using the recursive learning method for the Indian test data. The final prediction comes from an ensemble of the predictions of the different models. This method with two countries, Spain and Bangladesh, is seen to achieve the best performance amongst all the combinations as well as compared to other traditional regression models.

Introduction

The COVID-19 pandemic has placed India's relatively limited healthcare resources under tremendous pressure. Owing to the nation wide lockdown starting from 24 March 2020, imposed by the Indian government, the number of COVID-19 cases in the first wave were limited. However, with the opening up of the cities, its transport systems and allowance of various festivities, India has faced a very severe second wave with the peak at almost 0.4 million new cases a day. Although the peak seems to be over, the actual number of people to be affected in the coming days is very difficult to determine. There had been an exponential rise¹ of new cases in the second wave and the third wave and as of 24 July 2022, there are close to 0.15 million active COVID-19 cases in India.

COVID-19 has disrupted demand projections, which help merchants and providers of consumer products and services determine how much to purchase or create, where to purchase goods, and how much to promote or sell. During the early stages of the pandemic, abrupt curfews and a migration to working from home prompted panic purchases of various food products and household supplies. Some things were sold out, while others remained on the shelf. Insecurity abounds today on several levels. Certain items, such as toilet paper and frozen foods, are still scarce. Food retailers are stocking seasons' worth of basic necessities rather than days' worth in order to best prepare for winter season, when there could be a return of illnesses and people are anticipated to stay at home. Amidst all this, there are speculations of what the next pandemic would be. The COVID-19 pandemic has showed each country their limitations.

According to a worldwide trend, we can simply state that our existing medical capacity cannot meet the high health demands caused by the coronavirus pandemic². Infectious diseases often rise to pandemic level whenever the risk factors simultaneously happen. The circumstances can affect the availability of hospital beds, ICU beds, fans, PPEs and qualified medical staff across the country. It is therefore challenging for the authorities to supply all sectors of society with the required healthcare services. Indian medical system had similarly collapsed during the second wave of COVID-19 infection due to the high hospital admission rate.

If there was a prediction system available which could project a better estimate of the number of affected people much

earlier on, then the authorities could have maintained stocks according to that. Research has been done on various aspects of COVID-19 like COVID-19 detection³⁻⁵, number of cases prediction etc. Researchers have proposed various models for predicting the COVID-19 cases. Various mathematical models have been used to predict and to understand the spread of the disease⁶. Auto-regressive Integrated Moving Average (ARIMA) has been used⁷⁻⁹ as the standard model to predict the behaviour of the infection curves in different countries. The model was able to capture the total case statistics because the number of total cases is seen to follow a standard exponential curve. On the contrary, the number of new cases each day is highly uncertain and involves a lot many variables and therefore, it is much more difficult to handle using ARIMA^{2,10}. Researchers have explored the use of support vector machines¹¹, adaptive neuro-fuzzy inference system¹², exponential and non-linear growth models¹³ to predict the daily and active COVID-19 cases respectively. Recurrent neural networks (RNNs) have also been tested and have become state-of-the-art models for predicting the daily number of cases. Researchers¹⁴ have compared various RNN models like Long Short Term Memory (LSTM)¹⁵⁻¹⁸, Gated Recurrent Unit (GRU) and Bi-LSTMs. They have observed that these models are more robust than ARIMA or Support Vector Regression (SVR)^{19,20}. Transfer learning and ensemble modelling have also been applied to study the statistics of daily COVID-19 cases²¹ and it has shown to perform even better than the standard LSTM-RNNs. Ensemble learning in conjunction with transfer learning has, however, not been tested before. It might capture trends of multiple countries and take advantage of the knowledge of infection spread trends that the test country has not experienced before.

As we have witnessed during the peaks of the infection waves that the medical supplies were generally distributed to an area based on the real-time daily new cases in that particular area/ hospital. It takes some time to get these supplies and meanwhile the patients might be critical²². Therefore, a predictive method for the estimation of COVID-19 statistics for multiple days in advance, can provide a better framework for medical logistics. A predictive model in the similar direction has been proposed in this article. Some researchers²³⁻²⁵ have attempted the multi-step prediction of cumulative COVID-19 cases. However, multi-step prediction of daily new cases, daily fatalities and total active cases simultaneously is difficult due to their chaotic nature. The proposed model alleviates these problems and gives a multi-variable (3 COVID-19 parameters: New Cases, New Deaths and Active Cases) multi-day prediction for the daily statistics.

The proposed method draws motivation from the concept of recursive learning. The aim of recursive learning is to establish a model that can learn to fill in the missing parts in its input. In prediction tasks, the recursive model trains on its predictions²⁶. This provides the input with feedback from the output. The proposed method has been tested to give predictions for seven days ahead case. To achieve this, the proposed method uses a combination of several learning methodologies and integrates them into one. It has been shown that such a combined model is more efficient at providing multi-day forecasts than the existing standard models.

The contribution of this work can be highlighted along the following lines:

- Introduction of a transfer learning scenario for incorporating COVID-19 spread behavior in different countries.
- Recursive learning for 7-day prediction i.e., using the predictions recursively for new predictions.
- Combination of the predictions from different models using a weighted ensemble.

Rest of the article is organised as follows: The preliminaries for the proposed method is outlined in details in Section 2 with special emphasis on transfer learning and recursive learning employed. Section 3 describes the proposed method. Section 4 contains the experimental results, along with a comparison with existing standard methods. Section 5 contains a discussion of the results, along with statistical significance testing of the models. Section 6 concludes this article with the scope for future research.

Preliminaries

This section describes the dataset and the basic building blocks which have been used in the proposed model.

Dataset

The data for this study has been taken from the database of Worldometers website²⁷. This website provides COVID-19 related data including new cases, new deaths, active cases, total tests etc. for 222 different countries all along the period of the pandemic. To predict the multi-day ahead COVID-19 cases in India we have considered data related to daily new cases, daily new deaths, and active cases for six different countries: the USA, Brazil, Spain, Bangladesh, Australia and India for the period from 15 February 2020 to 16 June 2022.

Recurrent Neural Networks

The current solution relies on gated recurrent units (GRUs) as the basic building blocks for multi-day prediction of COVID-19 parameters. Gated recurrent units and long short-term memory (LSTMs), as prevalent members of recurrent neural networks (RNNs), have the innate ability to capture trends and seasonality in time-series data²⁸. They are, therefore, the go-to methods for time series predictions.

GRUs are able to solve the exploding and vanishing gradient problem that is common for vanilla RNNs²⁹. With its reset gate and update gate, a GRU is able to decide what information to keep from the previous state and what information to pass on to the next state. This gives GRUs the ability to keep relevant information from much earlier in the sequence while removing information that is no longer relevant for the task at hand. For a more detailed description of GRU, one can refer to²⁹.

GRUs are one of the simplest recurrent neural network models and have been used for time series prediction tasks in multiple domains, e.g., for traffic flow prediction³⁰, energy load forecasting³¹, stock market forecasting³², air pollution forecasting³³. It has also been used for COVID-19 prediction with the help of deep learning based models¹⁴.

The present work of multi-day COVID-19 prediction has been done using GRUs as the basic building blocks in order to harness its effective sequence modelling function and also to prevent over-fit in our relatively small dataset.

LSTMs are also used for time-series prediction in multiple domains³⁴⁻³⁷. In the present work, we have compared the performances of GRUs with that of LSTMs, considering both as the basic building blocks.

