Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
Spatio-temporal variation of Covid-19 health outcomes in India using deep learning based models

Asif Iqbal Middya, Sarbani Roy ∗

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

ARTICLE INFO

Keywords:
Covid-19
Deep learning
Spatio-temporal variation

ABSTRACT

Deep learning methods have become the state of the art for spatio-temporal predictive analysis in a wide range of fields, including environmental management, public health, urban planning, pollution monitoring, and so on. Despite the fact that a variety of powerful deep learning-based models can address various problem-specific issues in different research domains, it has been found that no single optimal model can outperform everywhere. Now, in the last two years, various deep learning-based studies have provided a variety of best-performing techniques for predicting COVID-19 health outcomes. In this context, this study attempts to perform a case study that investigates the spatio-temporal variation in the performance of deep-learning-based methods for predicting COVID-19 health outcomes in India. Various widely applied deep learning models namely CNN (convolutional neural network), RNN (recurrent neural network), Vanilla LSTM (long short-term memory), LSTM Autoencoder, and Bidirectional LSTM are considered to investigate their spatio-temporal performance variation. The effectiveness of the models is assessed using various metrics based on COVID-19 mortality time-series from 36 states and union territories of India.

1. Introduction

Spatio-temporal data analysis is becoming increasingly important due to the increasing availability of large scale data in various fields of research. Deep learning models have recently been used in spatio-temporal predictive analysis in a variety of applications, including pollution monitoring (e.g., particulate matter prediction), climate monitoring (e.g., precipitation forecasting), public health monitoring (e.g., COVID-19 health outcome prediction), and so on. However, due to the spatio-temporal variations in the data of target variables, the models’ performance varies greatly both spatially and temporally. In the aforementioned context, the primary objective of this research is to investigate spatio-temporal variation in the performance of deep learning-based methods for estimating COVID-19 health outcomes in India.

At the beginning of 2020, the world health organization announced the COVID-19 outbreak a public health emergency of global concern. So far, COVID-19 has been associated with over 256 million confirmed cases and about 5.2 million fatalities globally. Many countries seek to protect their citizens’ health by imposing mitigation strategies including social distancing, quarantines, travel restrictions, hard and soft locks, as well as postponing and cancellations of activities. In this context, several deep learning-based models for predicting COVID-19 health outcomes have been introduced in the existing works to aid in healthcare resources management and mitigation planning. Specifically, the accurate prediction could be helpful to decide the goals, guiding principles, and strategies to mitigate future consequences due to COVID-19.

Various deep learning based previous studies (Mohimont et al., 2021; Chimmula and Zhang, 2020; Zeroual et al., 2020; Said et al., 2021; Shastri et al., 2020) from around the world have yielded various top-performing methods for forecasting COVID-19 health outcomes. It is observed that there is no single best deep learning based forecasting model that can outperform everywhere. This is because there is a wide variation that exist in the COVID-19 health outcome time-series. It is worth noting that the variation in COVID-19 health outcome time-series depends on various risk factors that vary greatly across the geographical space. Moreover, several literature (Middya and Roy, 2021; Sannigrahi et al., 2020) highlights that there exists a geographically varying nature of the association between COVID-19 health outcomes and the complex risk factors. Hence, it is expected that there should have some variation in the models’ performance. In the aforementioned context, this study attempts to explore the spatio-temporal variation in the effectiveness of the deep learning based predictive methods for predicting covid-related health outcomes.

It is worth noting that the deep learning techniques can learn and extract features from raw and imperfect data on their own. Specifically,
these techniques could be helpful in a time series forecasting scenario for automatically learning the temporal dependency from the data. Moreover, advanced deep learning techniques, e.g., LSTM (long short-term memory) (Middya and Roy, 2022; Nath et al., 2021; Das et al., 2021) and its variants, are highly capable in detecting patterns from input data that span lengthy periods of time. These techniques can help to address the problem by removing the requirement for extensive feature engineering activities, data scaling techniques, and differentiating to make the data stationary. The classical timeseries forecasting methods cannot perform well if there exist complex associations between the variables, noisy data, missing data, and irregular temporal patterns in data. In other words, to perform successfully, classical timeseries forecasting methods often require clean and comprehensive data sets. On the other hand, the deep learning based techniques are resistant to noise in the input data, and can even learn and forecast in the existence of missing data.

