Prediction of COVID-19 epidemic situation via fine-tuned IndRNN

Zhonghua Hong [Corresp., 1, 2] , **Ziyang Fan** [1] , **Xiaohua Tong** [Corresp., 2] , **Ruyan Zhou** [1] , **Haiyan Pan** [1] , **Yun Zhang** [1] , **Yanling Han** [1] , **Jing Wang** [1] , **Shuhu Yang** [1] , **Hong Wu** [1] , **Jiahao Li** [1]

[1] College of Information Technology, Shanghai Ocean University, Shanghai, China

[2] College of Surveying and Geo-Informatics, Tongji University, Shanghai, China

Corresponding Authors: Zhonghua Hong, Xiaohua Tong
Email address: zhhong@shou.edu.cn, xhtong@tongji.edu.cn

The COVID-19 pandemic is the most serious catastrophe since the Second World War. To predict the epidemic more accurately under the influence of policies, a framework based on Independently Recurrent Neural Network (IndRNN) with fine-tuning are proposed for predict the epidemic development trend of confirmed cases and deaths in the United Stated, India, Brazil, France, Russia, China, and the world to late May, 2021. The proposed framework consists of four main steps: data pre-processing, model pre-training and weight saving, the weight fine-tuning, trend predicting and validating. It is concluded that the proposed framework based on IndRNN and fine-tuning with high speed and low complexity, has great fitting and prediction performance. The applied fine-tuning strategy can effectively reduce the error by up to 20.94% and time cost. For most of the countries, the MAPEs of fine-tuned IndRNN model were less than 1.2%, the minimum MAPE and RMSE were 0.05%, and 1.17, respectively, by using Chinese deaths, during the testing phase. According with the prediction and validation results, the MAPEs of the proposed framework were less than 6.2% in most cases, and it generated lowest MAPE and RMSE values of 0.05% and 2.14, respectively, for deaths in China. Moreover, Policies that play an important role in the development of COVID-19 have been summarized. Timely and appropriate measures can greatly reduce the spread of COVID-19; untimely and inappropriate government policies, lax regulations, and insufficient public cooperation are the reasons for the aggravation of the epidemic situations. The code is available at *https://github.com/zhhongsh/COVID19-Precdiction.* And the prediction by IndRNN model with fine-tuning are now available online (*http://47.117.160.245:8088/IndRNNPredict*)

1 # Prediction of COVID-19 epidemic situation via fine-
2 tuned IndRNN

5 Zhonghua Hong[1,2,*], Ziyang Fan[1], Xiaohua Tong[2,*], Ruyan Zhou[1], Haiyan Pan[1], Yun Zhang[1],
6 Yanling Han[1], Jing Wang[1], Shuhu Yang[1], Hong Wu[1], Jiahao Li[1]

8 [1] College of Information Technology, Shanghai Ocean University, 999 Huchenghuan Road,
9 Shanghai, 201306, China
10 [2] College of Surveying and Geo-Informatics, Tongji University, 1239 Siping Road, Shanghai,
11 200092, China

13 Corresponding Author:
14 Zhonghua Hong and Xiaohua Tong

16 ## Abstract
17   The COVID-19 pandemic is the most serious catastrophe since the Second World War. To
18 predict the epidemic more accurately under the influence of policies, a framework based on
19 Independently Recurrent Neural Network (IndRNN) with fine-tuning are proposed for predict
20 the epidemic development trend of confirmed cases and deaths in the United Stated, India,
21 Brazil, France, Russia, China, and the world to late May, 2021. The proposed framework
22 consists of four main steps: data pre-processing, model pre-training and weight saving, the
23 weight fine-tuning, trend predicting and validating. It is concluded that the proposed framework
24 based on IndRNN and fine-tuning with high speed and low complexity, has great fitting and
25 prediction performance. The applied fine-tuning strategy can effectively reduce the error by up to
26 20.94% and time cost. For most of the countries, the MAPEs of fine-tuned IndRNN model were
27 less than 1.2%, the minimum MAPE and RMSE were 0.05%, and 1.17, respectively, by using
28 Chinese deaths, during the testing phase. According with the prediction and validation results,
29 the MAPEs of the proposed framework were less than 6.2% in most cases, and it generated
30 lowest MAPE and RMSE values of 0.05% and 2.14, respectively, for deaths in China. Moreover,
31 policies that play an important role in the development of COVID-19 have been summarized.
32 Timely and appropriate measures can greatly reduce the spread of COVID-19; untimely and
33 inappropriate government policies, lax regulations, and insufficient public cooperation are the
34 reasons for the aggravation of the epidemic situations. The code is available at
35 *https://github.com/zhhongsh/COVID19-Precdiction.* And the prediction by IndRNN model with
36 fine-tuning are now available online (*http://47.117.160.245:8088/IndRNNPredict*)

38 ## Introduction

39   The coronavirus disease 2019 (COVID-19) is brought on by infection from Severe Acute
40   Respiratory Syndrome (SARS) Coronavirus 2, and the reports of related cases were first released
41   by Wuhan, Hubei Province, China in December 2019 (*Zhou et al., 2020*). The widespread
42   COVID-19 epidemic is a serious threat, and has become one of the most challenging global
43   catastrophes facing mankind since the Second World War (*Chinadaily, 2020*). On March 11th,
44   2020, the global development of COVID-19 was assessed by the World Health Organization
45   (WHO) as having met the characteristics of a pandemic (*World Health Organization, 2020a*).
46   The mortality rate of COVID-19 is estimated to be between 2% and 5%, i.e., lower than that of
47   SARS and Middle East Respirator Syndrome (*Gasmi et al., 2020*; *Wu, Chen & Chan, 2020*).
48   However, COVID-19 has a higher infection rate than bat-like SARS, and its pathogenicity is
49   between that of SARS and bat-like SARS (*Benvenuto et al., 2020*). The heart may be damaged
50   and develop myocardial inflammation after recovery from COVID-19 (*Puntmann et al., 2020*).
51   In addition, the widespread of the disease also slow down the national and global, along with
52   aggravating unemployment and hunger (*Mckibbin & Fernando, 2020*; *Abcnews, 2020a*).
53   On January 15th, 2020, the Chinese Center for Disease Control and Prevention (China CDC)
54   initiated a first-level emergency response (*Li et al., 2020a*). A series of policies began to be
55   implemented since January 23th, 2020 (*Yang et al., 2020*), and the epidemic had stabilized by
56   March 2020. In contrast, the epidemic situation overseas are not very optimistic. The number of
57   infected people continues to maintain substantial increase of tens of thousands. According to the
58   report by the WHO on April 25th, 2020, the global epidemic has infected more than 146 million
59   people, with 3,092,497 deaths, and the epidemic in the Americas, Europe, and South-East Asia
60   are the worst, especially in India and the United States (*World Health Organization, 2020b*).
61   Predicting the development trend based on the increase of cases is useful for the adjustment of
62   epidemic prevention policy. However, in the current epidemic prediction work, the models used
63   are complex and slow (*Yang et al., 2020*; *Bandyopadhyay & Dutta 2020*), and some methods are
64   fast but not effective (*Huang et al., 2020*). In addition, the number of cumulative increases is not
65   stable but variable, particularly, a sudden rapid or modest increase affects the stability of
66   modelling and therefore the accuracy of predicting future trends. Thus, it is necessary to propose
67   a model with coexistence of high precision, high speed, and low model complexity to predict and
68   analyze the development tendency of COVID-19 with more efficiency and accuracy, and to
69   summarise the meaningful and positive policies.
70   In this paper, a framework combining IndRNN model and fine-tuning strategy is proposed.
71   Different from the existing models in the prediction of COVID-19, the fine-tuning strategy is
72   added to reduce error and time. The proposed framework consists of four steps. First, the original
73   data are pre-processed, including normalization and sequence data generation. Then, The
74   IndRNN model is used to iteratively train to learn information or features hidden in COVID-19
75   sequence data, and weight parameters is obtained. Thirdly, the fine-tuning strategy is added to
76   load the recently updated data into the model with existing weight parameters for partial
77   parameter adjustment. Finally, the model that has learned the characteristics or information is
78   used to predict the trend, and the prediction results are compared with the real data for

79   verification, and the policy events and social economic influences under the epidemic situation
80   are analyzed. The contributions in this paper as follows:
81   1) A framework based on IndRNN and fine-tuning strategy, which can be adept at longer
82   sequence data, was used in COVID-19 epidemic prediction with more accuracy and high speed.
83   IndRNN is used to retain effective information and better learn the characteristics of changeable
84   sequence data in COVID-19 due to the advantage of independent neurons prevents gradient
85   explosion and disappearance phenomenon. The fine-tuning strategy is utilized to further
86   improves the accuracy in a short time, and it avoids the consumption brought by retraining after
87   data update.
88   2) The fitting performance of the proposed framework based on IndRNN and fine-tuning are
89   verified among the cumulative cases of India by comparing with the Long-Short-Term-Memory
90   (LSTM), bidirectional LSTM (Bi-LSTM), Gated-Recurrent-Unit (GRU), Stacked_Bi_GRU,
91   Convolutional Neural Network_LSTM (CNN_LSTM), Deep_CNN models. And the prediction
92   accuracy of the model is validated by the error between the prediction results and the real data of
93   the next four weeks in the cumulative case data of the United State, India, Brazil, France, Russia,
94   China, and world.
95   3) The growth of the cumulative population is combined to analyze the positivity of epidemic
96   policies and activities. The phenomenon is found that the same proactive policies have been
97   implemented inconsistently in different countries because of lax regulation and inadequate public
98   cooperation.
99   The rest of the paper is organized as follows: Section II describes the related work of this
100   research direction. Section III introduces the data sources and study area, the proposed
101   framework, and performance metrics; the experimental results are presented in Section Ⅳ. In
102   Section V, impact of fine-tuning on the prediction and validation accuracy, and the influence of
103   policies are discussed. Finally, some conclusions are drawn in Section Ⅵ.
104

Related work

106   There are two main methods for forecasting the epidemic development of COVID-19:
107   mathematical model and deep learning model.
108   A typical mathematical model in epidemic dynamics is the Susceptible-Exposed-Infectious-
109   Removed (SEIR) model, using mathematical formulas to reflect the relationships between the
110   flows of people at four states: susceptible, exposed, infectious, and recovered (*Fang, Nie &*
111   *Penny, 2020*). The SEIR model was used to effectively predict the peaks and sizes of COVID-19
112   epidemiological data with sufficient fitting performance (*Yang et al., 2020*; *Fang, Nie & Penny,*
113   *2020*). A modified SEIR model also showed a good effect for predicting the peaks and sizes
114   (*Yang et al., 2020*). The peak deviation of the another modified SEIR model in predicting
115   epidemiological data in China was 3.02% (*Fang, Nie & Penny, 2020*). However, SEIR focuses
116   on predicting trends in sensitive, exposed, infected, and recovered groups (*Yang et al., 2020*;
117   *Fang, Nie & Penny, 2020*), rather than cumulative confirmed and death cases. Moreover, it is
118   necessary to comprehensively consider the changes in some parameter values as affected by the
119   changes in epidemic policies and regional differences, such as the effective reproductive number,

120   number of contacts, and infection rate. And these parameters are not easy to obtain, and are
121   uncertain.
122   Deep learning automatically extracts the features from the data and builds the model without
123   the need for other specific parameters. This method generates a series of sequence data from
124   COVID-19 epidemiological data and looks for regular changes in the sequence data. Some deep
125   learning models based on LSTM was used to predict COVID-19 trend, and the results
126   demonstrated that LSTM has good prospects for predicting the trend; however, the fitting effect
127   needs to be improved (*Yang et al., 2020*; *Zandavi, Rashidi & Vafaee, 2020)*, too many internal
128   parameters of LSTM increase the complexity of model training, and it can only learn forward.
129   The Bi-LSTM, which adds bidirectional learning capability on LSTM, and the GRU, which
130   makes gates simplification on the internal structure of LSTM, are used by *Shahid, Zameer &*
131   *Muneeb (2020)* for epidemic prediction. Moreover, a stacked-Bi-GRU model was applied for
132   COVID-19 trend forecast, owing to the learning adequacy of the bidirectional cyclic network
133   *(Bandyopadhyay & Dutta 2020)*. LSTM and GRU solve the problem of gradient vanishing
134   explosion of RNN, but there is a gradient attenuation problem of layer. In addition, the
135   entanglement of neurons in the same layer makes their behavior difficult to interpret.
136   The existence of these multiple gate operations in the recurrent unit of the RNN variants,
137   which is LSTM, Bi-LSTM, GRU, makes the parameters complicated. CNN does not need many
138   parameters in virtue of the advantage of weight sharing, in comparison with these recurrent
139   networks. A deep CNN model (*Huang et al., 2020*) and a CNN-LSTM model (*Dutta,*
140   *Bandyopadhyay & Kim, 2020*) were proposed for analysing and predicting the number of
141   confirmed cases in China. However, the training speed of deep CNN is fast, but the effect is not
142   significantly improved. And CNN_LSTM model combines CNN and LSTM, but increases the
143   model complexity.
144   In the above deep learning network model, high precision, high speed, and low model
145   complexity cannot coexist. And the cumulative epidemic data is time series, so recurrent
146   networks are more suitable for processing such serialized data in comparing with CNN. Similar
147   to LSTM, GRU, Bi-LSTM, IndRNN is one of the variants of RNN, which can learn longer
148   sequence data, and it has no redundant gate operations, fewer parameters, and can be more easily
149   trained (*Li et al., 2018*). Simultaneously, IndRNN is designed to solve the problems of gradient
150   disappearance and explosion, and can be utilized to process sequence data like LSTM, GRU and
151   Bi-LSTM. IndRNN was used to learn relationships between plant gene sequences, overcoming
152   the uncertainty of artificially acquired traits, and has higher accuracy than LSTM (*Zhang et al.,*
153   *2020*). Therefore, Indrnn is adopted as the basic model in this paper to improve the accuracy of
154   epidemic prediction.
155   However, the epidemic data are updated daily, and the network weights need to be retrained
156   after obtaining the data of the last few weeks, which are time-consuming. Fine-tuning can
157   transfer the trained network model parameters to the required network for partial parameter
158   adjustments, without the need to train from scratch. *Tajbakhsh et al. (2016)* used a fine-tuned
159   pre-trained CNN network for medical image analysis and found that the effect was better than

