Engineering Applications of Artificial Intelligence 122 (2023) 106157

Contents lists available at ScienceDirect

Engineering Applications of Artificial Intelligence

journal homepage: www.elsevier.com/locate/engappai

Improved LSTM-based deep learning model for COVID-19 prediction using optimized approach

Luyu Zhou [a,b,*,1], Chun Zhao [a,1], Ning Liu [b], Xingduo Yao [b], Zewei Cheng [b]

[a] Department of Pharmacy, College of Biology, Hunan University, Changsha, Hunan 410082, China
[b] Institute for Translational Medicine, The Affiliated Hospital of Qingdao University, College of Medicine, Qingdao University, Qingdao 266021, China

ARTICLE INFO

Keywords:
Forecast
COVID-19
Long short-term memory (LSTM)
Prediction

ABSTRACT

Individuals in any country are badly impacted both economically and physically whenever an epidemic of infectious illnesses breaks out. A novel coronavirus strain was responsible for the outbreak of the coronavirus sickness in 2019. Corona Virus Disease 2019 (COVID-19) is the name that the World Health Organization (WHO) officially gave to the pneumonia that was caused by the novel coronavirus on February 11, 2020. The use of models that are informed by machine learning is currently a major focus of study in the field of improved forecasting. By displaying annual trends, forecasting models can be of use in performing impact assessments of potential outcomes. In this paper, proposed forecast models consisting of time series models such as long short-term memory (LSTM), bidirectional long short-term memory (Bi-LSTM), generalized regression unit (GRU), and dense-LSTM have been evaluated for time series prediction of confirmed cases, deaths, and recoveries in 12 major countries that have been affected by COVID-19. Tensorflow1.0 was used for programming. Indices known as mean absolute error (MAE), root means square error (RMSE), Median Absolute Error (MEDAE) and r2 score are utilized in the process of evaluating the performance of models. We presented various ways to time-series forecasting by making use of LSTM models (LSTM, BiLSTM), and we compared these proposed methods to other machine learning models to evaluate the performance of the models. Our study suggests that LSTM based models are among the most advanced models to forecast time series data.

1. Introduction

The outbreak and spread of the epidemic in many parts of the world have become a global epidemic, posing a huge threat to the health of people in all countries (Nawaz et al., 2020). Through a series of prevention and control measures in various places, such as frequent hand washing, wearing masks in public, showing health codes, and passing big data when going out, the epidemic situation in China has been better controlled, but the epidemic situation abroad is still relatively severe, so using historical data of new crown pneumonia to predict the development trend of the epidemic is of great significance to the formulation of reasonable intervention and prevention and control measures (Aamir et al., 2021; Bhatti et al., 2022).

Different studies provide strong academic support for the fight against the epidemic from multiple perspectives (Bedi et al., 2020). Cases along the world are increasing day by day and it requires much attention as Fig. 1 shows the results of worldwide cases. Existing methods can be divided into statistical methods, kinetic methods, and machine learning methods (Fildes, 1992). The statistical method is suitable for use in the case of incomplete information, which predicts the overall trend through the situation of some samples, and some samples have a large difference from the overall propagation trend (Bunn and Wright, 1991), so the prediction error of this method is large and cannot be accurate.

The trend of the spread of the epidemic has changed. The classical mathematical model of the kinetic method is SIR (susceptible infected recovered) model (Syage, 2020; Tan et al., 2020) and SEIR (susceptible-exposed-infected-removed) model (Li and Muldowney, 1995). The kinetic method has a good prediction of the early transmission trend of the epidemic, but it is unable to make an accurate estimate of the spread of the virus in the open flow environment, nor can it make the hypothetical disease transmission capacity and the probability of cure the constants consistent with the actual situation, so it is impossible to make a long-term accurate analysis of the epidemic trend with the increase of new crown pneumonia data, machine learning has shown great superiority (Comito et al., 2019; Bhatti et al., 2019, 2018). Xavier et al. proposed a method with limited data, and used the least-squares criterion with gradient descent algorithm to predict the trend of the number of confirmed COVID-19 patients by nonlinear regression of the data (Xavier, 2020), but this method requires the artificial addition of

* Corresponding author at: Department of Pharmacy, College of Biology, Hunan University, Changsha, Hunan 410082, China.
E-mail address: zhouluyu@hnu.edu.cn (L. Zhou).
[1] These authors contributed equally to this work.

https://doi.org/10.1016/j.engappai.2023.106157
Received 19 December 2022; Received in revised form 8 March 2023; Accepted 13 March 2023
Available online 16 March 2023

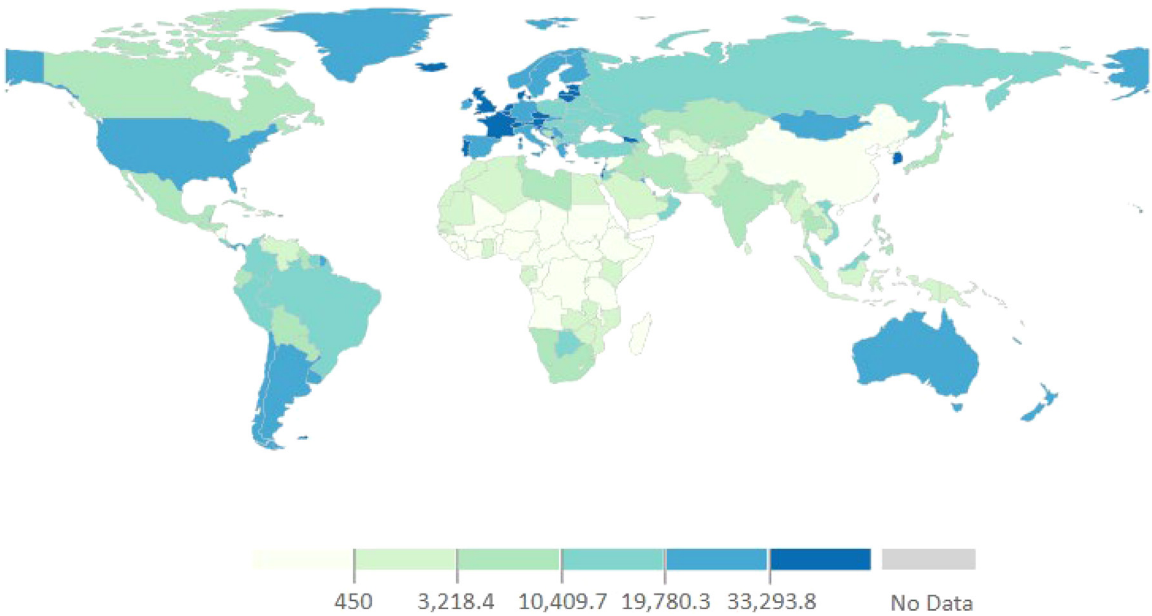


Fig. 1. COVID-19 cases worldwide.

time characteristics to ensure prediction accuracy. In order to solve the above problems, different studies (Benvenuto et al., 2020; Bhatti et al., 2021; Khan and Gupta, 2020; Kufel, 2020) used the autoregressive integrated moving average model (ARIMA) to predict the development of the new crown pneumonia epidemic, which has high time sequence requirements for the data, does not need to artificially add time features, but the nonlinear fitting ability is not strong, and the prediction effect decreases as the number of data increases. In order to solve the above problems, different studies (Badri et al., 2022; Chimmula and Zhang, 2020; Shahid et al., 2020) established a (LSTM) based on deep learning, which achieved model fitting and prediction, which improved the accuracy of short-term new crown pneumonia prediction to a certain extent, but had the following shortcomings:

- Did not achieve a more accurate long-term prediction of the development trend of new crown pneumonia;
- LSTM neural network has a large demand for data volume, and a simple data acquisition method is not provided;
- The overfitting problem caused by the large number of parameters and the complexity of the model in the deep neural network is not considered.

LSTM neural networks have many improvement mechanisms in trend prediction, which are roughly divided into two categories:

- Improvements for the LSTM neural network itself;
- Introduce other methods to improve the LSTM algorithm.

Fahim et al. (2021) improve LSTM neural network using attention mechanism t, when the network is trained, the output is fed back to the input so that its secondary training achieves the purpose of improving the generalization ability. Kim et al. used LSTM with Convolutional Neural Networks (CNN) for the prediction of energy consumption by getting spatial correlation of intersection to explore the time series characteristics of data flow through LSTM model (Kim and Cho, 2019), and the extracted spatio-temporal features are fused to achieve short-term traffic prediction. Wu et al. (2019) used ensemble empirical mode decomposition, EEMD, to build a multi-level LSTM predictive model to improve the accuracy of the model prediction. Bi-LSTM model used in this paper is another approach to improvement of Uni-directional LSTM. The historical data of new crown pneumonia is time-series data, and the LSTM neural network is good at processing time-series data, and can adjust the specific number of days predicted according to the experimental effect to achieve the purpose of predicting the development trend of new crown pneumonia.

Due to the large amount of historical data on covid-19, the traditional manual collection method is no longer applicable, so this article uses web crawler technology to obtain relevant data from the World Bank data site for use in this study. Considering that there is a strong time series between the historical data of new crown pneumonia, this paper uses the LSTM neural network that is good at processing time-series data as a basic model to predict the epidemic trend. Compared with the ARIMA model, this model has stronger nonlinear fitting capabilities and does not require the artificial addition of time features, maximizing the nonlinear relationship between time series data. In order to avoid overfitting problems during multilayer network training, this paper constructs a multilayer LSTM neural network and introduces layer to inactivate neurons according to random probabilities. Finally, the accuracy of this method was verified by using the cumulative confirmed cases, existing confirmed cases and a cumulative number of cured people in different countries. Our work is devoted to the development of novel coronavirus prediction models, consisting of time series models such as long-term short-term memory (LSTM), bidirectional long-term short-term memory (Bi-LSTM), generalized regression unit (GRU), and dense LSTM, that were evaluated to predict the time series of confirmed cases, deaths, and recoveries in the 12 leading countries affected by COVID-19. Our study suggests that LSTM is one of the most advanced models to forecast time series data.

2. Literature review

This section highlights the different methods of forecasting using in different studies.