There are several existing deep learning based forecasting methods for COVID-19 health outcome forecasting. For instance, in Mohimont et al. (2021), Mohimont et al. employed temporal CNN to forecast various health outcomes, e.g., number of confirmed cases, deaths, etc. By using LSTM (long short-term memory) networks, Chimmula and Zhang (2020) predicted the COVID-19 transmission in Canada. Moreover, some previous works (Zeroual et al., 2020; Sait et al., 2021; Shastri et al., 2020) explore various LSTM variants such as Bi-LSTM, Stacked LSTM, etc for covid-19 prognosis. However, to the best of our knowledge, none of the existing literature focuses on investigating spatio-temporal variation in the performance of deep learning based models. This study attempts to explore such spatio-temporal variation in the models’ performance and also identify spatial and temporal distribution of the methods for predicting covid-related health outcomes in India. The major contributions of this work can be summarized in the following points.

• Explore spatio-temporal variability pattern in the performance of deep learning based forecasting models (namely CNN, RNN, Vanilla LSTM, LSTM Autoencoder, and Bidirectional LSTM) for predicting COVID-19 health outcomes in India.

• Thoroughly evaluate the performance of the methods based on various performance indicators for COVID-19 mortality time-series data of 36 states and union territories of India for various forecast horizons.

• Finding spatial distribution of best-performing models for predicting Covid-19 mortality data for different temporal resolutions (e.g., one week, two weeks, etc.).

The article is structured as follows. Section 3 provides the problem definition and the presented approach. In Section 4, experimental setting, evaluation metrics, and results are provided. Future directions and conclusion are provided in Section 5.

2. Related work

This section discusses existing works that are primarily concerned with spatiotemporal data modelling of COVID-19 health outcomes. Specifically, in the following paragraphs, we present related works, their limitations/gaps, and how our work attempts to bridge the gaps.

Several previous works capture patterns of COVID-19 progression using mathematical, machine learning, and deep learning techniques (John et al., 2022). Kavouras et al. (2022) studied the efficacy of employing deep learning techniques to properly model COVID-19 transmission. However, they focus on low-granular (country-level) data modelling and analysis over EU (European Union) regions. Furthermore, they created models that can only predict COVID-19 health outcomes on a daily basis (i.e., short forecast horizon). Nikparvar et al. (2021) conducted spatio-temporal forecasting of COVID-19 health outcomes in US counties. However, they only used vanilla LSTM model and did not verify the effectiveness of their proposed model with the advanced deep learning models, e.g., other LSTM variants. In da Silva et al. (2021), the authors used spatial and time-dependent features to perform monitoring and real-time spatio-temporal prediction of COVID-19 in Brazil. They only focuses on shallow machine learning based models (e.g., random forests, support vector machines, etc.) and eventually end up with low predictive accuracy. Bhimala et al. (2021) used the meteorological parameter integrated deep learning technique to estimate COVID-19 incidents for various states of India. However, they only consider a vanilla LSTM model for relatively short forecast horizon. Conventional epidemiological methods (e.g., the Susceptible–exposed–infectious–removed (SEIR) method and variations) (Tang et al., 2020; Zhao et al., 2020; Fanelli and Piazza, 2020) have also been used to quantify viral dissemination and predict the influence of policy actions on infection rates. However, these methods suffers from various shortcomings. For instance, these methods are reliant on a large number of hypothesized input factors. Since SEIR methods are so sensitive to variations in these input factors, their predicted accuracy might suffer significantly.

In contrast to earlier research, this study focused on data modelling to investigate if advanced deep learning models can better capture the spatiotemporal variability of COVID-19 health outcomes. Specifically, a variety of advanced deep learning models namely CNN, RNN, Vanilla LSTM, LSTM Autoencoder, and Bidirectional LSTM are built to accurately perform state-level high granular COVID-19 variability analysis in India. This research work also focuses on developing models with a relatively longer forecast horizon (maximum forecast horizon of 21 days, i.e., 3 weeks). Moreover, compared to the existing works, the advanced deep learning models are rigorously evaluated on multiple forecast horizons to verify the models’ scalability.

3. Problem statement and proposed approach

3.1. Problem statement

Suppose \( S = \{S_1, S_2, S_3, \ldots, S_n\} \) is a set of states (=28) and union territories (=8) of India. Given a total of \( n = 28 + 8 = 36 \) different COVID-19 death time-series datasets corresponding to the states and union territories. Let, \( D_i = \{V_{i,1}, V_{i,2}, V_{i,3}, \ldots, V_{i,m}\} \) denotes the time series dataset of daily COVID-19 death count for \( S_i \in S \). Here, \( V_{i,j} \) is the number of COVID-19 deaths on \( j \)th day in \( S_i \). Suppose, \( M = \{M_1, M_2, \ldots, M_k\} \) is a set of popular deep learning based time-series forecasting models that could be used for predicting future values of COVID-19 health outcomes. The goal of this research work is to determine the optimal set of forecasting models \( M^{(o)} \subseteq M \) and a mapping between \( S \) and \( M^{(o)} \), i.e., \( \sigma : \{p : S_i \in S\} \rightarrow \{q : M_k \in M^{(o)}\} \). More specifically, the objective is to find a state-wise spatial mapping of optimal deep learning based methods for estimating the number of covid-related deaths in India. Moreover, this study attempts to explore spatial variability in the performance of the optimal forecasting models.

