

基于组合预测模型的疫情确诊人数预测

庞梦吟¹, 王海宁¹, 万通明¹, 马 苗^{1,2*}

(1. 陕西师范大学 计算机科学学院, 陕西 西安 710119;

2. 空天地海一体化大数据应用技术国家工程实验室, 陕西 西安 710129)

摘 要: 公共卫生安全事件常常是牵动着一个地区、一个国家,乃至整个世界的重大问题。2019 年底爆发的新型冠状病毒 (Corona Virus Disease 2019, COVID-19) 疫情迅速席卷了很多国家,及时了解疫情确诊人数变化,对协助挖掘肺炎疫情的传播模式和特征规律十分重要。针对此问题,提出了一种组合预测模型来预测此次新型冠状病毒肺炎累计确诊人数数量。首先从国家卫生健康委员会等权威机构所发布的数据中获取历史累计确诊人数,然后再以 Logistic 模型和长短期记忆深度学习网络模型 (Long Short-Term Memory, LSTM) 的预测结果为基础,选取一定时间段的累计确诊人数序列来训练线性组合参数,得到最终的组合预测模型,最后通过 RMSE 等预测性能评价指标对比所提模型和 Logistic、LSTM 和 SEIR 等三种模型的预测性能。实验结果表明,所提模型的 RMSE 值为 10.101 7, MAE 值为 7.633 6, MAPE 值为 0.008 3%,其准确性和拟合效果均优于其他模型,能够为后续的疫情预测和防控工作提供技术支撑。

关键词: Logistic 模型; LSTM 模型; 组合预测模型; 深度学习; 公共卫生安全事件

中图分类号: TP183

文献标识码: A

文章编号: 1673-629X(2022)11-0198-06

doi:10.3969/j.issn.1673-629X.2022.11.029

Epidemic Data Prediction Based on Combined Prediction Model

PANG Meng-yin¹, WANG Hai-ning¹, WAN Tong-ming¹, MA Miao^{1,2*}

(1. School of Computer Science, Shaanxi Normal University, Xi'an 710119, China;

2. National Engineering Laboratory for Integrated Aero-space-ground-ocean Big Data Application Technology, Xi'an 710129, China)

Abstract: Public health security incidents are often major issues that affect a region, a country, and even the entire world. The outbreak of corona virus disease 2019 (COVID-19) quickly swept across many countries, and timely understanding of the changes in the number of confirmed cases of the epidemic is very important to assist in the discovery of the transmission pattern and characteristics of the pneumonia epidemic. In response to this problem, we propose a combined prediction model to predict the cumulative number of confirmed cases of COVID-19. Firstly, the historical cumulative number of diagnoses is obtained from the data released by the National Health Commission and other authoritative institutions. Then based on prediction results of the Logistic model and the long short-term memory deep learning network model (LSTM), the sequence of the cumulative confirmed number of people in a certain period of time is selected to train the linear combination parameters to obtain the final combined prediction model. Finally, the proposed model and the three models of Logistic, LSTM and SEIR are compared in terms of performance according to the prediction performance evaluation indicators such as RMSE. The experimental results show that the RMSE of the proposed model is 10.101 7, the MAE is 7.633 6, and the MAPE is 0.008 3%. Its accuracy and fitting effect are better than other models, which can provide technical support for subsequent epidemic prediction and prevention and control.

Key words: Logistic prediction model; LSTM; combined prediction model; deep learning; public health and safety incidents

0 引 言

公共卫生安全事件常常是牵动着一个地区、一个国家,乃至整个世界的重大问题。新冠肺炎疫情的发

生和蔓延,对中国乃至全球经济产生了巨大的冲击和深远影响。在对抗新冠肺炎疫情的过程中,已经有不少的科研人员研究了疫情传播及其发展趋势,为疫情

收稿日期:2021-05-12

修回日期:2021-09-13

基金项目:陕西省大学生创新创业训练计划项目(S202010718178);陕西师范大学中央高校公共卫生与健康科研专项引导基金(GK202007029);空天地海一体化大数据应用技术国家工程实验室开放课题基金项目(20200201)