2.1. Traditional methods of forecasting

The problem of forecasting has always attracted the attention of many experts and scholars around the globe and has always been a research hotspot. Different scholars have also done a lot of research on improving the accuracy of forecasting, including some classic forecasting methods. Such as regression analysis and time series analysis, as well as ARMA, gray model and so on. Most of these methods are based on linear analysis. Although they have many advantages, such as the algorithm logic is clear and easy to understand, and the running speed is fast, they lack the nonlinear fitting ability, and the COVID-19 data is a nonlinear multivariate time series (Dairi et al., 2021). Traditional It is often difficult to achieve high accuracy when the method is applied to the problem of COVID-19 forecasting because data fluctuations are sudden due to spreading out of cases. In recent years, the development of AI technology has led to more and more intelligent methods being proposed. New methods such as expert systems, machine learning, and fuzzy reasoning have been applied in the field of COVID-19 forecasting (Comito and Pizzuti, 2022; Jamshidi et al., 2022). An artificial neural network can take nonlinear factors into account, and it has gradually become a popular forecasting method (Ayoobi et al., 2021). Based on artificial neural network, many methods related to time series forecasting have been derived, including many researches on COVID-19 forecasting.

The time series prediction method starts from the continuity of the development of things. The development of things has its own regularity and continuity, and can basically follow its own inherent laws in the development process, as long as there is no major essential change in the effective conditions of its dependence (Lara-Benitez et al., 2021). In light of the situation, it is possible for things to fundamentally continue to develop continually in accordance with their rules. The time series forecasting method involves taking sequence data samples and recording them at predetermined intervals, performing statistics and analysis on historical time series data, constructing a mapping relationship between future data and historical data, and using historical data to predict future data. All of these steps are carried out by the time series forecasting method. The autoregressive model (AR), the moving average model, the integral autoregressive sliding model (ARIMA), and the autoregressive moving average model are the primary components of the classic approaches to time prediction (Tyass et al., 2022). Some scholars have proposed a short-term load forecasting method that combines time series algorithm and fuzzy logic, and considers external factors such as temperature as the factors affecting COVID-19 (Hadjira et al., 2021).

The regression analysis method analyses the data samples and establishes a correct mathematical model on this basis to infer the future predicted data value and divides it into one variable according to the number of independent variables (Saha et al., 2021). According to whether the relationship between independent variables and dependent variables is linear or nonlinear, regression and multiple regression can be broken down into linear regression models and nonlinear regression models (Mehrolia et al., 2021). Linear regression models are the more common type of regression model. In regression analysis, the approach that uses the fewest squares is the one that is most frequently employed between load-influencing factors and load values. The regression analysis method begins by identifying a number of different factors that can have an effect on the predicted object. Next, a suitable function is used to express the relationship between all of the factors and the predicted object. Finally, data samples are used to estimate the parameters of the model and to predict the object (Irfan et al., 2022). Error checking, making predictions after the model is determined. Regression analysis methods are also often combined with time series models, such as ARIMA models. In the regression model, the selection of influencing factors is very important. While not omitting important influencing factors, it is necessary to keep a small number of factors to keep the model concise (Yilanci and Pata, 2022). In the COVID-19 forecasting problem, the factors affecting the cases value, such as weather factors, close contacts and other factors, are analysed as independent variables, and the cases data is used as the dependent variable to establish a regression model. The prediction model of the regression analysis method is relatively simple, easy to understand, and has the advantages of faster operation speed. However, in COVID-19 forecasting, there are many load influencing factors and most of them have a nonlinear relationship with the cases value. Rubbaniy et al. (2021) used regression analysis to predict the COVID-19 data using wavelet methods. Some other scholars combine wavelet analysis and regression analysis method, first use wavelet transform to decompose short-term power load data, and then use regression analysis method for prediction respectively (Mach et al., 2021; Pan et al., 2021).

The advantage of grey model method is its effectiveness in solving problems with uncertain factors. Professor Deng Julong, a Chinese researcher, is credited with being the one who originally suggested using the grey model method in 1982 (Julong, 1982). The discipline of systems research has been given access to new scientific theories and methodologies as a result of the grey system theory (Huang and Jane, 2009; Liu and Guo, 2010). After processing the time series, the grey model approach seeks to decrease or eliminate the randomness of the sequence to a given amount, and then generate time series data that has strong regularity. In the COVID-19 instances forecasting problem, some of the influencing elements are determinable, but some of the influencing factors are not particularly determinable, and the system is regarded as a gray system when it comes to carrying out cases forecasting (Sahin and Sahin, 2020). Establishing a model in such a system where there are unknown components is the goal of the grey model. It is a dynamic model that is capable of both adding new information and eliminating information that is no longer relevant. It is an advantage of the grey model method, which is extensively used in a variety of fields, in the power load forecasting problem since certain uncertain components are included in the problem.

2.2. Artificial intelligence methods

A growing number of researchers are using artificial intelligence technology into the process of power load forecasting as a result of the ongoing development of technologies related to artificial intelligence. The model of the artificial neural network is able to imitate the structure of the human brain while simultaneously simplifying and abstracting it. It possesses an activating function and is made up of numerous layers of neuronal connections. It possesses powerful processing capabilities for complicated data, in particular for the nonlinear data processing that is required by classical load forecasting methods. Weak issues offer clear advantages, a strong capacity to deal with nonlinear data, and inherent advantages in the actual prediction process. Weak problems also have a great ability to deal with nonlinear data. The error back-propagation (BP) algorithm is the one that is employed the most frequently and is the one that is most common among many different neural networks (Martinez-Alvarez et al., 2020). The classic BP algorithm, on the other hand, has a very long iteration period and a very slow convergence speed. The difficulty of the BP algorithm's sluggish convergence speed is solved by the literature (Ceylan, 2021; Zivkovic et al., 2021), which optimizes the BP method using the PSO algorithm. This overcomes the problem. Dahiya et al. (2022) developed a BP-ANN model for the purpose of scientific prediction, and they were successful in doing so. Back propagation neural networks (BPNN) were used by Wu et al. (2020) to estimate the peak load forecast in Taiwan. The experimental findings revealed that the prediction accuracy of BPNN was greatly improved when compared with the classic regression model. Leng et al. (2017) created a neural network model for the results of pre-classification of the data set by fuzzy cluster analyses, and eventually achieved the prediction results by means of pattern recognition. This model was based on the findings of the pre-classification of the data set. As a result of the progression of research, an increasing number of academics have generated more outstanding models by enhancing the models that were originally developed. The idea of a hybrid kernel is incorporated into the extreme learning machine model by Xie et al. (2020), which effectively enhances the prediction accuracy. Combining particle swarm optimization (PSO) and Elman neural network (ENN), using particle swarm optimization algorithm for hyperparameter optimization in ENN, Heydari et al. (2022) proposed a model called PSO-ENN for short-term power load forecasting. Experiments are compared with the basic model to prove its effectiveness.

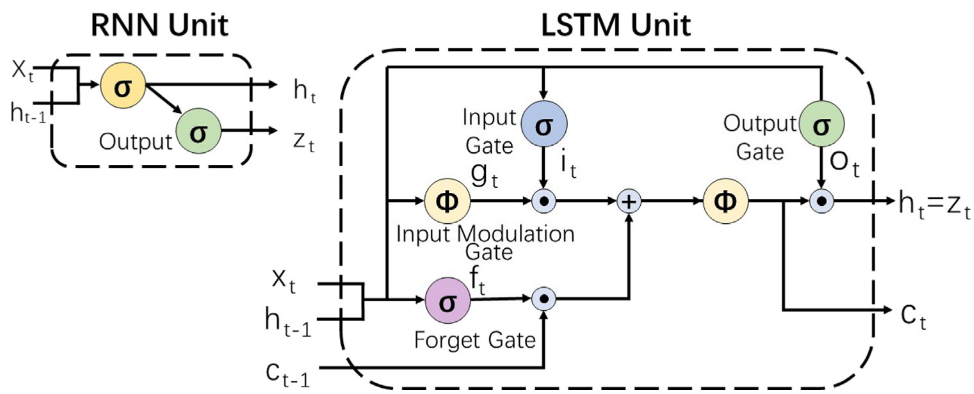


Fig. 2. LSTM and RNN comparison.

Aslan et al. (2021) extracts the features of the load data using a CNN, and they bring out the advantages of CNN in nonlinear data, which can extract more advanced features.

Neural networks and machine learning are universal in structural deformation prediction, and have been widely used in practical engineering. Wang and Adeli (2015) proposed a method for predicting the nonlinear relationship between bridge displacement and its cause by means of wavelet neural network (WNN), which can determine the optimal number of nodes in the implicit layer of the network, with good local characteristics, strong learning ability and arbitrary function approximation ability, and realize the accurate fitting and prediction of bridge deformation, but the random initialization of parameters will lead to slow or even non-convergence of WNN. Liu et al. (2021a) used BP neural network model to predict the displacement of piles in foundation pits, which accurately reflected the actual deformation trend. (AlOmar et al., 2020) the maximum surface settlement was predicted by wavelet-artificial neural network, and different wavelets were used as activation functions to predict the maximum surface settlement caused by tunnel excavation, so that the prediction error value was reduced. However, the sample and model parameters have a greater influence on the model, and it is easy to fall into local optimization during training. Jiang et al. (2020) used LSTM to predict periodic displacements of landslides. Fang et al. (2021) proposed a prediction model based on ARMA-LSTM, which solved the problem that a single prediction algorithm could not solve the linear and nonlinear components in the dam deformation data at the same time. This model uses the ARIMA model to predict the linear components in the dam deformation data, and the LSTM model to predict the nonlinear components in the dam deformation data, which accurately reflects the dam deformation trend. The results show that LSTM can remember historical information and accurately predict future structural data, but the model only uses forward information to predict the current data, ignores the influence of the change law of structural deformation data after the time point on the current data, and cannot fully tap the time characteristic information of the monitoring data, resulting in the prediction accuracy needs to be further improved.

Deep learning technology has seen significant advancement in recent years, thanks in large part to the arrival of the era of big data and the continued development of technology related to artificial intelligence. Although machine learning models, such as k-nearest neighbour (KNN) were also used in predicting Covid-19, deep learning provides apparent advantages in application settings such as larger data samples, more dimensional data features, and more complicated issues since it combines the underlying features to generate higher-level abstract features (Chumachenko et al., 2022; Pahar et al., 2021). These advantages may be seen in the following sentences. The Recurrent Neural Network (RNN) model (Al-Rakhami et al., 2020) is able to analyse time series data and has seen significant application in the field of natural speech processing (Tarwani and Edem, 2017; Yin et al., 2017). The usage of RNN models and models resulting from their advancements in the field of time series prediction has garnered the attention of an increasing number of academics in recent years. Models such as LSTM and GRU (gated neural unit) are based on RNN. The RNN model in its most fundamental form can be improved upon by using a variant of the model. RNN, which is a standard model for recurrent neural networks, is plagued by the issue of gradient disappearance; nevertheless, LSTM provides an excellent solution to this problem. The contextual information of time series data can be memorized by LSTM through the utilization of gated neural units. It is composed of three distinct types: an input gate, an output gate, and a forgetting gate. Long-term characteristics of the sequence can be extracted by LSTM thanks to the gate mechanism. In the context of solving time series problems, LSTM has seen significant application. In order to verify the effectiveness of LSTM, Chimmula and Zhang (2020) proposed a model for COVID-19 forecasting based on LSTM and LGBM, and achieved good forecasting results. Islam et al. (2020) proposed a method combining CNN and LSTM for short-term COVID-19 cases prediction, and the high-level features of the cases are extracted by CNN, which improves the accuracy when the LSTM input is too long.