160   training all data from scratch CNN. *Boyd, Czajka & Bowyer (2020)* demonstrated that fine-
161   tuning existing network weights to extract iris features was more accurate. Therefore, a fine-
162   tuning strategy is added in this study to further improve the accuracy and speed of prediction.
163   Combined with the above, in order to achieve rapid and more accurate epidemic prediction, a
164   framework based on IndRNN and fine-tuning to predict COVID-19 epidemiological data is
165   proposed in this paper.
166

Materials & Methods

Data sources and study area

169   The study area of this research is shown in Fig. 1. Totally, six countries are involved, which are
170   China, the United States, India, Brazil, France and Russia. China is the country where the first
171   COVID-19 has been reported in the world. The epidemiological data of COVID-19 in China
172   between January 18th, 2020 and May 22th, 2021 were downloaded from
173   *http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml* (*National Health Commission of the People's*
174   *Republic of China, 2020)*. The United States, India, Brazil, France and Russia are the most
175   severely affected countries in the world during this period. The data of these top five countries
176   were download from a dataset based on global COVID-19 epidemic statistics, as published by
177   *https://github.com/CSSEGISandData/COVID-19* (*Johns Hopkins University in the United States,*
178   *2020)*. In addition, the cumulative overseas epidemic data also analyzed, the data were retrieved
179   from *https://voice.baidu.com/act/newpneumonia/newpneumonia/?from=osari_pc_3* (*Baidu,*
180   *2020*). And the Chinese and overseas COVID-19 data were combined to generate the numbers of
181   cumulative confirmed cases and cumulative deaths worldwide. It should be noted that all of the
182   three websites share the same definition of cumulative confirmed and death cases, which ensures
183   the consistency of the data. The detailed information of the data used in this study is presented in
184   Table 1.
185

Methods

187   The flowchart of the proposed framework is illustrated in Figure 2. It mainly consists of four
188   steps as follows: (1) the cumulative confirmed and death cases obtained from various websites
189   are first preprocessed to generate sequence data of size in a uniform range, and the sequence data
190   are divided into training data, fine-tuning data, and testing data; (2) the deep learning model
191   trains and learns the features using the training data, then the weights are obtained. The training
192   process is the process of establishing the optimal relationship between the training data and the
193   training labels, in which the training data is the input and the training label is the perfect output.
194   Through continuous iterative training, the characteristics of training data are found, and the
195   parameters in the network model are adjusted to continuously fit the actual output and training
196   labels; (3) the fine-tuning data is loaded with the weight file, and the parameters are adjusted by
197   fine-tuning model, continuously reducing the error between the fine-tuning data and the fine-
198   tuning labels; (4) though applying the model built from the previous training and fine-tuning, the
199   testing data are used for testing, later trends are predicted, and validated using true data. The

200   testing data are input for testing by applying the model with weights, to determine the gap
201   between the actual output of test data and the test labels. After that, the model is used for direct
202   trend prediction, and the prediction results are compared with the real updated data to judge the
203   prediction performance of the proposed framework.
204   Specially, seven deep learning models which are LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU,
205   CNN_LSTM, Deep_CNN, and IndRNN are used in this research for model comparison. As
206   shown in Figure 3, the structures of the LSTM, Bi-LSTM, GRU, and IndRNN models are
207   approximately the same. The flatten layer in LSTM, Bi-LSTM and GRU models is used to
208   convert three-dimensional data into two-dimensional data. The IndRNN model does not add the
209   Flatten layer because the output of the IndRNN layer is 2-dimensional data. The details of the
210   proposed framework are described below.
211

212 **Pre-processing**
213 Pre-processing includes MinMaxScaler operation and sequence data generation. To keep the data
214 at the same order of magnitude and facilitate characteristic analysis and model convergence, the
215 cumulative confirmed cases and deaths data are scaled to 0~1 after MinMaxScaler operation
216 based on the minimum and maximum values, and the original proportion of the data is retained.
217   Sequence data generation is a crucial step, and it is the premise of sequence model training.
218 The sequence data can reflect changes in which no regularity can be traced in single data.
219 Therefore, the individual data were organised into sequential data. As shown in Figure 4, the
220 dimensionality of the individual data for n weeks was $(n,1)$, a window with a sequence length of
221 x and step size of 1 was initially selected for sliding orderly through the weekly data. All of the
222 data contained in the window was considered as a data sequence, and the next week data outside
223 the window was the corresponding label data. After the window slid through all of the weekly
224 data, the sequence data with dimension of $(n - x, x)$ was generated. It is noteworthy that the
225 content of sequence data represents features. For example, when the sequence length is 3, the
226 data from the first to the third form a sequence, and the data from the fourth day is the label data;
227 next, the data of days 2 to 4 form a sequence, and the data of day 5 is label data. The sequence
228 data generation is completed until the label data reaches the last data of the input data.
229

230 **LSTM and Bi-LSTM**
231 In RNN, too many cyclic units may lead to the loss of previously learned rule information and
232 the existence of long-term dependence, and can lead to the problems such gradient disappearance
233 or gradient explosion (*Goodfellow, Bengio & Courville, 2016*). LSTM is mainly used to learn
234 and overcome the long-term dependence of RNN (*Hochreiter & Schmidhuber, 1997*). The
235 internal units of the LSTM include the memory cell, forget gate, input gate, and output gate
236 (*Goodfellow, Bengio & Courville, 2016*). In the LSTM internal unit, the memory cell carries the
237 necessary information transfers between the LSTM internal circulation units, thereby solving the
238 problem of gradient disappearance, and learning long-term dependence (*Olah, 2015*). The
239 sigmoid layers in the three gates constrain the value between 0 and 1, so as to determine which

240 information should be saved or forgotten (*Goodfellow, Bengio & Courville, 2016*; *Olah, 2015*).
241 When the information flow enters the circulation unit, the operation is as follows (*Olah, 2015*):

$$\text{forget}_t = \sigma(W_{\text{forget}} \cdot [h_{t-1}, x_t] + b_{\text{forget}}) \tag{1}$$

$$input_t = \sigma(W_{input} \cdot [h_{t-1}, x_t] + b_{input}) \tag{2}$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C) \tag{3}$$

$$C_t = forget_t * C_{t-1} + input_t * \tilde{C}_t \tag{4}$$

$$output_t = \sigma(W_{output} \cdot [h_{t-1}, x_t] + b_{output}) \tag{5}$$

$$h_t = output_t * \tanh(C_t) \tag{6}$$

248 In the above, t stands for the moment, $forget_t$, $input_t$, and $output_t$ represent the output after the
249 activation function σ in the forget gate, input gate and output gate, respectively. W_{forget}, W_{input},
250 W_{output}, W_C, b_{forget}, b_{input}, b_{output}, b_C symbolize the weight and bias offset of the three gates
251 and memory cell, respectively. h_{t-1} is the output at the previous moment, x_t is the input at the
252 current moment. The updated memory cell C_t is obtained by adding $forget_t * C_{t-1}$ which
253 represents unnecessary information to be discarded, and $input_t * \tilde{C}_t$ which represents the
254 information to be updated. Finally, the information in the memory unit C_t is controlled by the
255 output gate to return the final output h_t for this LSTM cell.
256   LSTM can only predict the output based on the previous content, but the later information will
257 also help understand the current text. Therefore, the Bi-LSTM proposed by *Schuster & Paliwal*
258 *(1997)* performs bidirectional input and makes full use of the context information. The forward
259 input is the sequence input at time t and the output at time $t - 1$, and the backward input is the
260 sequence input at time t and the output at time t+1. The final output is a combination of forward
261 output $h_{forword} = (hf_0, hf_1, ..., hf_{n-1})$ and reverse output $h_{backword} = (hb_{n-1}, hb_{n-2}, ..., hb_0)$.
262
263 **GRU**
264 The difference between GRU and LSTM is that the GRU changes the forgetting gate and input
265 gate in LSTM to update gate, and reduces the amount of gate controls (*Chung et al., 2014*). The
266 GRU transmits information directly through the hidden state, rather than through the memory
267 cell. The input information of the current moment and the output information of the previous
268 moment are firstly determined by the update gate whether to be updated. The reset gate is then
269 used to control whether the information is set to 0, i.e., to determine the amount of information to
270 be retained in the candidate information. Finally, it is up to update gate to control how much
271 information in the output at the previous time is forgotten and how much information is added to
272 the hidden information, and the retained information forms the output of the GRU recurrent unit
273 at the current moment. The main formulas of GRU internal structure are as follows (*Chung et al.,*
274 *2014*):

$$update_t = \sigma(W_{update} \cdot [h_{t-1}, x_t]) \tag{7}$$

$$reset_t = \sigma(W_{reset} \cdot [h_{t-1}, x_t]) \tag{8}$$

$$\tilde{h}_t = \tanh(W \cdot [reset_t * h_{t-1}, x_t]) \tag{9}$$

278 $$h_t = (1 - update_t) * h_{t-1} + update_t * \tilde{h}_t \tag{10}$$

279

280 Where, $update_t$, $reset_t$ represent the output after the activation function σ in the update gate

281 and reset gate, respectively. $\tilde{h}_t$ is candidate information, and h_t represents the output at the

282 current moment t.

283

284 **IndRNN**

285 IndRNN solves the problem of neuron independence in RNN and the gradient attenuation that

286 appears in LSTM and GRU, and can handle longer sequences. In addition, gate computation is

287 added to LSTM and GRU, which increases the computational complexity of the network layer.

288 In IndRNN, there are few parameters, and the neuron only receives input from this moment and

289 its own state input from the previous moment at any given moment, making each neuron

290 independent (*Li et al., 2018*). The state update formula of the hidden layer is (*Li et al., 2018*):

291 $$h_t = \sigma(Wx_t + u \odot h_{t-1} + b) \tag{11}$$

292 Here, W and u are the weights of a certain neuron at the current and previous moment,

293 respectively.

294

295 **Fine-tuning**

296 Fine-tuning refers to copying the weight of the trained network model to the network that needs

297 to be used, and continuing to train and adjust part of the weight (*Tajbakhsh et al. 2016*). The

298 advantage of fine-tuning is that it can achieve better results in a shorter time than training the

299 network from scratch, owing to pre-training using a large amount of data (*Tajbakhsh et al. 2016*).

300 The fine-tuning model in this paper is consistent with the pre-training model, as shown in Figure

301 3. Firstly, the fine-tuning model copies and transfers the weight of the pre-training model to itself.

302 Then, the parameters of the net layer before the dense layer in the Stacked_Bi_GRU,

303 CNN_LSTM, Deep_CNN models, and the LSTM layer, Bi-LSTM layer, GRU layer, IndRNN

304 layer, are frozen, respectively, and do not participate in the following training. The parameters of

305 the dense layer and the activation layer of each fine-tuning model are in an active state, waiting

306 for adjustment. Finally, import fine-tuning data to start fine-tuning training. The error between

307 feature sequence data and label data is gradually narrowed by iterative training with the full

308 utilization of the features or knowledge acquired by the pre-training model.

309

310 **Assessment metrics**

311 To evaluate the effectiveness of the model, the RMSE and MAPE were used to access the fitting

312 performance between the output (prediction data) and label data of each sequence data. The

313 quations are as follows:

314 $$RMSE(L,h) = \sqrt{\frac{1}{m}\sum_{i=1}^{m}(L_i - h_i)^2} \tag{12}$$

315 $$MAPE(L,h) = \frac{100}{m}\sum_{i=1}^{m}\left|\frac{L_i - h_i}{L_i}\right| \tag{13}$$

316 Here, m denotes the number of sequential data, and h_i and L_i represent the output of each

317 sequence i after testing in model and the corresponding label data, respectively. The smaller the

318 value of RMSE and MAPE, the better the prediction effect, which means that the error between

319 prediction data and label data is smaller.