作者简介:庞梦吟(2000-),女,研究方向为计算机科学与技术;通讯作者:马 苗(1977-),女,教授,研究方向为图像处理。

防控政策的制定提供宏观的数据依据,节省决策时间。

现有的传染病预测方法已有很多,如流行控制图法、灰色模型、求和自回归移动平均模型(Auto Regressive Integrated Moving Average Model, ARIMA)、Logistic 回归模型、神经网络模型等。针对新冠肺炎疫情问题,研究人员进行了深入研究,并取得了一系列成果。例如,张琳采用一般增长模型分 3 个阶段非线性拟合新冠肺炎疫情中确诊、疑似和密切接触人数,拟合结果与实证数据高度吻合^[1]。喻孜等人基于时变参数-SIR 模型评估和预测 COVID-19 疫情,预测在 2020 年 2 月 9 日确诊人数达到高峰后,将呈现下降趋势^[2]。Li 等人建立梯度推进决策树(Gradient Boosting Decision Tree, GBDT)、Logistic 回归(Logistic Regression, LR)模型和简化 LR 模型预测 COVID-19 的死亡率,其结果表明 GBDT 模型显示出最高的敏感性和特异性,且死亡率的预测性能优于 LR 模型^[3]。盛华熊等人对武汉封城阶段运用经典的 SIR 模型和差分递推方法分析来预测疫情,理论值和实际值能较好吻合。对武汉封城前的自由传染阶段运用 Logistic 模型进行预测和比较,且模型对其他地区疾病传播的分析与预测具有广泛适用性^[4]。李伟伟等人利用动力学模型(Susceptible Exposed Infectious Removed, SEIR)拟合 COVID-19 的再生系数曲线,并对湖北省以及国外疫情进行预测分析,其结果表明湖北省的疫情传播相对于国外得到更好控制^[5]。王瑞等人提出运用 Dropout 技术的 LSTM 神经网络预测模型,构建更多层数的 LSTM 神经网络预测模型,并在隐藏层采用 Dropout 技术,能够有效解决过拟合问题,较精准地预测疫情趋势^[6]。Takasar 等人采用 SEIQR 模型通过参数校正预测了巴基斯坦的情况并对未来情况进行了讨论^[7]。Yudistira 等人考虑 COVID-19 病例增长的多种因素,使用可解释卷积-LSTM 的视觉归因观察高贡献因素,比其他模型的预测误差都低^[8]。季伟东等人分析粒子群优化算法的寻优特点,改进了寻优速度和寻优目的,采用 100 天新增确诊人数数据,将改进算法与其他算法对比,验证其算法的良好性能^[9]。Wibowo 等人采用 Logistic 回归模型预测印尼新冠肺炎疫情,实验结果误差较小^[10]。Song Jialu 等人提出一种通过考虑时变模型参数来动态预测新冠肺炎扩展的新方法,用最大似然估计理论估计时变模型参数,提出扩展卡尔曼滤波器来估计新冠肺炎扩散情况^[11]。Shahid 等人采用 ARIMA、SVR、LSTM、Bi-LSTM 等模型对 COVID-19 影响的十个主要国家进行预测,其结果表明预测模型的性能优劣排序为 Bi-LSTM>LSTM>SVR>ARIMA^[12]。

受已有研究成果启发,由于 Logistic 模型常常被

应用于描述种群、传染病增长以及人口预测等领域,因此利用 Logistic 曲线的特性选择该模型,通过调参拟合疫情的历史变化趋势,并预测未来的疫情数据。但是在本次疫情传播过程中,受外界诸多因素影响较大,同时疫情确诊人数数据具有时间序列特征,而 LSTM 深度学习网络模型擅长解决与时间序列高度相关的问题,可以用来学习一些人为干扰因素,因此该文利用国家卫生健康委员会、世界卫生组织等权威机构公开报道的数据,以历史累计确诊人数为基础,采用上述两个模型的组合来分析预测之后疫情累计确诊人数;最后,通过评价指标值对各模型预测结果进行分析比较,结果表明组合预测模型相对其他模型的预测效果更好,为后续的疫情预测和防控工作提供技术支撑。