3. LSTM forecasting models

In this section used models are discussed with way of their implementation.

3.1. Long short-term memory network (LSTM)

LSTM is a variant of RNN. By setting multiple gate structures, LSTM can record long-term and short-term information at the same time, and solve the problem of gradient disappearance (that is, loss of previous information) that may occur during the training of the RNN model. The structure of the classic RNN model and LSTM model is shown in Fig. 2. It can be seen that the structure of the LSTM model is more complex than that of the classic RNN. There is only one tanh activation function in the classic RNN unit; while the LSTM unit contains three different gate structures of the control function is the forget gate, the input gate, and the output gate. These gate structures are used to learn some useful long-term information and discard some meaningless information. LSTM is often used for semantic classification, trajectory prediction and action classification.

In Fig. 2, X_t is the input sequence, that is, the system state monitoring information for fault prediction; h_t is the output of the hidden layer, that is, the learning result of each LSTM unit; f_t is the forgetting gate; i_t is the input gate; O_t is the output gate; σ is the sigmoid activation function; tanh is the activation function; W is the weight matrix; $*$ is the point pair product.

(1) Forgetting gate: LSTM processes time-series data in a specified order. The information contained in the data in a certain period of time

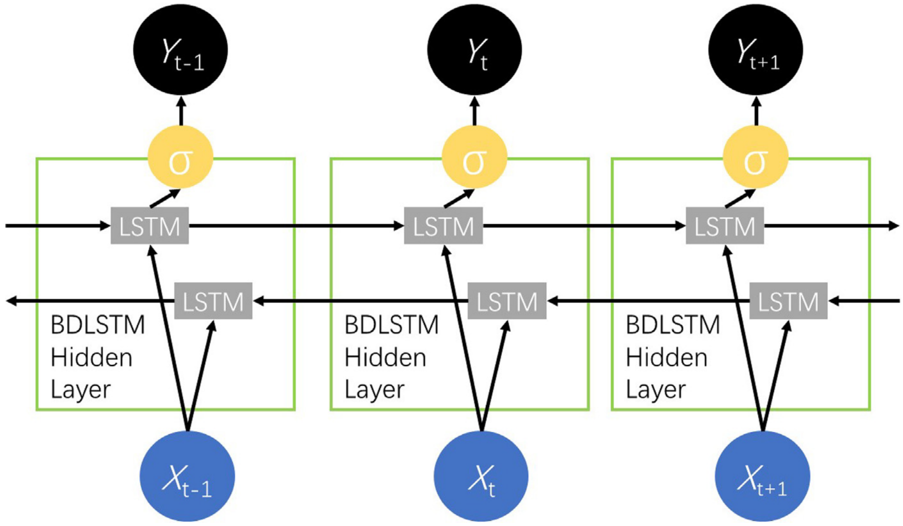


Fig. 3. LSTM model.

is useful and useless. The function of the forgetting gate is to decide which information to retain and which information to ignore directly. The output of the forget gate is shown in (1).

$$f_t = \sigma\left(\boldsymbol{W}_f \cdot [h_{t-1}, x_t] + b_f\right) \tag{1}$$

In the formula: W_f is the weight of the forget gate; b_f is the bias of the forget gate.

(2) Input gate: The information enters the input gate after the selection of the forget gate. The function of the input gate is to determine which parameters need to be updated and how to update them. The output of the input gate is shown in (2), (3) and (4).

$$i_t = \sigma\left(\boldsymbol{W}_i \cdot [h_{t-1}, x_t] + b_i\right) \tag{2}$$

$$C_t = \tanh\left(\boldsymbol{W}_C \cdot [h_{t-1}, x_t] + b_C\right) \tag{3}$$

$$C_t = f_t * C_{t-1} + i_t * C_t \tag{4}$$

In the formula: W_i and W_C are the corresponding weights respectively; b_i and b_C are the corresponding biases respectively; C_t is the current cell state value.

(3) Output gate: The information reaches the output gate after being screened by the forget gate and the input gate. The function of the output gate is to decide which information to output. The output of the output gate is shown in (5) and (6).

$$O_t = \sigma\left(W_O \cdot [h_{t-1}, x_t] + b_O\right) \tag{5}$$

$$h_t = O_t \cdot \tanh\left(C_t\right) \tag{6}$$

where: W_o is the weight of the output gate; b_o is the bias of the output gate; h_t is the output value of the current unit.

The LSTM model as shown in Fig. 3. The advantage of LSTM is that it can refer to the information of the previous unit when processing the information of this unit, but the disadvantage is that it does not refer to the information of the subsequent unit. However, for some problems, it is not only related to the previous unit information, but also related to the subsequent unit information, such as the state parameters reflecting the system performance degradation process, due to inaccurate observations or data uncertainty in the sensor acquisition process, the previous time and parameters collected at subsequent moments are helpful to confirm the accuracy of the parameters at the current moment. As a result, the bidirectional long–short-term memory network (Bi-LSTM) (Shen et al., 2021; Xia et al., 2021; Zeng et al., 2021), which can simultaneously learn forward and backward data regularities, came into being.

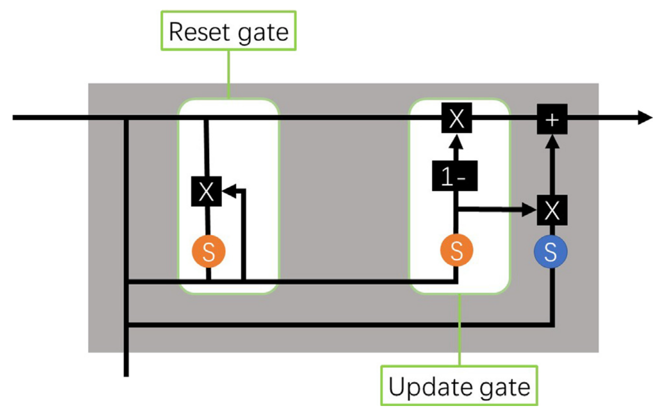


Fig. 4. Bi-LSTM network structure.

3.2. Bidirectional long short-term memory network (Bi-LSTM)

Bi-LSTM is a variant of LSTM network that can learn both information from previous moments as well as information from future moments (Chen et al., 2021; Liu et al., 2021b). Bi-LSTM solves the problem that only one-way timing is considered in LSTM calculation by LSTM calculation in both forward and backward directions at the same time, and on the other hand, it improves the problem that the weight in LSTM calculation is greatly affected by timing. The Bi-LSTM network structure is shown in Fig. 4 and comparison with LSTM is shown in Fig. 5.

The hidden layer of Bi-LSTM will save the vector $\overrightarrow{\boldsymbol{h}}$ obtained by forward inference and the vector $\overleftarrow{\boldsymbol{h}}$ obtained by reverse inference, and the calculation methods are based on one-way LSTM:

$$\overrightarrow{\boldsymbol{h}} = \mathrm{LSTM}\left(x_t, \overrightarrow{\boldsymbol{h}}_{t-1}\right) \tag{7}$$

$$\overleftarrow{\boldsymbol{h}} = \mathrm{LSTM}\left(x_t, \overleftarrow{\boldsymbol{h}}_{t-1}\right) \tag{8}$$

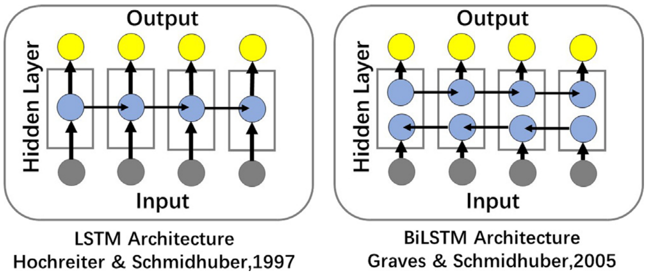


Fig. 5. Comparison between Bi-LSTM and LSTM network structure.

In Bi-LSTM, $\vec{h}$ and $\overleftarrow{h}$ are combined together to form the final output y_t:

$$y_t = g\left(W_{\vec{h}y}\vec{h}_t + W_{\overleftarrow{h}y}\overleftarrow{h}_t + b_y\right) \tag{9}$$

Assuming that a set of data monitored by the database is $X_N = [x_1, x_2, x_3, \ldots, x_N]$ (N is the number of sampling points), by analysing the structural deformation data and the characteristics of the Bi-LSTM model, the first M monitoring data is determined as the prediction model training The last $N - M$ monitoring data is the test set of the prediction model.

$$X_{tr} = \begin{bmatrix} x_1 & x_2 & \cdots & x_L \\ x_2 & x_3 & \cdots & x_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{M-L} & x_{(M-L)\times L} & \cdots & x_{M-1} \end{bmatrix}_{M-L+1} \tag{10}$$

$$Y_{tr} = \begin{bmatrix} x_{L+1} \\ x_{L+2} \\ \vdots \\ x_M \end{bmatrix}_{(M-L)\times 1} \tag{11}$$

$$X_T = \begin{bmatrix} x_{M-L+1} & x_{M-L+2} & \cdots & x_M \\ x_{M-L+2} & x_{M-L+3} & \cdots & x_{M-L+1} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N-L} & x_{N-L+1} & \cdots & x_M \end{bmatrix}_{(N-M)\times L} \tag{12}$$

Among them, L is the number of nodes in the input layer of the model, X_{tr} is the training input, Y_{tr} is the target output, and X_T is the test input.

For the model forward LSTM unit, the ith input $X_{tr}(i)$ of the training set is substituted into Eqs. (10) to (12) to obtain a set of state outputs $\{h_i^f, h_{i+1}^f, \ldots, h_{i+L}^f\}$.

For the model reverse LSTM unit, the ith input $X_{tr}(i)$ of the training set is substituted into Eqs. (10) to (12) in reverse order to obtain a set of state outputs $\{h_{i+L}^b, h_{i+L-1}^b, \ldots, h_i^b\}$.

The obtained two sets of state outputs have the same feature dimension, and the two are spliced to obtain $H_i \in R^{2\times L}$:

$$H_i = \left\{\left\{h_i^f, h_{i+L}^b\right\}, \left\{h_{i+1}^f, h_{i+L-1}^b\right\}, \ldots, \left\{h_{i+L}^f, h_i^b\right\}\right\} \tag{13}$$

The spliced vector is processed through the output layer activation function to obtain the ith training output $\tilde{y}_i$.