3.2. Proposed approach

Fig. 1 provides a detailed framework for exploring spatio-temporal variation in the performance of deep learning based methods for estimating state-level COVID-19 mortality in India. This section demonstrates how state-level COVID-19 mortality time-series data is acquired, preprocessed, batched, and subsequently utilized for model training and performance evaluation.

3.2.1. Dataset description

On the basis of daily time-series data from India, location-specific optimal forecasting models are found. The COVID-19 death time series data are collected for a total of 36 states and union territories of India. More specifically, 36 time-series datasets (one for each of the states and the union territories) are constructed. The raw COVID-19 data are collected up to November 01, 2021 from two sources namely (i)
Fig. 1. Overall approach of spatio-temporal variation in the performance of deep learning based methods for estimating state-level COVID-19 mortality in India.

MOHFW website\(^1\) (ii) COVID19INDIA website\(^2\). COVID19INDIA is a crowd-sourcing project that documents COVID-19 data from India’s states and union territories. Each of the 36 datasets contains about 576 days’ data samples. Geographical distributions of state-level death counts are presented in Fig. 2. More than 46% of the states report more than 10000 COVID-19 related death cases. The state of Maharashtra is the most severely affected by COVID-19, with more than 140000 deaths.

3.2.2. Data pre-processing and batching

This section provides the details of the pre-processing of the raw time-series data and batching of the pre-processed data. The COVID-19 mortality time-series for each state and UTs is smoothed with an exponential function to lessen the jaggedness of the curve. Eq. (1) provides the mathematical form of exponential the smoothing function.

\[
Z_k = \delta X_k + (1 - \delta) Z_{k-1}, \quad k > 0
\]  

where \( \delta \) represents the smoothing factor and \( \delta \in (0, 1) \). The raw COVID-19 mortality time series data is denoted by \( \{X_k\} \) starts at \( k = 0 \) and the smoothed output is represented as \( \{Z_k\} \). Note that Eq. (1) considers that \( Z_0 = X_0 \).

The data are divided into train and test sets once it has been smoothed. The time-series data cannot be shuffled since the order is determined by ‘timestamp’. As a result, the test set is made up of the last 150 days of data, which constitutes around 25% of the whole dataset. After that, we fit a standard scaler to the training data only, then use that scaler to standardize both the training and test data. Standardization can be expressed mathematically in Eq. (2).

\[
u = \frac{x - \bar{x}}{\sigma}
\]

where, \( x \) denotes data point, \( \bar{x} \) and \( \sigma \) represent mean and standard deviation respectively.

Next, the time-series sequence data are converted into the supervised form which we call the batching process. Let us consider, \( Seq = \{v_1, v_2, v_3, \ldots, v_n\} \) is a sequence of data values found after standardizing the mortality time-series. The sequence of data can be split into several input-output vectors, where \( n \) consecutive values are considered as an input vector and next \( t \) consecutive values are considered as an output vector (i.e., target vector). If the input and corresponding output vectors are denoted by \( x \) and \( t \) respectively, the scaled data after batching can be represented as \( L = [(x_1, t_1), (x_2, t_2), \ldots, (x_r, t_r)] \), where, \( x_r = [v_{i1}, v_{i+1}, v_{i+2}, \ldots, v_{i+t}] \) and \( t_r = [v_{i+t+1}, v_{i+t+2}, v_{i+t+3}, \ldots, v_{i+2t}] \). Note that, \( z = r - (n + t - 1) \). The details of the batching process are provided in Algorithm 1.

3.2.3. Deep learning models

This section discusses various deep learning models that are employed in predicting state-level COVID-19 health outcomes in India. Spatial variation in the models’ performance could then be used in finding the geographical distribution of best-performing models.