Transfer Learning

Transfer learning is the scenario where a pre-trained model for one particular problem is applied to a second, different but related problem³⁸. Transfer learning tries to take advantage of what has already been learned in a problem and applies it to improve the generalization in another related problem.

The domain in which the model is trained is called the source domain and the domain in which the model is applied is called the target domain³⁹. The source and the target domain may be different enough but need to have some sort of a relation. The predictive model in the source domain needs to be similar to that of target domain in order for transfer learning to work. Transfer learning is mainly applied in such target domains where sufficient labeled data is not available⁴⁰.

Transfer learning has been applied in the COVID-19 scenario for different tasks. It has been used for classification of COVID-19 from non-COVID-19 patients by using chest CT images⁴¹, for face-mask detection in public areas⁴², COVID-19 cases and death forecasts using LSTMs⁴³ etc.

In the present work, transfer learning has been chosen for the task of COVID-19 case prediction to learn from the experiences of countries affected by COVID-19. Countries with different circumstances, different climates, different measures for infection control are chosen as the source domain and COVID-19 cases prediction for India is done as the target domain. In one of our previous works²¹, it is seen that transfer learning has given better results for next day prediction for COVID-19 cases using LSTMs. In this present work, we are exploiting transfer learning with recursion for multi-day ahead prediction. The details are given below.

Recursive Learning

The GRU model built in this work is able to predict the next day parameter after looking at the parameters over a period of past days (called the look-back period). As mentioned, in order to achieve a multi-day prediction of COVID-19 cases, a recursive learning methodology is adopted as shown in Figure 1.

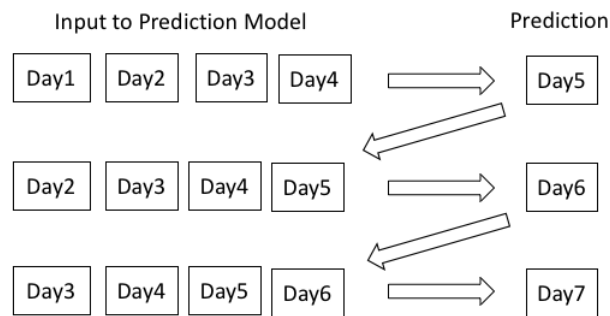


Figure 1. Recursive learning used in COVID-19 prediction

In a recursive way, the predicted output of the model is fed back to the input in the next step to obtain the subsequent prediction. As an example, in Figure 1, data for day 1 to day 4 is used as input to the model to predict the data for day 5. In the

next step, the data for day 5 is added to the input and the data for day 2 to day 5 is taken as input to the model to predict the output for day 6. This process is repeated recursively till the required days of prediction is obtained.

This recursive learning methodology uses the sliding window approach with the previous predictions being used as a part of the input to do future predictions. It is intuitive that the performance will vary with the change in look-back period.

In this work, this recursive learning methodology works for 7 steps to predict the COVID-19 cases for the next 7-days. However, this process of recursion can be used for prediction of COVID-19 cases for any number of days in advance.

Ensemble Learning

Through ensemble learning one could exploit the unique abilities of multiple models in an integrated manner by combining the results obtained from various models. In this work, the results obtained through recursive learning approach initiated from various transfer-learned models (trained on data from respective countries) are ensembled to obtain final predictions. Several ensemble techniques exist in literature^{44,45}. In the present approach, we have proposed a weighted ensemble technique.

Performance Metric

To assess the performance of the proposed approach, two metrics have been used for this study. The first is the relative mean squared error (R-MSE) and the second one is relative mean absolute error (R-MAE). Rather than considering actual value, calculating the error as a fraction of the actual value is seen to be effective. As the error value is compared relative to the actual value of the parameter, hence the term relative has been used for the standard error metrics of mean squared error and mean absolute error.

These metrics are defined as follows:

$$Relative - MSE(R - MSE) = \frac{\sum_{i=1}^d \left(\frac{original(i) - predicted(i)}{original(i)} \right)^2}{d} \quad (1)$$

$$Relative - MAE(R - MAE) = \frac{\sum_{i=1}^d \left| \frac{original(i) - predicted(i)}{original(i)} \right|}{d} \quad (2)$$

where, $predicted(i)$ is the predicted value on the i^{th} day and $original(i)$ is the actual value on the i^{th} day. d is the total number of days involved. Note that, R-MSE better reflects the error for higher deviations than that of R-MAE as it penalizes higher deviations with a greater error value.

For each of the three COVID-19 parameters predicted (daily new cases, daily new deaths and total active cases), these two errors values (as an average of all the prediction errors over the test set) have been shown in the results section.

Proposed Method

The proposed model is an ensemble of four different models pre-trained on data from four different countries (The United States of America (USA), Brazil, Spain and Bangladesh) in order to predict the COVID-19 daily new cases, daily new deaths and active cases for India for the next 7-days. The idea was to learn the infection spread in the worst affected country from each of the seven continents.

The USA and Brazil were obvious choices from the North American and South American continents due to high number of COVID-19 infections. The USA has witnessed the most number of cases. The first wave in the USA was prolonged and the second wave has resulted in an increased number of deaths per day. The third and fourth waves also have a similar pattern. The USA has not witnessed a plateau in the total number of cases after the first wave. However, the daily new cases have decreased considerably after the peak of the third wave. Brazil has the third highest number of cases and currently the death rate is 3,015 per million population, which is one of the highest in the world. However, the total cases curve in Brazil has a prolonged second wave and severe third wave. Spain has been chosen from the European continent due to its well marked first, second and third waves of infection as compared to the initial plateau of infections in Italy. Bangladesh has been chosen from the Asian continent to incorporate similar climatic conditions and being a neighbouring country of the test country, India. South Africa was first taken into consideration from the African continent, however it was not incorporated into the model as the number of cases in the waves of infection have been quite low as compared to the other countries selected. Australia was also taken into consideration, but was not introduced in the model due to the very late nature of infection with it still being in the first wave of infections. To incorporate population density into account, the data of each country is divided by the corresponding population density.

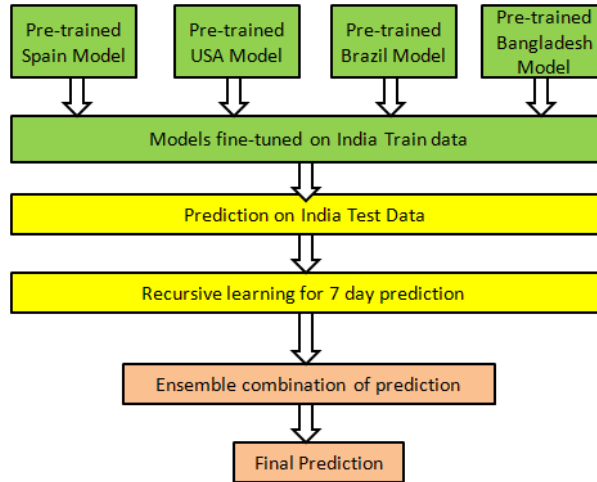


Figure 2. Flowchart of the proposed method

India witnessed a decline in the number of daily cases which suggested the ending of the third wave. However, cases have started increasing in some parts of the country again, signalling a possible fourth wave. This obviously brings us to the crucial question about the condition in India about the subsequent waves of infection. Now, since all of the four countries have shown different trends, it is not sure as to which path Indian trend would follow. This is why all possible combinations of these four countries were taken into account. More countries could have been taken for pre-training, but that would have added more complexity to the model.