The training and testing process of the model divided in different ratios. During model training, in order to prevent the model from overfitting, regularization is added to the hidden layer, and the model uses the gradient descent algorithm to update the network weight matrix and bias vector. All parameters are initialized; then Bi-LSTM extracts the forward and backward information of the time series as the input of the fully connected layer, calculates the output of the forward hidden layer and the backward hidden layer respectively, and obtains the weights through splicing and linear operations. Then use the loss function backpropagation to update the model parameters to obtain the optimal solution; finally, use the test set data to test the trained model, and compare the model prediction results with the real data. The prediction accuracy of the model is evaluated by the prediction error.

The specific steps are as follows:

Step 1. The experimental sedimentation data is divided into training sets and test sets, and the data are normalized.

Step 2. Initialize the model parameters, set the number of nodes in the input layer of the model, the number of forward and backward hidden layers and nodes, the learning step and the number of training rounds;

Step 3. The model is trained by using the training set data to determine the optimal solution of the model.

Step 4. The trained model is tested by test set data, and the prediction accuracy of the model is evaluated by using prediction error.

3.3. GRUs

GRU is a variant of LSTM. Compared with LSTM (Wang et al., 2022), GRU simplifies the structure of the network, as shown in the figure, as shown in the figure, he only has update gate (zt) and reset gate (rt) Two logic gates. It combines the forget gate and the input gate into a single update gate, which is used to control how much the information in the previous moment is brought into the information of the current moment. The reset gate is used to control the degree of ignoring the information of the previous moment, the smaller the value, the more ignoring. The final model is simpler than the standard LSTM model and a very popular variant.

$$\begin{aligned} z_t &= \text{sigmoid}\left(W_z \cdot \left[h_{t-1}, x_t\right]\right) \\ r_t &= \text{sigmoid}\left(W_r \cdot \left[h_{t-1}, x_t\right]\right) \\ \tilde{h}_t &= \tanh\left(W \cdot \left[r_t * h_{t-1}, x_t\right]\right) \\ h_t &= \left(1 - z_t\right) * h_{t-1} + z_t * \hat{h}_t \end{aligned} \tag{14}$$

In formula: zt—updaters; rt—reset gate; W—weight matrix; ht—hidden layer state.

4. Materials and methods

This section describes the evaluation criteria for model predictive performance, results and discussion about those results. Flowchart of study flow is shown in Fig. 6. Tensorflow1.0 was used for programming.

4.1. Validation methods

In this article RMSE, Mean Absolute Percentage Error (MAPE), Median Absolute Error (MEDAE) and coefficient of determination R^2 (Coefficient of determination) are used as predictive performance metrics, as in other articles (Kumari and Singh, 2022; Zhang et al., 2022).

(1) Root Mean Square Error (RMSE)

Root Mean Square Error (RMSE) see Equation below. The root mean square error is the square of the difference between the predicted value and the true value and the square root is found on the basis of the ratio to the number of observation samples N. The smaller the value of the root mean square error, the more effective the model is Good. RMSE will have different evaluation criteria for datasets with different number data values and different dimensions. The original number depends on the size of the set data value, the evaluation criteria of RMSE are also different, and in general, the larger the value of the original data set, the larger the value of RMSE that can be accepted. RMSE is generally only suitable for evaluation on different models of the same data set comparison of models. In Eq. (15), x_i—the *ith* actual data,

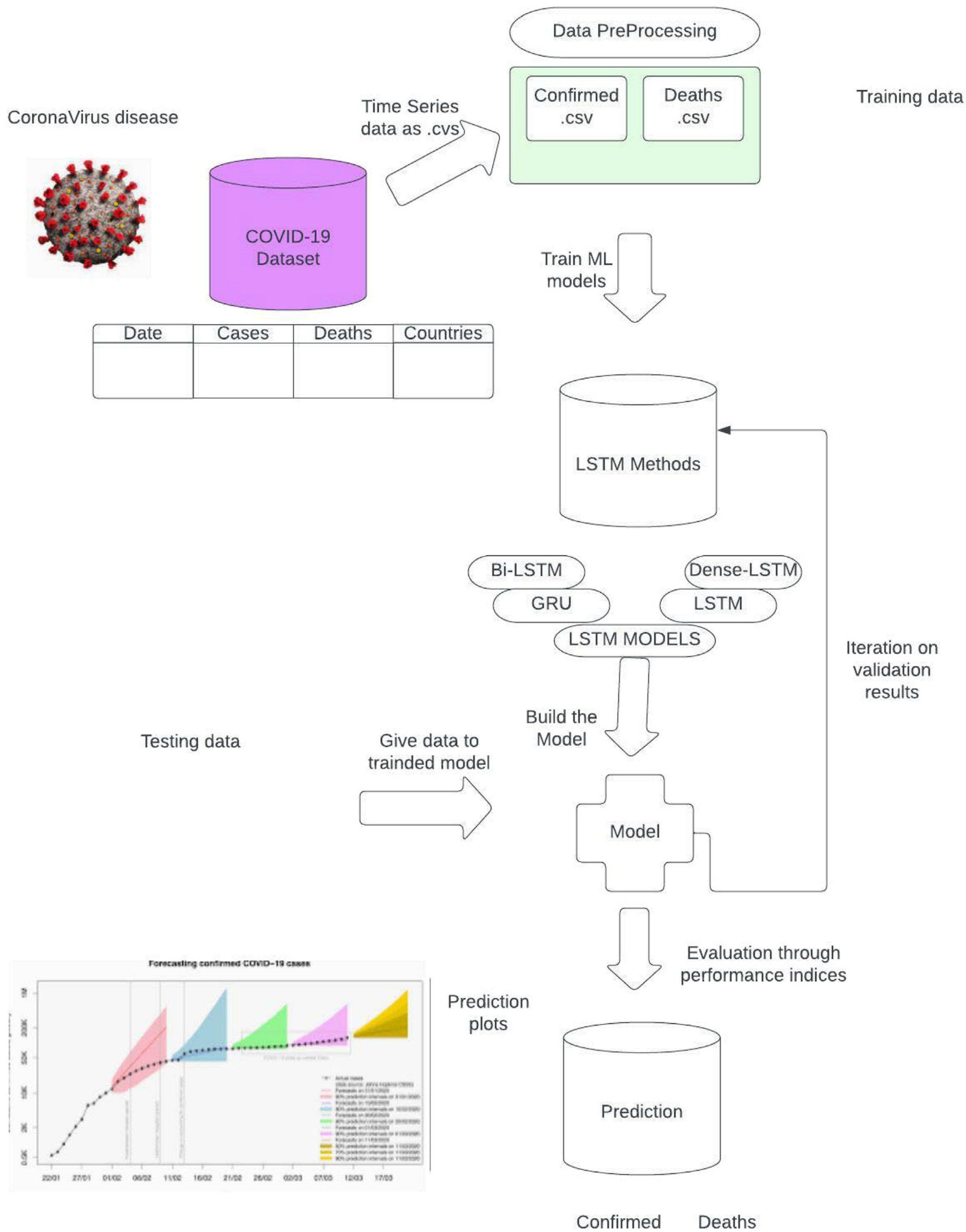


Fig. 6. Flowchart of complete study.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(x_i - \hat{x}_i\right)^2} \tag{15}$$

I, x—ith prediction, N—data length.

(2) Average Absolute Percentage Error (MAPE)

The MAPE is shown in Eq. (16). The average absolute percentage error is the predicted value. The ratio of the difference from the true value to the true value is absolutely valued, and then divided by the number of samples to get a percentage. The smaller the MAPE

Values represent better model effect.

Eq. (12), x_i—the ith actual data,

$$MAPE = \frac{1}{N}\sum_{i=1}^{N}\left|\frac{x_i - \widehat{x}_i}{x_i}\right| \times 100\% \tag{16}$$

I, x—ith prediction, N—data length.

Table 1
Overall cases worldwide for COVID-19 till June 2022.

S.No	Country	Death	Cases
1	United States	998977	83851879
2	India	524708	43185049
3	Brazil	667005	31159335
4	France	145150	28731486
5	Germany	139388	26498346
6	United Kingdom	178866	22338950
7	Russia	379657	18358459
8	Italy	167019	17514589
9	Turkey	98969	15073722
10	Spain	106797	12403245
11	Vietnam	43081	10726045
12	Argentina	128973	9276618

Table 2
Parameter settings for LSTM models.

Algorithm	Parameters	Values
GRU	C	3.0
	epsilon	0.0000001
	degree	3
	tolerance	0.000001
LSTM	Total Layers	3
	Neurons #	(16, 32, 64, 128)
	Rate of Learning	0.001
	Optimization Approach	Adam
	Size of the Batch	10
	Total repetition (Epochs)	300
	Step down time	3
Bi-LSTM	Total Layers	3
	Neurons #	(16, 32, 64, 128)
	Rate of Learning	0.001
	Optimization Approach	Adam
	Size of the Batch	10
	Total repetition (Epochs)	300
	Step down time	3

(3) Coefficient of determination

The coefficient of determination is shown in formula (17).

$$R^2 = 1 - \frac{\sum \left(Y\_\text{actual} - Y\_\text{predict}\right)^2}{\left(Y\_\text{actual} - Y\_\text{mean}\right)^2} \tag{17}$$

Formula (17): R^2—determinant, Y_actual—true value, Y_predict—predicted value, Y_mean—true The average of the values.

(4) Median Absolute Error (MEDAE)

The MEDAE is shown in formula (18).

$$MEDAE = \frac{1}{N} \sum_{i=1}^{N} \left|x_i - median(x)\right| \tag{18}$$

4.2. Dataset

Data has been taken from the WHO international website for daily COVID-19 cases (https://covid19.who.int/data) and shown in Table 1. Total of 12 countries are selected for prediction because these are the countries with most of the COVID-19 cases. The .csv file of confirmed cases, death cases and recovered cases of all countries is used column wise. Data has been selected from 6th January to 2020 to 6th June 2022 however the first cases date of each country is varied according to situation of that country. The data is pre-processed before it is given to ML models for training. Partitioning for dataset is set to 70% and 30% for training and testing. Parameter setting for the machine learning models are shown in Table 2.

5. Results and discussion

This section describes the evaluation criteria for model predictive performance, results and discussion about those results. Flowchart of study flow is shown in Fig. 6.

