FOCUS

Prediction of the COVID-19 infectivity and the sustainable impact on public health under deep learning algorithm

Weiwei Wang[1] · Jinghui Cai[2] · Jiali Xu[3] · Yuxiang Wang[4] · Yulin Zou[5]

Abstract

The aim is to explore the development trend of COVID-19 (Corona Virus Disease 2019) and predict the infectivity of 2019-nCoV (2019 Novel Coronavirus), as well as its impact on public health. First, the existing data are analyzed through data pre-processing to extract useful feature factors. Then, the LSTM (Long-Short Term Memory) prediction model in the deep learning algorithm is used to predict the epidemic situation in Hubei Province, outside Hubei nationwide, and the whole country, respectively. Meanwhile, the impact of intervention time changes on the epidemic situation is compared. The results show that the prediction results are almost consistent with the actual values. Specifically, Hubei Province abolishes quarantine restrictions after the Spring Festival holiday, and the first COVID-19 peak is reached in late February, while the second COVID-19 peak has been reached in early March. Finally, the cumulative number of diagnoses reaches 85,000 cases, with an increase of 15,000 cases compared with the nationwide cases outside Hubei under the continuous implementation of prevention and control measures. Under the prediction of the proposed LSTM model, if the nationwide implementation of prevention and control interventions is postponed by 5 days, the epidemic will peak in early March, and the cumulative number of diagnoses will be about 200,000; and if the intervention measures are implemented five days earlier, the epidemic will peak in mid-February, with a cumulative number of diagnoses of approximately 40,000. Meanwhile, the proposed LSTM model predicts the RMSE values of the epidemic situation in Hubei Province, outside Hubei nationwide, and the whole country as 34.63, 75.42, and 50.27, respectively. Under model comparison analysis, the prediction error of the proposed LSTM model is small and has better applicability over similar algorithms. The results show that the LSTM model is effective and has high performance in infectious disease prediction, and the research results can provide scientific and effective references for subsequent related research.

Keywords Deep learning algorithm · 2019-nCoV · Infectivity · Prediction model · Public health

1 Introduction

Coronaviruses are a large family of viruses. The novel coronavirus belongs to the β-coronavirus of the coronavirus family. The virus also includes severe acute respiratory syndrome coronavirus (SARS-CoV) and the Middle East respiratory syndrome coronavirus (MERS-CoV) (Rothe et al. 2020) (Wang et al. 2020). In December 2019, pneumonia caused by 2019-nCoV (2019 Novel Coronavirus) breaks out in Wuhan, Hubei Province. Then, it spreads to all provinces of China as well as many countries and regions. The World Health Organization (WHO)

declares it a global public health emergency (Zhao et al. 2020). Knowing and characterizing the virus in time are crucial to fight the epidemic. The gene morphology of 2019-nCoV is significantly different from SARS-CoV and MERS-CoV. It is more likely to grow in primary human airway epithelial cells and has strong infectivity in people (Jin et al. 2020). The 2019-nCoV is mainly transmitted through the respiratory tract droplet and direct contact, which is generally susceptible to the crowd. It can cause symptoms, such as fever, cough, fatigue, or myalgia, and half of the patients can have trouble breathing (Zhang et al. 2020). Most patients are mildly ill, with a good prognosis. A few patients are critically ill and even die. The presence of COVID-19 (Corona Virus Disease 2019) poses a serious threat to human health. As of March 31, the cumulative number of diagnoses worldwide has reached 806,847 cases,

Extended author information available on the last page of the article

with 36,985 deaths. The data are still rising (Tang et al. 2020). Under such a severe infection situation and risk, effective control and prediction of COVID-19 have become the focus of worldwide research.

With technological development, artificial intelligence (AI) has been applied to various research fields and is playing a huge role in the prediction research of COVID-19. At present, with the development of the epidemic, researchers have continuously deepened their understanding of COVID-19, and their mastery of the epidemiological characteristics of the disease is constantly improving (Huang et al. 2020). Effective and timely monitoring, prediction of COVID-19, as well as characterization and quantification of disease incidence, are key components of prevention and remission strategies. It can save lives, reduce pain, and minimize the impact. In the past research, some scholars used deep learning technology to realize the prediction of infectious diseases (Alakus & Turkoglu 2020) (Shorten et al. 2021) (Panwar et al. 2020). Deep learning is an analysis method that can be used to perform tasks that are difficult to use in conventional analysis methods. With the continuous development of deep learning, more deep learning algorithm models have been actively applied in various fields (Oakden-Rayner et al. 2017) (Kang and Kang 2017). Therefore, the method of infectious disease prediction using deep learning helps design effective models.

In summary, the deep learning algorithm is integrated into the prediction of COVID-19 and applied to the public health field to monitor the incidence of COVID-19 and prevents its further outbreak. The proposal can provide a reference for the prediction of 2019-nCoV infection, which is greatly significant.

2 Related work

2.1 Research on COVID-19

The outbreak of COVID-19 has a specific impact on the whole world and has seriously affected people's lives and economic activities (Andreadakis et al. 2020). The research on COVID-19 has also become the focus of researchers in various industries. Meanwhile, the prediction and analysis of COVID-19 can effectively analyze the epidemic prevention status of COVID-19, provide scientific and effective references for the prevention and control of COVID-19, and provide effective data support for various economic activities and the policy formulations, which is of great practical significance (Klopfenstein et al. 2020).