1 组合预测模型

单个预测模型进行预测时会存在一些缺陷,如信息源的不全面、对模型设定形式敏感等。自 Bates 和 Granger 首次提出组合预测方法以来,因其能有效提高预测精度,受到国内外预测工作者的重视。组合预测是指应用两种和两种以上的预测模型对传染病进行预测,综合利用单个模型所提供的信息,以适当的加权平均形式得出组合预测模型^[10],尽可能地提高精度。组合预测模型有线性组合模型、最优线性组合模型、贝叶斯组合模型、转换函数组合模型、计量经济与系统动力学组合模型等,该文采用的是线性组合模型。

1.1 Logistic 模型

Logistic 模型,又可称为逻辑回归模型,是一种广义的线性回归分析模型^[13],它是皮埃尔·弗朗索瓦·韦吕勒在研究它与人口增长的关系时命名的。在疫情传播过程中,刚开始的环境阻力比较小,因此增长率比较大;之后随着“封城”、隔离、戴口罩等措施的实施,环境阻力增大,增长率减小,传播规律与 Logistic 模型类似,故可以用 Logistic 模型来预测未来疫情数据,其公式为:

$$P(t) = \frac{KP_0 e^{rt}}{K + P_0(e^{rt} - 1)} \quad (1)$$

其中, t 为时间; K 为容量,即增长到最后, $P(t)$ 能达到的极限; P_0 为初始容量,即 $t=0$ 时刻的数量; r 为增长速率, r 越大则增长越快,越快逼近 K 值, r 越小增长越慢,越慢逼近 K 值。该文建立 Logistic 模型后,采用非线性最小二乘法对 K 、 P_0 、 r 进行参数估计,并以此对未来数据进行预测。

Logistic 模型起初阶段大致是指数增长;然后随着开始变得饱和,增加变慢;最后达到成熟时,增加停止。Logistic 模型常用于数据挖掘、疾病自动诊断、经济预测等领域。例如,探讨引发疾病的危险因素,并根据危

险因素预测疾病发生的概率等^[13]。

1.2 LSTM 模型

LSTM 由 Hochreiter 和 Schmidhuber 在 1997 首先提出,是为解决一般循环神经网络 (Recurrent Neural Network, RNN) 存在的长期依赖问题而专门设计出来

的一种时间循环神经网络^[14]。LSTM 通过增加输入门限、遗忘门限和输出门限,使得在自循环的过程中权重是变化的,从而避免梯度消失或者梯度膨胀的问题。由于其独特的设计结构, LSTM 适合于处理和预测时间序列中有间隔和延迟的重要事件。

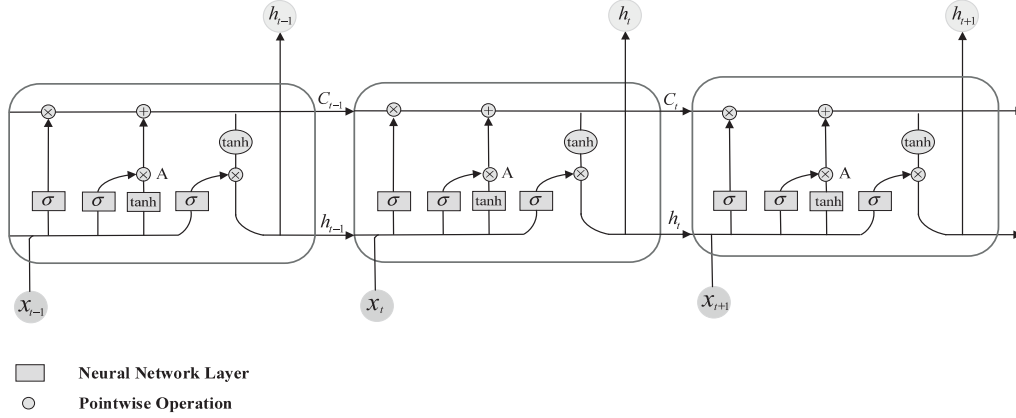


图 1 LSTM 单元结构流动图

图 1 为 LSTM 单元结构流动图,其中 A 模块为 LSTM 的单元结构图。假定 h_i 为第 i 时刻状态的输出值, x_i 为第 i 时刻的输入值, W 表示相应门的权重矩阵, b 表示相应门的偏置, LSTM 的具体流程为:

(1) 遗忘门限:将前一时刻的网络输出和当前时刻的网络输入结合起来,然后进行线性变换后经过 sigmoid 激活函数,并将结果映射到 0 ~ 1 作为记忆衰减系数 f_i ,其中 1 表示“完全接受”,0 表示“完全忽略”,公式如下:

$$f_i = \sigma(W_f \cdot [h_{i-1}, x_i] + b_f) \quad (2)$$

(2) 输入门限:确定需要在细胞状态中保存哪些新信息。首先,采用类似于计算记忆衰减系数的方法得到输入门限的系数 i_i ,之后通过线性变换和 tanh 激活函数得到当前状态学习到的记忆 \tilde{C}_i ,这两步公式为:

$$i_i = \sigma(W_i \cdot [h_{i-1}, x_i] + b_i) \quad (3)$$

$$\tilde{C}_i = \tanh(W_c \cdot [h_{i-1}, x_i] + b_c) \quad (4)$$

(3) 细胞状态更新:对前一时刻状态 C_{i-1} 乘以其衰减系数 f_i ,加上该时刻下学到的记忆 \tilde{C}_i 乘以其衰减系数 i_i ,得到状态更新值,公式为:

$$C_i = f_i \times C_{i-1} + i_i \times \tilde{C}_i \quad (5)$$

(4) 输出门限:首先采用类似于计算记忆衰减系数的方法得到输出门限的系数 o_i ,之后将细胞状态输入 tanh 后与该系数相乘得到输出值,这两步公式为:

$$o_i = \sigma(W_o \cdot [h_{i-1}, x_i] + b_o) \quad (6)$$

$$h_i = o_i \times \tanh(C_i) \quad (7)$$

LSTM 模型常用于文本生成、机器翻译、语音识别、生成图像描述和视频标记等领域,同时 LSTM 模

型也常被应用于机械故障诊断和预测领域,如处理机械设备的振动信号、航班延误预测、金融时间序列预测、加油站时序数据异常检测等问题。

1.3 组合预测模型

基于以上两种模型,该文将 Logistic 模型和 LSTM 模型的预测结果通过线性回归网络^[15]训练权重参数并进行预测,以期得到更好的预测效果。

组合预测模型定义为:

$$\hat{y} = x_1 w_1 + x_2 w_2 + b \quad (8)$$

其中,模型输出 \hat{y} 是对真实累计确诊人数的预测, x_1 为 LSTM 模型得到的结果, x_2 为 Logistic 模型得到的结果, w_1 和 w_2 是权重, b 是偏差。

损失函数为:

$$l(w_1, w_2, b) = \frac{1}{n} \sum_{i=1}^n l^{(i)}(w_1, w_2, b) = \frac{1}{n} \sum_{i=1}^n \frac{1}{2} (x_1^{(i)} w_1 + x_2^{(i)} w_2 + b - y^{(i)})^2 \quad (9)$$

在模型训练中,找出一组模型参数,记作 w_1^*, w_2^*, b^* ,采用小批量随机梯度下降来优化参数,使训练样本的平均损失最小,公式为:

$$w_1^*, w_2^*, b^* = \underset{w_1, w_2, b}{\operatorname{argmin}} l(w_1, w_2, b) \quad (10)$$