Tables 3 and 4 shows the prediction performance of the LSTM based models for cumulative cases and death cases prediction. The results of the performance prediction are varied in each dataset. Observing the values of RMSE and MAE, for some countries and even for some feature, one predicts better and for others, another model gives better results. The MEDAEs are all positive and very close to zero, which suggested good performances of models. As shown in Table 3, MAE value for cumulative cases using LSTM model, GRU model, Bi-LSTM model and Dense-LSTM model are similar. The same is true for MSE value and R^2 value. Graphical representation of Table 3 results is shown in Fig. 7.

To sum up, the Bi-LSTM model has more advantages than the LSTM models in the prediction of high-frequency data. The main reason is that the Bi-LSTM model regards the data sequence as a random sequence, and the dependencies of this group of random variables reflect it due to the continuity of the original data in time, and has a certain law of its own changes, more accurate prediction values can be obtained through regression analysis. This characteristic is very similar to the characteristics of worldwide COVID-19 cases data, so the prediction accuracy of short-term high-frequency data is high. Other LSTM and GRU models are compatible with the full-period average and moving average, and does not discard past data, but only gives a gradually weakening degree of influence, that is, as the data moves away, it is given a weight that gradually converges to zero. This feature makes the model more accurate in predicting medium and long-term data by predicting high-frequency data and accumulating it. Therefore, in the actual work process, the two models can be used in combination, with better effect and stronger practicability.

After the outbreak of the epidemic, major media, websites, and software will update the latest progress of the world epidemic in real time. People can consider carefully when choosing travel, which provides a good premise guarantee for people's travel safety. This is the product of the current big data era. Different countries study and determine the type of virus in a relatively short period of time, which allows us to formulate more scientific and targeted prevention and control measures, and also provides prerequisites for the development of vaccines. The research and development of vaccines is a major project integrating talents, innovation, and technology. Countries developed effective vaccines that can be used by the people of the whole country. This is the starting stone for different countries to become innovative country. In the construction of the public health system in the future, we must put innovation in a more prominent position, so that our country can keep up with the times and lead the times in this era when the concept of a community with a shared future for mankind is becoming more and more popular, so as to save the lives of people around the world. Make new and greater contributions to safety and physical health.

At present, most literatures assume that emergency demand is not sensitive to time factors. However, in actual rescue, the demand for medical resources changes dynamically with time, and the allocation of medical materials in the early stage will affect the later demand. In addition, the existing literature mainly focuses on the research of infectious disease transmission model and emergency rescue, and there are not many literatures that integrate the two aspects into consideration. Table 5 shows the stages of COVID-19 with cases.

The LSTM method is not only one of the deep learning's most primary methods, but also widely used to forecasting, which has been used in many fields of medicine. Our research incorporated some of the latest deep learning tools. However, there remains some limitations. For example, LSTMs are more complicated than traditional RNNs. And LSTMs needs more training data to achieve an excellent performance (Zhang et al., 2020). In the case of insufficient data, over-fitting is

Table 3
Forecasting approach using LSTM, GRU, BiLSTM and Dense-LSTM models for cumulative cases.

Countries	LSTM				GRU				Bi-LSTM				Dense-LSTM			
	Loss	MAE	MSE	R^2	Loss	MAE	MSE	R^2	Loss	MAE	MSE	R^2	Loss	MAE	MSE	R^2
United States	0.00110	0.02110	0.00110	0.79000	9.22500	0.00730	0.02998	0.82000	0.27070	0.00670	0.27000	0.83000	1.71400	0.00820	0.00017	0.83000
Vietnam	0.00140	0.02630	0.00145	0.81000	0.00280	0.01860	0.00280	0.69000	0.01790	0.05040	0.01790	0.78000	0.00530	0.02550	0.00532	0.78000
UK	0.00023	0.00850	0.00023	0.83000	0.00016	0.00670	0.00016	0.72000	0.00066	0.01670	0.00066	0.79000	0.00051	0.01580	0.00051	0.79000
Turkey	0.00020	0.00960	0.00020	0.72000	0.00009	0.00700	0.00009	0.83000	0.00010	0.00670	0.00010	0.79000	0.00008	0.00630	0.00083	0.79000
Spain	0.00076	0.01530	0.00076	0.83000	0.00014	0.00810	0.00014	0.78000	0.00020	0.00670	0.00020	0.81000	0.00023	0.00630	0.00023	0.81000
Russia	0.00010	0.00620	0.00010	0.87000	0.00002	0.00350	0.00002	0.79000	0.00009	0.00530	0.00009	0.82000	0.00000	0.00091	0.00000	0.78000
Italy	0.00054	0.01100	0.00054	0.82000	0.00046	0.00800	0.00046	0.79000	0.00093	0.01270	0.00093	0.69000	0.00051	0.01460	0.00051	0.79000
India	0.00006	0.00570	0.00006	0.69000	0.00003	0.00460	0.00003	0.81000	0.00008	0.00720	0.00008	0.72000	0.00008	0.00730	0.00008	0.79000
Germany	0.02070	0.04780	0.02070	0.72000	0.00020	0.00670	0.00020	0.83000	0.00940	0.03610	0.00944	0.81000	0.00079	0.01100	0.00079	0.81000
France	0.00110	0.01680	0.00108	0.83000	0.00056	0.01310	0.00056	0.72000	0.00063	0.01350	0.00063	0.83000	0.00050	0.01580	0.00050	0.83000
Brazil	0.00018	0.00890	0.00018	0.78000	0.00003	0.00390	0.00003	0.83000	0.00004	0.00440	0.00004	0.72000	0.00005	0.00510	0.00005	0.79000
Argentina	0.00003	0.00410	0.00003	0.79000	0.00004	0.00520	0.00004	0.87000	0.00003	0.00420	0.00003	0.83000	0.00004	0.00550	0.00004	0.81000

Countries	LSTM MEDAE	GRU MEDAE	Bi-LSTM MEDAE	Dense-LSTM MEDAE
United States	0.02164	0.00681	0.00681	0.00753
Vietnam	0.02641	0.01848	0.05071	0.02452
UK	0.00773	0.00668	0.01597	0.01607
Turkey	0.00995	0.00786	0.00695	0.00575
Spain	0.01628	0.00773	0.00627	0.00573
Russia	0.00715	0.00313	0.00628	0.00165
Italy	0.01131	0.00801	0.01318	0.01483
India	0.00491	0.00489	0.00770	0.00714
Germany	0.04805	0.00691	0.03696	0.01036
France	0.01715	0.01377	0.01333	0.01561
Brazil	0.00861	0.00415	0.00426	0.00610
Argentina	0.00313	0.00565	0.00440	0.00629

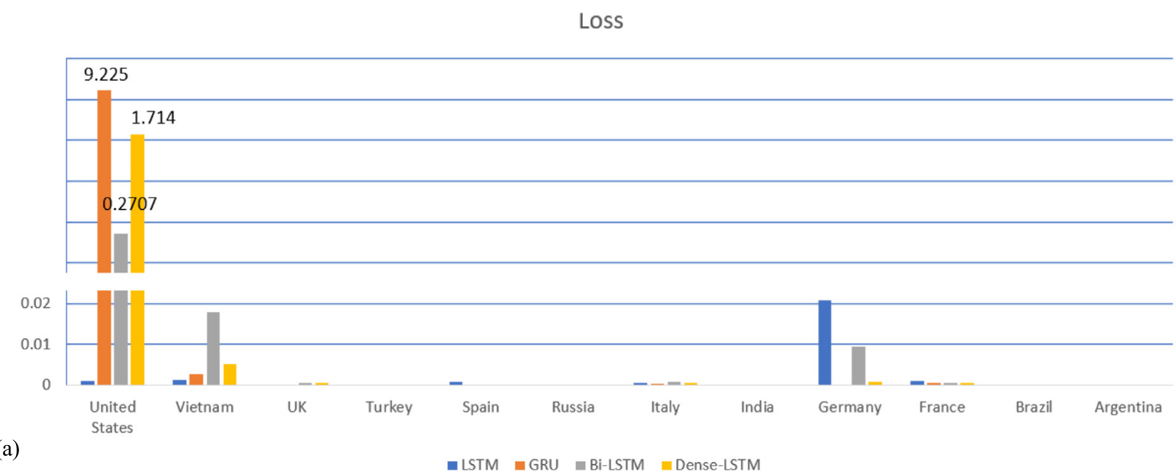
Table 4
Forecasting approach using LSTM, GRU, Bi-LSTM and Dense-LSTM models for death cases.

Countries	LSTM				GRU				Bi-LSTM				Dense-LSTM			
	Loss	MAE	MSE	R^2	Loss	MAE	MSE	R^2	Loss	MAE	MSE	R^2	Loss	MAE	MSE	R^2
United States	0.00102	0.01962	0.00102	0.73470	8.57925	0.00679	0.02788	0.76260	0.25175	0.00623	0.25110	0.77190	1.59402	0.00763	0.00016	0.77190
Vietnam	0.00130	0.02446	0.00135	0.75330	0.00260	0.01730	0.00260	0.64170	0.01665	0.04687	0.01664	0.72540	0.00493	0.02372	0.00495	0.72540
UK	0.00021	0.00791	0.00021	0.77190	0.00015	0.00623	0.00015	0.66960	0.00062	0.01553	0.00062	0.73470	0.00047	0.01469	0.00047	0.73470
Turkey	0.00018	0.00893	0.00018	0.66960	0.00008	0.00651	0.00008	0.77190	0.00010	0.00623	0.00010	0.73470	0.00008	0.00586	0.00077	0.73470
Spain	0.00071	0.01423	0.00071	0.77190	0.00013	0.00753	0.00013	0.72540	0.00018	0.00623	0.00018	0.75330	0.00022	0.00586	0.00022	0.75330
Russia	0.00009	0.00577	0.00009	0.80910	0.00002	0.00326	0.00002	0.73470	0.00008	0.00493	0.00008	0.76260	0.00000	0.00085	0.00000	0.72540
Italy	0.00050	0.01023	0.00050	0.76260	0.00043	0.00744	0.00043	0.73470	0.00087	0.01181	0.00087	0.64170	0.00048	0.01358	0.00048	0.73470
India	0.00005	0.00530	0.00005	0.64170	0.00003	0.00428	0.00003	0.75330	0.00008	0.00670	0.00008	0.66960	0.00007	0.00679	0.00007	0.73470
Germany	0.01925	0.04445	0.01925	0.66960	0.00019	0.00623	0.00019	0.77190	0.00874	0.03357	0.00878	0.75330	0.00074	0.01023	0.00074	0.75330
France	0.00102	0.01562	0.00100	0.77190	0.00052	0.01218	0.00052	0.66960	0.00059	0.01256	0.00059	0.77190	0.00046	0.01469	0.00046	0.77190
Brazil	0.00017	0.00828	0.00017	0.72540	0.00003	0.00363	0.00003	0.77190	0.00004	0.00409	0.00004	0.66960	0.00005	0.00474	0.00005	0.73470
Argentina	0.00003	0.00381	0.00003	0.73470	0.00004	0.00484	0.00004	0.80910	0.00002	0.00391	0.00002	0.77190	0.00004	0.00512	0.00004	0.75330