320

321 **Results**

322 This paper aims to achieve prediction of epidemic trends with more accurate by proposed

323 framework based on IndRNN and fine-tuning strategy on COVID-19 epidemiology data. The

324 deadline for training data is March 13th, 2021, the data from 3/20/2021 to 4/24/2021 is used for

325 fine-tuning and testing, and we have taken cases from 1/22/2021 to 5/10/2021 for validation. The

326 deep learning environment of our experiment was mainly built based on the Ubuntu 16.04

327 system environment, which encoded by python Keras with the support of Intel Core i9-10920X

328 CPU @3.50GHz×24.

329 Three tasks are completed in this paper: 1) model comparison by LSTM, Bi-LSTM, GRU,

330 Stacked_Bi_GRU, CNN_LSTM, Deep_CNN, and IndRNN models, 2) the fine-tuned IndRNN

331 model is utilized to predict the number of confirmed cases in the United Sates, India, Brazil,

332 France, Russia, China and the world, and verify their accuracy, and 3) the growth status of the

333 cumulative cases in 6 countries in combination with current policies are analyzed. All results in

334 this section use the values in Table 2. The experimental procedure for each model follows the

335 steps in the method. Moreover, during fine-tuning, the numbers of layers in the seven models

336 that do not participate in training are set as shown in Table 3.

337

338 **Comparison of models**

339 The performance of seven different models which are LSTM, Bi-LSTM, GRU,

340 Stacked_Bi_GRU, CNN_LSTM, Deep_CNN, and IndRNN models are compared based on the

341 COVID-19 statistical data of India. The reason for this selection lays on the fact that India is the

342 worst-affected country under the current situation, the weekly increase of cases is dramatically

343 higher than other countries, with vast changing data. Before adding the updated data, the six

344 countries and the world used in this paper have at least 55 weeks of data. To find the most

345 suitable sequence length, and to ensure that the number of training data is more than fine-tuning

346 data, Indian dataset was used for training by LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU,

347 CNN_LSTM, Deep_CNN, and IndRNN models with sequence length between 5 and 45 weeks

348 (separated by 4 weeks, i.e. one month) as single data cannot become sequence data. The best

349 results were obtained when the sequence length was 45.

350 After normalization and sequence data generation, the dimensions of the sequence data

351 generated by using weekly cumulative confirmed and death cases in India are (20, 1, 45) and (14,

352 1, 45), and the corresponding label data are 20 and 14 respectively. The last 6 of them are used

353 for fine-tuning and testing, and the other data are used for training.

354 These 7 models use Indian COVID-19 data to save the weights after pre-training, respectively.

355 Under the condition of freezing the network layer before the fully connected layer, the pre-

356 trained models load the corresponding weights for iterative fine-tuning by the fine-tuning data. In
357 this process, to find the appropriate amount of fine-tuning data, 1, 3, and 5 pieces of fine-tuning
358 data are used for fine-tuning, and the number of corresponding testing data are 5, 3, and 1
359 respectively.
360 Figure 5, Table 4, and Table 5 show the comparison results. Where, in Table 4, "f" and "t" in
361 the column "Split" represent the number of fine-tuning data and testing data, respectively. As can
362 be seen from the table and the figure, IndRNN model shows the best performance for both with
363 no-fine-tuning and fine-tuning models, compared to LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU,
364 CNN_LSTM, and Deep_CNN models. Where, IndRNN model had the least MAPE and RMSE
365 before and after fine-tuning, the lowest total runtime when using cumulative diagnosis data, the
366 second lowest total runtime when using cumulative death data, and minimum number of total
367 parameters. The MAPE and RMSE of LSTM, BI-LLSTM and GRU are similar, and the effect
368 ranking of the remaining models is Deep_CNN, Stacked_Bi_GRU, CNN_LSTM, respectively.
369 After fine-tuning, the RMSE and MAPE of the seven models are all significantly decreased,
370 especially for IndRNN, the MAPE of which is decreased by 0.27%~12.36%, 0.12%~14.29%,
371 0.3%~10.91%, 0.69%~8.89%, 2.21%~11.84%, -1.59%~13.79%, and 0.36%~20.94%,
372 respectively. The best test results were obtained when five fine-tuning data were used to fine-
373 tune the parameters of the pre-trained IndRNN model. It is proved that fine-tuning plays a very
374 important role in reducing errors. In addition, IndRNN also exhibits superiority in computational
375 efficiency. That is also true for that with fine-tuning. Figure 6 shows the fitting performance of
376 the IndRNN model. The predicted value is very close to the true value and the two lines are
377 almost overlapped, which demonstrates the effectiveness of the fine-tuned IndRNN model in
378 predicting the development of COVID-19 cases.
379
380 **Predictive performance analysis of the COVID-19 epidemic situations in six countries**
381 In this section, the proposed fine-tuned IndRNN model is used to predict the COVID-19
382 epidemic trend in the cumulative confirmed and death cases of the China, the United States,
383 India, Brazil, France and Russia with 5 fine-tuning data. The prediction period is one month,
384 according to the fact that the impact of policies or events generally exhibits in the future month.
385 And then the actual data is utilized to validate the predict result. The development tendency of
386 the COVID-19 is analyzed, the significant phenomenon is then detected and the reasons that
387 responsible for this is further explored.
388 Figure 7 shows the development tendency and prediction result of COVID-19 cumulative
389 confirmed and death cases in China, and the general development stages of the epidemic in
390 China are summarized in the Table S1. As can be seen from the figure, the cumulative confirmed
391 cases in China experienced dramatic changes. It firstly starts to rise sharply in mid-January 2020,
392 while the number of deaths continues to rise, and the epidemic has entered the outbreak stage
393 (*The State Council Information Office of the People's Republic of China, 2020*). During this
394 period, the Chinese government issued a timely emergency response, e.g., expanding the
395 laboratories capacities for nucleic acid detection, the construction of 'Huoshenshan',

396 'Leishenshan', and Fangcang shelter hospitals for accommodating more patients, medical
397 observation was conducted on close contacts, people remained at home as much as possible, and
398 wearing masks, avoiding gatherings, and keeping a physical distance of 1 meter when going out
399 were strictly required (*Li et al., 2020b*; *Chinese Center for Disease Control and Prevention,*
400 *2020a*). The China CDC issued prevention guidelines for people of different ages and places to
401 strengthen the awareness of safety (*Chinese Center for Disease Control and Prevention, 2020b*).
402 As a result, the epidemic has been promptly controlled, and the cumulative confirmed cases in
403 China has gradually flattened and stabilized by late April 2020. The COVID-19 deaths in China
404 have been basically flat since the end of April 2020, and have remained at 4,636 since January
405 25th, 2021. It is demonstrated that the epidemic in China has been adequately controlled since the
406 late April 2020.
407   Figure 8 illustrates the development trend of cumulative confirmed and death cases in the
408 United States, India, Brazil, France, Russia. Different from that of China, the growth of
409 cumulative cases of the top five countries is relatively slow before July 2020. After that, it began
410 to increase sharply, especially for the United States and India. The rudimentary development of
411 the COVID-19 epidemic in the United States is summarized in the Table S2. After the first
412 COVID-19 patient in the United States was recorded on January 20th, 2020 (*Holshue et al.,*
413 *2020*), the following measures were taken: non-Americans who had visited China in the past 14
414 days were banned from entering on February 2th, 2020 (*National Immigration Administration,*
415 *2020*), cruise ship ban, and masks were promoted by the CDC (*Centers for Disease Control and*
416 *Prevention, 2020*). Nevertheless, there were still a campaign of mass protests and campaign
417 rallies without protective measures (*Schuchat, 2020*). As a result, the account of confirmed cases
418 in the United States continued to increase significantly after surpassing China on March 26th,
419 2020. In addition, the 'opening up' measures promoted by the American government
420 (*Goodmorningamerica, 2020a*), and the opening of schools and holding of the rally about the
421 presidential campaign since August 2020 (*USA TODEY, 2020*), promote the intensification of the
422 epidemic, hence the cumulative confirmed cases in the United Stated start to rise sharply in
423 October 2020.
424   The India has no obvious growth before July 2020, in which the 'city closure' measures
425 implemented by the Indian government at the end of March have played a key role (*Xinhuanet,*
426 *2020a*). But some areas have been re-blocked, and the number of people suffering from the
427 COVID-19 in India has increased extraordinary, because of an outbreak rebound caused by the
428 increase of outdoor activities after a gradual unsealing since June 2020 (*Xinhuanet, 2020a*;
429 *Abcnews, 2020b*). Many massive rallies had been held in India since March 2021, including
430 political rallies and festival (*Bhuyan, 2021*), and the number of cumulative confirmed cases has
431 surged.
432   Brazil, France, and Russia do not show a significant increase in general, but the number of
433 cases in Brazil is higher than that in France and Russia. Brazil implemented a social distancing
434 policy, but some people neglected to wear masks at rallies (*The paper, 2020*), and professional
435 football matches were held in Rio, Brazil on June 18th, 2020 (*Xinhuanet, 2020b*), these events

436 promoted the development of the epidemic. France implemented a lockdown on March 17th,
437 2020, and gradually unblocked it from May 11th, 2020. However, the amount of cumulative
438 diagnoses in France starts to rise in October 2020, by reason of the emergence of cases in school
439 and surges in cases in many areas, hence the French government closed cities for a second time
440 on October 30th, 2020 (*Gouvernement.fr., 2020*; *CNN, 2020*); The Russian government did not
441 take timely measures to prevent the European epidemic, causing a large number of imported
442 cases (*Kankanews, 2020*); moreover, the holding of a military parade on June 24th, 2020
443 (*Xinhuanet, 2020c*), has become one of the factors driving the curve of Russia upward.
444   According to Figures 7 and 8, the trend forecasts in six countries are basically in line with the
445 curve trend. Table 6 is the testing and validation accuracy of cumulative confirmed and death
446 COVID-19 cases in five countries by IndRNN model with fine-tuning. As indicated from the
447 table, in most cases, the MAPE of testing and validation are less than 1.2%, and 6.2%,
448 respectively. In the table, it generates lowest MAPE and RMSE values of 0.04% and 2.00 in
449 testing results, and 0.05% and 1.17 in verification results, respectively, for deaths in China,
450 which shows the effectiveness of fitting and prediction by the proposed framework.
451
452 **Verification of prediction accuracy in global diagnosis**
453 The IndRNN model was used to predict epidemic trends by training the cumulative confirmed
454 weekly cases for the global from February 15th, 2020 to April 24th, 2021. Figure 9 displays the
455 prediction diagram. By May 22th, 2021, the number of global cumulative confirmed cases may
456 surpass 171.4 million, and deaths may reach 3.5 million, and the deviations between these results
457 and the real data are 2.81%, and 2.30%, respectively.
458

Discussion

460 Among the seven models introduced in Section 2, the IndRNN model with fine-tuning strategy
461 achieved the optimal result. In different models, fine-tuning can continuously adjust the
462 parameters in a relatively short time. Compared with the result without fine-tuning, the error is
463 decreased by up to 20.94%. It is confirmed that the fine-tuning takes full advantage of the
464 characteristics obtained from the pre-trained model, and this strategy plays a good role in
465 increasing accuracy.
466   According to the experimental results of LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU,
467 CNN_LSTM, Deep_CNN, and IndRNN models, the IndRNN model takes the second least time,
468 but its RMSE and MAPE are the smallest. Among the seven networks, the errors of the stacked
469 networks are increased compared with the simple networks, and the phenomenon of overfitting
470 occurs, which is because the stacked networks over-interpret information, resulting in the weak
471 generalization ability of the models; compared with CNN, the variant network (i.e. LSTM, Bi-
472 LSTM, GRU, IndRNN) of RNN has higher accuracy, due to the memory function of the internal
473 structure of RNN series network, which is able to extract features from sequences of different
474 moments. This ability to process time series is more consistent with the characteristics of
475 COVID-19 epidemiological data changing over time. Among the variant networks of RNN, the

476 accuracy of LSTM, Bi-LSTM and GRU is not much different, which is inconsistent with the
477 result that Bi-LSTM is better than LSTM and GRU found by *Shahid, Zameer & Muneeb (2020)*.
478 The reason is that the Bi-LSTM has insufficient learning ability on cumulative data of India with
479 variable growth rate. By contrast, IndRNN has the best performance as the non-interference of
480 independent neurons with each other, it is more adaptable to the changeable epidemic data, and
481 is capable of better extract and transmit information hidden in the COVID-19 sequence data.
482 Therefore, it can be considered that the IndRNN model can understand the regular pattern of
483 epidemic data, has good learning capabilities, and achieves the coexistence of high precision and
484 low time consumption.
485   The MAPE of the testing by IndRNN model with fine-tuning is less than 1.2%, except for 5.89%
486 of the cumulative confirmed cases in France and 2.99% of the cumulative deaths in India, and
487 the MAPE of the validation by this framework is less than 6.2%, except for 11.53% of the
488 cumulative confirmed cases in France and 13.92% of the cumulative deaths in India. This is due
489 to the sudden increase in the cases and the short duration of the growth trend, which affects the
490 ability of model to obtain stable features and predict trend.
491   The more fine-tuning data is used, and the more error is reduced after fine-tuning, owing to
492 enough fine-tuning data can make full use of the features extracted by the pre-trained model.
493 Concretely, the reduced errors of LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU, CNN_LSTM,
494 Deep_CNN, and IndRNN models after fine-tuning are 0.27%~12.36%, 0.12%~14.29%,
495 0.3%~10.91%, 0.69%~8.89%, 2.21%~11.84%, -1.59%~13.79%, 0.36%~20.94%, respectively.
496 About Deep_CNN model, there is a slight increase in the error after fine-tuning when the number
497 of fine-tuning data is 1, and one possible reason is that too few fine-tuning parameters and data
498 affect the effect of parameter adjustment. The degree of error reduction is affected by the ability
499 of the model to learn features and the fine-tuned model to learn features from the pre-trained
500 model, and the different changes in different data, as well as the amount of fine-tuning data.
501 Among them, the fine-tuning model using 5 fine-tuning data makes full use of features obtained
502 from pre-trained IndRNN model, and get best result.
503   The model speed is proportional to the number of parameters, but not to the fitting effect. The
504 parameters of IndRNN model are the least by reason of no gate calculations in internal units in
505 compare with LSTM, GRU, Bi-LSTM units. However, as the three gates and memory cell in the
506 recurrent unit of the LSTM model are more than the GRU model with two gates, and Bi-LSTM
507 model has bidirectional LSTM layer in the internal structure, hence, LSTM model has more
508 parameters than GRU model, and the parameters of Bi-LSTM model are about twice that of the
509 LSTM model. Stacked_Bi_GRU and CNN_LSTM have a large number of parameters due to the
510 stack of multiple network layers, which affects the running speed. Among the seven networks,
511 Deep_CNN is the model with the fewest fine-tuning parameters because of a small number of
512 dense layer parameters caused by characteristic of shared parameters, hence the fine-tuning
513 speed is the fastest. However, the test RMSE of this model is up to 9.22% higher than that of
514 IndRNN, which may be because the reduction of parameters affects the effect of feature
515 extraction of sequence data.