In COVID-19 epidemic prediction, Santosh (2020) constructed a COVID-19 epidemic prediction model based on machine learning and statistical theory. The results showed that the model could be applied to more complex situations (Santosh 2020). Mandal et al. (2020) explored the dynamic model of COVID-19 to predict and control COVID-19 (Mandal et al. 2020). Sanchez-Caballero et al. (2020) built a COVID-19 prediction model to provide a scientific and effective reference for inviting prevention and control policies (Sanchez-Caballero et al. 2020). The above studies indicate that in terms of COVID-19 model prediction, prediction accuracy improvement is the key. Moreover, emerging model prediction methods should be explored to further expand the research content.

2.2 Literature review of deep learning in epidemiology prediction

Deep learning is composed of the input layer, the hidden layer, and the output layer. Meanwhile, it is a multi-layer structure of the forward feedback neural network model. The function in the model has nonlinear characteristics and fitting ability, which is an important foundation for the development of information technology. Compared with traditional intelligent technology, deep learning technology can extract more valuable features with large amounts of training data, thereby improving the accuracy of data prediction and classification. Therefore, deep learning has obvious advantages in forecasting epidemic diseases.

At present, deep learning technology has become the focus of epidemic prediction research. Deep learning has shown good applicability in disease prediction research. Ayyoubzadeh et al. (2020) predicted the incidence of COVID-19 in Iran (Ayyoubzadeh et al. 2020) based on the concepts of data mining and deep learning. Fong et al. (2020) designed a COVID-19 prediction model based on hybrid deep learning and fuzzy rules (Fong et al. 2020). The above studies suggest that deep learning has very good applicability in the prediction of the COVID-19 epidemic, and the related research has very important theoretical significance and practical application value.

3 Methods

3.1 Key technologies for prediction of 2019-nCoV

Combined with the prediction of traditional viral infectious diseases, the deep learning method is applied to the field of 2019-nCoV prediction. The specific key technologies are data analysis, data preprocessing, and prediction model construction, as shown in Fig. 1. Data preprocessing is the first step in the entire prediction process. Thereupon, the infectivity of 2019-nCoV is predicted, and its impact on public health is analyzed. First, data with a specific format are chosen for the deep learning prediction algorithm

Fig. 1 Key technologies in the prediction process of 2019-nCoV infection
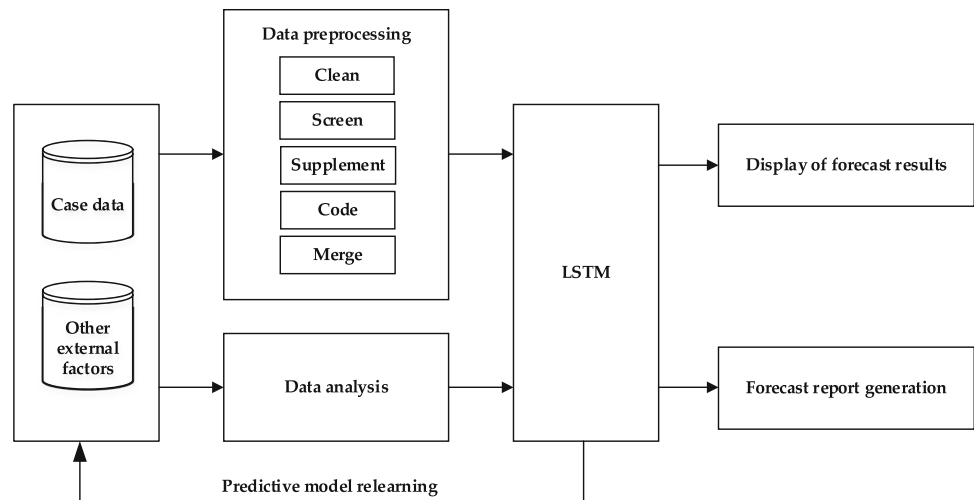
(Zeroual et al. 2020). Afterward, according to the characteristics of the data, one-hot-coding is performed for part of the data. Therefore, data preprocessing is a prerequisite for model research. According to the characteristics of the analyzed data, Python scripts are developed to facilitate and intellectualize the data pre-processing. Hence, no manual participation is needed in the online learning process of the subsequent prediction algorithm. Data analysis follows the data preprocessing process. In the preprocessing, noise in the data set is removed, and the missing values of the data are checked (Arora et al. 2020). Also, the required data are selected and encoded. Further, data redundancy attributes, such as the patient's name and contact information in the infection case table are deleted (Desai et al. 2020). Additionally, the missing data should be completed, and one-hot-coding should be performed for the disease coding data and area coding data. Some data analysis steps can also be carried out before data preprocessing. Through data analysis, the general incidence of viral infection is analyzed from different angles and dimensions. The generation and display of prediction result pairs are the core of the research (Bhattacharya et al. 2021). More accurate and reliable prediction methods are found through the analysis results of large amounts of data to visualize prediction results and online learning of prediction models. Here, the 2019-nCoV prediction model algorithm is optimized through continuous adjustment of the combination of data features, prediction time step, and prediction model parameters (Amyar et al. 2020). Finally, one or more reliable 2019-nCoV prediction models with high prediction accuracy are formed.

3.2 Data preprocessing

It has been argued that a healthy person will enter the incubation period after being infected with 2019-nCoV.