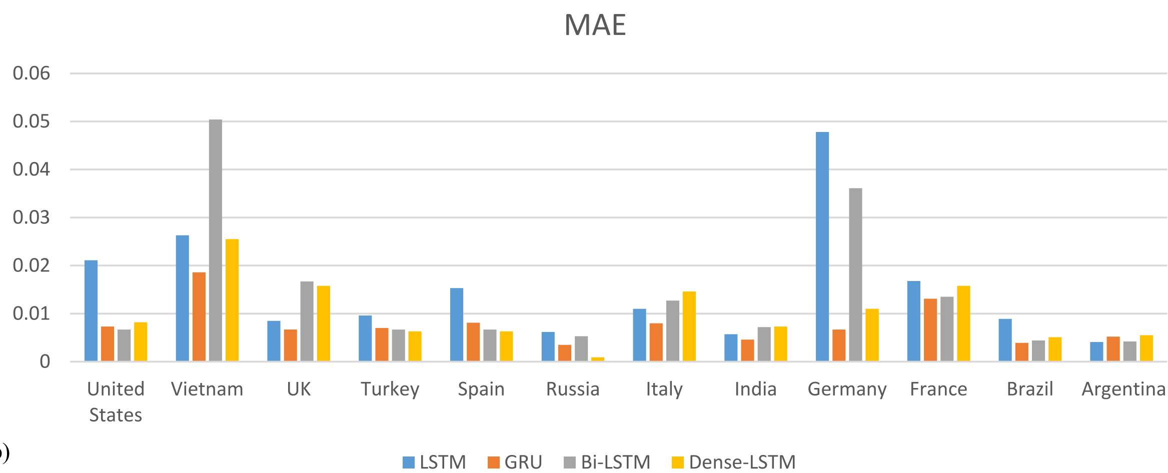
Countries	LSTM MEDAE	GRU MEDAE	Bi-LSTM MEDAE	Dense-LSTM MEDAE
United States	0.01882	0.00779	0.00566	0.00760
Vietnam	0.02505	0.01654	0.04710	0.02467
UK	0.00853	0.00668	0.01502	0.01374
Turkey	0.00988	0.00553	0.00689	0.00565
Spain	0.01466	0.00843	0.00642	0.00555
Russia	0.00527	0.00374	0.00504	0.00037
Italy	0.01081	0.00750	0.01227	0.01259
India	0.00533	0.00330	0.00759	0.00740
Germany	0.04518	0.00701	0.03402	0.01118
France	0.01598	0.01209	0.01243	0.01458
Brazil	0.00794	0.00463	0.00486	0.00390
Argentina	0.00369	0.00512	0.00388	0.00545

Table 5
COVID-19 cases stages.

Stage	Produce	Break out	Break out	Decline
Time	End of December 2019 - January 22, 2020	January 23, 2020 - February 10, 2020	February 11, 2020 - February 22, 2020	February 23, 2020 - present
Representative event	Epidemic, confirmed human-to-human transmission	Wuhan imposes traffic control	Closed management of the community	Gradual resumption of work and production
The number of infected people	The daily number of newly infected people is gradually increasing, and the cumulative number of infected people is less	Rapid increase in daily new and cumulative infections	The daily number of newly infected people is gradually decreasing, and the cumulative number of infected people is still high	The daily number of new and cumulative cases of infections has continued to decrease, and finally cleared to zero
Epidemic Information Status	The number of infected people is uncertain, and the degree of information uncertainty is high	Information such as contagiousness and incubation period is still uncertain and is being updated continuously	Information uncertainty is reduced	Relatively low level of information uncertainty



(a)



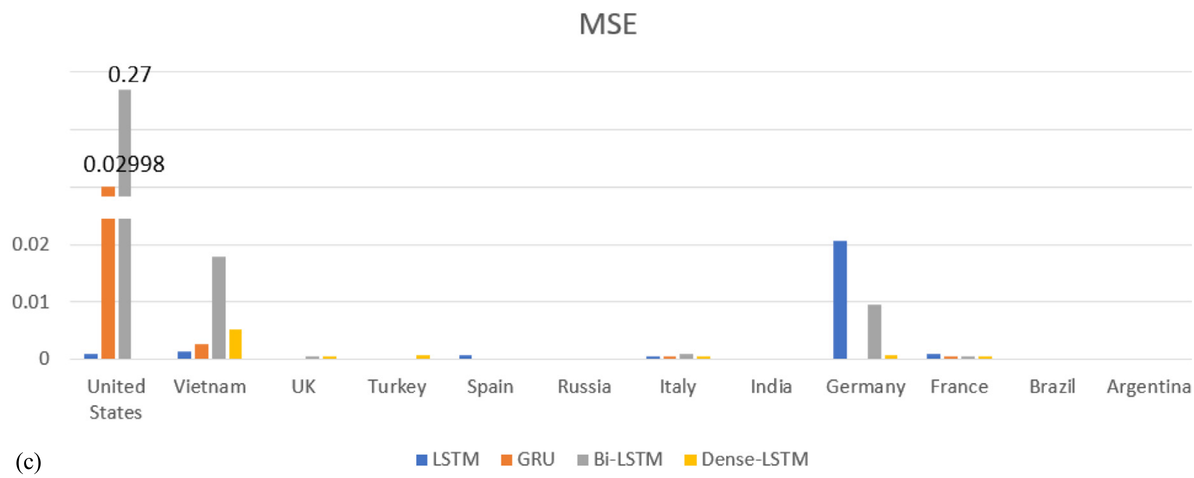
(b)

Fig. 7. Analysis of results of cases of COVID-19 in Table 3 (a) Loss (b) MAE (c) MSE (d) R^2 (e) MEDAE.

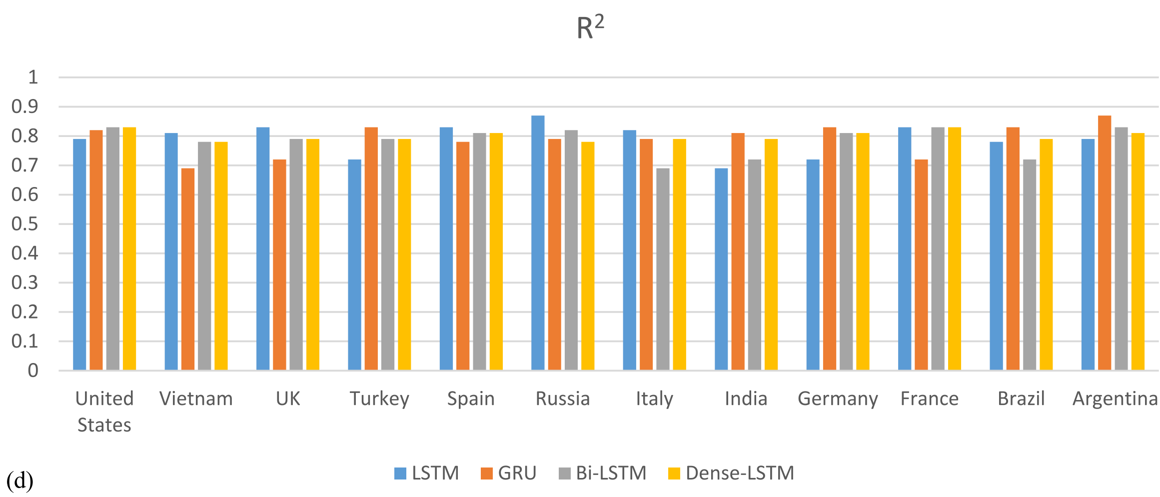
an easily occurred problem. Comparing to other data sets used in deep learning, COVID-19 has far less data. Although the complexity of the model can be reduced by reducing the number of LSTM hidden layers, alleviating the over-fitting problem to some extent, the effect of such a solution is still limited in the case of data shortage. Also, this work only made short term forecasts. It also does not include the prediction of the ending of the pandemics, which might require more complicated models. Despite the limitations, LSTM based models are the most suitable model for time series data. Koç and Türkoğlu reported a LSTM network has a high level of accuracy, that can be used to predict the number of cases in the COVID-19 (Koc and Turkoglu, 2022). LSTM networks were also used in different countries to forecast new COVID cases and deaths (Gautam, 2022). In this paper, a prediction model of emergency material demand is constructed based on the infectious disease model, and a time-varying demand and LSTM sequential decision model are respectively established to provide a scientific and effective prediction method for the actual emergency rescue work. Predicting infectious disease trends is a daunting and complex task. From the perspective of data mining, the difficulty is to build a most accurate model with very little data. Therefore, this study focused on different LSTM based forecasting models, which combines traditional infectious disease prediction methods and deep learning prediction methods: the GRU model reveals the autocorrelation of infectious disease development, while the LSTM model shows the intrinsic factors of data fluctuations. Therefore, the approach proposed in this paper not only preserves the general trend of epidemic development, but also captures occasional fluctuations. In addition, different countries migration data is included in the proposed model, and the diversified data helps to retrieve other characteristics related to the epidemic and accurately build the model. In many studies, hybrid models are widely used to improve performance.