Convolutional neural network (CNN)

\(1\) http://www.mohfw.gov.in/
\(2\) http://www.covid19india.org/
The CNN (Albawi et al., 2017; Zhao et al., 2017) is usually used to extract features from two-dimensional data and is commonly utilized for image processing. However, because of its ability to learn both local and global characteristics from sequence data, CNNs are also used in several time-series applications. In this context, a one-dimensional CNN (1D-CNN) is employed for forecasting COVID-19 health outcomes from COVID-19 mortality time-series data. CNN’s major feature learning layers are the convolutional and pooling layers. The following is the procedure (i.e., Eq. (3)) for performing a 1D convolutional operation

$$ (x * s)[i] = \sum_{j=0}^{m-1} x[i-j] \cdot s[j-m] $$

(3)

where, $x$ and $s$ are the input vector and the kernel respectively; $m$ is the kernel size. Across the feature maps produced by the convolution layer, the pooling is utilized to decrease the size of the feature map.

Recurrent neural network (RNN)

For sequence prediction tasks, RNNs (Hewamalage et al., 2021) are the most often utilized neural network architecture. RNN consists several RNN units. Elman developed the ERNN (Elman RNN) cell, which is a type of basic recurrent unit (Elman, 1990). An ERNN cell could be mathematically expressed using Eqs. (4) and (5).

$$ J_i = \sigma \left( K_i \cdot x + Q_i \cdot J_{i-1} + b_i \right) $$

(4)

$$ z_i = \tanh \left( Q_o \cdot J_i + b o \right) $$

(5)

where the input and output of the cell at $i$th time step are $x_i$ and $z_i$ respectively, $J_i$ indicates hidden state. For the hidden state, $Q_i$ and $K_i$ represent the weight matrices, and $b_i$ represents the bias vector. Similarly, $Q_o$ and $b_o$ are the cell output’s weight matrix and bias vector, respectively. Here, $\sigma(\cdot)$ represents the sigmoid function. Note that the present hidden state is determined by the earlier time step’s hidden state as well as the present input. For the lengthy sequences, the RNNs suffer from the exploding gradient and vanishing gradient issues. This means that RNN will not be capable to carry long-range dependencies into the future.

Vanilla LSTM

Many variants of the basic RNN have been produced over time to overcome its flaws (Hewamalage et al., 2021). LSTM model, which is able to learn long-range dependencies, is possibly the most popular branch of RNN. An LSTM cell has two states: the internal cell state and the hidden state, which are different from the standard RNN cell. Eq. (6) - Eq. (12) describes an LSTM cell mathematically

$$ i = \sigma \left( K_i \cdot x + Q_i \cdot J_{i-1} + b_i \right) $$

(6)

$$ o = \sigma \left( K_o \cdot x + Q_o \cdot J_{i-1} + b_o \right) $$

(7)

$$ f = \sigma \left( K_f \cdot x + Q_f \cdot J_{i-1} + b_f \right) $$

(8)

The capability of LSTM to identify patterns in time series data over several time-series applications. In this context, a one-dimensional CNN (1D-CNN) is employed for forecasting COVID-19 health outcomes from COVID-19 mortality time-series data. CNN’s major feature learning layers are the convolutional and pooling layers. The following is the procedure (i.e., Eq. (3)) for performing a 1D convolutional operation

$$ (x * s)[i] = \sum_{j=0}^{m-1} x[i-j] \cdot s[j-m] $$

(3)

where, $x$ and $s$ are the input vector and the kernel respectively; $m$ is the kernel size. Across the feature maps produced by the convolution layer, the pooling is utilized to decrease the size of the feature map.

Recurrent neural network (RNN)

For sequence prediction tasks, RNNs (Hewamalage et al., 2021) are the most often utilized neural network architecture. RNN consists several RNN units. Elman developed the ERNN (Elman RNN) cell, which is a type of basic recurrent unit (Elman, 1990). An ERNN cell could be mathematically expressed using Eqs. (4) and (5).

$$ J_i = \sigma \left( K_i \cdot x + Q_i \cdot J_{i-1} + b_i \right) $$

(4)

$$ z_i = \tanh \left( Q_o \cdot J_i + b o \right) $$

(5)

where the input and output of the cell at $i$th time step are $x_i$ and $z_i$ respectively, $J_i$ indicates hidden state. For the hidden state, $Q_i$ and $K_i$ represent the weight matrices, and $b_i$ represents the bias vector. Similarly, $Q_o$ and $b_o$ are the cell output’s weight matrix and bias vector, respectively. Here, $\sigma(\cdot)$ represents the sigmoid function. Note that the present hidden state is determined by the earlier time step’s hidden state as well as the present input. For the lengthy sequences, the RNNs suffer from the exploding gradient and vanishing gradient issues. This means that RNN will not be capable to carry long-range dependencies into the future.