Training of each of these models has been done using a sliding window technique with a look back period of 14 days. This look back period is the number of days in the past the model looks at while making the prediction of cases for the next day. This way the relation between the number of cases on successive days can be learned by the model for giving better predictions. The look back period has been varied from 7 to 19 in order to find the best look back period for the 7-day ahead prediction task.

The proposed model is shown in Figure 2. It consists of three main steps which are discussed in the subsequent sections.

Step 1: Train-Test Splitting of Input Data

The data from the period 15 February 2020 to 31 December 2021 has been used for training the models for the individual four countries. This period has been chosen to take 80% of the total data for training. Indian data for the period 15 February 2021 to 31 December 2021 has been used for fine-tuning the transfer learning models before testing them on Indian data for the period 16 January 2022 to 16 June 2022. The remaining period of 01 January 2022 to 15 January 2022 for the Indian data has been used for cross-validation in the ensemble weighted averaging as mentioned in the subsequent paragraphs.

Step 2: Forming the Model

Two RNN models (LSTMs and GRUs) have been taken independently as the building blocks for the proposed method. The proposed RNN models with parameters and their values chosen are given in Table 1. As mentioned earlier, the GRU and the LSTM models are trained and tested in order to do performance comparison of the two RNN types.

Parameter	Values
Layers	3
Number of Neurons	30,80,50
Learning Rate	0.001
Optimizer	Adam

Table 1. Values of the parameters of the RNN models

Step 3: Transfer Learning

As stated earlier, the proposed model consists of all sixteen possible combinations of four RNN networks of two types (GRUs and LSTMs), each of which is pre-trained on data from four different countries. To pre-train each model, the data of each of the countries is taken from 15 February 2020 to 31 December 2021, as mentioned earlier. The models built on the individual countries need to be fine-tuned on Indian data in order to take into account the recent trend of COVID-19 infections in the target country, India. Therefore, the pre-trained models have been fine-tuned on Indian data for the same period 15 February 2020 to 31 December 2021.

This method of pre-training followed by fine-tuning introduces a transfer learning⁴⁶ ability in the individual models. Li et. al.⁴⁷ has shown that transfer learning can improve forecasting models that are based on deep learning. They have built a source domain of 12 countries, combining their data, and tried to predict the confirmed cases per million for the target countries. However, the scope of the study is limited by the prediction of just one COVID-19 parameter and also being tested on a shorter time period (31/12/2019 to 31/05/2020). Whereas, in the proposed approach, three COVID-19 parameters are predicted over an advance period of 7 days using an ensembled approach of transfer-learnt recursive model on a larger time period.

Step 4: Recursive Learning

To obtain 7-day ahead predictions, we have incorporated recursive learning in the proposed model. If the COVID-19 parameters for the next 7-days is to be predicted at the n^{th} day, in the first step, the COVID-19 parameters for the next day ($(n+1)^{th}$ day) will be predicted by the model. This prediction for the $(n+1)^{th}$ day is then fed back to the input to make the new input frame for predicting the COVID-19 parameters for the $(n+2)^{th}$ day. This process is repeated 7 times to get the predicted values for the 7-days ($(n+1)^{th}$ day to $(n+7)^{th}$ day). An example of the process involved is shown in Figure 1. This method can be applied to predict the COVID-19 parameters for a different time period by changing the period setting. However, to do a stable comparison, the prediction period has been set at 7-days for this study. Experiments are also done to study the effect of varying the look-back period used for prediction.

Step 5: Ensemble

As there is diversity in the situations of the 4 different countries involved in the model (e.g., lock down measure, vaccination strategy, climatic conditions etc), different weights need to be given to them for further prediction. The weight values are considered based on errors obtained through validation data.

To do so, once the predictions are obtained for subsequent 7-days, the predictions from the combination of models are aggregated using weighted averaging. The weights are calculated based on the relative mean squared error (R-MSE) obtained from model cross-validation on validation data. 15 days of data for India (01 January 2022 to 15 January 2022) is kept aside for this validation task.

The relative mean squared error (R-MSE) for the validation data is calculated using the equation(1). The weights (w_i) for the Model i are given by equation(3), and the final prediction on any date D is given by equation(4). We have also compared our proposed ensemble method with an equally weighted ensemble technique.

$$w_i = \frac{1/R - MSE(Model\ i)}{\sum_{j=1}^n 1/R - MSE(Model\ j)}, \quad (3)$$

where, $R - MSE(Model\ i)$ is the relative *mse* obtained for the i^{th} model on validation data, n is the number of models involved in the ensemble. The transfer learnt model with less error is given a higher weight as per equation(3).

$$prediction(D) = \sum_{i=1}^n w_i * prediction_i(D), \quad (4)$$

where, $prediction_i(D)$ is the prediction by the i^{th} model for date D. This is a weighted summation of the predictions from the transfer learnt models.

Experimental Results

In order to predict the cases on any date, we need to use a look-back period. This period, in this case, is the number of days our model looks at for doing the predicting task. Finding the optimum value of the look-back period is crucial for the proposed method. This is because, depending on the look-back period, the performance of the models varies rapidly.

It is to be noted that, since we are relying on a recursive learning based multi-day prediction, where the predicted values are used as inputs for the subsequent predictions, we cannot afford to take the look-back to be smaller than the number of days in

Method	New Cases				New Deaths				Active Cases			
	R-MSE		R-MAE		R-MSE		R-MAE		R-MSE		R-MAE	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Spain	0.0043	0.0025	0.0586	0.0224	0.0048	0.0032	0.0654	0.0283	0.0016	0.001	0.0258	0.0189
USA	0.0052	0.0027	0.0709	0.0188	0.0065	0.0018	0.0886	0.0166	0.0015	0.001	0.0209	0.015
Brazil	0.0038	0.0013	0.0528	0.0231	0.0032	0.0012	0.0436	0.0137	0.001	0.0004	0.0174	0.0115
Bangladesh	0.0035	0.0024	0.0477	0.0262	0.0037	0.0025	0.0504	0.0212	0.0018	0.0013	0.0265	0.0198
Spain-Bangladesh	0.0013	0.0009	0.0177	0.0127	0.0021	0.0011	0.0286	0.0249	0.0009	0.0003	0.017	0.011
Brazil-Bangladesh	0.0025	0.0016	0.035	0.015	0.0033	0.0012	0.0450	0.0138	0.0011	0.0005	0.0189	0.0128
Spain-Brazil	0.0028	0.0014	0.0382	0.0144	0.0026	0.0013	0.0357	0.0148	0.001	0.0007	0.0189	0.0126
Spain-USA	0.0032	0.0023	0.0442	0.0194	0.0029	0.0021	0.0395	0.0203	0.0014	0.001	0.0221	0.0159
USA-Brazil	0.0038	0.0015	0.0518	0.0139	0.0038	0.0012	0.0518	0.0131	0.001	0.0008	0.0176	0.0117
USA-Bangladesh	0.0026	0.0017	0.0354	0.0201	0.0034	0.0018	0.0463	0.0171	0.0015	0.001	0.0222	0.0157
Brazil-USA-Bangladesh	0.0027	0.0017	0.0368	0.0154	0.0036	0.0012	0.0491	0.0137	0.0011	0.001	0.0187	0.0127
Spain-Brazil-Bangladesh	0.0016	0.0009	0.0218	0.0162	0.0027	0.0013	0.0374	0.0156	0.0011	0.0008	0.0199	0.0137
Spain-Brazil-USA	0.002	0.001	0.0273	0.0151	0.0028	0.0013	0.0382	0.0147	0.0011	0.001	0.0188	0.0128
Spain-USA-Bangladesh	0.0019	0.0011	0.0259	0.0205	0.0025	0.002	0.0341	0.0202	0.0014	0.0008	0.0228	0.0165
USA-Brazil-Bangladesh-Spain	0.002	0.0013	0.0273	0.0163	0.0025	0.0013	0.0348	0.0154	0.0011	0.0006	0.0197	0.0135
USA-Brazil-Bangladesh-Spain - Australia	0.0024	0.0014	0.0312	0.0243	0.0029	0.002	0.0388	0.0285	0.0015	0.0006	0.0231	0.0269
India	0.0037	0.0021	0.0519	0.0253	0.0042	0.0034	0.0492	0.0310	0.0014	0.001	0.0233	0.0192
SVR (RBF)	1.4627	NA	1.1538	NA	0.1351	NA	0.4286	NA	0.7528	NA	0.9521	NA
SVR (Poly)	0.8453	NA	0.9521	NA	0.0948	NA	0.2794	NA	0.4821	NA	0.7364	NA
ARIMA	0.2351	NA	0.4928	NA	0.0537	NA	0.2263	NA	0.1302	NA	0.4826	NA
Facebook Prophet	0.0673	NA	0.2972	NA	0.0185	NA	0.1492	NA	0.0464	NA	0.2273	NA