6. Conclusion

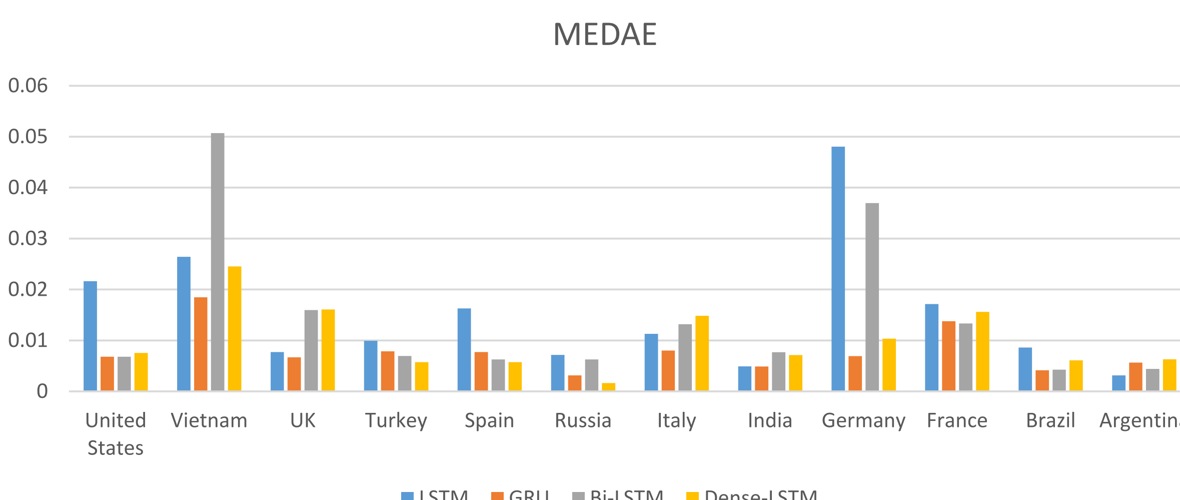
In order to better explore the impact of different prevention and control intensities on the epidemic, in the early stage of the epidemic in various countries, the changes in the effective cases number of the 12 countries were calculated, and the prevention and control effect of the epidemic in different countries under the prevention and control policies of different intensities was analysed. The GRU, LSTM, Dense-LSTM and Bi-LSTM were constructed to model and predict the daily new confirmed case data of the main epidemic countries, and the prediction effect of each model was compared by error, and the results showed that the LSTM models had the highest prediction error and the highest prediction accuracy for the daily new confirmed case data in the 12 countries. The LSTM neural network with more layers is used for the prediction of cumulative diagnosis, existing diagnosis, and cumulative cure in countries, and on this basis, dropout technology is introduced to random probability inactivation of neurons, which effectively avoids the problem of overfitting of neural networks and fully explores the temporal and nonlinear relationship of data. The cumulative number of confirmed cases, existing confirmed cases, and cumulative cured numbers nationwide have verified that the use of the LSTM model to predict the trend of new crown pneumonia is completely feasible and has a high degree of accuracy.



(c)



(d)



(e)

Fig. 7. (*continued*).

CRediT authorship contribution statement

Luyu Zhou: Conceptualization, Methodology, Validation, Formal analysis, Resources, Data curation, Writing – original draft, Supervision, Review & editing, Revising the manuscript. **Chun Zhao:** Validation, Formal analysis, Review & editing, Revising the manuscript. **Ning Liu:** Validation, Formal analysis. **Xingduo Yao:** Validation, Formal analysis. **Zewei Cheng:** Validation, Formal analysis.

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

All data are included in this article.

Acknowledgements

This work was supported by National Natural Science Foundation of China (81870236); Hunan University Scientific research start-up funds, China; Innovation and entrepreneurship project for college students, China (202011065318). All authors have read and agreed to the published version of the manuscript.

References

Aamir, M., Li, Z., Bazai, S., Wagan, RA., Bhatti, UA., Nizamani, MM., Akram, S., 2021. Spatiotemporal change of air-quality patterns in hubei province-A pre- to post-COVID-19 analysis using path analysis and regression. Atmosphere 12, http://dx.doi.org/10.3390/atmos12101338.

Al-Rakhami, MS., Islam, MM., Islam, MZ., Asraf, A., Sodhro, AH., Ding, W., 2020. Diagnosis of COVID-19 from X-rays using combined CNN-RNN architecture with transfer learning. MedRxiv:2020.2008.2024.20181339.

AlOmar, MK., Hameed, M., AlSaadi, MA., 2020. Multi hours ahead prediction of surface ozone gas concentration: Robust artificial intelligence approach. Atmos. Pollut. Res. 11, 1572–1587. http://dx.doi.org/10.1016/j.apr.2020.06.024.

Aslan, MF., Unlersen, MF., Sabanci, K., Durdu, A., 2021. CNN-based transfer learning-BiLSTM network: A novel approach for COVID-19 infection detection. Appl. Soft Comput. 98, http://dx.doi.org/10.1016/j.asoc.2020.106912.

Ayoobi, N., Sharifrazi, D., Alizadehsani, R., Shoeibi, A., Gorriz, JM., Moosaei, H., Khosravi, A., Nahavandi, S., Chofreh, AG., Goni, FA., Klemes, JJ., Mosavi, A., 2021. Time series forecasting of new cases and new deaths rate for COVID-19 using deep learning methods. Results Phys. 27, http://dx.doi.org/10.1016/j.rinp.2021.104495.

Badri, AK., Heikal, J., Terah, YA., Nurjaman, DR., 2022. Decision-making techniques using LSTM on antam mining shares before and during the COVID-19 pandemic in Indonesia. APTISI Trans. Manag. (ATM) 6, 167–180.

Bedi, JS., Dhaka, P., Vijay, D., Aulakh, RS., Gill, JPS., 2020. Assessment of air quality changes in the four metropolitan cities of India during COVID-19 pandemic lockdown. Aerosol Air Qual. Res. 20, 2062–2070. http://dx.doi.org/10.4209/aaqr.2020.05.0209.

Benvenuto, D., Giovanetti, M., Vassallo, L., Angeletti, S., Ciccozzi, M., 2020. Application of the ARIMA model on the COVID-2019 epidemic dataset. Data Brief 29, http://dx.doi.org/10.1016/j.dib.2020.105340.

Bhatti, UA., Huang, M., Wang, H., Zhang, Y., Mehmood, A., Di, W., 2018. Recommendation system for immunization coverage and monitoring. Hum. Vacc. Immunother. 14, 165–171. http://dx.doi.org/10.1080/21645515.2017.1379639.

Bhatti, UA., Huang, M., Wu, D., Zhang, Y., Mehmood, A., Han, H., 2019. Recommendation system using feature extraction and pattern recognition in clinical care systems. Enterp. Inform. Syst. 13, 329–351. http://dx.doi.org/10.1080/17517575.2018.1557256.

Bhatti, UA., Yan, Y., Zhou, M., Ali, S., Hussain, A., Huo, Q., Yu, Z., Yuan, L., 2021. Time series analysis and forecasting of air pollution particulate matter (PM2.5): An SARIMA and factor analysis approach. IEEE Access 9, 41019–41031. http://dx.doi.org/10.1109/access.2021.3060744.

Bhatti, UA., Zeeshan, Z., Nizamani, MM., Bazai, S., Yu, Z., Yuan, L., 2022. Assessing the change of ambient air quality patterns in Jiangsu Province of China pre-to post-COVID-19. Chemosphere 288, http://dx.doi.org/10.1016/j.chemosphere.2021.132569.

Bunn, D., Wright, G., 1991. Interaction of judgemental and statistical forecasting methods: issues & analysis. Manage. Sci. 37, 501–518.

Ceylan, Z., 2021. Short-term prediction of COVID-19 spread using grey rolling model optimized by particle swarm optimization. Appl. Soft Comput. 109, http://dx.doi.org/10.1016/j.asoc.2021.107592.

Chen, J., Pi, D., Wu, Z., Zhao, X., Pan, Y., Zhang, Q., 2021. Imbalanced satellite telemetry data anomaly detection model based on Bayesian LSTM. Acta Astronaut. 180, 232–242. http://dx.doi.org/10.1016/j.actaastro.2020.12.012.

Chimmula, VKR., Zhang, L., 2020. Time series forecasting of COVID-19 transmission in Canada using LSTM networks. Chaos Solitons Fractals 135, http://dx.doi.org/10.1016/j.chaos.2020.109864.

Chumachenko, D., Meniailov, I., Bazilevych, K., Chumachenko, T., Yakovlev, S., 2022. Investigation of statistical machine learning models for COVID-19 epidemic process simulation: Random forest, K-nearest neighbors, gradient boosting. Computation 10, http://dx.doi.org/10.3390/computation10060086.

Comito, C., Forestiero, A., Papuzzo, G., 2019. A clinical decision support framework for automatic disease diagnoses. In: IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining. ASONAM, Vancouver, Canada, pp. 933–936.

Comito, C., Pizzuti, C., 2022. Artificial intelligence for forecasting and diagnosing COVID-19 pandemic: A focused review. Artif. Intell. Med. 128, http://dx.doi.org/10.1016/j.artmed.2022.102286.

Dahiya, S., Soni, P., Nadappattel, HS., Fraz, M., 2022. A hybrid approach of ANN-PSO technique for anomaly detection. In: Proceedings of Second Doctoral Symposium on Computational Intelligence. DoSCI 2021, Springer, pp. 757–767.

Dairi, A., Harrou, F., Zeroual, A., Hittawe, MM., Sun, Y., 2021. Comparative study of machine learning methods for COVID-19 transmission forecasting. J. Biomed. Inform. 118, http://dx.doi.org/10.1016/j.jbi.2021.103791.

Fahim, A., Tan, Q., Mazzi, M., Sahabuddin, M., Naz, B., Ullah Bazai, S., 2021. Hybrid LSTM self-attention mechanism model for forecasting the reform of scientific research in Morocco. Comput. Intell. Neurosci. 2021, http://dx.doi.org/10.1155/2021/6689204.

Fang, Z., Dowe, DL., Peiris, S., Rosadi, D., 2021. Minimum message length in hybrid ARMA and LSTM model forecasting. Entropy 23, http://dx.doi.org/10.3390/e23121601.

Fildes, R., 1992. The evaluation of extrapolative forecasting methods. Int. J. Forecast. 8, 81–98.

Gautam, Y., 2022. Transfer learning for COVID-19 cases and deaths forecast using LSTM network. ISA Trans. 124, 41–56. http://dx.doi.org/10.1016/j.isatra.2020.12.057.

Hadjira, A., Salhi, H., Hafa, F.El., 2021. A comparative study between ARIMA model, holt-winters–no seasonal and fuzzy time series for new cases of COVID-19 in Algeria. Am. J. Public Health 9, 248–256.

Heydari, A., Majidi Nezhad, M., Astiaso Garcia, D., Keynia, F., De Santoli, L., 2022. Air pollution forecasting application based on deep learning model and optimization algorithm. Clean Technol. Environ. Policy 24, 607–621. http://dx.doi.org/10.1007/s10098-021-02080-528.

Huang, KY., Jane, C-J., 2009. A hybrid model for stock market forecasting and portfolio selection based on ARX, grey system and RS theories. Expert Syst. Appl. 36, 5387–5392. http://dx.doi.org/10.1016/j.eswa.2008.06.103.

Irfan, M., Razzaq, A., Suksatan, W., Sharif, A., Elavarasan, RM., Yang, C., Hao, Y., Rauf, A., 2022. Asymmetric impact of temperature on COVID-19 spread in India: Evidence from quantile-on-quantile regression approach. J. Therm. Biol. 104, http://dx.doi.org/10.1016/j.jtherbio.2021.103101.

Islam, MZ., Islam, MM., Asraf, A., 2020. A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. Inform. Med. Unlocked 20, 100412-100412. http://dx.doi.org/10.1016/j.imu.2020.100412.