采用小批量随机梯度下降来优化参数,将模型各参数作如下迭代:

$$\begin{cases} w_1 \leftarrow w_1 - \frac{\eta}{|\beta|} \sum_{i \in B} \frac{\partial l^{(i)}(w_1, w_2, b)}{\partial w_1} \\ w_2 \leftarrow w_2 - \frac{\eta}{|\beta|} \sum_{i \in B} \frac{\partial l^{(i)}(w_1, w_2, b)}{\partial w_2} \\ b \leftarrow b - \frac{\eta}{|\beta|} \sum_{i \in B} \frac{\partial l^{(i)}(w_1, w_2, b)}{\partial b} \end{cases} \quad (11)$$

其中, $|\beta|$ 表示每个小批量中的样本个数, η 称为学习率, η 过大可能会导致震荡而无法收敛, η 过小会导

致收敛的速度慢,因此选择一个恰当的 η 值至关重要。

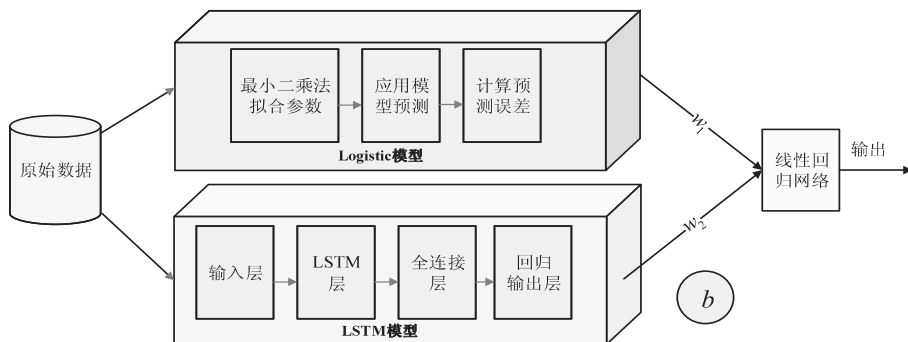


图2 组合预测模型结构

图2为所建立的组合预测模型结构,其中LSTM模型有四层网络:输入层、LSTM层、全连接层、回归输出层,其中设定LSTM层有200个隐含单元;Logistic模型和LSTM模型预测结果作为线性回归网络的两个输入,通过线性回归网络训练该组合预测模型后,输出的结果为该组合预测模型的预测结果。

2 模型求解与对比分析

该文采用国家卫生健康委员会公布的2020年1月13日至2020年11月13日的累计确诊人数,将前276天(即2020年1月13日至2020年10月14日)的数据作为训练集来训练模型,之后的30天(即2020年10月15日至2020年11月13日)作为测试集来测试模型的效果。并利用评价指标对组合预测模型与LSTM模型、Logistic模型、SEIR模型预测进行对比分析。

2.1 模型求解

基于Python语言和PyTorch平台将Logistic模型与LSTM模型的预测结果作为输入层,将其结果进行归一化后输入到线性回归网络中,通过网络不断迭代对组合预测模型的参数进行训练,得到的模型参数为:

$$w_1 = 1.000\ 5, w_2 = -0.001\ 5, b = -0.001\ 8 \quad (12)$$

即最终得到的组合预测模型为:

$$\hat{y} = 1.000\ 5x_1 - 0.001\ 5x_2 - 0.001\ 8 \quad (13)$$

其中, x_1 为LSTM模型预测数据归一化的值, x_2 为Logistic模型预测数据归一化的值,得到的 \hat{y} 反归一化后为组合预测模型预测的累计确诊人数。

图3为组合预测模型的预测结果,横坐标为天数(1月13日为开始),纵坐标为确诊人数。实线部分为真实数据,虚线部分为预测的数据。图3表明,在前期疫情刚刚爆发,没有人为因素、政策等措施的控制,疫情的累计确诊人数急剧增长,在“封城”等一些措施实施后,疫情的累计确诊人数的增长速率减小,在中后期确诊人数趋于稳定,但仍在缓慢增长。