Vanilla LSTM

Many variants of the basic RNN have been produced over time to overcome its flaws (Hewamalage et al., 2021). LSTM model, which is able to learn long-range dependencies, is possibly the most popular branch of RNN. An LSTM cell has two states: the internal cell state and the hidden state, which are different from the standard RNN cell. Eq. (6) - Eq. (12) describes an LSTM cell mathematically

$$ i = \sigma \left( K_i \cdot x + Q_i \cdot J_{i-1} + b_i \right) $$

(6)

$$ o = \sigma \left( K_o \cdot x + Q_o \cdot J_{i-1} + b_o \right) $$

(7)

$$ f = \sigma \left( K_f \cdot x + Q_f \cdot J_{i-1} + b_f \right) $$

(8)
0.2 are used for max pooling and the dropout layers respectively. The output of ConvB2 is flattened by the flatten layer which is followed by two dense layers. In the first dense layer, the number of units ($N_u$) is 48. On the other hand, in the second dense layer, $N_u$ is the value of the forecast horizon (i.e., the number of output steps). Fig. 3(b) presents the architecture of the LSTM auto-encoder. The encoder part contains two LSTM layers ($L_1$ and $L_2$) with 100 and 50 units respectively. Similarly, the decoder part has two LSTM layers ($L_1$ and $L_2$) with 50 and 100 units. There is a repeat vector layer between the encoder and decoder parts. Finally, similar to CNN, there are two dense layers with the same $N_u$ values. The architecture of the RNN based predictive model is presented in Fig. 3(c). It made up of an RNN layer followed by a dropout layer and two Dense layers. Here, $N_u$ in the RNN layer and the dropout rate in the dropout layer are $N_u = 80$ and $D_t = 0.2$ respectively. Fig. 3(d) presents the architecture of the vanilla LSTM. It has a single LSTM layer with 100 units followed by a dropout layer ($D_t = 0.2$) and two dense layers. Finally, the bi-directional LSTM model’s architecture is shown in Fig. 3(e). There are two bi-directional LSTM layers each with $N_u$ of 70. These layers are followed by an LSTM layer (with $N_u = 50$), a dropout layer ($D_t = 0.2$), and two dense layers.

As previously stated, approximately 25% of the most recent COVID-19 mortality data is included in the test set, while the remaining 75% is used for training purposes. When creating learning models, hyper-parameter optimization is a critical step. The manual process of trying random combinations takes a long time because identifying the set of optimal parameters is a time-consuming procedure. To combat this, a parameter tuning technique known as grid-search is applied to various parameter combinations to identify the optimal set of parameter values. The architecture of the vanilla LSTM and RNN based predictive models are developed using the following major Python libraries: (i) numpy (version 1.19.5), (ii) pandas (version 1.1.5), (iii) keras (version 2.7.0), (iv) sklearn (0.23.1). Moreover, for the visualization of spatial data, QGIS (Quantum GIS) is utilized.

4. Evaluation

4.1. Experimental setup

The numerical experiments are conducted under a Windows environment on a PC with 3.30 GHz Intel(R) Xeon(R) CPU E3-1226 v3 processor and 128 GB of RAM. Python (https://docs.python.org/3.7/) is used to implement all of the deep learning-based COVID-19 health outcome prediction models and various analyses. The predictive models are developed using the following major Python libraries: (i) numpy (version 1.19.5) (ii) pandas (version 1.1.5) (iii) keras (version 2.7.0) (iv) sklearn (0.23.1). Moreover, for the visualization of spatial data, QGIS (Quantum GIS) is utilized.

4.2. Evaluation metrics

Various evaluation metrics (namely mean absolute error, root mean squared error, and coefficient of determination) are used to assess the effectiveness of the models for predicting COVID-19 mortality counts. The mathematical form of root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination ($R^2$) are expressed in Eq. (20), Eq. (21), and Eq. (22) respectively.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{T}(O_i - E_i)^2}{T}}$$

$$MAE = \frac{1}{T} \sum_{i=1}^{T} |O_i - E_i|$$

$$R^2 = 1 - \frac{RSS}{TSS}$$
Fig. 4. Spatial distribution of MAE values for different models. (a) LSTM Auto-Encoder (b) Bi-LSTM (c) CNN (d) vanilla LSTM (e) RNN.

Fig. 5. Spatial distribution of RMSE values for different models. (a) LSTM Auto-Encoder (b) Bi-LSTM (c) CNN (d) vanilla LSTM (e) RNN.
A.I. Middya and S. Roy

Fig. 6. Box and swarm plots of MAE values for various predictive models over different forecast horizons (a) one week (b) two weeks (c) three weeks.