Table 2. Comparison of results i.e. Mean and Standard Deviation(STD) of 20 runs, from all possible combinations of four countries considered in this study. The GRU models have been fine-tuned on Indian data. Results are also shown for SVR, ARIMA and Facebook Prophet models.

the multi-day prediction task. This will result in the last few predictions (of the recursive learning methodology) being made only on predictions and not on any actual data. We have experimented with a wide range of look-back periods and a value of 14 gives the best results for all the three variables.

The values of R-MSE and R-MAE (averaged over 20 runs) obtained for all 16 combinations using the proposed model (using both GRUs and LSTMs) with those for the support vector regression (SVR), auto-regressive integrated moving average (ARIMA) and Facebook Prophet models are shown in the Tables 2 and 3. The results are shown separately for the 3 COVID-19 parameters predicted in this study (new cases, new deaths and active cases). For establishing the efficacy of fine-tuning on Indian data, results obtained for the above mentioned methods, using GRUs, without fine-tuning are also put in Table 4. The results put in this table are the average of 20 independent runs. It is seen that the combination model of Spain and Bangladesh gives the best results for multi-day forecasting of all the three predicted variables.

Discussion

It is clearly visible from Tables 2 and 3 that the results are similar for both GRUs and LSTMs, with fine-tuning on Indian data. Also, all the models seem to improve with the fine-tuning on Indian data as depicted in Tables 2 and 3 with respect to Table 4. As expected, in transfer learning, if a model pre-trained on data from the source domain is fine-tuned on data from the transfer domain, the results seem to improve with respect to the model just pre-trained on the data from source domain⁴⁶.

Since GRUs and LSTMs yield similar results as seen in Tables 2 and 3, GRUs have been preferred over LSTMs for the present experimentation as they have lesser number of parameters with respect to LSTMs⁴⁸. Hence, for the rest of this discussion, results obtained only with the GRU networks are analyzed further.

New Cases

For the single country models, Bangladesh model gives the best results, followed by Brazil, Spain and USA. The model built using the data from Bangladesh is able to predict the trend in Indian data in a very accurate way. For the combination of two-country models, the presence of Bangladesh (with better trend tracking behaviour for Indian data) influences the performance in a positive way. Spain-Bangladesh combination gives the best result with the R-MSE of 0.0013 and R-MAE of 0.0177, and yields the best result amongst all the combinations. For the combination of more than two-country models, the performance does not improve further than the Spain-Bangladesh model. Overall, the two-country combination of Spain-Bangladesh model gives the best performance.

Method	New Cases				New Deaths				Active Cases			
	R-MSE		R-MAE		R-MSE		R-MAE		R-MSE		R-MAE	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Spain	0.0042	0.0038	0.0584	0.0263	0.0063	0.0031	0.0665	0.027	0.0024	0.0011	0.0268	0.0151
USA	0.0054	0.0035	0.0710	0.0201	0.0064	0.0023	0.0889	0.015	0.0009	0.00042	0.0201	0.0162
Brazil	0.0041	0.0028	0.0526	0.0328	0.0030	0.0021	0.0440	0.017	0.0014	0.001	0.0186	0.0151
Bangladesh	0.0034	0.0023	0.0476	0.0317	0.0049	0.0015	0.0513	0.026	0.0016	0.0005	0.0260	0.019
Spain-Bangladesh	0.0015	0.0012	0.0178	0.0129	0.0027	0.0017	0.0299	0.017	0.0013	0.0005	0.0176	0.012
Brazil-Bangladesh	0.0027	0.0018	0.0348	0.012	0.0046	0.0029	0.0461	0.0217	0.0015	0.0007	0.0191	0.0132
Spain-Brazil	0.0028	0.0013	0.0383	0.0156	0.0027	0.0016	0.0347	0.0174	0.0016	0.0009	0.0203	0.0171
Spain-USA	0.0029	0.0016	0.0441	0.0173	0.0040	0.0023	0.0400	0.0214	0.0014	0.0009	0.0219	0.0153
USA-Brazil	0.0040	0.0024	0.0518	0.0169	0.0045	0.0028	0.0528	0.0142	0.0020	0.00064	0.0190	0.0141
USA-Bangladesh	0.0025	0.002	0.0354	0.0263	0.0036	0.0019	0.0469	0.0325	0.0026	0.0009	0.0241	0.0182
Brazil-USA-Bangladesh	0.0027	0.0018	0.0370	0.0157	0.0042	0.0016	0.0505	0.0229	0.0016	0.0008	0.0195	0.0133
Spain-Brazil-Bangladesh	0.0017	0.0011	0.0219	0.0131	0.0035	0.0028	0.0361	0.0151	0.0021	0.0006	0.0213	0.0188
Spain-Brazil-USA	0.0017	0.0011	0.0272	0.0183	0.0027	0.0019	0.0388	0.0216	0.0020	0.0011	0.0200	0.014
Spain-USA-Bangladesh	0.0020	0.0017	0.0258	0.0193	0.0039	0.0028	0.0357	0.0317	0.0026	0.0015	0.0236	0.0158
USA-Brazil-Bangladesh-Spain	0.0023	0.0015	0.0275	0.0181	0.0027	0.0015	0.0363	0.0273	0.0015	0.0009	0.0209	0.013
USA-Brazil-Bangladesh-Spain - Australia	0.0025	0.0014	0.0311	0.0212	0.003	0.0017	0.0421	0.0253	0.0019	0.0007	0.0292	0.0192
India	0.0035	0.0033	0.0427	0.0323	0.0047	0.0029	0.0535	0.0382	0.0022	0.0013	0.0225	0.0120

Table 3. Comparison of results i.e. Mean and Standard Deviation(STD) of 20 runs, from all possible combinations of four countries considered in this study. Similar configuration LSTM models, fine-tuned on Indian data, have been used in this case.