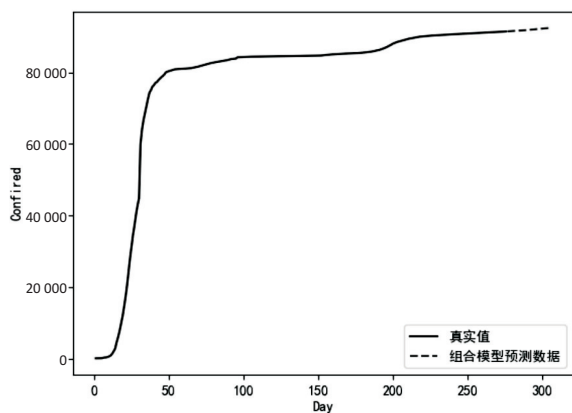


图3 组合预测模型预测结果

2.2 对比分析

组合预测模型、LSTM模型、Logistic模型、SEIR模型在2020年10月15日-2020年11月13日的累计确诊人数预测结果与真实数据如表1所示。从表1中可以看出,组合预测模型得到的结果最接近于真实数据,LSTM模型次之,Logistic模型和SEIR模型的结果与真实数据相差较远。Logistic模型的多日预测结果差距并不大,这是因为Logistic模型的特征为后期阶段达到容量后数据基本不变,因此在该时期预测结果较差。由于在后期国家对疫情的传播进行了有效的控制使得累计确诊人数增长缓慢,同时感染概率、康复概率、潜伏者的患病概率等因素值无法精确评估,因此在该时期SEIR模型的预测效果相对较差,预测结果与真实数据相差较大。

对于回归模型的评价指标有以下三种:

(1) 均方根误差(Root Mean Squared Error, RMSE)。

$$RMSE(y, \hat{y}) = \sqrt{\frac{1}{n} \sum_{i=1}^n \|y_i - \hat{y}_i\|^2} \quad (14)$$

(2) 平均绝对误差(Mean Absolute Error, MAE)。

$$MAE(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i| \quad (15)$$

(3) 平均绝对百分比误差(Mean Absolute

Percentage Error, MAPE)。

表 1 不同模型预测 2020 年 10 月 15 日-2020 年 11 月 13 日的累计确诊人数

日期	真实数据	组合预测模型	LSTM 模型	Logistic 模型	SEIR 模型
2020 年 10 月 15 日	91 436	91 429.89	91 449.26	91 257.96	91 397.472 5
2020 年 10 月 16 日	91 460	91 457.46	91 476.79	91 257.97	91 397.472 6
2020 年 10 月 17 日	91 490	91 486.17	91 505.34	91 257.97	91 397.472 6
2020 年 10 月 18 日	91 507	91 515.54	91 535.32	91 257.97	91 397.472 6
2020 年 10 月 19 日	91 546	91 536.83	91 557.52	91 257.97	91 397.472 6
2020 年 10 月 20 日	91 565	91 567.57	91 588.16	91 257.97	91 397.472 6
2020 年 10 月 21 日	91 588	91 593.72	91 614.33	91 257.98	91 397.472 6
2020 年 10 月 22 日	91 621	91 616.09	91 637.32	91 257.98	91 397.472 6
2020 年 10 月 23 日	91 653	91 644.73	91 665.96	91 257.98	91 397.472 6
2020 年 10 月 24 日	91 675	91 677.59	91 698.82	91 257.98	91 397.472 6
2020 年 10 月 25 日	91 701	91 704.09	91 726.24	91 257.98	91 397.472 6
2020 年 10 月 26 日	91 725	91 728.23	91 751.15	91 257.98	91 397.472 6
2020 年 10 月 27 日	91 772	91 752.08	91 774.98	91 257.98	91 397.472 7
2020 年 10 月 28 日	91 821	91 791.34	91 813.64	91 257.98	91 397.472 7
2020 年 10 月 29 日	91 852	91 841.15	91 863.77	91 257.98	91 397.472 7
2020 年 10 月 30 日	91 893	91 880.27	91 905.2	91 257.99	91 397.472 7
2020 年 10 月 31 日	91 921	91 916.71	91 943.44	91 257.99	91 397.472 7
2020 年 11 月 1 日	91 955	91 947.52	91 974.54	91 257.99	91 397.472 8
2020 年 11 月 2 日	92 015	91 979.37	92 006.06	91 257.99	91 397.472 8
2020 年 11 月 3 日	92 045	92 030.62	92 056.72	91 257.99	91 397.472 8
2020 年 11 月 4 日	92 077	92 073.34	92 100.88	91 257.99	91 397.472 8
2020 年 11 月 5 日	92 121	92 105.46	92 135.4	91 257.99	91 397.472 8
2020 年 11 月 6 日	92 164	92 143.00	92 173.25	91 257.99	91 397.472 8
2020 年 11 月 7 日	92 195	92 186.58	92 216.2	91 257.99	91 397.472 8
2020 年 11 月 8 日	92 242	92 223.52	92 253.85	91 257.99	91 397.472 8
2020 年 11 月 9 日	92 271	92 264.96	92 296.21	91 257.99	91 397.472 8
2020 年 11 月 10 日	92 299	92 300.03	92 332.07	91 257.99	91 397.472 8
2020 年 11 月 11 日	92 336	92 328.67	92 361.39	91 257.99	91 397.472 8
2020 年 11 月 12 日	92 372	92 361.67	92 394.12	91 257.99	91 397.472 8
2020 年 11 月 13 日	92 404	92 398.55	92 430.7	91 257.99	91 397.472 8