Fig. 7. Box and swarm plots of RMSE values for various predictive models over different forecast horizons (a) one week (b) two weeks (c) three weeks.

Fig. 8. Box and swarm plots of $R^2$ values for various predictive models over different forecast horizons (a) one week (b) two weeks (c) three weeks.

where, $O_i$ and $E_i$ are $i$th observed and the predicted value, $T$ is the total number of predictions, sum of squares of residuals is represented by $RSS$, and total sum of squares is represented by $TSS$.

4.3. Spatial variability in the models' performance

For different states and union territories, deep learning based forecasting techniques are built using state-level COVID-19 mortality data. The geographical distribution of MAE (mean absolute error) and RMSE (root mean squared error) values for different states and union territories are provided in Figs. 4 and 5 respectively. The accuracies are computed for a 3-week forecast horizon. It is observed that for the north-western (primarily the states of Rajasthan and Punjab) and south-eastern (primarily the states of Andhra Pradesh, Telangana, and Odisha) parts of India, the CNN-based predictive models perform better in terms of MAE and RMSE values. Almost all of the models achieve better accuracy in the northern and eastern parts. In the southern parts (primarily the states of Kerala and Tamil Nadu) the LSTM auto-encoder performs relatively better compared to other methods with respect to both the MAE and RMSE values. In comparison to the other models, the Bi-LSTM model performs better in the western regions (primarily the states of Gujarat, Maharashtra, and Karnataka).

The distribution of the models’ COVID-19 health outcome predictive accuracies are also observed to vary significantly over the forecast horizon. Fig. 6, Fig. 7, and Fig. 8 present the distribution of MAE, RMSE, and $R^2$ values over three forecast horizons (1 week, 2 weeks, and 3 weeks). The accuracy distribution is visualized using box plots that overlap with the corresponding swarm plots. The box plots show various distribution summaries namely minimum value, first quartile, median value, third quartile, and the maximum value. Note that a swarm plot corresponding to a box plot indicates where the data point falls in the distribution summaries. For the box plots of MAE (Fig. 6) and RMSE (Fig. 7) values, it is clear that all the methods perform well for a shorter horizon. For a shorter forecast horizon, the interquartile range (IQR) for different models is smaller. It implies that the predictive accuracy across most of the states does not differ significantly for shorter forecast horizons. The increase in IQR due to the increase of forecast horizon is less in the case of Bi-LSTM model. Moreover, for both the MAE and RMSE, the visual analysis reveals that the models like Bi-LSTM and CNN achieve better median accuracies across the forecast horizons. Fig. 8 shows relatively high $R^2$ values are by Bi-LSTM and CNN models with less IQR.

The observation versus prediction plots of COVID-19 death count for some of the highly covid-affected states is provided in Fig. 9. For a particular state, the plots are provided only for the best-performing model. The results show that the estimated COVID-19 death counts are closely spread along the $x = y$ line relative to the observed values in the majority of cases. The models are unable to predict higher values for some states and union territories (e.g., Maharashtra, Delhi, etc.).

4.4. Spatial distribution of optimal models

This section presents the spatially varying distribution of best-performing models for predicting COVID-19 death counts in India.
There is no single best method that can outperform across different states in India due to the wide variation in COVID-19 mortality time-series. Fig. 10 shows the distribution of the state-specific optimal models in India for different forecast horizons (1 week, 2 weeks, and 3 weeks). It is interesting to observe how the distribution of best-performing models changes over different forecast horizons. Fig. 10(a) presents the models' distribution when the forecast horizon is 1 week. Note that in the case of a shorter horizon, both the basic models such as CNN and RNN dominates for most of the states. For instance, CNN mainly dominate in the southern, northern and eastern parts of India. On the other hand, RNN mainly dominates in the central and western parts of India. However, in the case of a longer forecast horizon, some advanced architecture like Bi-LSTM, LSTM auto-encoder outperform instead of RNN. Bi-LSTM and CNN achieve significantly better results across the states in all three forecast horizons. Bi-LSTM provides optimal performance in approximately 14%, 39%, and 34% of the states for a forecast horizon of 1 week, 2 weeks, and 3 weeks respectively. On the other hand, CNN achieves optimal performance in approximately 42%, 28%, and 39% of the states for a forecast horizon of 1 week, 2 weeks, and 3 weeks respectively.