Jamshidi, MB., Roshani, S., Talla, J., Lalbakhsh, A., Peroutka, Z., Roshani, S., Parandin, F., Malek, Z., Daneshfar, F., Niazkar, HR., 2022. A review of the potential of artificial intelligence approaches to forecasting COVID-19 spreading. AI 3, 493–511.

Jiang, H., Li, Y., Zhou, C., Hong, H., Glade, T., Yin, K., 2020. Landslide displacement prediction combining LSTM and SVR algorithms: A case study of shengjibao landslide from the three Gorges Reservoir Area. Appl. Sci. Basel 10, http://dx.doi.org/10.3390/app10217830.

Julong, D., 1982. Grey control system. J. Huazhong Univ. Sci. Technol. 3 (18).

Khan, FM., Gupta, R., 2020. ARIMA and NAR based prediction model for time series analysis of COVID-19 cases in India. J. Saf. Sci. Resil. 1, 12–18.

Kim, T-Y., Cho, S-B., 2019. Predicting residential energy consumption using CNN-LSTM neural networks. Energy 182, 72–81. http://dx.doi.org/10.1016/j.energy.2019.05.230.

Koc, E., Turkoglu, M., 2022. Forecasting of medical equipment demand and outbreak spreading based on deep long short-term memory network: the COVID-19 pandemic in Turkey. Signal Image Video Process. 16, 613–621. http://dx.doi.org/10.1007/s11760-020-01847-5.

Kufel, T., 2020. ARIMA-based forecasting of the dynamics of confirmed Covid-19 cases for selected European countries. Equilib. Q. J. Econ. Econ. Policy 15, 181–204. http://dx.doi.org/10.24136/eq.2020.009.

Kumari, S., Singh, SK., 2022. Machine learning-based time series models for effective CO2 emission prediction in India. Environ. Sci. Pollut. Res. http://dx.doi.org/10.1007/s11356-022-21723-8.

Lara-Benitez, P., Carranza-Garcia, M., Riquelme, JC., 2021. An experimental review on deep learning architectures for time series forecasting. Int. J. Neural Syst. 31, http://dx.doi.org/10.1142/s0129065721300011.

Leng, X.z., Wang, J., Ji, H., Wang, Q.g., Li, H., Qian, X., Li, F., Yang, M., 2017. Prediction of size-fractionated airborne particle-bound metals using MLR, BP-ANN and SVM analyses. Chemosphere 180, 513–522. http://dx.doi.org/10.1016/j.chemosphere.2017.04.015.

Li, MY., Muldowney, JS., 1995. Global stability for the SEIR model in epidemiology. Math. Biosci. 125, 155–164. http://dx.doi.org/10.1016/0025-5564(95)92756-5.

Liu, B., Guo, S-D., 2010. Attentions on grey system theories by China scholars — Based on literature metrology during 1982–2009. J. Grey Syst. 22, 137–146.

Liu, Z-Q., Peng, M-C., Sun, Y-C., 2021b. Estimation of driver lane change intention based on the LSTM and Dempster–Shafer evidence theory. J. Adv. Transp. 2021, 1–11.

Liu, Q., Yang, C-Y., Lin, L., 2021a. Deformation prediction of a deep foundation pit based on the combination model of wavelet transform and gray BP neural network. Math. Probl. Eng. 2021, http://dx.doi.org/10.1155/2021/2161254.

Mach, L., Bedrunka, K., Kuczuk, A., Szewczuk-Stepien, M., 2021. Effect of structural funds on housing market sustainability development-correlation, regression and wavelet coherence analysis. Risks 9, http://dx.doi.org/10.3390/risks9100182.

Martinez-Alvarez, F., Asencio-Cortes, G., Torres, JF., Gutierrez-Aviles, D., Melgar-Garcia, L., Perez-Chacon, R., Rubio-Escudero, C., Riquelme, JC., Troncoso, A., 2020. Coronavirus optimization algorithm: A bioinspired metaheuristic based on the COVID-19 propagation model. Big Data 8, 308–322. http://dx.doi.org/10.1089/big.2020.0051.

Mehrolia, S., Alagarsamy, S., Solaikutty, VM., 2021. Customers response to online food delivery services during COVID-19 outbreak using binary logistic regression. Int. J. Consum. Stud. 45, 396–408. http://dx.doi.org/10.1111/ijcs.12630.

Nawaz, SA., Li, J., Bhatti, UA., Mehmood, A., Shoukat, MU., Bhatti, MA., 2020. Advance hybrid medical watermarking algorithm using speeded up robust features and discrete cosine transform. PLoS One 15, http://dx.doi.org/10.1371/journal.pone.0232902.

Pahar, M., Klopper, M., Warren, R., Niesler, T., 2021. COVID-19 cough classification using machine learning and global smartphone recordings. Comput. Biol. Med. 135, http://dx.doi.org/10.1016/j.compbiomed.2021.104572.

Pan, J., Yao, Y., Liu, Z., Meng, X., Ji, JS., Qiu, Y., Wang, W., Zhang, L., Wang, W., Kan, H., 2021. Warmer weather unlikely to reduce the COVID-19 transmission: An ecological study in 202 locations in 8 countries. Sci. Total Environ. 753, http://dx.doi.org/10.1016/j.scitotenv.2020.142272.

Rubbaniy, G., Khalid, AA., Samitas, A., 2021. Are cryptos safe-haven assets during Covid-19? Evidence from wavelet coherence analysis. Emerg. Mark. Finance Trade 57, 1741–1756. http://dx.doi.org/10.1080/1540496x.2021.1897004.

Saha, S., Al-Rifai, RH., Saha, S., 2021. Diabetes prevalence and mortality in COVID-19 patients: a systematic review, meta-analysis, and meta-regression. J. Diabetes Metab. Disord. 20, 939–950. http://dx.doi.org/10.1007/s40200-021-00779-2.

Sahin, U., Sahin, T., 2020. Forecasting the cumulative number of confirmed cases of COVID-19 in Italy, UK and USA using fractional nonlinear grey Bernoulli model. Chaos Solitons Fractals 138, http://dx.doi.org/10.1016/j.chaos.2020.109948.

Shahid, F., Zameer, A., Muneeb, M., 2020. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM. Chaos Solitons Fractals 140, http://dx.doi.org/10.1016/j.chaos.2020.110212.

Shen, M-L., Lee, C-F., Liu, H-H., Chang, P-Y., Yang, C-H., 2021. Effective multinational trade forecasting using LSTM recurrent neural network. Expert Syst. Appl. 182, http://dx.doi.org/10.1016/j.eswa.204.115199.

Syage, JA., 2020. A statistical and dynamical model for forecasting COVID-19 deaths based on a hybrid asymmetric gaussian and SEIR construct. medRxiv.

Tan, W., Bian, R., Yang, W., Hou, Y., IEEE, 2020. Analysis of 2019-nCoV epidemic situation based on modified SEIR model and DSGE algorithm. In: 5th International Conference on Information Science, Computer Technology and Transportation. ISCTT, Shenyang, Peoples R China, pp. 369–376.

Tarwani, KM., Edem, S., 2017. Survey on recurrent neural network in natural language processing. Int. J. Eng. Trends Technol. 48, 301–304.

Tyass, I., Bellat, A., Raihani, A., Mansouri, K., Khalili, T., 2022. Wind speed prediction based on seasonal ARIMA model. In: E3S Web of Conferences. EDP Sciences, p. 00034.

Wang, N., Adeli, H., 2015. Self-constructing wavelet neural network algorithm for nonlinear control of large structures. Eng. Appl. Artif. Intell. 41, 249–258. http://dx.doi.org/10.1016/j.engappai.2015.01.018.

Wang, S., Geng, M., Yu, C., Cai, J., 2022. Improved behavioral modeling using augmented LSTM networks for ultra-broadband mmwave PA. Microw. Opt. Technol. Lett. http://dx.doi.org/10.1002/mop.33297.

Wu, Y-X., Wu, Q-B., Zhu, J-Q., 2019. Improved EEMD-based crude oil price forecasting using LSTM networks. Physica A 516, 114–124. http://dx.doi.org/10.1016/j.physa.2018.09.120.

Wu, D., Zhang, D., Liu, S., Jin, Z., Chowwanonthapunya, T., Gao, J., Li, X., 2020. Prediction of polycarbonate degradation in natural atmospheric environment of China based on BP-ANN model with screened environmental factors. Chem. Eng. J. 399, http://dx.doi.org/10.1016/j.cej.2020.125878.

Xavier, . A., 2020. A C++ code for predicting COVID-19 cases by least-squares fitting of the Logistic model. Pre-print available on Research Gate (https://www.researchgate.net/).

Xia, J., Feng, Y., Lu, C., Fei, C., Xue, X., 2021. LSTM-based multi-layer self-attention method for remaining useful life estimation of mechanical systems. Eng. Fail. Anal. 125, http://dx.doi.org/10.1016/j.engfailanal.2021.105385.

Xie, K., Yi, H., Hu, G., Li, L., Fan, Z., 2020. Short-term power load forecasting based on Elman neural network with particle swarm optimization. Neurocomputing 416, 136–142. http://dx.doi.org/10.1016/j.neucom.2019.02.063.

Yilanci, V., Pata, UK., 2022. COVID-19, stock prices, exchange rates and sovereign bonds: A wavelet-based analysis for Brazil and India. Int. J. Emerg. Mark. http://dx.doi.org/10.1108/ijoem-09-2021-1465.

Yin, W., Kann, K., Yu, M., Schütze, H., 2017. Comparative study of CNN and RNN for natural language processing. arXiv preprint arXiv:1702.01923.

Zeng, W., Li, J., Quan, Z., Lu, X., 2021. A deep graph-embedded LSTM neural network approach for airport delay prediction. J. Adv. Transp. 2021, http://dx.doi.org/10.1155/2021/6638130.

Zhang, J., Li, Y., Xiao, W., Zhang, Z., 2020. Non-iterative and fast deep learning: Multilayer extreme learning machines. J. Franklin Inst. B 357, 8925–8955.

Zhang, J., Zhao, Y., Shone, F., Li, Z., Frangi, AF., Xie, SQ., Zhang, Z-Q., 2022. Physics-informed deep learning for musculoskeletal modelling: Predicting muscle forces and joint kinematics from surface EMG. IEEE Trans. Neural Syst. Rehabil. Eng..

Zivkovic, M., Bacanin, N., Venkatachalam, K., Nayyar, A., Djordjevic, A., Strumberger, I., Al-Turjman, F., 2021. COVID-19 cases prediction by using hybrid machine learning and beetle antennae search approach. Sustainable Cities Soc. 66, http://dx.doi.org/10.1016/j.scs.2020.102669.