Method	New Cases				New Deaths				Active Cases			
	R-MSE		R-MAE		R-MSE		R-MAE		R-MSE		R-MAE	
	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD	Mean	STD
Spain	0.0065	0.0035	0.1806	0.0254	0.0102	0.0026	0.2010	0.034	0.0102	0.001	0.2007	0.0194
USA	0.0071	0.0047	0.1943	0.018	0.011	0.0015	0.2149	0.018	0.0117	0.001	0.2182	0.023
Brazil	0.0042	0.0023	0.1248	0.0235	0.0099	0.0019	0.1685	0.0194	0.0088	0.0008	0.1628	0.0182
Bangladesh	0.0039	0.0014	0.1824	0.0241	0.0096	0.0024	0.2180	0.028	0.0107	0.0017	0.2232	0.021
Spain-Bangladesh	0.0021	0.0013	0.0758	0.0152	0.0044	0.001	0.1203	0.027	0.0074	0.0007	0.1287	0.014
Brazil-Bangladesh	0.0028	0.0015	0.1229	0.018	0.0068	0.0032	0.1495	0.0149	0.0097	0.0005	0.1652	0.0159
Spain-Brazil	0.0031	0.0017	0.1957	0.0123	0.0065	0.0019	0.2036	0.0193	0.0099	0.001	0.2222	0.0292
Spain-USA	0.0035	0.0026	0.1592	0.0182	0.0075	0.0028	0.2026	0.0321	0.0077	0.0009	0.2037	0.0172
USA-Brazil	0.0037	0.0013	0.2608	0.0157	0.0067	0.0015	0.2749	0.0184	0.0111	0.00072	0.2281	0.0315
USA-Bangladesh	0.0032	0.0027	0.2166	0.0232	0.0073	0.0021	0.2594	0.0452	0.0086	0.002	0.2485	0.0273
Brazil-USA-Bangladesh	0.003	0.0023	0.0948	0.0192	0.006	0.0014	0.1715	0.0241	0.0080	0.003	0.1894	0.0153
Spain-Brazil-Bangladesh	0.0022	0.0012	0.0969	0.0185	0.0057	0.0017	0.1375	0.0174	0.0076	0.0001	0.1409	0.0142
Spain-Brazil-USA	0.0029	0.0015	0.2039	0.011	0.0082	0.0017	0.2316	0.0182	0.0094	0.0013	0.2377	0.0182
Spain-USA-Bangladesh	0.0028	0.0014	0.0979	0.032	0.005	0.0024	0.1277	0.0328	0.0085	0.0009	0.1489	0.0172
USA-Brazil-Bangladesh-Spain	0.0024	0.0016	0.0990	0.0249	0.0052	0.0018	0.1321	0.0182	0.0083	0.0007	0.1511	0.0142
USA-Brazil-Bangladesh-Spain - Australia	0.0029	0.0018	0.1820	0.038	0.0064	0.0025	0.1585	0.0271	0.0093	0.0006	0.1709	0.0249

Table 4. Comparison of results i.e. i.e. Mean and Standard Deviation(STD) of 20 runs, from all possible combinations of four countries considered in this study. The GRU models have not been fine-tuned on Indian data.

New Deaths

For the single country models, Brazil model gives the best results, followed by Bangladesh, Spain and USA. Bangladesh is again one of the better models, similar to the case of new cases. For the combination of two-country models, Spain-Bangladesh model gives the best results with a R-MSE of 0.0021 and R-MAE of 0.0286. Addition of data from the other two countries is not able decrease this error any further.

Active Cases

Prediction of active cases gives the best results as compared to the other two parameters. For the single country models, Brazil gives the best results, followed by USA, Spain and Bangladesh. For all the other models, the results improve only marginally for the Spain-Bangladesh model with a R-MSE of 0.0009 and R-MAE of 0.017. Rest of the results are all similar with no further improvement in the error metric.

Analysis

For the two-country models, the Spain-Bangladesh combination (highlighted by light gray in Table 2), gives a lower R-MSE than Bangladesh alone. There is an improvement in the result when pre-trained Spain model is combined with the pre-trained Bangladesh model.

Once the combination of Spain-Bangladesh model has been built, further addition of models built with the other two countries (Brazil, USA) data does not reduce the error any further. This may be due to the following facts: Brazil has a high number of cases and has similar geography like that of India, the infections spread happened later than that of India.

Bangladesh has similar climatic conditions and people's behaviour like that of India. Also, the percentage of people vaccinated with respect to the total population is similar in India and Bangladesh. As a result infection trend in Bangladesh has a positive impact in predicting infection trend in India. It may also be noted that, both India and Spain were vigilant at the start of the pandemic and had imposed strict infection control measures like lockdowns, social distancing etc. One such study⁴⁹ also corroborated this and compared the spread of COVID-19 infection in Spain and India by analysing the policy implications using epidemiological and social media data. Spain was one of the early COVID-19 infected countries, which is already at the end of fourth wave of infections. Whereas, India is at the end of the third wave. Also the spread increases sharply and then falls rapidly in both India and Spain. Such similar characteristics might be responsible for the low error in predictions obtained for the models built with Spain data.

It is to be noted that, the India model shown in the Table 2 is one where the GRU model is built with India data and then trained on India data i.e., it does not involve any transfer learning. It can be clearly seen from Tables 2 and 4 that the transfer learnt models are better predictors of all the three COVID-19 parameters. Models built with support vector regression (SVR), both with polynomial and RBF kernels, are unable to predict the COVID-19 parameters with a good level of accuracy. Same is the case with ARIMA and Facebook Prophet⁵⁰. Different nature of infection spread in different waves is difficult to take into account for SVR and ARIMA based predictions. This is seen from the high errors of prediction obtained using these models. Predicting daily new deaths is especially very uncertain. Presence of comorbidity, age etc. play a significant role in new deaths.

Hence, the proposed method, with the advantage of transfer learning from Bangladesh and Spain data combined, is able to predict the number of daily new cases, daily new deaths and active cases with the least error.