516   Through the policies mentioned above, we divide these policies into active measures and
517 inactive measures according to whether they contribute to the mitigation of the epidemic (Table
518 7). We conclude that, on the final results of epidemic prevention and control, the timely release
519 of prevention and control policies by the government is the most important. Local
520 implementation and cooperation of residents are also crucial. The COVID-19 epidemic of China
521 gradually declines to a stable state in April 2020 as good performance in these respects. Some
522 countries have adopted the same policy as China but the effect is not obvious, e.g. the United
523 States, India, owing to the deficiency of implementation of some regions and collaboration of
524 residents, such as no mask, organizing of the rally. The short implementation time of blockade
525 measures in some countries promoted the recurrence of the epidemic, for instance, India. Policy
526 relaxation and economic restarts should be conducted after the epidemic situation is relatively
527 stable, rather than prematurely. Large-scale activities with crowd gatherings should be cancelled
528 as much as possible, owing to the existence of a large potential infection rate. At the resident
529 level, some countries have promulgated corresponding measures, but citizens have not strictly
530 implemented them. Some citizens resist the 'closing the city' measure, for they think it conflicts
531 with personal freedom. Some people have negative emotions about the severe epidemic and
532 determination of them to protest together is insufficient, affecting the policy response. It is
533 essential to actively respond to the relevant prevention and control policies issued by the state,
534 and not to violate the prohibition without authorisation. Although the current epidemic situation
535 in some countries is gradually stable, keeping vigilant, paying attention to personal protection,
536 and reducing long-distance travel remain important.
537   To ensure that the training data is more than the fine-tuning data to get good results, if the time
538 interval is long, it is necessary to download the continuously updated data for retraining again,
539 and then use the subsequent data for fine-tuning, so as to achieve the goal of not needing to re-
540 train the data in a short time and improving the accuracy.
541

542 **Conclusions**
543 In this work, deep learning models were utilized to research the development of COVID-19 in
544 China, the top five countries and the world. The framework based on IndRNN and fine-tuning
545 consists of four steps: data preprocessing, pre-training the model and saving the weight, fine-
546 tuning the weight, trend prediction and validation. The development tendency of COVID-19 was
547 analyzed, predicted, and validated. Some conclusions are draw as follows:
548 (1) The validity of the proposed framework is verified by comparing it with LSTM, Bi-
549   LSTM, GRU, Stacked_Bi_GRU, CNN_LSTM, and Deep_CNN. The result demonstrated
550   that IndRNN model shows the best performance, and has low complexity. Compared
551   with no-fine-tuning, the fine-tuned IndRNN model can effectively the reduce the
552   prediction errors by up to 20.94% and the time cost. For most of the countries, the MAPE
553   of IndRNN model with fine-tuning was less than 1.2%, and the lowest MAPE and RMSE
554   values of 0.04% and 2.00 in testing results were gendered, for deaths in China, which
555   indicated the effectiveness of the proposed framework.

556 (2) According the prediction and validation results, the MAPE of IndRNN model with fine-
557   tuning was less than 6.2% in most cases, and this framework obtained a minimum MAPE
558   and RMSE of 0.05% and 1.17, respectively, by using Chinese deaths. The deviations
559   between predicted cumulative confirmed and death cases on the world in late May 2021
560   and the real data were 2.81%, and 2.30%, respectively, which confirmed the predictive
561   performance of the proposed framework.
562 (3) Policies play an important role in the development of COVID-19. Timely and appropriate
563   measures can greatly reduce the spread of COVID-19. Some countries have adopted the
564   same positive policy but the effect is not ideal due to the untimely and inappropriate
565   government policies, low implementation ability and the coordination degree of the
566   public. The positive measures include release of emergency response, sealing off cities
567   and closing entertainment venues, prohibiting activities, establishment of temporary
568   hospital, border inspection, cruise ship ban, the release of prevention and control
569   guidelines, the vigorously production of medical materials, improvement of the detection
570   ability of COVID-19. And the negative measures are relaxing restrictions too early for
571   opening, timely policy implementation, hosting of great events, and face Masks are not
572   mandatory. Additionally, untimely and inappropriate government policies, lax
573   regulations, and insufficient public cooperation are the reasons for the aggravation of the
574   epidemic situations.
575

Acknowledgements

577 We appreciate the World Health Organization (WHO), National Health Commission of the
578 people's Republic of China, and the Johns Hopkins University in the United States for providing
579 us with the use some of information or data.

580

References

582 **Zhou P, Yang X, Wang X, Hu B, Zhang L, Zhang W, Si H, Zhu Y, Li B, Huang C, Chen H,**
583   **Chen J, Luo Y, Guo H, Jiang R, Liu M, Chen Y, Shen X, Wang X, Zheng X, Zhao K,**
584   **Chen Q, Deng F, Liu L, Yan B, Zhan F, Wang Y, Xiao G, Shi Z. 2020.** A pneumonia
585   outbreak associated with a new coronavirus of probable bat origin. *Nature* **579:** 270–273 DOI
586   10.1038/s41586-020-2012-7.
587 **Chinadaily. 2020.** The Secretary-General of the United Nations calls on everyone to unite: the
588   new crown epidemic is the most serious global crisis facing humanity since World War II.
589   *Available at https://world.chinadaily.com.cn/a/202004/01/WS5e83fa64a3107bb6b5-*
590   *7aa057.html* (accessed 31August 2020).
591 **World Health Organization. 2020a.** WHO Director-General's opening remarks at the media
592   briefing on COVID-19 - 11 March 2020. *Available at*
593   *https://www.who.int/zh/dg/speeches/detail/who-director-general-s-opening-remarks-at-the-*
594   *media-briefing-on-co-vid-19---11-march-2020* (accessed 13 August 2020).

595 **Gasmi A, Noor S, Tippairote T, Dadar M, Menzel A, Bjørklund G. 2020.** Individual risk
596   management strategy and potential therapeutic options for the COVID-19 pandemic. *Clinical*
597   *Immunology* **215**: 108409 DOI 10.1016/j.clim.2020.108409.

598 **Wu Y, Chen C, Chan Y. 2020.** The outbreak of COVID-19: an overview. Journal of the
599   Chinese Medical Association **83**: 217-220 DOI 10.1097/JCMA.0000000000000270.

600 **Benvenuto D, Giovanetti M, Ciccozzi A, Spoto S, Angeletti S, Ciccozzi M. 2020.** The
601   2019-new coronavirus epidemic: Evidence for virus evolution. *Journal of Medical Virology*.
602   **92:** 455-459 DOI 10.1002/jmv.25688.

603 **Puntmann VO, Carerj ML, Wieters I, Fahim M, Arendt C, Hoffmann J, Shchendrygina A,**
604   **Escher F, Vasa-Nicotera M, Zeiher AM, Vehreschild M, Nagel E. 2020.** Outcomes of
605   Cardiovascular Magnetic Resonance Imaging in Patients Recently Recovered From
606   Coronavirus Disease 2019 (COVID-19). *JAMA Cardio*l **5(11):**1265-1273 DOI
607   10.1001/jamacardio.2020.3557.

608 **Mckibbin WJ, Fernando R. 2020.** The Global Macroeconomic Impacts of COVID-19: Seven
609   Scenarios. *CAMA Working Paper* no. **19/2020** DOI 10.2139/ssrn.3547729.

610 **Abcnews. 2020a.** UN: Virus could push 14 million into hunger in Latin America. *Available at*
611   *https://abcnew-s.go.com/Health/wireStory/virus-push-14-million-hunger-latin-america-*
612   *70919402* (accessed 13 August 2020).

613 **Li Q, Guan X, Wu P, Wang X, Zhou L, Tong Y, Ren R, Leung KSM, Lau EHY, Wong JY,**
614   **Xing X, Xiang N, Wu Y, Li C, Chen Q, Li D, Liu T, Zhao J, Liu M, Tu W, Chen C, Jin**
615   **L, Yang R, Wang Q, Zhou S, Wang R, Liu H, Luo Y, Liu Y, Shao G, Li H, Tao Z, Yang**
616   **Y, Deng Z, Liu Boxi, Ma Z, Zhang Y, Shi G, Lam TTY, Wu JT, Gao GF, Cowling BJ,**
617   **Yang B, Leung GM, Feng Z. 2020a.** Early Transmission Dynamics in Wuhan, China, of
618   Novel Coronavirus–Infected Pneumonia. *The New England Journal of Medicine* **382:** 1199-
619   1207 DOI 10.1056/NEJMoa2001316.

620 **Yang Z, Zeng Z, Wang K, Wong S, Liang W, Zanin M, Liu P, Cao X, Gao Z, Mai Z, Liang**
621   **J, Liu X, Li S, Li Y, Ye F, Guan W, Yang Y, Li F, Luo S, Xie Y, Liu B, Wang Z, Zhang**
622   **S, Wang Y, Zhong N, He J. 2020.** Modified SEIR and AI prediction of the epidemics trend of
623   COVID-19 in China under public health interventions. *Journal of Thoracic Disease* **12:** 165-
624   174 DOI 10.21037/jtd.2020.02.64.

625 **World Health Organization. 2020b.** Coronavirus disease (COVID-19) Weekly
626   Epidemiological Update. *Available at https://www.who.int/publications/m/item/weekly-*
627   *epidemiological-update-on-covid-19---27-april-2021* (accessed 30 May 2021).

628 **Fang Y, Nie Y, Penny M. 2020.** Transmission dynamics of the COVID-19 outbreak and
629   effectiveness of government interventions: A data-driven analysis. *Journal of Medical*
630   *Virology* **92:** 645-659 DOI 10.1002/jmv.25750.

631 **Zandavi SM, Rashidi TH, Vafaee F. 2020.** Forecasting the Spread of Covid-19 Under Control
632   Scenarios Using LSTM and Dynamic Behavioral Models. *Preprint* **2020:**200512270 DOI
633   abs/2005.12270.

634   **Shahid F, Zameer A, Muneeb M. 2020.** Predictions for COVID-19 with Deep Learning
635   Models of LSTM, GRU and Bi-LSTM. *Chaos Solitons & Fractals* **140:**110212 DOI
636   10.1016/j.chaos.2020.110212.

637   **Bandyopadhyay SK, Dutta S. 2020.** Artificial Intelligence Based Study on Analyzing of Habits
638   and with History of Diseases of Patients for Prediction of Recurrence of Disease Due to
639   COVID-19. 2020. *Preprint* **2020:** 2020080542 DOI:10.20944/preprints202008.0542.v1.

640   **Huang CJ, Chen YH, Ma Y, Kuo PH. 2020.** Multiple-Input Deep Convolutional Neural
641   Network Model for COVID-19 Forecasting in China. *Preprint* **2020:** 20041608
642   DOI:10.1101/2020.03.23.20041608.

643   **Dutta S, Bandyopadhyay SK, Kim TH. 2020.** CNN-LSTM Model for Verifying Predictions of
644   Covid-19 Cases. *Asian Journal of Computer Science and Information Technology* **5(4):**25-32
645   DOI:10.9734/AJRCOS/2020/v5i430141.

646   **Pathan RK, Biswas M, Khandaker MU. 2020.** Time Series Prediction of COVID-19 by
647   Mutation Rate Analysis using Recurrent Neural Network-based LSTM Model. *Chaos Solitons*
648   *& Fractals* **138:**110018 DOI:10.1016/j.chaos.2020.110018.