After the incubation period, he will enter the onset period. After the onset period, the patients are left to three situations: self-healing, quarantine, and death (Brunese et al. 2020). For a better understanding of the infectivity of 2019-nCoV, the feature factors that affect the incidence should be extracted through data analysis. Then, the data are preprocessed according to the extracted feature factors combined with other external factors (such as population density) that affect the incidence. Then, through comparative experiments, a method more suitable for the prediction of 2019-nCoV is found. The constants that control all links of infection are the initial number of infections, the rate of viral transmission, the latency of the virus, the hospital response time, hospital beds, and mobility intentions. The data used here come from the public data of the National Health Commission, provincial and municipal health commissions, provincial and municipal governments, as well as Hong Kong, Macao, and Taiwan official channels, from *January 20 to March 31, 2020*. The data preprocessing process includes four major parts: data cleaning, data filtering, data supplementation, and data encoding and merging. First, the data cleaning is performed to minimize the data storage space for more rapid and effective subsequent preprocessing. Then, the unknown part of the encoded data is removed from the data set, and the discontinuous time-series data are completed. Usually, no disease areas are filled with zeros to form complete time-series data. Then, one-hot-coding processing is performed on the encoded data to merge all the data.

3.3 Long-Short Term Memory (LSTM) prediction model

LSTM is a variant of deep learning (Recurrent Neural Network) RNN, and deep learning is a branch of machine learning Q. LSTM adds a memory unit between each

neuron unit in the hidden layer, which can control the time-series information. Each time when transferring in each unit of the LSTM hidden layer, several controllable gates (forget gate, input gate, and output gate) can be used to control the memory and forgetting degree of the previous and current data information (Fan et al. 2021). Therefore, LSTM can remember information for a longer time, and its practical application is wider (Hua et al. 2019). Here, the LSTM is applied to the prediction of the incidence of COVID-19 and verification of its effectiveness and accuracy.

LSTM is a variant of deep learning RNN. RNN adds the connection between the units of the hidden layer based on the ordinary neural network (Zhao et al. 2019). The weight can transfer the value of the last time-series to the current neuron. The data information of the last time-series can still affect the data information of the current time-series, to play a memory function (Fischer and Krauss 2018). Although RNN has a certain memory, it cannot remember data that are too far ahead or too far behind. Therefore, LSTM appears based on ordinary RNN. When transferring in each unit of the LSTM hidden layer, the memory and forgetting degree of the previous and the current data information can be controlled through the controllable gate. Therefore, compared with ordinary RNN, LSTM increases the controllability of memory function (Hua et al. 2019). In ordinary RNN, the neuron unit has only a simple repeating structure. Although LSTM has the same neuron structure, the results of neurons within the LSTM neural network are relatively more complex. The key to LSTM is the state of neuronal cells as well as the input and output of neuronal cells (Tsiouris et al. 2018). The role of the forget gate is to decide which information can be retained and which information can be discarded. The specific model of the forget gate is shown in Eq. (1).

$$f_t = \sigma\left(W_f \bullet [h_{t-1}, x_t] + b_f\right) \tag{1}$$

In Eq. (1), W_f represents the weight matrix connected to the forget gate. h_{t-1} represents the output of the last time node of the hidden layer. x_t represents the output of the input layer at time t. b_f represents the offset value of the forget gate. The following equation can be deduced by analogy.

After the forget gate determines which information to be retained and which to be forgotten, the input gate will determine which information to be updated. The tanh layer of the input layer determines what is used for updating. This part will be added to the neuron cell state. The input gate combines the two parts to update the neuron cell state. The specific expressions of the input model read:

$$i_t = \sigma(W_i \bullet [h_{t-1}, x_t] + b_i) \tag{2}$$

$$C_t = \tanh(W_C \bullet [h_{t-1}, x_t] + b_C) \tag{3}$$

The old state C_t and the state f_t of the forget gate are multiplied. The information that needs to be discarded is removed. The product of the input state i_t of the input gate and $\tilde{C}_t$ is added, that is, a new candidate value that needs to be updated. The expression of the final input state reads:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \tag{4}$$

The result of the tanh layer and the output of the sigmoid is multiplied to determine the final output part. The expression of output state reads:

$$o_t = \sigma(W_o \bullet [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \times \tanh(C_t) \tag{6}$$

The 2019-nCoV prediction should be analyzed using the features related to the occurrence and spread of viral infections. In the field of virus infection prediction, historical data are an indispensable data set. The RNN strategy gradient is improved to establish a link between the number of historical virus infections and the impact features, as well as the number of current virus infections and the impact features. The ordinary neural network is improved to LSTM so that the input information earlier and the current input information are connected (Qin et al. 2019). LSTM solves the problem of gradient disappearance and gradient explosion in the training process of the RNN. Compared with ordinary RNN, LSTM is easier to simulate complex functions. But the training time of the LSTM model will be longer. Through the training summary, the LSTM suitable for 2019-nCoV is obtained. Its structure is shown in Fig. 2.
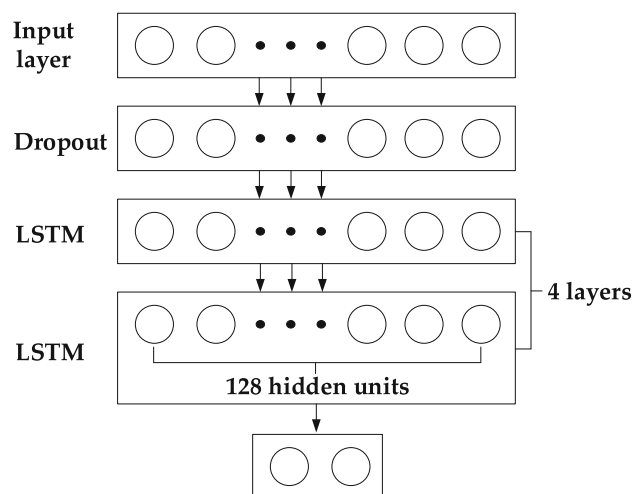