### 4.5. Discussion

The existing studies (Middya and Roy, 2021) reveal that the COVID-19 health outcomes depend on complex factors including socio-economic, environmental pollution, meteorological parameters, local strategies for mitigation, and public health measures. The impact of these factors varies greatly across the geographical space in the countries like India. Consequently, it is not unexpected that the nature of COVID-19 mortality time-series varies across different states. Due to the geographically varying nature of the association between COVID-19 health outcomes and the complex risk factors, the predictive models’ performance differs significantly across the geographic space.

The results of various deep learning models highlight that they could be adopted by most of the states to forecast COVID-19 health outcomes in advance. Specifically, deep learning methods like Bi-LSTM and CNN produce overall better predictive accuracies compared to the others. However, the findings show that selecting an appropriate forecasting model is influenced not only by the geographic location but also by the forecast horizon for which the prediction is made. For example, in terms of the shorter horizon (e.g., 1 week), CNN dominates in India’s southern, northern, and eastern regions, whereas RNN dominates in the country’s central and western regions. In 65% of the states and union territories, the CNN and RNN models perform better. However, the basic RNN does not perform well over longer time horizons, such as 2 or 3 weeks, because it is unable to carry long-term dependencies into the future. For a longer forecast horizon (see Fig. 10 (b, c)), Bi-LSTM provides better accuracies compared to RNN. It could also be verified by the box and swarm plots of Figs 6, 7, and 8. The IQR along with the median MAE and RMSE values in Fig. 6(b, c) and Fig. 7(b, c) are smaller in the case of Bi-LSTM compared to RNN. Furthermore, when compared to the basic RNN, Bi-LSTM has a significantly higher median $R^2$ value.
5. Conclusion

This paper focuses on finding the spatial distribution of best performing deep learning based techniques to predict Covid-19 health outcomes over India. Specifically, the use of widely applied deep learning techniques for such forecasting at the state level, as well as the geographic variability of their predictive performance, are investigated. To find the state-specific optimal deep learning model and the spatial variability of their performances, a total of 36 time-series datasets (one for each of India’s states and union territories) are used. Various deep learning based time-series forecasting models including CNN, RNN, Vanilla LSTM, LSTM Autoencoder, and Bidirectional LSTM are investigated in this work. It is found that there is no single best method that can outperform across different states in India due to the wide variation in COVID-19 mortality time-series data. Additionally, the best-performing models’ geographical distribution varies significantly across forecast horizons. Models like CNN and RNN dominate for shorter horizons, but models like Bi-LSTM and CNN dominate for longer horizons in most states.

We would like to explore some possible extensions of this work in the future. For instance, one of the major research questions that could be addressed is how the relationship between COVID19 mortality and risk factors will change in the future.

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.

Data availability

Data will be made available on request.

Acknowledgments

This research work is supported by the project entitled “Modelling geographically varying relationships of COVID-19 outcomes with socioeconomic and environmental factors for predictive risk analysis” funded by the Department of Science & Technology, NIT-Division, Government of India. The research work of Asif Iqbal Middy is also supported by UGC-NET Junior Research Fellowship (UGC-Ref. No.: 3684/INET-JULY 2018) provided by the University Grants Commission, Government of India.