649   **Li S, Li W, Cook C, Zhu C, Gao Y. 2018.** Independently Recurrent Neural Network (IndRNN):
650   Building A Longer and Deeper RNN. *2018 IEEE/CVF Conference on Computer Vision and*
651   *Pattern Recognition* 5457-5466 DOI 10.1109/CVPR.2018.00572.

652   **Zhang P, Meng J, Luan Y, Liu C. 2020.** Plant miRNA–lncRNA Interaction Prediction with the
653   Ensemble of CNN and IndRNN. *Interdisciplinary Sciences: Computational Life Sciences* **12:**
654   82–89 DOI 10.1007/s12539-019-00351-w.

655   **Tajbakhsh N, Shin JY, Gurudu SR, Hurst RT, Kendall CB, Gotway MB, Liang J. 2016.**
656   Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning?.
657   *IEEE Transactions on Medical Imaging* **35:**1299-1312 DOI 10.1109/TMI.2016.2535302.

658   **Boyd A, Czajka A, Bowyer K. 2020.** Deep Learning-Based Feature Extraction in Iris
659   Recognition: Use Existing Models, Fine-tune or Train From Scratch?. *Preprint*
660   2020:200208916 *Available at* https://arxiv.org/abs/2002.08916.

661   **National Health Commission of the people's Republic of China. 2020.** Report of the epidemic
662   situation (in Chinese). *Available at* http://www.nhc.gov.cn/xcs/yqtb/list_gzbd.shtml (accessed
663   31 August 2020).

664   **Johns Hopkins University in the United States. 2020.** Open data sets. *Available at*
665   *https://github.com/CSSGISandData/CO-VID-19* (accessed 31 August 2020).

666   **Baidu. 2020.** Real-time big data report on the epidemic (in Chinese). *Available at*
667   *https://voice.baidu.com/act/newpneumonia/newpneumonia/?from=osari_pc_3* (accessed 31
668   August 2020).

669   **Goodfellow I, Bengio Y, Courville A. 2016.** *Deep Learning* 367-415. Available at
670   *http://www.deeplearningbook.org*.

671   **Hochreiter S, Schmidhuber J. 1997.** Long Short-Term Memory. *Neural Computation* **9:** 1735-
672   1780 DOI 10.1162/n-eco.1997.9.8.1735.

673   **Olah C. 2015.** Understanding LSTM Networks. *Available at http://colah.github.io/posts/2015-*
674   *08-Understanding-LSTMs/* (accessed 14 August 2020).

675   **Schuster M, Paliwal KK. 1997.** Bidirectional recurrent neural networks. *IEEE Transactions on*
676   *Signal Processing* **45(11):** 2673-2681 DOI 10.1109/78.650093.

677   **Chung J, Gulcehre C, Cho K, Bengio Y. 2014.** Empirical Evaluation of Gated Recurrent
678   Neural Networks on Sequence Modeling. *Preprint* **2014:** 14123555. *Available at*
679   *https://arxiv.org/abs/1412.3555.*

680   **The State Council Information Office of the People's Republic of China. 2020.** Full Text:
681   Fighting COVID-19: China in Action. *Available at*
682   *http://www.scio.gov.cn/zfbps/32832/Document/1681809/1681809.htm* (accessed 17 October
683   2020)

684   **Li, Z. Chen Q, Feng L, Rodewald L, Xia Y, Yu H, Zhang R, An Z, Yin W, Chen W, Qin Y,**
685   **Peng Z, Zhang T, Ni D, Cui J, Wang Q, Yang X, Zhang M, Ren X, Wu D, Sun X, Li Y,**
686   **Zhou L, Qi X, Song T, Gao GF, Feng Z, Li Z, Chen Q, Feng L, Rodewald L, Xia Y, Yu**
687   **H, Zhang R, Yin W, Ni D, Qin Y, Zhang T, Cui J, Wang Q, Yang X, Zhang M, Chen W,**
688   **Peng Z, Ren X, An Z, Wu D, Sun X, Li Y, Zhou L, Qi X, Gao GF, Feng Z, Song T, Luo**
689   **H, Yin Z, Wang L, Ma C, Li S. 2020b.** Active case finding with case management: the key
690   to tackling the COVID-19 pandemic. *The Lancet* **396:** 63-70 DOI 10.1016/S0140-
691   6736(20)31278-2.

692   **Chinese Center for Disease Control and Prevention. 2020a.** Guidelines for COVID-19
693   epidemiological investigations. *Available at*
694   *http://weekly.chinacdc.cn/fileCCDCW/journal/article/ccdcw/2020/19/PDF/Annex-1.pdf*
695   (accessed 14 August 2020).

696   **Chinese Center for Disease Control and Prevention. 2020b.** COVID-19 Prevention and
697   Control Guide (First Edition) (in Chinese). *Available at*
698   *http://www.chinacdc.cn/jkzt/crb/zl/szkb_11803/jszl_2275/202002/t20200201_212138.html*
699   (accessed 14 August 2020).

700   **Holshue ML, DeBolt C, Lindquist S, Lofy KH, Wiesman J, Bruce H, Spitters C, Ericson K,**
701   **Wilkerson S, Tural A, Diaz G, Cohn A, Fox L, Patel A, Gerber SI, Kim L, Tong S, Lu X,**
702   **Lindstrom S, Pallansch MA, Weldon WC, Biggs HM, Uyeki TM, Pillai SK. First Case of**
703   **2019 Novel Coronavirus in the United States. 2020.** *New England Journal of Medicine* **382:**
704   929-936 DOI 10.1056/NEJMoa2001191.

705   **National Immigration Administration. 2020.** Remind on recent immigration control measures
706   of relevant countries (in Chinese). *Available at*
707   *https://www.nia.gov.cn/n794014/n1050181/n1050489/c1214826/content.html* (accessed 14
708   August 2020).

709   **Centers for Disease Control and Prevention. 2020.** Coronavirus disease 2019 (COVID-19).
710   Recommendation for cloth face covers. Atlanta, GA: US Department of Health and Human
711   Services. *Available at https://www.cdc.gov/coronavirus/2019-ncov/prevent-getting-sick/cloth-*
712   *face-cover.html.*

713 **Schuchat A. 2020.** Public Health Response to the Initiation and Spread of Pandemic COVID-19
714   in the United States, February 24–April 21, 2020. *Morbidity and Mortality Weekly Report* **69:**
715   551-556. *Available at https://www.cdc.gov/mmwr/volumes/69/wr/mm6918e2.htm*.
716 **Goodmorningamerica. 2020a.** Hospitalizations rise as Trump pushes reopening. *Available at*
717   *https://www.goodmorningamerica.com/news/video/hospitalizations-rise-trump-pushes-*
718   *reopening-72048700* (accessed 14 August 2020).
719 **USA TODEY. 2020.** Trump's campaign made stops nationwide. Coronavirus cases surged in his
720   wake in at least five places.
721   *https://www.usatoday.com/story/news/investigations/2020/10/22/trumps-campaign-made-*
722   *stops-nationwide-then-coronavirus-cases-surged/3679534001/* (accessed 2 June 2021).
723 **Xinhuanet. 2020a.** Epidemic rebounds, "Silicon Valley of India" and other places are "closed"
724   again (in Chinese). *Available at http://www.xinhuanet.com/world/2020-*
725   *07/16/c_1210703949.htm* (accessed 14 August 2020).
726 **Abcnews. 2020b.**The Latest: India surpasses 150,000 virus cases. *Available at*
727   *https://abcnews.go.com/Health/wireStory/la-test-india-reports-biggest-jump-virus-cases-*
728   *70876706* (accessed 14 August 2020).
729 **Bhuyan A. 2021.** Experts criticise India's complacency over COVID-19. The Lancet 397:1611-
730   1612 DOI 10.1016/S0140-6736(21)00993-4.
731 **The paper. 2020.** Brazil's epidemic timeline: President takes the lead in ignoring epidemic
732   prevention measures and finally infects the COVID-19 (in Chinese). *Available at*
733   *https://www.thepaper.cn/newsDetail_forward_8174050* (accessed 14 August 2020).
734 **Xinhuanet. 2020b.** The first professional football match at Rio, Brazil after the epidemic (in
735   Chinese). *Available at http://sports.xinhuanet.com/c/2020-06/19/c_1126136210.htm* (accessed
736   14 August 2020).
737 **Gouvernement.fr. 2020.**The Government's actions. *Available at*
738   *https://www.gouvernement.fr/info-coronavirus/les-actions-du-gouvernement* (accessed 7
739   November 2020).
740 **CNN. 2020.** ICUs are nearing capacity in this French city. And it's only September. *Available at*
741   *https://edition.cnn.com/2020/09/16/europe/france-bordeaux-coronavirus-hospitals-*
742   *intl/index.html* (accessed 7 November 2020).
743 **Kankanews. 2020.** The number of confirmed cases soared to 20,000. Why did the Russian
744   epidemic suddenly worsen? (in Chinese). *Available at http://www.kankanews.com/a/2020-04-*
745   *14/0039217322.shtml* (accessed 14 August 2020).
746 **Xinhuanet. 2020c.** Hot Q&A: Two highlights of the Russian military parade on Victory Day (in
747   Chinese). *Available at http://www.xinhuanet.com/world/2020-06/24/c_1126154264.htm*
748   accessed 14 August 2020).
749 **Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, Zhang L, Fan G, Xu J, Gu X, Cheng Z, Yu**
750   **T, Xia J, Wei Y, Wu W, Xie X, Yin W, Li H, Liu M, Xiao Y, Gao H, Guo L, Xie J, Wang**
751   **G, Jiang R, Gao Z, Jin Q, Wang J, Cao B. 2020.** Clinical features of patients infected with

752   2019 novel coronavirus in Wuhan, China. *The Lancet* **395:** 497-506 DOI 10.1016/S0140-
753   6736(20)30183-5.

754   **Human Resources and Social security Department of Hubei Province. 2020.** The decision on
755   awarding Dingyu Zhang and Jixian Zhang for great achievements (in Chinese). *Available at*
756   *http://m.hbdysh.cn/article/25927* (accessed 14 August 2020).

757   **Hubei Provincial People's Government. 2020a.** Announcement of Wuhan City's COVID-19
758   Epidemic Prevention and Control Headquarters (in Chinese). *Available at*
759   *http://www.hubei.gov.cn/zhuanti/2020/gzxxgzbd/zxtb/202001/t20200123_2014402.shtml*
760   (accessed 14 August 2020).

761   **People.cn. 2020.** Provide strong legal protection for epidemic prevention and control (in
762   Chinese). *Available at http://yuqing.people.com.cn/n1/2020/0211/c429781-31580633.html*
763   (accessed 14 August 2020).

764   **Chinanews. 2020.** From "Huoshenshan", "Leishenshan" to the "Fangcang" hospital, race against
765   the epidemic at the speed of China! (in Chinese). *Available at*
766   *http://www.chinanews.com/sh/shipin/cns/2020/02-05/news846922.shtml* (accessed 14 August
767   2020).

768   **Hubei Provincial People's Government. 2020b.** At 0:00 today, Wuhan "unblocked" (in
769   Chinese). *Available at http://www.hubei.gov.cn/2019/tpy-w/202004/t20200408_2207205.shtml*
770   (accessed 14 August 2020).

771   **China Central Television. 2020.** The National Health Commission reminds again during the
772   May Day holiday: time-sharing appointments for peak-shift trips. *Available at*
773   *http://news.cctv.com/2020/05/02/ARTIlsOonxEyXZQdSG4ioeV6200502.shtml* (accessed 14
774   August 2020).

775   **Jernigan DB. 2020.** Update: public health response to the coronavirus disease 2019 outbreak—
776   United States, February 24, 2020. *Morbidity and Mortality Weekly Report* **69**: 216-219 DOI
777   10.15585/mmwr.mm6908e1 (accessed 14 August 2020).

778   **Abcnews. 2020c.** Mass protests could lead to another wave of coronavirus infections. *Available*
779   *at https://abcnews.go.com/-US/mass-protests-lead-wave-coronavirus-*
780   *infections/story?id=70997184* (accessed 14 August 2020).

781   **Goodmorningamerica. 2020b.** Trump reacts after low turnout at Tulsa rally. *Available at*
782   *https://www.goodmorningamerica.com/news/video/trump-reacts-low-turnout-tulsa-rally-*
783   *71379690* (accessed 14 August 2020).

784   **Xinhuanet. 2020d.** Spotlight: More U.S. states mandate wearing masks in public as COVID-19
785   cases continue to rise. *Available at http://www.xinhuanet.com/english/2020-*
786   *06/25/c_139166307.htm* (accessed 14 October 2020).

787   **Xinhuanet. 2020e.** "Restart" epidemic briefing. Trump calls for wearing masks (in Chinese).
788   *Available at http://m.xinhuane-t.com/world/2020-07-23/c_1210716033.htm* (accessed 14
789   October 2020).

790   **Xinhuanet. 2020f.** Only one week after school starts, the number of confirmed cases in U.S.

791   universities surges. *Available at http://www.xinhuanet.com/video/2020-*

792   *08/26/c_1210770179.htm* (accessed 31 August 2020).