Fig. 2 LSTM structure diagram

3.4 Experimental scheme

Here, the parameter selection method is adopted from the deep learning method. In the case of other fixed parameters, the number of hidden layers is 4, the number of neurons in each hidden layer is 128, and the batch size is 32, and the dropout is 0. The prediction models with variable and fixed parameters are trained on the data, and the obtained models are compared to determine the best prediction model. The LSTM neural network structure is constructed to predict the incidence of 2019-nCoV. The performance and the accuracy of the prediction results of the LSTM model are verified through epidemic prediction experiments in Hubei Province, outside Hubei nationwide, and the whole country. The LSTM model is used to predict the development of the epidemic situation in Hubei Province when the quarantine restrictions are removed after the Spring Festival holiday. Also, the impact of changes in the implementation time of prevention and control interventions (postponed 5 days and advanced 5 days) on the national epidemic situation is compared. The prediction rate of the prediction model is evaluated by indexes, such as the root mean square error (RMSE), mean absolute scale error (MASE), and mean absolute error (MAE). RMSE is a commonly used evaluation method for the difference between the predicted value and the actual value. MAE can effectively reflect the prediction error, MASE can reflect the absolute mean error in the sample, which are usually used in other applications and the prediction of infectious diseases. The expressions of the three indexes read:

$$RMSE = \sqrt{\frac{(y'_t - y_t)^2}{n}} \tag{7}$$

$$MAE = \frac{\sum |y'_t - y_t|}{n} \tag{8}$$

$$q_j = \frac{e_j}{\frac{1}{T-m}\sum_{t=m+1}^{T} |y_t - y_{t-m}|} \tag{9}$$

$$MASE = mean(|q_j|) \tag{10}$$

In Eqs. (7), (8), (9), (10), y'_t represents the predicted value of the model, y_t denotes the actual value, n indicates the number of observation time points, e_i stands for the difference between the predicted point and the actual point at the moment i, y_t means the actual value at moment t, m refers to the seasonal cycle, and $y_{t\text{-}m}$ is the actual value at moment t-m.

4 Results

4.1 Analysis of model training

First, the LSTM model is trained. The training set is shown in Fig. 3.

Figure 3a shows that with the increase in the iterations, the loss value of the model training gradually decreases. Overall, the loss value of the LSTM model fluctuates around 0.2 and maintains a relatively stable state. Figure 3b displays that with the increase in iterations, the test accuracy of the model gradually increases. When the number of iterations is about 500, the test accuracy remains
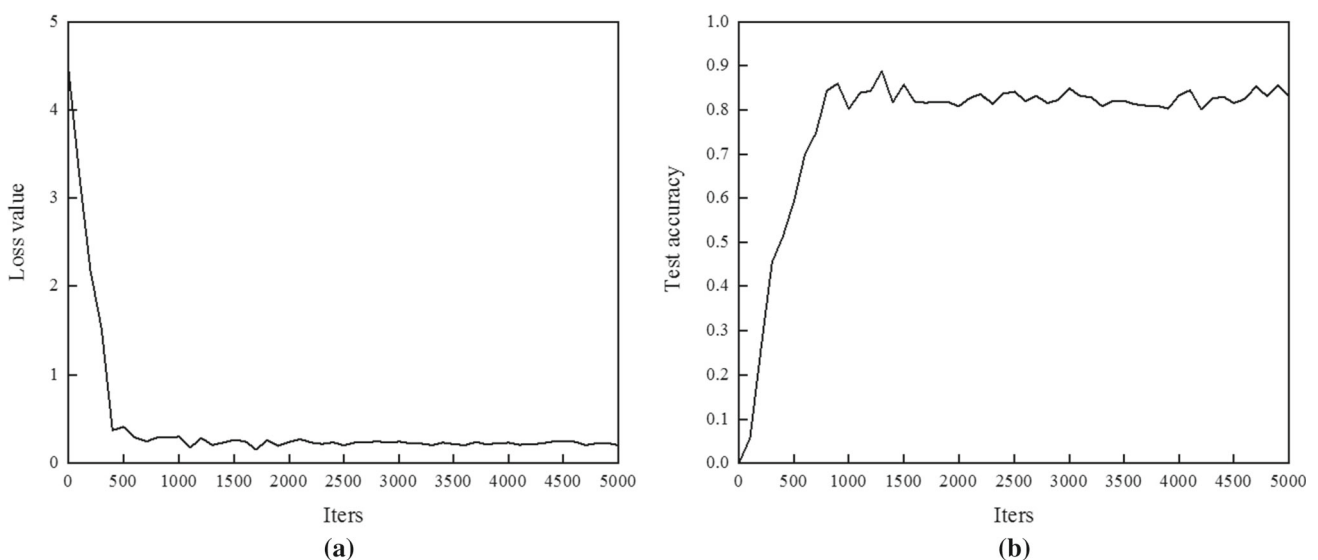

Fig. 3 Training result analysis of the model (**a**: model loss value. **b**: model test accuracy)

above 0.8. The experimental results show that the LSTM model has high accuracy in training.