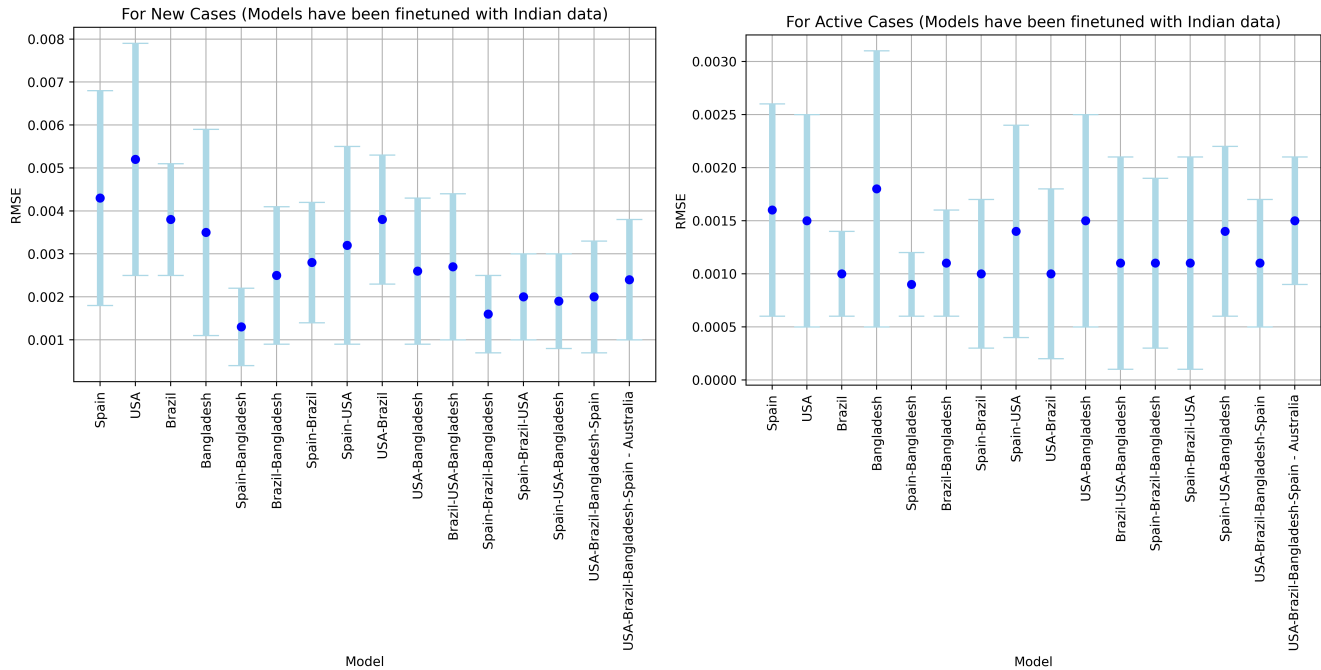
Standard deviation (STD) of the results are also studied. STD values for 20 runs for the prediction of new cases and active cases for all the models fine-tuned on India data, are given in light blue in Figures 3a and 3b respectively, showing clearly the Spain-Bangladesh combination to be the best model.

Ensemble Type	New Cases		New Deaths		Active Cases	
	R-MSE	R-MAE	R-MSE	R-MAE	R-MSE	R-MAE
Proposed Weighted Ensemble	0.0013	0.0177	0.0021	0.0286	0.0009	0.017
Equally Weighted Ensemble	0.0049	0.0635	0.0055	0.0791	0.0021	0.0384

Table 5. Comparison of results for Spain-Bangladesh model (with fine-tuning), using the proposed weighted ensemble and equally weighted ensemble

The variation of prediction error with look-back period for new cases and active cases are shown in Figures 4a and 4b. The prediction error is seen to be the least for the look-back period of 14 for all the three cases.

As mentioned earlier in the present method, a weighted ensemble is used for combining the models built with individual countries. Performance of the weighted ensemble method is compared with that of an equally weighted ensemble (where simple average is taken for combining the models) and the results for the best model (Spain-Bangladesh) is shown in Table 5. It is seen that our proposed method with weighted ensemble performs better than the equally weighted ensemble approach.



(a) Average (over 20 runs) R-MSE after fine-tuning (blue dot) along with the standard deviation (light blue bars) for new cases prediction for all the models (b) Average (over 20 runs) R-MSE after fine-tuning (blue dot) along with the standard deviation (light blue bars) for active cases prediction for all the models

Figure 3. R-MSE and STD over 20 runs for all the models

Statistical Significance Testing

The proposed method was tested on a total of 167 test sets, with each set having a duration of 7 days (i.e., 7 day prediction). In order to do a statistical significance testing of the best model (Spain-Bangladesh model) obtained in our comparisons, it has been statistically tested against each of the other single country and two-country models using the Wilcoxon signed rank test.

More than two country models were not used for calculating the statistical significance as they can be thought of as extensions of the two-country models. Only results for the prediction errors of new cases has been used in this testing.

Wilcoxon signed rank test is a non-parametric test for doing hypothesis testing of paired dependent samples. More details on Wilcoxon signed rank test can be found in⁵¹.

For doing the statistical significance testing, 167 number of R-MSE values (for the 167 test sets) for two models involved are treated as 167 number of paired observations. Wilcoxon signed rank test has been used over paired t-test as the difference between the observations are seen to be not normally distributed. Here,

Null hypothesis - H_0 : There is no difference between the model and Spain-Bangladesh model

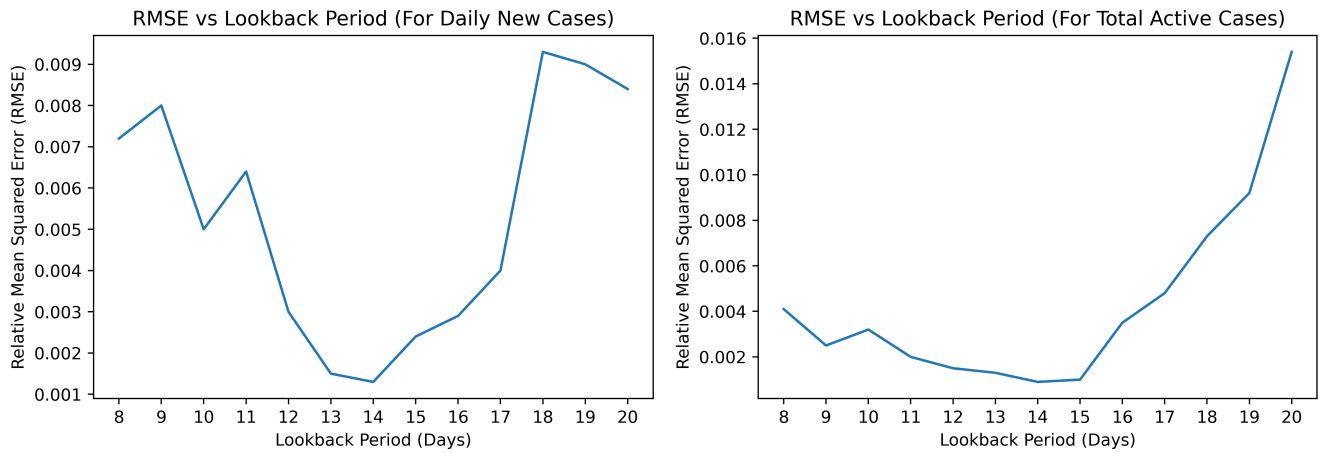
Alternate hypothesis - H_1 : There is a difference between the model and Spain-Bangladesh model

Two tailed hypothesis with a significance level of 0.05 has been used for the testing. The Z-score for each of the one-country and two-country models when compared with Spain-Bangladesh model is given in Table 6. All the Z-scores in the Table 6 are less than -1.96 which is the critical Z-score for a two-tailed test at a significance level of 0.05. Hence, it can be concluded that the Spain-Bangladesh model is statistically significant from the other models in consideration.

Conclusion and Future Work

Prediction of various COVID-19 parameters, 7 days in advance, can really be a game changer for better handling of pandemics. Proper prediction can help better plan the allocation of scarce medical resources like life-saving medicines, oxygen, and hospital beds. Understanding of the infection spread, well in advance, can also lead to better administrative decisions.