$$\text{MAPE}(y, \hat{y}) = \frac{1}{n} \sum_{i=1}^n \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\% \quad (16)$$

式(14)~(16)中, \hat{y} 为预测值, y 为真实值。在以上三种评价指标中, RMSE、MAE 和 MAPE 的值越小, 模型的预测效果越好。

根据这三个指标, 所提出的组合预测模型与其他三种模型的指标值如表 2 所示。

表 2 表明, Logistic 模型与 SEIR 模型的 RMSE 值、MAE 值和 MAPE 值更大, 说明预测效果较差; 而 LSTM 模型的 RMSE 值、MAE 值和 MAPE 值较小, 说明预测效果较好; 组合预测模型较其他模型误差最小, 说明预测效果最好。

表 2 组合预测模型与单一模型的评价指标

模型	RMSE	MAE	MAPE/%
LSTM 模型	19.768 2	18.381 7	0.049 8
Logistic 模型	587.227 0	501.267 5	0.548 9
SEIR 模型	577.434 6	493.260 6	0.535 7
组合预测模型	10.101 7	7.633 6	0.008 3

3 结束语

为克服单一模型预测精度低的缺点, 在 Logistic 模型和 LSTM 模型预测的基础上, 建立了基于线性回归网络进行训练的组合预测模型, 并预测了之后的累

计确诊人数,其中所采用的 LSTM 模型使用前一时间步的观测值来更新网络状态,使得到的预测值更加准确。采用三种评价指标将组合预测模型与 LSTM 模型、Logistic 模型、SEIR 模型进行评估对比,结果表明所建立的组合预测模型误差更小,预测效果更准确,可以为疫情预测提供科学有效的参考,同时为之后可能发生的公共卫生安全事件提供准确的预测模型。

参考文献:

- [1] 张琳. 新冠肺炎疫情传播的一般增长模型拟合与预测[J]. 电子科技大学学报, 2020, 49(3): 345-348.
- [2] 喻孜, 张贵清, 刘庆珍, 等. 基于时变参数-SIR 模型的 COVID-19 疫情评估和预测[J]. 电子科技大学学报, 2020, 49(3): 357-361.
- [3] LI S, LIN Y, ZHU T, et al. Development and external evaluation of predictions models for mortality of COVID-19 patients using machine learning method[J]. Neural Computing and Applications, 2021(11): 1-10.
- [4] 盛华雄, 吴琳, 肖长亮. 新冠肺炎疫情传播建模分析与预测[J]. 系统仿真学报, 2020, 32(5): 759-766.
- [5] 李伟伟, 杜蓉, 陈曙东, 等. 新型冠状病毒肺炎传播特性分析与疫情发展趋势预测[J]. 厦门大学学报: 自然科学版, 2020, 59(6): 1025-1033.
- [6] 王瑞, 闫方, 逯静, 等. 运用 Dropout-LSTM 模型的新冠肺炎趋势预测[J]. 电子科技大学学报, 2021, 50(3): 414-421.
- [7] HUSSAIN T, OZAIR M, ALI F, et al. Sensitivity analysis and optimal control of COVID-19 dynamics based on SEIQR model[J]. Results in Physics, 2021, 22(8): 103956.
- [8] YUDISTIRA N, SUMITRO S B, NAHAS A, et al. Learning where to look for COVID-19 growth: multivariate analysis of COVID-19 cases over time using explainable convolution LSTM[J]. Applied Soft Computing, 2021, 109(1): 107469.
- [9] 季伟东, 徐浩天, 林平. 自适应变异粒子群优化算法及在新冠肺炎疫情传播预测中的应用[J]. 小型微型计算机系统, 2021, 42(3): 472-477.
- [10] WIBOWO F W, WIHAYATI. Prediction modelling of COVID-19 outbreak in Indonesia using a Logistic Regression model[J]. Journal of Physics: Conference Series, 2021, 1803(1): 012015.
- [11] SONG Jialu. Maximum likelihood-based extended Kalman filter for COVID-19 prediction[J]. Chaos, Solitons & Fractals, 2021, 146: 110922.
- [12] SHAHID F, ZAMEER A, MUNEEB M. Predictions for COVID-19 with deep learning models of LSTM, GRU and Bi-LSTM[J]. Chaos, Solitons & Fractals, 2020, 231(1): 110212.
- [13] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
- [14] 梁天新, 杨小平, 王良, 等. 记忆神经网络的研究与发展[J]. 软件学报, 2017, 28(11): 2905-2924.
- [15] 吕晓玲, 宋捷. 大数据挖掘与统计机器学习[M]. 北京: 中国人民大学出版社, 2016.
- [16] 杨华青. 基于好奇心的营销传播应用研究[D]. 沈阳: 辽宁大学, 2016.
- [17] 王亚奇, 王静. 考虑好奇心理机制的动态复杂网络谣言传播研究[J]. 山东大学学报: 理学版, 2017, 52(6): 99-104.
- [18] 朱冠桦, 蒋国平, 夏玲玲. 社交网络上从众现象对谣言传播影响的研究[J]. 计算机科学, 2016, 43(2): 135-139.
- [19] 林芹, 郭东强. 优化 SIS 模型的社交网络舆情传播研究: 基于用户心理特征[J]. 情报科学, 2017, 35(3): 53-56.
- [20] 王小初, 万佑红. 在线网络中具有反沉默螺旋现象的谣言传播模型[J]. 南京邮电大学学报: 自然科学版, 2016, 36(4): 95-100.
- [21] 迟钰雪, 刘怡君. 逆反心理的网络舆情传播机制研究[J]. 系统工程学报, 2019, 34(5): 610-620.
- [22] 黄骐, 陈春萍, 罗跃嘉, 等. 好奇心的机制及作用[J]. 心理科学进展, 2021, 29(4): 723-736.
- [23] 杨华青. 基于好奇心的营销传播应用研究[D]. 沈阳: 辽宁大学, 2016.
- [24] 赵志朋. 以网络谣言的传播观察从众心理的发展趋势——基于疫情期间的讨论[J]. 牡丹江师范学院学报: 社会科学版, 2020(5): 74-84.
- [25] 唐绪军. 中国新媒体发展报告 No. 12(2021)[M]. 北京: 社会科学文献出版社, 2021.
- [26] 万佑红, 王小初. 考虑从众效应的谣言传播模型[J]. 计算机应用, 2016, 36(9): 2381-2385.
- [27] CAI K, XIE H, LUI J. Information spreading forensics via sequential dependent snapshots[J]. IEEE/ACM Transactions on Networking, 2018, 26(1): 478-491.
- [28] 丁学君, 蒋曼, 田勇, 等. 在线社交网络中考虑抑制代价的谣言澄清策略研究[J]. 系统工程理论与实践, 2021, 41(5): 1119-1137.

(上接第 197 页)

(1): 58-67.