793

Figure 1

China and top five countries of confirmed COVID-19 cases worldwide.

The data on the left represents the amount of cumulative confirmed cases in China and five countries with severe epidemic situation as of April 24, 2021.

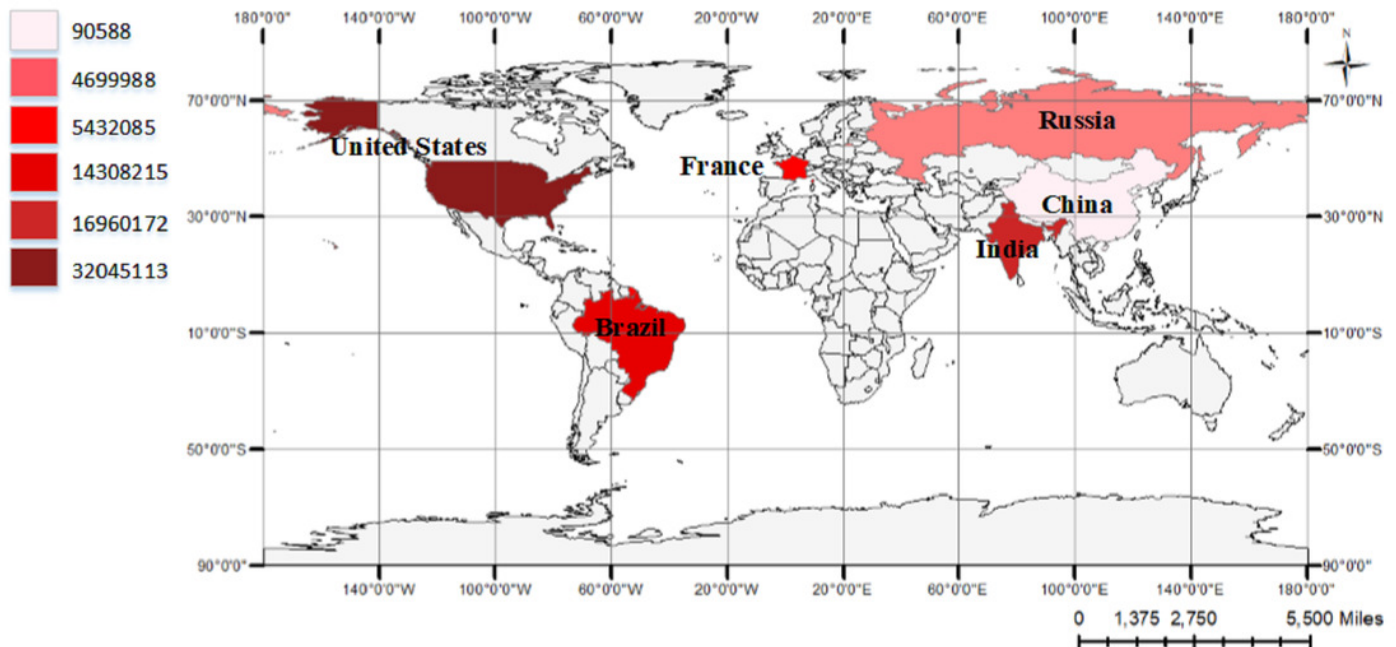



Figure 2

Flow chart of the proposed scheme.

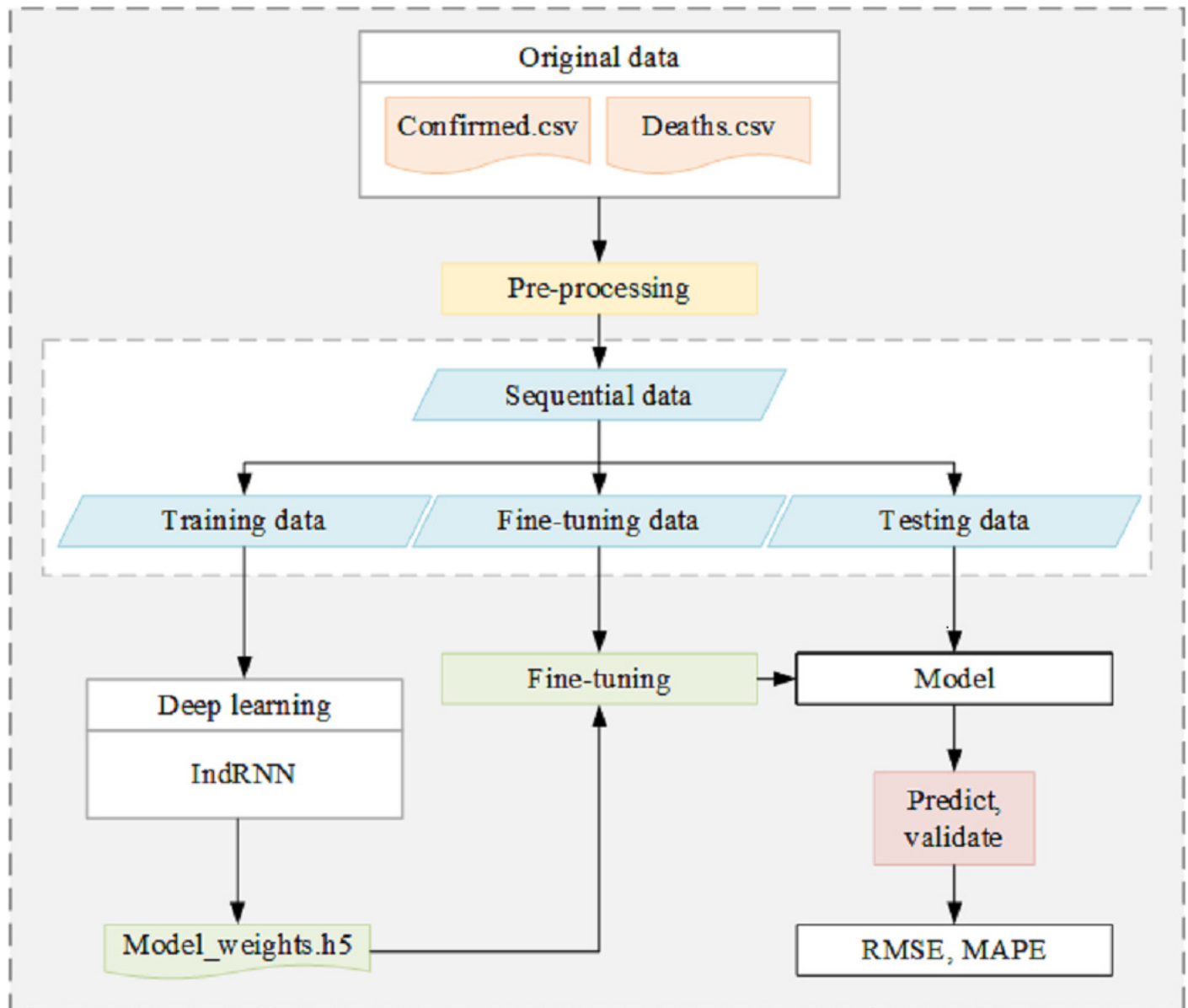



Figure 3

The structure of LSTM, Bi-LSTM, GRU and IndRNN model.

The purple boxes are the network layers that need to be fine-tuned.

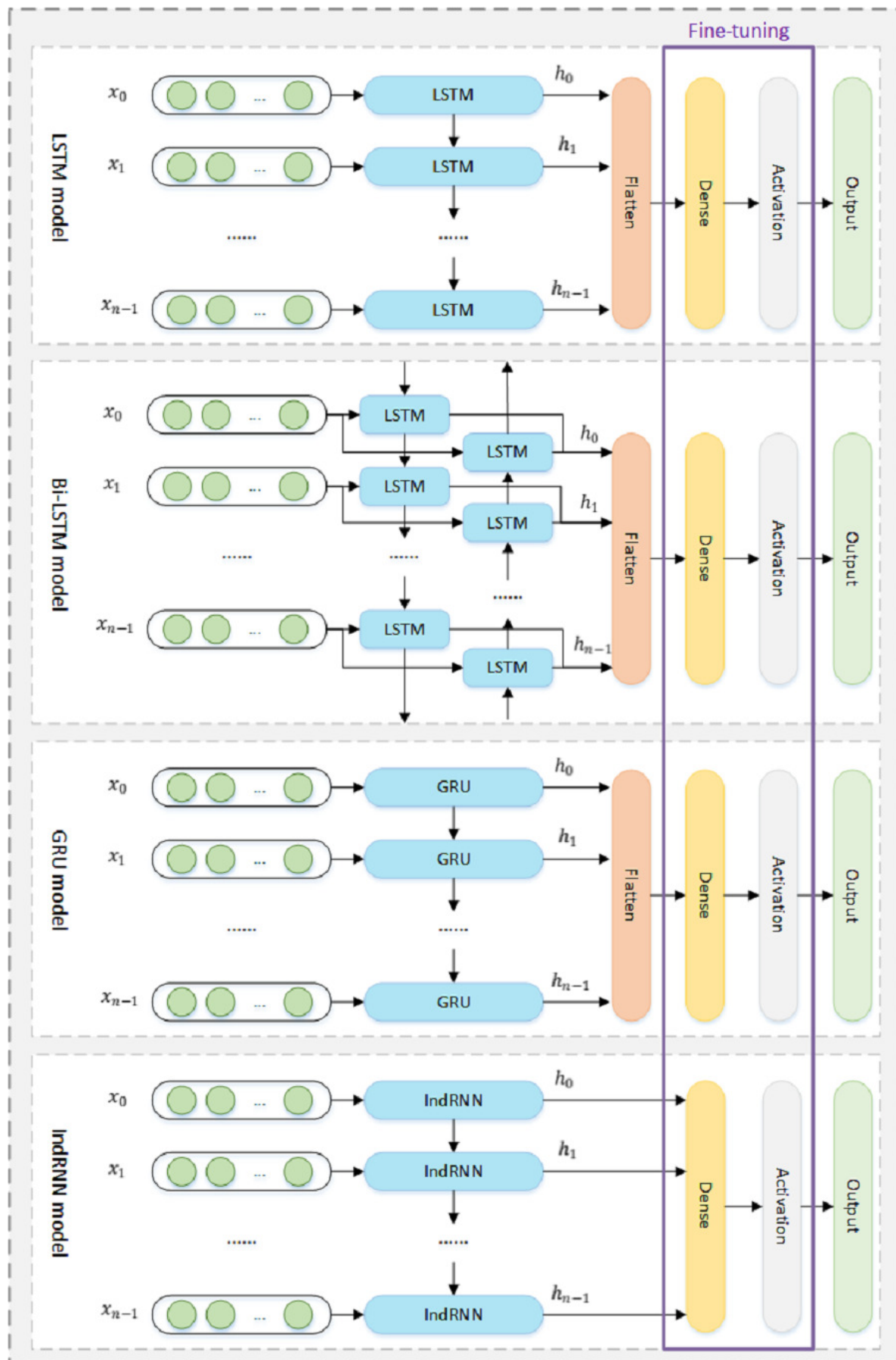



Figure 4

Sequence Data Generation.

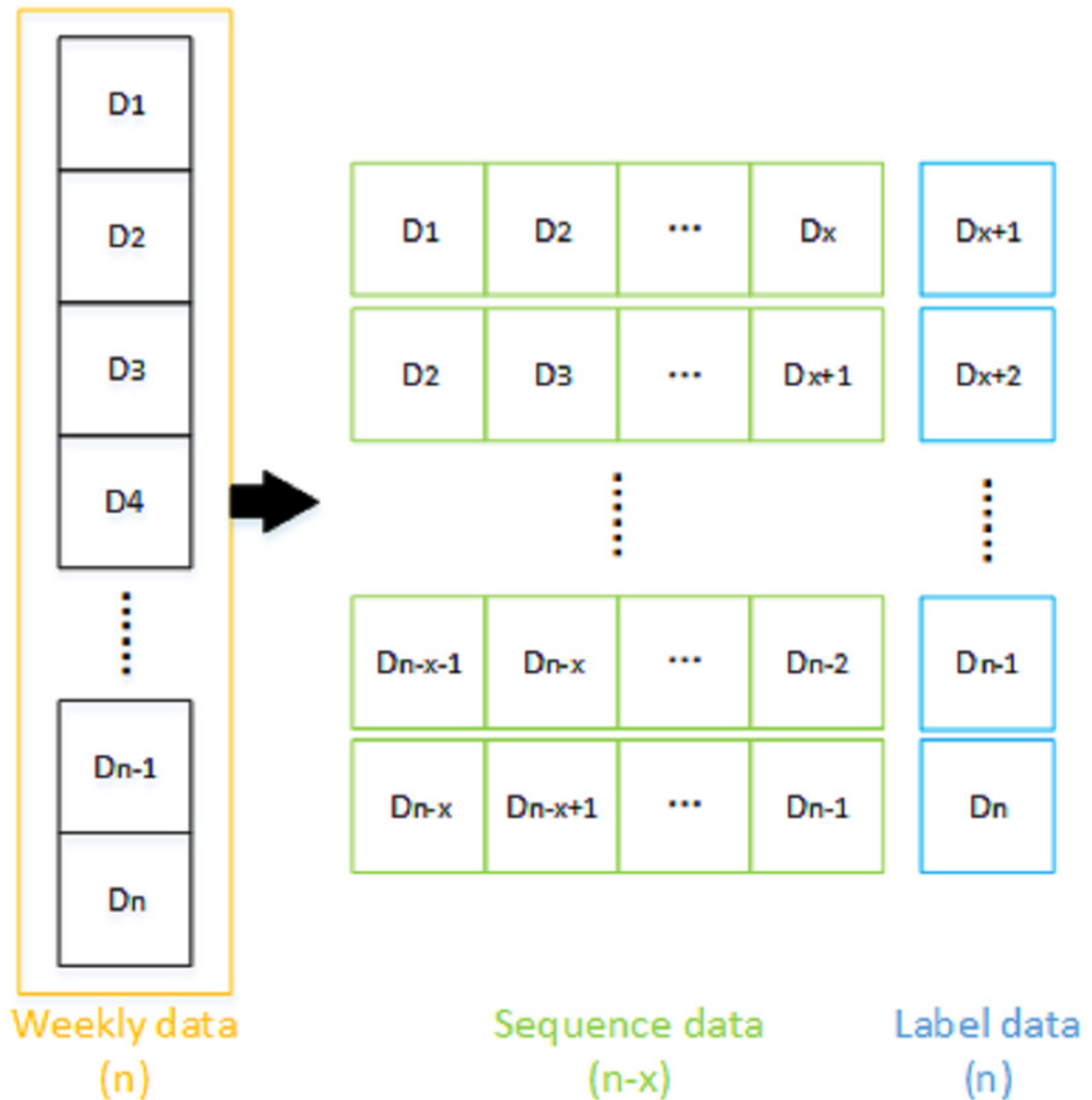



Figure 5

Comparison among LSTM, Bi-LSTM, GRU, IndRNN in terms of RMSE.