References

Albawi, S., Mohammed, T.A., Al-Zawi, S., 2017. Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET). IEEE, http://dx.doi.org/10.1109/icengtechnol.2017.8308186.
Bhimala, K.R., Patra, G.K., Mopuri, R., Matheneni, S.R., 2021. Prediction of COVID-19 cases using the weather integrated deep learning approach for India. Transboundary. Emerg. Dis. 69 (3), 1349–1363. http://dx.doi.org/10.1111/bted.14102.
Chimnula, V.K.R., Zhang, L., 2020. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos Solitons Fractals 135, 109864. http://dx.doi.org/10.1016/j.chaos.2020.109864.
Das, R., Middy, A.I., Roy, S., 2021. High granular and short term time series forecasting of so2/so3 (PM1.25) so2 air pollutant - A comparative review. Artif. Intell. Rev. 55 (2), 1253–1287. http://dx.doi.org/10.1007/s10462-021-09991-1.
Elman, J.L., 1990. Finding structure in time. 14, (2), pp. 179–211.
Fanelli, D., Piazza, F., 2020. Analysis and forecast of COVID-19 spreading in China, Italy and France. Chaos Solitons Fractals 134, 109761.
Hewamalage, H., Bergmeir, C., Bandara, K., 2021. Recurrent neural networks for time series forecasting: Current status and future directions. Int. J. Forecast. 37 (1), 388–427. http://dx.doi.org/10.1016/j.ijforecast.2020.06.008.
Hinton, G.E., Salakhutdinov, R.R., 2006. Reducing the dimensionality of data with neural networks. Science 313 (5786), 504–507. http://dx.doi.org/10.1126.science.1127647.
John, C.C., Ponnumamy, V., Chandrasankar, S.K., R, N., 2022. A survey on mathematical, machine learning and deep learning models for COVID-19 transmission and diagnosis. IEEE Rev. Biomed. Eng. 15, 325–340. http://dx.doi.org/10.1109/rbme.2021.3069213.
Kavouros, I., Kaselimi, M., Protopoulosakidis, E., Bakulon, N., Doulamis, N., Doulamis, A., 2022. COVID-19 spatio-temporal evolution using deep learning at a European level. Sensors 22 (10), 3658. http://dx.doi.org/10.3390/s22103658.
Middy, A.I., Roy, S., 2021. Geographically varying relationships of COVID-19 mortality with different factors in India. Sci. Rep. 11 (1), http://dx.doi.org/10.1038/s41598-021-86987-5.
Middy, A.I., Roy, S., 2022. Pollutant specific optimal deep learning and statistical model building for air quality forecasting. Environ. Pollut. 301, 118972. http://dx.doi.org/10.1016/j.envpol.2022.118972.
Mohimont, L., Chemchem, A., Alin, F., Krajeccki, M., Steffenel, L.A., 2021. Convolutional neural networks and temporal CNNs for COVID-19 forecasting in France. Appl. Intell. 53 (12), 8794-8809. http://dx.doi.org/10.1007/s10489-021-02358-6.
Nath, P., Saha, P., Middy, A.I., Roy, S., 2021. Long-term time-series pollution forecast using statistical and deep learning methods. Neural Comput. Appl. 33 (19), 12551–12570. http://dx.doi.org/10.1007/s00521-021-05901-2.
Nikparvar, B., Rahman, M.M., Hatami, F., Thili, J.-C., 2021. Spatio-temporal prediction of the COVID-19 pandemic in US counties: modeling with a deep LSTM neural network. Sci. Rep. 11 (1), http://dx.doi.org/10.1038/s41598-021-11197-3.
Said, A.B., Erradi, A., Aly, H.A., Mohamed, A., 2021. Predicting COVID-19 cases using bidirectional LSTM on multivariate time series. Environ. Sci. Pollut. Res. 28 (40), 56643-56652. http://dx.doi.org/10.1007/s11356-021-14286-7.
Sanmigrabi, S., Pillai, F., Basu, B., Basu, A.S., Molter, A., 2020. Examining the association between socio-demographic composition and COVID-19 fatalities in the European region using regression approach. Sustainable Cities Soc. 62, 102418. http://dx.doi.org/10.1016/j.scs.2020.102418.
Shastry, S., Singh, K., Kumar, S., Kour, P., Mansoota, V., 2020. Time series forecasting of Covid19 using deep learning models: India-USA comparative case study. Chaos Solitons Fractals 140, 110227. http://dx.doi.org/10.1016/j.chaos.2020.110227.
da Silva, C.C., de Lima, C.L., da Silva, A.C.G., Silva, E.L., Marques, G.S., de Araújo, L.J.B., Junior, L.A.A., de Souza, S.B.J., de Santana, M.A., Gomes, J.C., et al., 2021. Covid-19 dynamic monitoring and real-time spatio-temporal forecasting. Front. Public Health 9.
Tang, B., Bragazzi, N.L., Li, Q., Tang, S., Xiao, Y., Wu, J., 2020. An updated estimation of the risk of transmission of the novel coronavirus (2019-ncov). Infect. Dis. Innov. 5, 248–255.
Zeralou, A., Harrou, F., Dairi, A., Sun, Y., 2020. Deep learning methods for forecasting COVID-19 time-series data: A comparative study. Chaos Solitons Fractals 140, 110121. http://dx.doi.org/10.1016/j.chaos.2020.110121.
Zhao, S., Lin, Q., Ran, J., Musa, S.S., Yang, G., Wang, W., Lou, Y., Gao, D., Yang, L., He, D., et al., 2020. Preliminary estimation of the basic reproduction number of novel coronavirus (2019-ncov) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak. Int. J. Infect. Dis. 92, 214–217.
Zhao, B., Lu, H., Chen, S., Liu, J., Wu, D., 2017. Convolutional neural networks for time series classification. J. Syst. Eng. Electron. 28 (1), 162–169. http://dx.doi.org/10.1007/s11356-021-14286-7.