4.2 Epidemic situation prediction results in Hubei Province

The prediction model divides the transmission process of COVID-19 into two stages. In the first stage, the lack of awareness of the specific conditions of the disease and limited awareness of the epidemic led to the lack of effective precautions against the transmission of the disease. At this stage, the number of infections continues to rise and the scale of the epidemic continues to expand, triggering a pandemic of COVID-19. In the second stage, great importance has been attached to disease control and prevention, and a more accurate understanding of the disease transmission path has been obtained. Thus, the population will gradually take effective and scientific prevention and control measures, so that the basic regeneration number continues to decline until it falls below 1. Also, it continues for a period to achieve the eradication of the disease. Figure 4 shows the trend of epidemic situation prediction in Hubei Province. From February 12th to 13th, during this period, Hubei Province included clinically diagnosed cases in the statistics of confirmed cases, resulting in a sharp increase in the number of cases. Then, the cumulative number of diagnoses gradually stabilized in late February. It is predicted that the epidemic situation in Hubei Province will come to an end in late April, and the total number of infected patients will be about 70,000 (Fig. 4a). The predicted data of the model are almost consistent with the actual data.

It is assumed that after the Spring Festival holiday, Hubei Province allows the normal movement of people and removes the quarantine restrictions. The province will reach the first epidemic peak in late February and the second peak in early March. The final scale of the epidemic will reach 85,000 cases, an increase of 15,000 cases compared to the epidemic scale of the continuous implementation of prevention and control measures (Fig. 4b).

Figure 4 implies that the difference between the predicted value and the actual value of COVID-19 is small in the two months of investigation. The predicted value is slightly higher than the actual value. The proposed model has good application performance.

4.3 Prediction results of domestic epidemic

On January 23, 2020, with the development of the COVID-19 epidemic, Wuhan has announced the "closure of the city." From that day on, Hubei and even China have taken unprecedented measures to prevent and control the COVID-19 epidemic. These measures include large-scale quarantine, strict personnel movement control, and extensive monitoring of suspected cases. The prediction results of the domestic epidemic situation are shown in Fig. 5.

Figure 5a and b represents the infection situation outside Hubei Province and the national infection situation, respectively. The results indicate that the infection situation of provinces and cities outside Hubei Province has stabilized in mid-February, and the national epidemic prevention and control has stabilized in late February. However, due to overseas imported cases, the number of people infected by the epidemic increased slightly in mid-
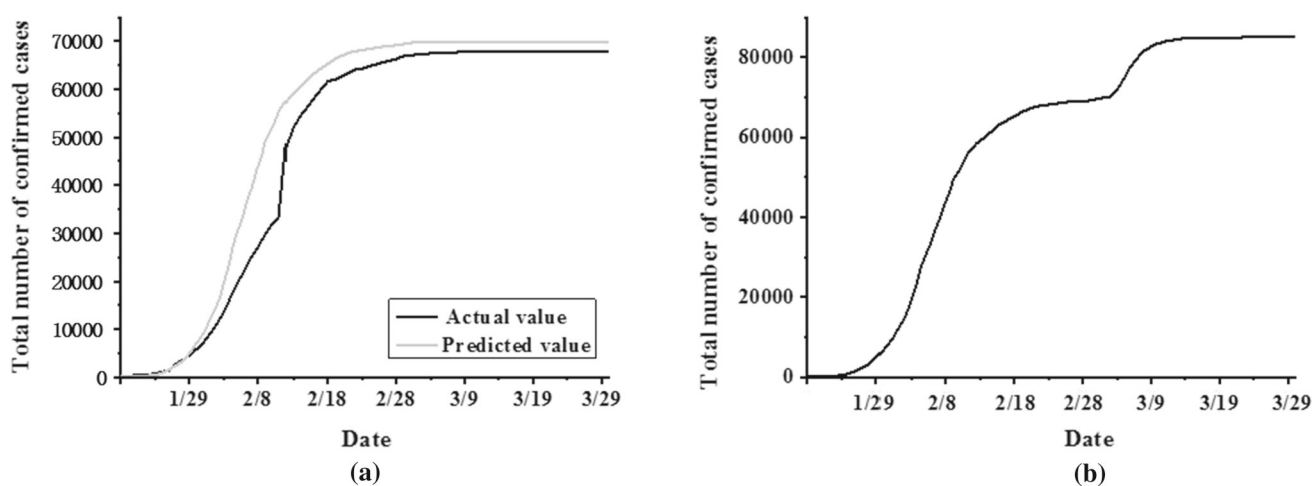