Prediction of cumulative confirmed COVID-19 cases in India. The sequence length of 45 weeks is used in the model.

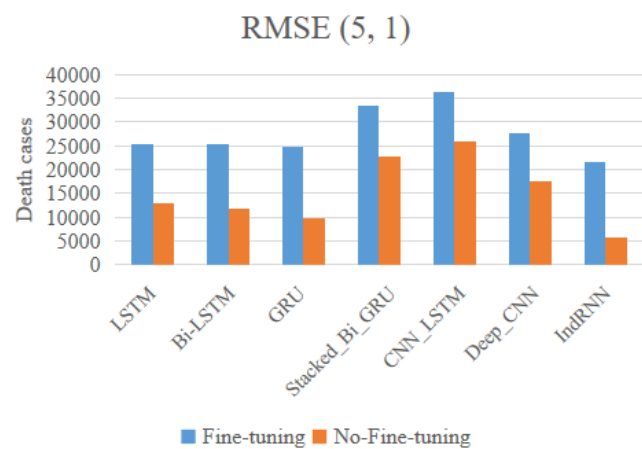
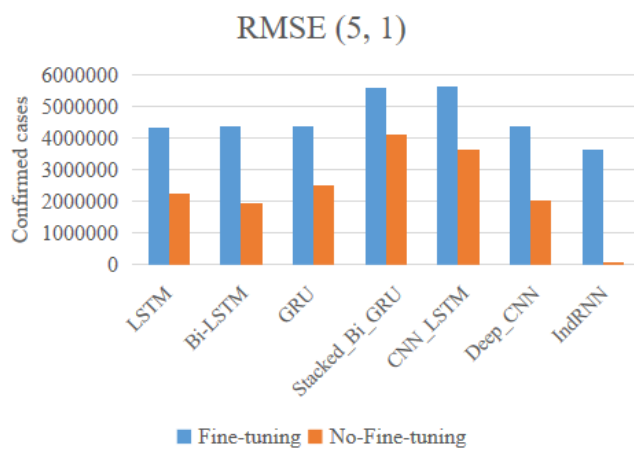
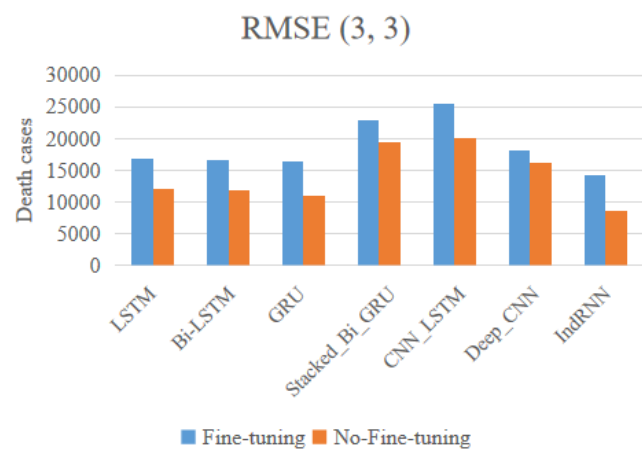
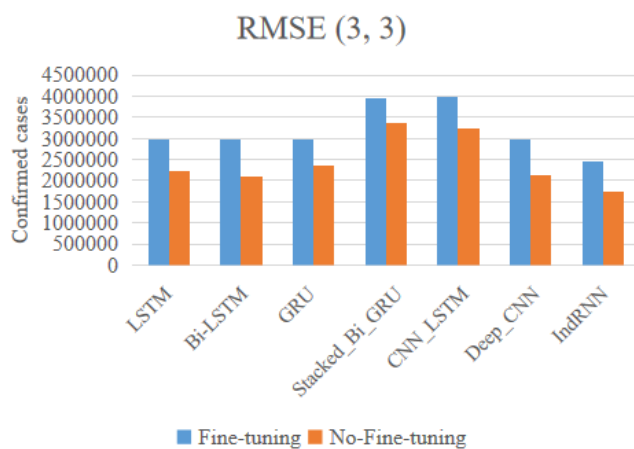
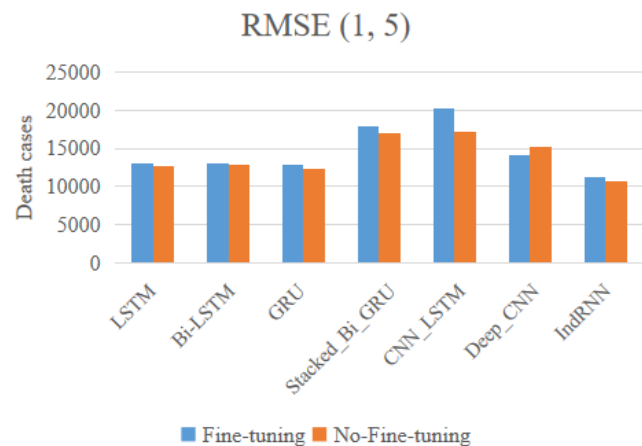
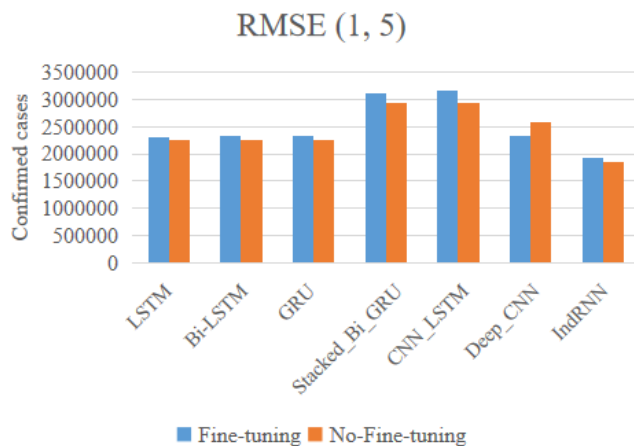



Figure 6

The fitting performance of the IndRNN model with fine-tuning.

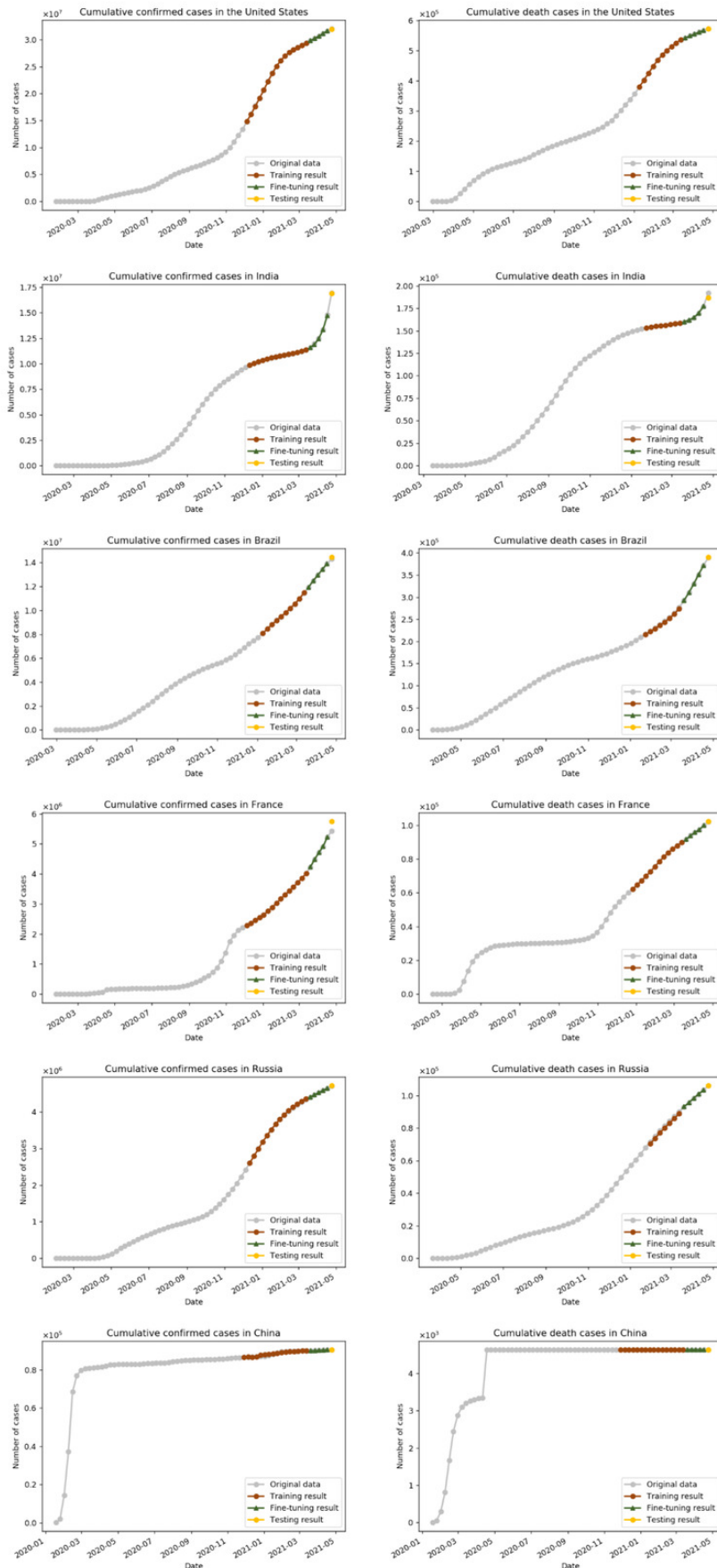



Figure 7

Prediction of cumulative COVID-19 cases in China by IndRNN model with fine-tuning.

The solid line is the actual value and the dashed line is the predicted value.

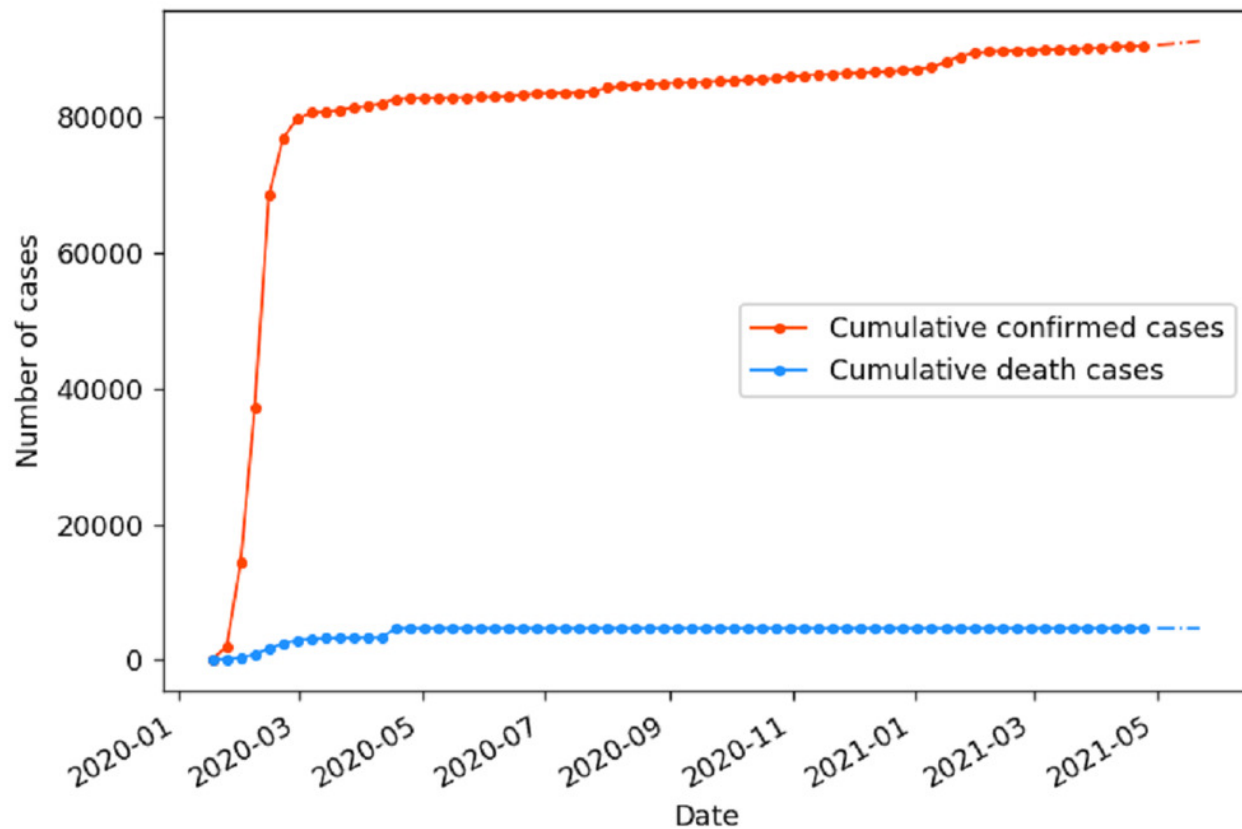



Figure 8

Prediction of cumulative COVID-19 cases in five countries by IndRNN model with fine-tuning.