The proposed method for the prediction of new cases, new deaths and active cases for COVID-19 for India with the combination of Spain-Bangladesh model outperforms the other combinations as well as other traditional regression models considered in our experiment. This is because the proposed method leverages the capabilities of both transfer learning and ensemble learning while taking into account the excellent sequence modelling abilities of GRUs. The multi-day ahead prediction using recursive learning provides an added benefit of knowing the COVID-19 statistics multiple days ahead. The proposed



(a) Prediction error (R-MSE) for new cases vs Look-back period (b) Prediction error (R-MSE) for active cases vs Look-back period

Figure 4. Prediction error (R-MSE) vs Look-back period for Spain-Bangladesh model

Model	Z-Score
Spain	-9.59
USA	-10.81
Brazil	-10.79
Bangladesh	-7.63
Brazil - Bangladesh	-6.29
Spain - Brazil	-9.38
Spain - USA	-8.28
USA - Brazil	-10.85
USA - Bangladesh	-7.28

Table 6. Wilcoxon signed rank test Z-score of the models with respect to Spain-Bangladesh Model

method has currently been tested only on India data and can be extended for any country considering relevant source countries as the basis for transfer learning.

This study is data driven and does not explicitly take into account the steps adopted by the government to handle the pandemic. The work is experimental in nature by considering various combinations of source countries to evaluate the behaviour of the models. Better prediction could be made if such policy decisions are incorporated into the model.

The proposed method can be extended for individual states of India by incorporating information of other Indian states or other countries having similar features. Similarly, this could also be explored for predicting the cases for other countries. Individual waves of infection could also be analysed by using the transfer learning phenomenon from other waves of infection.

Data availability

The data used for this study was taken from the database of Worldometers website²⁷. Data used for this study is available from the corresponding author S.G. on request.

References

1. Mondal, S. & Ghosh, S. Fear of exponential growth in Covid-19 data of India and future sketching. *MedRxiv* DOI: <https://doi.org/10.1101/2020.04.09.20058933> (2020).
2. Pandey, G., Chaudhary, P., Gupta, R. & Pal, S. Seir and regression model based COVID-19 outbreak predictions in India. *arXiv preprint arXiv:2004.00958* (2020).
3. Mahanty, C., Kumar, R., Asteris, P. G. & Gandomi, A. H. COVID-19 patient detection based on fusion of transfer learning and fuzzy ensemble models using CXR images. *Appl. Sci.* **11**, DOI: [10.3390/app112311423](https://doi.org/10.3390/app112311423) (2021).

4. Mahanty, C., Kumar, R. & Patro, S. G. Internet of medical things-based COVID-19 detection in CT images fused with fuzzy ensemble and transfer learning models. *New Gener. Comput.* **40**, 1125–1141, DOI: [10.1007/s00354-022-00176-0](https://doi.org/10.1007/s00354-022-00176-0) (2022).
5. Mahanty, C., Kumar, R., Mishra, B. K. & Barna, C. COVID-19 detection with X-ray images by using transfer learning. *J. Intell. Fuzzy Syst.* **43**, 1717–1726, DOI: [10.3233/JIFS-219273](https://doi.org/10.3233/JIFS-219273) (2022).
6. Sanchez-Caballero, S., Selles, M. A., Peydro, M. A. & Perez-Bernabeu, E. An efficient COVID-19 prediction model validated with the cases of China, Italy and Spain: Total or partial lockdowns? *J. Clin. Medicine* **9**, 1547 (2020).
7. Achterberg, M. A. *et al.* Comparing the accuracy of several network-based COVID-19 prediction algorithms. *Int. J. Forecast.* **38**, 489–504, DOI: <https://doi.org/10.1016/j.ijforecast.2020.10.001> (2020).
8. Hyndman, R. J. & Khandakar, Y. Automatic time series forecasting: The forecast package for R. *J. Stat. Software, Articles* **27**, 1–22, DOI: <https://doi.org/10.18637/jss.v027.i03> (2008).
9. Singh, R. K. *et al.* Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model. *JMIR Public Heal. Surveillance* **6**, e19115 (2020).
10. Mandal, M. *et al.* A model based study on the dynamics of COVID-19: Prediction and control. *Chaos, Solitons & Fractals* **136**, 109889 (2020).
11. Singh, V. *et al.* Prediction of COVID-19 corona virus pandemic based on time series data using support vector machine. *J. Discret. Math. Sci. Cryptogr.* **23**, 1583–1597, DOI: <https://doi.org/10.1080/09720529.2020.1784535> (2020).
12. Al-qaness, M. A. *et al.* Efficient artificial intelligence forecasting models for COVID-19 outbreak in Russia and Brazil. *Process. Saf. Environ. Prot.* **149**, 399–409, DOI: <https://doi.org/10.1016/j.psep.2020.11.007> (2021).
13. Mahanty, C. *et al.* Prediction of COVID-19 active cases using exponential and non-linear growth models. *Expert. Syst.* **39**, e12648, DOI: <https://doi.org/10.1111/exsy.12648> (2022).
14. Shahid, F., Zameer, A. & Muneeb, M. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. *Chaos, Solitons & Fractals* **140**, 110212 (2020).
15. Alazab, M. *et al.* COVID-19 prediction and detection using deep learning. *Int. J. Comput. Inf. Syst. Ind. Manag. Appl.* **12**, 168–181 (2020).
16. Elsheikh, A. H. *et al.* Deep learning-based forecasting model for COVID-19 outbreak in Saudi Arabia. *Process. Saf. Environ. Prot.* **149**, 223–233, DOI: <https://doi.org/10.1016/j.psep.2020.10.048> (2021).
17. Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural Comput.* **9**, 1735–1780, DOI: <https://doi.org/10.1162/neco.1997.9.8.1735> (1997).
18. Nikparvar, B., Rahman, M. M., Hatami, F. & Thill, J.-C. Spatio-temporal prediction of the COVID-19 pandemic in US counties: modeling with a deep LSTM neural network. *Sci. Reports* **11**, 21715, DOI: [10.1038/s41598-021-01119-3](https://doi.org/10.1038/s41598-021-01119-3) (2021).
19. Saba, A. I. & Elsheikh, A. H. Forecasting the prevalence of COVID-19 outbreak in Egypt using nonlinear autoregressive artificial neural networks. *Process. Saf. Environ. Prot.* **141**, 1–8, DOI: <https://doi.org/10.1016/j.psep.2020.05.029> (2020).
20. Yamak, P. T., Yujian, L. & Gadosey, P. K. A comparison between ARIMA, LSTM, and GRU for time series forecasting. In *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence, ACAI 2019*, 49–55, DOI: <https://doi.org/10.1145/3377713.3377722> (Association for Computing Machinery, New York, NY, USA, 2019).
21. Chakraborty, D., Goswami, D., Ghosh, A., Chan, J. & Ghosh, S. Learning from others: A data driven transfer learning based daily new COVID-19 case prediction in India using an ensemble of LSTM-RNNs. In *The 12th International Conference on Advances in Information Technology, IAIT2021*, DOI: [10.1145/3468784.3470769](https://doi.org/10.1145/3468784.3470769) (Association for Computing Machinery, New York, NY, USA, 2021).
22. Moran, K. R. *et al.* Epidemic forecasting is messier than weather forecasting: the role of human behavior and internet data streams in epidemic forecast. *The J. Infect. Dis.* **214**, S404–S408 (2016).
23. Fritz, C., Dorigatti, E. & Rügamer, D. Combining graph neural networks and spatio-temporal disease models to improve the prediction of weekly COVID-19 cases in Germany. *Sci. Reports* **12**, 3930, DOI: [10.1038/s41598-022-07757-5](https://doi.org/10.1038/s41598-022-07757-5) (2022).
24. Rajesh, A., Pai, H., Roy, V., Samanta, S. & Ghosh, S. COVID-19 prediction for India from the existing data and SIR (D) model study. *medRxiv* DOI: <https://doi.org/10.1101/2020.05.05.20085902> (2020).