Fig. 4 Prediction results of the epidemic situation in Hubei Province: **a** cumulative number of confirmed cases in Hubei Province; **b** cumulative number of confirmed cases in Hubei Province when quarantine restrictions are removed
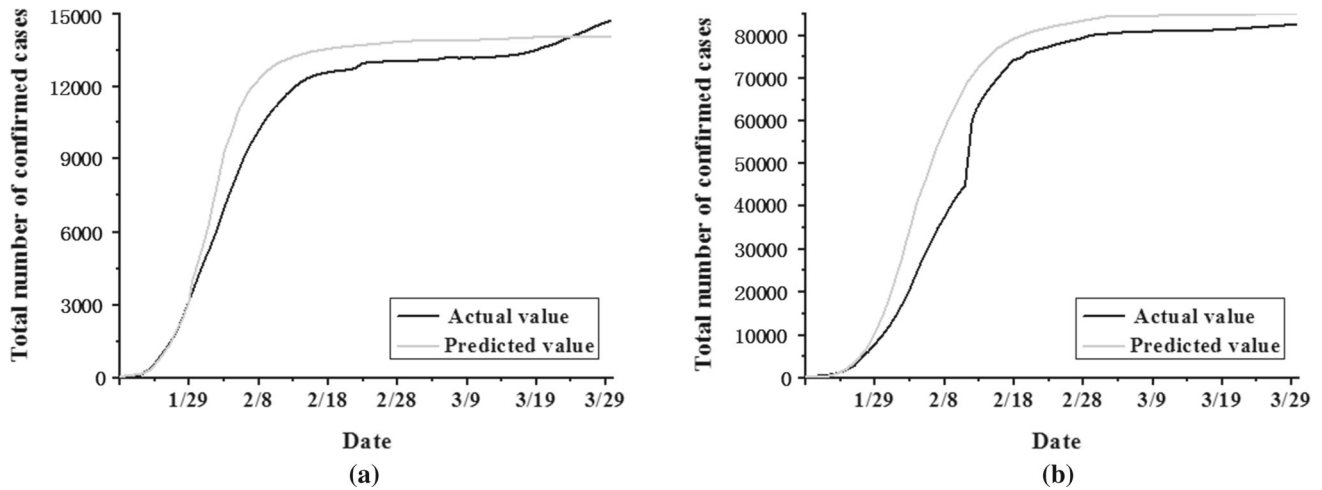
Fig. 5 Prediction results of domestic epidemic situation: **a** prediction results of the epidemic situation outside Hubei nationwide; **b** prediction results of national epidemic situation

March. The difference between the model prediction results and the real results is small, which also proves the accuracy of the model.

The results of the impact of intervention time changes on the epidemic situation are shown in Fig. 6.

Figure 6 shows that if the time for prevention and control interventions (January 23) is postponed by 5 days, the chances of the 2019-nCoV outspread will increase. The LSTM model predicts that the national epidemic situation will peak in early March and the cumulative number of diagnoses will reach 200,000. If the intervention measures are implemented 5 days earlier, the national epidemic will peak in mid-February, and the final epidemic scale will be 40,000 cases. Therefore, the proposed model can provide a

scientific and effective reference for comprehensive epidemic prevention and control decisions.

4.4 Comparative analysis of model prediction results

The prediction results of the LSTM model, ANN, and CNN are compared to further explore the prediction effect of the proposed LSTM model. Here, RMSE, MAE, and MASE are introduced as evaluation indexes. The specific results are shown in Fig. 7.

Figure 7 demonstrates that the RMSE, MAE, and MASE of the LSTM model are 6.903, 5.742, and 0.246, respectively, which are the smallest in all the comparison models. The RMSE, MAE, and MASE of the auto-
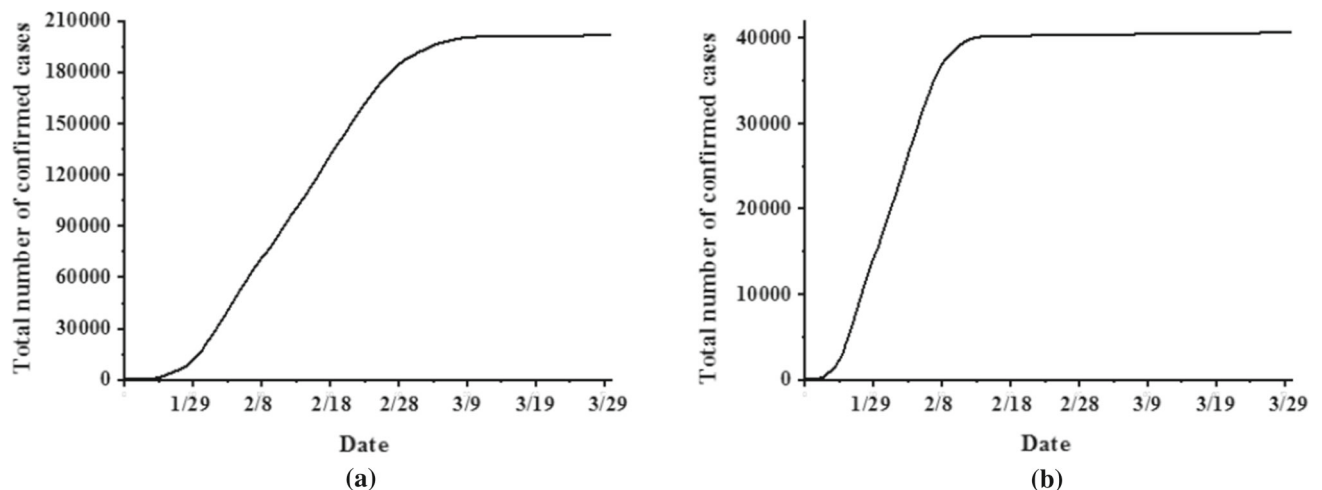