(A) Prediction of weekly cumulative confirmed cases. (B) Prediction of weekly cumulative death cases. The solid line is the actual data, and the dotted line is the forecast data.

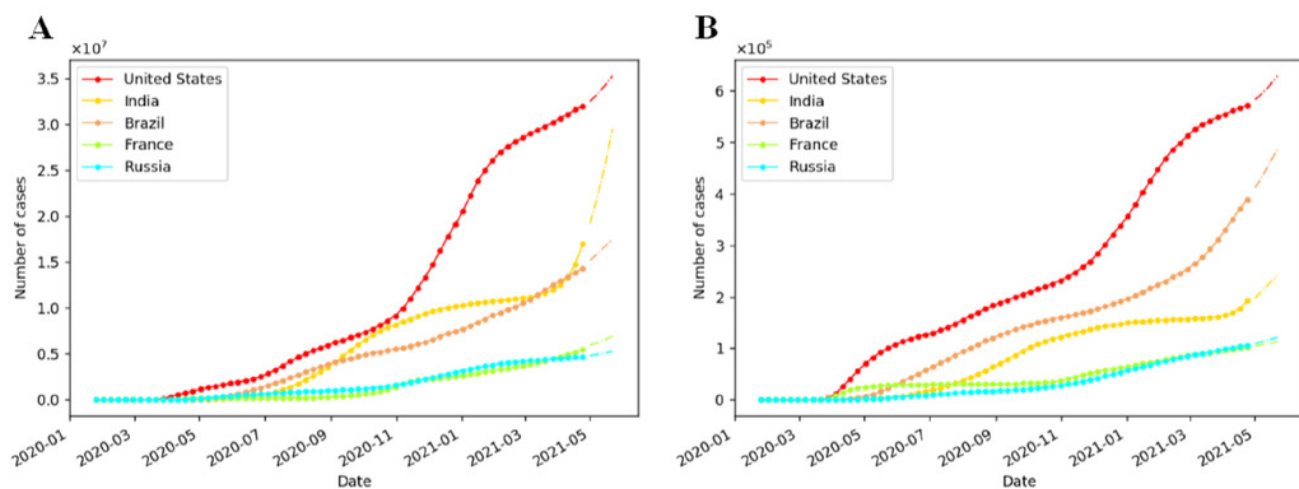
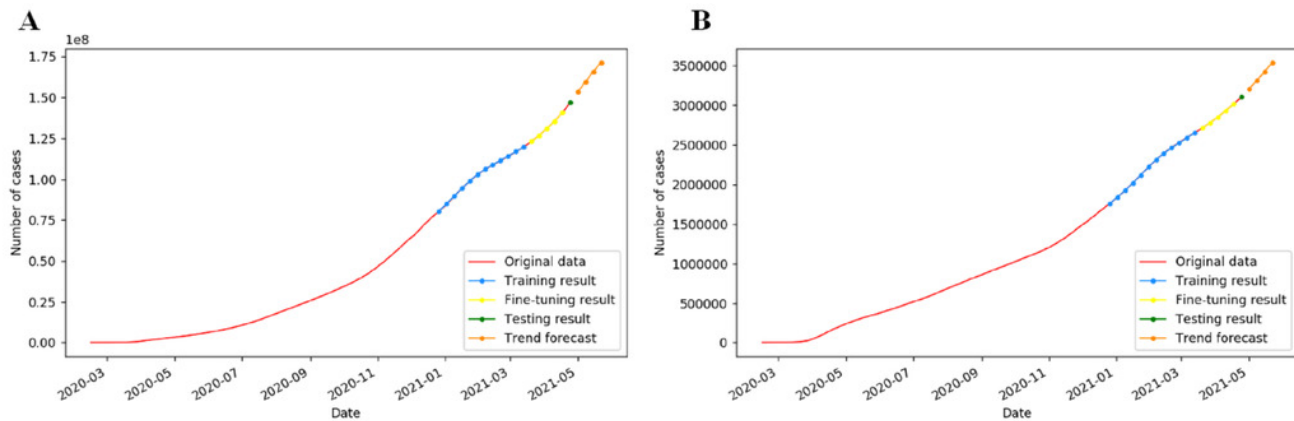



Figure 9

Prediction of global cumulative COVID-19 cases by IndRNN model with fine-tuning.

The solid line is actual data, the dotted line represents the predicted data.

Detailed information of the COVID-19 epidemiological data used in this study.

1 **Table 1.**

2 **Detailed information of the COVID-19 epidemiological data used in this study.**

Country /Region	Statistical period		Source	
	Confirmed	Deaths	Institution	Website
China	2020/01/18~ 2021/5/22	2020/01/18~ 2021/5/22	National Health Commission of the People's Republic of China	*http://www.nhc .gov.cn/xcs/yqt b/list_gzbd.sht ml*
United States	2020/01/25~ 2021/5/22	2020/02/29~ 2021/5/22	Johns Hopkins University in the United States	*https://github.c om/CSSEGISa ndData/COVI D-19*
India	2020/02/01~ 2021/5/22	2020/03/14~ 2021/5/22		
Brazil	2020/02/29~ 2021/5/22	2020/03/21~ 2021/5/22		
France	2020/03/07~ 2021/5/22	2020/02/15~ 2021/5/22		
Russia	2020/02/01~ 2021/5/22	2020/03/21~ 2021/5/22		
Global	2020/01/15~ 2021/5/22	2020/02/15~ 2021/5/22	National Health Commission of the People's Republic of China; Baidu	*http://www.nhc .gov.cn/xcs/yqt b/list_gzbd.sht ml*

Consistent parameters and their values in LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU, CNN_LSTM, Deep_CNN, and IndRNN models.

1   **Table 2.**

2   **Consistent parameters and their values in LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU,**

3   **CNN_LSTM, Deep_CNN, and IndRNN models.**

Parameters	Values
Loss	Mean Squared Error
Optimizer	Adam
Batch size	1
Epoch in training	3000
Epoch in testing	3000
Sequence length	45 weeks

The numbers of network layers of the model and frozen network layers when fine-tuning.

1 **Table 3.**

2 **The numbers of network layers of the model and frozen network layers when fine-tuning.**

Model	Quantity of Total layer	Quantity of frozen layer when fine-tuning
LSTM	5	3
Bi-LSTM	5	3
GRU	5	3
IndRNN	4	2
Stacked_Bi_GRU	14	10
CNN_LSTM	19	14
Deep_CNN	8	7

Comparison among LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU, CNN_LSTM, Deep_CNN, and IndRNN in terms of MAPE.

1 **Table 4.**

2 **Comparison among LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU, CNN_LSTM, Deep_CNN,**

3 **and IndRNN in terms of MAPE.**

4 Prediction of cumulative confirmed COVID-19 cases in India.

Data	Model	Training time(s)	Split (f, t)	MAPE (%)		Fine-tuning time(s)
				No-fine-tuning	Fine-tuning	
Confirmed cases	LSTM	733.38	1, 5	11.69	11.03	21.40
			3, 3	17.13	12.07	47.19
			5, 1	25.60	13.24	77.71
	Bi-LSTM	1296.12	1, 5	11.72	11.10	28.82
			3, 3	17.19	11.10	68.27
			5, 1	25.72	11.43	105.57
	GRU	581.84	1, 5	11.72	11.05	20.14
			3, 3	17.18	12.86	45.32
			5, 1	25.70	14.79	70.85
	Stacked_Bi_GRU	5291.32	1, 5	17.02	15.33	69.52
			3, 3	23.70	19.22	183.36
			5, 1	33.01	24.12	287.70
	CNN_LSTM	3006.03	1, 5	17.58	15.29	41.15
			3, 3	23.88	18.45	104.69
			5, 1	33.16	21.32	170.93
	Deep_CNN	261.37	1, 5	11.63	13.22	13.11
			3, 3	17.21	11.41	29.44
			5, 1	25.84	12.05	45.71
	IndRNN	248.72	1, 5	9.74	9.11	14.85
			3, 3	14.17	8.85	34.96
			5, 1	21.40	0.46	57.08
Confirmed deaths	LSTM	413.85	1, 5	5.29	5.02	20.93
			3, 3	7.94	5.19	48.49
			5, 1	13.21	6.77	74.72
	Bi-LSTM	747.53	1, 5	5.20	5.08	29.68
			3, 3	7.89	5.09	68.36
			5, 1	13.21	6.16	104.81
	GRU	333.79	1, 5	5.21	4.91	19.65
			3, 3	7.80	4.49	44.93
			5, 1	12.96	5.14	69.41
	Stacked_Bi_GRU	2976.06	1, 5	7.82	7.13	68.48

			3, 3	11.30	9.17	166.58
			5, 1	17.40	11.90	274.43
	CNN_LSTM	1703.61	1, 5	9.49	7.28	41.11
			3, 3	12.91	9.55	100.99
			5, 1	18.90	13.45	173.57
	Deep_CNN	143.17	1, 5	5.62	6.20	13.03
			3, 3	8.60	7.45	29.32
			5, 1	14.37	9.21	44.42
	IndRNN	142.34	1, 5	4.72	4.36	14.54
			3, 3	6.89	3.42	33.95
			5, 1	11.22	2.99	56.54

Comparison among LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU, CNN_LSTM, Deep_CNN, and IndRNN in terms of params.

1   **Table 5.**

2   **Comparison among LSTM, Bi-LSTM, GRU, Stacked_Bi_GRU, CNN_LSTM, Deep_CNN,**

3   **and IndRNN in terms of params.**

Model	Total params	Fine-tuning params
LSTM	89217	129
Bi-LSTM	178433	257
GRU	66945	129
Stacked_Bi_GRU	1110905	569
CNN_LSTM	688161	673
Deep_CNN	13121	65
IndRNN	6145	129

Testing and validation of cumulative confirmed and death COVID-19 cases in five countries by IndRNN model with fine-tuning.

1   **Table 6.**

2   **Testing and validation of cumulative confirmed and death COVID-19 cases in five countries**

3   **by IndRNN model with fine-tuning.**

Data	Country	Testing		Validation	
		RMSE	**MAPE (%)**	**RMSE**	**MAPE (%)**
Cumulative cases	United States	138,071.00	0.43	1355857.54	3.34
	India	77,668.00	0.46	1587170.63	4.30
	Brazil	164,181	1.15	1027451.51	6.18
	France	320,132.50	5.89	719450.18	11.53
	Russia	19,647.00	0.42	213593.59	3.77
	China	17.90	0.02	187.94	0.17
	Global	558,483	0.38	2953263.00	1.60
Cumulative deaths	United States	1273.06	0.22	25260.92	3.72
	India	5,752.81	2.99	39954.53	13.92
	Brazil	1,460.31	0.37	24974.75	4.90
	France	300.83	0.29	3236.43	2.36
	Russia	194.50	0.18	3333.14	2.47
	China	2.00	0.04	2.14	0.05
	Global	5,829.50	0.19	47980.75	1.17

Positive and negative events.

1   **Table 7.**

2   **Positive and negative events.**

Type	Event
positive	(1) Release of emergency response; (2) Seal off the city, close entertainment venues; (3) The prohibition of activities; (4) Establishment of temporary hospital; (5) Border inspection; (6) Cruise ship ban; (7) Prevention and control guidelines; (8) Vigorously produce medical materials, improvement of the detection ability of COVID-19.
negative	(1) Relax restrictions too early for opening; (2) Policy implementation is not timely; (3) Hosying of great events: demonstration, campaign, festival; (4) Face masks are not mandatory.