25. Ribeiro, M. H. D. M., da Silva, R. G., Mariani, V. C. & dos Santos Coelho, L. Short-term forecasting COVID-19 cumulative confirmed cases: Perspectives for Brazil. *Chaos, Solitons & Fractals* **135**, 109853, DOI: <https://doi.org/10.1016/j.chaos.2020.109853> (2020).
26. Yoo, H. & Pimmel, R. L. Short term load forecasting using a self-supervised adaptive neural network. *IEEE Transactions on Power Syst.* **14**, 779–784, DOI: <https://doi.org/10.1109/59.761912> (1999).
27. Worldometers. Worldometers.info. <https://www.worldometers.info/coronavirus/> (2020). Accessed: 2022-05-10.
28. Sherstinsky, A. Fundamentals of recurrent neural network (RNN) and long short-term memory (LSTM) network. *Phys. D: Nonlinear Phenom.* **404**, 132306 (2020).
29. Cho, K. *et al.* Learning phrase representations using RNN encoder-decoder for statistical machine translation (2014). [1406.1078](https://arxiv.org/abs/1406.1078).
30. Fu, R., Zhang, Z. & Li, L. Using LSTM and GRU neural network methods for traffic flow prediction. In *2016 31st Youth Academic Annual Conference of Chinese Association of Automation (YAC)*, 324–328, DOI: [10.1109/YAC.2016.7804912](https://doi.org/10.1109/YAC.2016.7804912) (2016).
31. Kumar, S., Hussain, L., Banarjee, S. & Reza, M. Energy load forecasting using deep learning approach-LSTM and GRU in spark cluster. In *2018 Fifth International Conference on Emerging Applications of Information Technology (EAIT)*, 1–4, DOI: [10.1109/EAIT.2018.8470406](https://doi.org/10.1109/EAIT.2018.8470406) (2018).
32. Althelaya, K. A., El-Alfy, E.-S. M. & Mohammed, S. Stock market forecast using multivariate analysis with bidirectional and stacked (LSTM, GRU). In *2018 21st Saudi Computer Society National Computer Conference (NCC)*, 1–7, DOI: [10.1109/NCG.2018.8593076](https://doi.org/10.1109/NCG.2018.8593076) (2018).
33. Tao, Q., Liu, F., Li, Y. & Sidorov, D. Air pollution forecasting using a deep learning model based on 1D CONVNETS and bidirectional GRU. *IEEE Access* **7**, 76690–76698, DOI: [10.1109/ACCESS.2019.2921578](https://doi.org/10.1109/ACCESS.2019.2921578) (2019).
34. Gers, F. A., Eck, D. & Schmidhuber, J. Applying LSTM to time series predictable through time-window approaches. In *Neural Nets WIRN Vietri-01*, 193–200 (Springer, 2002).
35. Karevan, Z. & Suykens, J. A. Transductive LSTM for time-series prediction: An application to weather forecasting. *Neural Networks* **125**, 1–9 (2020).
36. Li, Y., Zhu, Z., Kong, D., Han, H. & Zhao, Y. EA-LSTM: Evolutionary attention-based LSTM for time series prediction. *Knowledge-Based Syst.* **181**, 104785 (2019).
37. Yadav, A., Jha, C. & Sharan, A. Optimizing LSTM for time series prediction in Indian stock market. *Procedia Comput. Sci.* **167**, 2091–2100 (2020).
38. Torrey, L. & Shavlik, J. Transfer learning. In *Handbook of research on machine learning applications and trends: Algorithms, methods, and techniques*, 242–264 (IGI global, 2010).
39. Pan, S. J. & Yang, Q. A survey on transfer learning. *IEEE Transactions on Knowl. Data Eng.* **22**, 1345–1359 (2009).
40. Weiss, K., Khoshgoftaar, T. M. & Wang, D. A survey of transfer learning. *J. Big data* **3**, 1–40 (2016).
41. Pathak, Y., Shukla, P. K., Tiwari, A., Stalin, S. & Singh, S. Deep transfer learning based classification model for COVID-19 disease. *IRBM* **43**, DOI: <https://doi.org/10.1016/j.irbm.2020.05.003> (2020).
42. Horry, M. J. *et al.* COVID-19 detection through transfer learning using multimodal imaging data. *IEEE Access* **8**, 149808–149824 (2020).
43. Gautam, Y. Transfer learning for COVID-19 cases and deaths forecast using LSTM network. *ISA Transactions* DOI: <https://doi.org/10.1016/j.isatra.2020.12.057> (2021).
44. Dietterich, T. G. Ensemble methods in machine learning. In *International workshop on multiple classifier systems*, 1–15 (Springer, 2000).
45. Zhang, C. & Ma, Y. *Ensemble machine learning: Methods and applications* (Springer, 2012).
46. Yang, Q., Zhang, Y., Dai, W. & Pan, S. J. *Transfer learning* (Cambridge University Press, 2020).
47. Li, Y. *et al.* ALeRT-COVID: Attentive lockdown-aware transfer learning for predicting COVID-19 pandemics in different countries. *J. Healthc. Informatics Res.* **5**, 98–113, DOI: <https://doi.org/10.1007/s41666-020-00088-y> (2021).
48. ArunKumar, K., Kalaga, D. V., Mohan Sai Kumar, C., Kawaji, M. & Brenza, T. M. Comparative analysis of gated recurrent units (GRU), long short-term memory (LSTM) cells, autoregressive integrated moving average (ARIMA), seasonal autoregressive integrated moving average (SARIMA) for forecasting COVID-19 trends. *Alex. Eng. J.* **61**, 7585–7603, DOI: <https://doi.org/10.1016/j.aej.2022.01.011> (2022).

49. Asawa, P., Gaur, M., Roy, K. & Sheth, A. COVID-19 in Spain and India: Comparing policy implications by analyzing epidemiological and social media data, DOI: <https://doi.org/10.48550/arXiv.2010.14628> (2020).
50. Taylor, S. J. & Letham, B. Forecasting at scale. *The Am. Stat.* **72**, 37–45 (2018).
51. Woolson, R. F. Wilcoxon signed-rank test. *Wiley encyclopedia clinical trials* 1–3 (2007).

Acknowledgements

Support from the Science and Engineering Research Board (SERB) via sanctioning an ASEAN-India collaborative research project "A Multi Modal Approach to Medical Diagnosis Embedding Deep and Transfer Learning" (IMRC/AIST/DF/CRD/2019/000151 dated 29 June 2020) is gratefully acknowledged. Part of the research was also supported by Technology Innovation Hub (TIH) at Indian Statistical Institute, Kolkata, India.

Author contributions statement

D.C., D.G. and S.G. conceived the study and designed the algorithm. D.C. and D.G. extracted the data set, conducted the experiments, and drafted the manuscript. Results were analyzed by all the authors and further experiments were conducted to finalize the set of results. S.G., A.G., J.C. and L.W. enhanced and finalized the manuscript.

Competing interests

The authors have no competing financial or non-financial interests that are relevant to the content of this article.