Fig. 6 Results of the impact of changes in intervention time on the epidemic situation: **a** National implementation of prevention and control interventions is postponed by 5 days; **b** National implementation of prevention and control interventions is advanced by 5 days
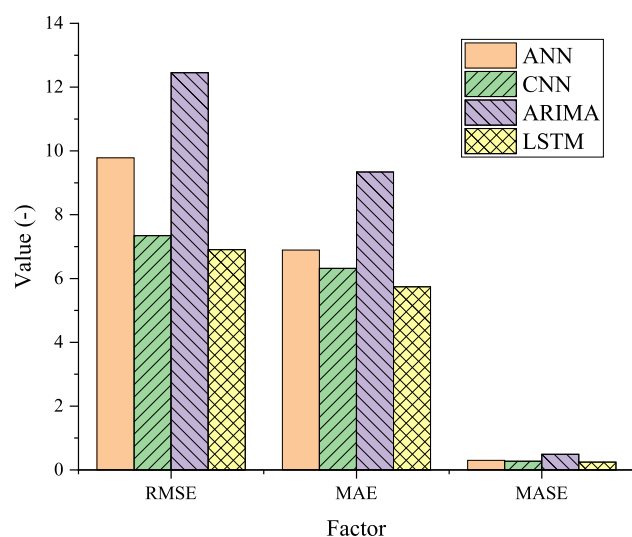
Fig. 7 Comparative analysis of LSTM model and other models

regression integrated moving average (ARIMA) model are 12.45, 9.34, and 0.489, respectively, which are the largest in all comparison algorithms. The results show that the proposed LSTM model has the best prediction effect, and the ARIMA model has the worst prediction effect. Thus, the proposed LSTM model can effectively and accurately predict 2019-nCoV, meet the needs of 2019-nCoV prevention and control, and achieve the expected requirements.

5 Discussion

The LSTM model is used to predict the 2019-nCoV infectivity and the impact on public health in Hubei Province, outside Hubei nationwide, and the whole country. The results show that the model prediction data are almost consistent with the actual data, indicating that the LSTM model has high accuracy. The RMSE of the LSTM model is small, indicating that the proposed LSTM model is suitable for the prediction of the 2019-nCoV infectivity. The LSTM model is used to predict the results of the impact of intervention time changes on the epidemic situation. The results show that if the prevention and control interventions are postponed by 5 days, the LSTM model predicts that the national epidemic will peak in early March and the cumulative number of diagnoses will reach 200,000. If the intervention measures are implemented five days earlier, the national epidemic will peak in mid-February, and the final epidemic scale will be 40,000 cases. The epidemic outside Hubei nationwide has not eventually caused a pandemic of COVID-19 similar to Hubei province because in addition to the continuous implementation of strict prevention and control quarantine measures in Hubei

Province, other provinces have invested many medical resources in Hubei Province for the construction of new hospitals and quarantine centers. It alleviates the shortage of local medical resources and reduces the risk of virus transmission, which plays a vital role in reducing the scale of the epidemic. The effective implementation of prevention and control measures has slowed the growth of the epidemic situation, which also shows the effectiveness of quarantine, as well as prevention and control measures. Continuous "early detection and early quarantine" measures can also effectively prevent the second-wave spread of the epidemic outside Hubei nationwide. Noticeably, however, as the epidemic situation of the COVID-19 in the international community gradually becomes severe, China also should pay attention to the prevention and control of imported cases.

In summary, the 2019-nCoV infectivity is predicted, and its impact on public health is analyzed based on the deep learning algorithm. Meanwhile, the effectiveness and high performance of the proposed LSTM model are verified in the prediction of the 2019-nCoV infection.

6 Conclusions

Here, the 2019-nCoV infectivity is predicted, and its impact on public health is analyzed based on the deep learning algorithm. The research results indicate that the prediction results of the LSTM model are almost consistent with the actual value, with high accuracy. The RMSE of the LSTM model is small, proving that the LSTM model is an effective prediction model for the 2019-nCoV infectivity. Therefore, the LSTM model has effectiveness and high performance in predicting the 2019-nCoV infectivity and the impact on public health. The results provide a reference for the prediction of 2019-nCoV infection. It has important application value and theoretical significance. However, there are some deficiencies. For example, the selection of feature factors is not comprehensive enough. Asymptomatic cases and imported cases are not considered. Therefore, in the follow-up study, the performance of the algorithm will be further improved, and more feature information will be considered to improve the prediction effect of the proposed algorithm so that the research algorithm has a more practical application value.

Author contributions Weiwei Wang, Jinghui Cai, and Jiali Xu contributed to conceptualization ; Yuxiang Wang was involved in formal analysis; Weiwei Wang contributed to writing—original draft preparation;Yulin Zou was involved in writing—review and editing;

Declarations

References

Alakus TB, Turkoglu I (2020) Comparison of deep learning approaches to predict COVID-19 infection. Chaos Soliton Fract 140:110120

Amyar A, Modzelewski R, Li H, Ruan S (2020) Multi-task deep learning-based CT imaging analysis for COVID-19 pneumonia: classification and segmentation. Comput Biol Med 126:104037

Andreadakis Z, Kumar A, Román RG, Tollefsen S, Saville M, Mayhew S (2020) The COVID-19 vaccine development landscape. Nat Rev Drug Discov 19(5):305–306

Arora P, Kumar H, Panigrahi BK (2020) Prediction and analysis of COVID-19 positive cases using deep learning models: a descriptive case study of India. Chaos Soliton Fract 139:110017

Ayyoubzadeh SM, Ayyoubzadeh SM, Zahedi H, Ahmadi M, Kalhori SRN (2020) Predicting COVID-19 incidence through analysis of google trends data in iran: data mining and deep learning pilot study. JMIR Public Hlth Sur 6(2):e18828

Bhattacharya S, Maddikunta PKR, Pham Q-V, Gadekallu TR, Chowdhary CL, Alazab M, Piran MJ (2021) Deep learning and medical image processing for coronavirus (COVID-19) pandemic: a survey. Sustain Cities and Soc 65:102589

Brownlee J (2017) Long short-term memory networks with python: develop sequence prediction models with deep learning: Machine Learning Mastery

Brunese L, Mercaldo F, Reginelli A, Santone A (2020) Explainable deep learning for pulmonary disease and coronavirus COVID-19 detection from X-rays. Comput Meth Prog Bio 196:105608

Desai SB, Pareek A, Lungren MP (2020) Deep learning and its role in COVID-19 medical imaging. Intell-Based Med. 3:100013

Fan D, Sun H, Yao J, Zhang K, Yan X, Sun Z (2021) Well, production forecasting based on ARIMA-LSTM model considering manual operations. Energy 220:119708

Fischer T, Krauss C (2018) Deep learning with long short-term memory networks for financial market predictions. Eur J Operat Res 270(2):654–669

Fong SJ, Li G, Dey N, Crespo RG, Herrera-Viedma E (2020) Composite Monte Carlo decision making under high uncertainty of novel coronavirus epidemic using hybridized deep learning and fuzzy rule induction. Appl Soft Comput 93:106282

Hua Y, Zhao Z, Li R, Chen X, Liu Z, Zhang H (2019) Deep learning with long short-term memory for time series prediction. IEEE Commun Mag 57(6):114–119

Huang C et al (2020) Clinical features of patients infected with 2019 novel coronavirus in Wuhan. China Lancet 395(10223):497–506

Jin YH et al (2020) A rapid advice guideline for the diagnosis and treatment of 2019 novel coronavirus (2019-nCoV) infected pneumonia (standard version). Military Med Res 7(1):1–23

Kang H-W, Kang H-B (2017) Prediction of crime occurrence from multi-modal data using deep learning. PloS one 12(4):e0176244

Klopfenstein T, Kadiane-Oussou N, Toko L, Royer P-Y, Lepiller Q, Gendrin V, Zayet S (2020) Features of anosmia in COVID-19. Méd Maladies Infect 50(5):436–439

Mandal M, Jana S, Nandi SK, Khatua A, Adak S, Kar T (2020) A model based study on the dynamics of COVID-19: prediction and control. Chaos Soliton Fract 136:109889

Oakden-Rayner L, Carneiro G, Bessen T, Nascimento JC, Bradley AP, Palmer LJ (2017) Precision radiology: predicting longevity using feature engineering and deep learning methods in a radiomics framework. Sci Rep 7(1):1–13

Panwar H, Gupta P, Siddiqui MK, Morales-Menendez R, Singh V (2020) Application of deep learning for fast detection of COVID-19 in X-Rays using nCOVnet. Chaos Solitons Fractals 138:109944

Qin Y et al (2019) Hybrid forecasting model based on long short term memory network and deep learning neural network for wind signal. App Energy 236:262–272

Rothe C et al (2020) Transmission of 2019-nCoV infection from an asymptomatic contact in Germany. New Engl J Med 382(10):970–971

Sanchez-Caballero S, Selles MA, Peydro MA, Perez-Bernabeu E (2020) An efficient COVID-19 prediction model validated with the cases of China, Italy and Spain: Total or partial lockdowns? J Clin Med 9(5):1547

Santosh K (2020) COVID-19 prediction models and unexploited data. J Med Sys 44(9):1–4

Shorten C, Khoshgoftaar TM, Furht B (2021) Deep Learning applications for COVID-19. J Big Data 8(1):1–54

Tang B, Wang X, Li Q, Bragazzi NL, Tang S, Xiao Y, Wu J (2020) Estimation of the transmission risk of the 2019-nCoV and its implication for public health interventions. J Clin Med 9(2):462

Tsiouris KM, Pezoulas VC, Zervakis M, Konitsiotis S, Koutsouris DD, Fotiadis DI (2018) A long short-term memory deep learning network for the prediction of epileptic seizures using EEG signals. Comp Bio Med 99:24–37

Wang M, Cao R, Zhang L, Yang X, Liu J, Xu M, Xiao G (2020) Remdesivir and chloroquine effectively inhibit the recently emerged novel coronavirus (2019-nCoV) in vitro. Cell Res 30(3):269–271

Zeroual A, Harrou F, Dairi A, Sun Y (2020) Deep learning methods for forecasting COVID-19 time-Series data: a Comparative study. Chaos Soliton Fract 140:110121

Zhang W, Du RH, Li B, Zheng X-S, Yang X-L, Hu B, Shi Z-L (2020) Molecular and serological investigation of 2019-nCoV infected patients: implication of multiple shedding routes. Emerg Microbes Infec 9(1):386–389

Zhao J, Deng F, Cai Y, Chen J (2019) Long short-term memory-fully connected (LSTM-FC) neural network for PM2. 5 concentration prediction. Chemosphere 220:486–492

Zhao L 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 Infec Di 92:214–217

Authors and Affiliations

Weiwei Wang[1] · Jinghui Cai[2] · Jiali Xu[3] · Yuxiang Wang[4] · Yulin Zou[5]

✉ Yulin Zou
  zouyulin1213@163.com

  Weiwei Wang
  0619036@zju.edu.cn

  Jinghui Cai
  legolascai@163.com

  Jiali Xu
  zhenlanda@126.com

  Yuxiang Wang
  ywang594@jh.edu

[1] College of Media and International Culture, Zhejiang University, Hangzhou 310058, China

[2] College of Information Science and Technology, Jinan University, Guangzhou 510000, China

[3] School of Mathematics, Shanghai University of Finance and Economics, Shanghai 200433, China

[4] Whiting School of Engineering, Johns Hopkins University, Baltimore 21218, USA

[5] The Third Clinical Medical College of China Three Gorges University, Gezhouba Central Hospital of Sinopharm, Yichang 443002, China