

# 基于智能算法优化的 CNN - LSTM 模型在 手足口病预测中的应用

周浩<sup>1,2</sup>, 董阿莉<sup>3</sup>, 李虹<sup>3</sup>, 康娅楠<sup>1,2</sup>, 杨启越<sup>1,2</sup>, 王星雨<sup>1,2</sup>, 白丽霞<sup>1,2,4</sup>

1. 山西医科大学公共卫生学院流行病学教研室, 山西 太原 030001; 2. 山西医科大学附属儿科医院;

3. 山西省疾病预防控制中心; 4. 山西省儿童医院(山西省妇幼保健院)

**摘要:**目的 分析 LSTM 模型、改进 CNN - BiLSTM 组合模型及智能算法优化在手足口病发病预测预警中的应用, 以探讨预测手足口病发病趋势的最优模型, 为相关部门制定防治措施提供参考。方法 收集山西省疾病预防控制中心 2009 年 1 月至 2019 年 12 月发布的山西省手足口病月度发病人数和《山西省统计年鉴 2008—2020》发布的年末常住人口数据, 据此测算出 2009 年 1 月至 2019 年 12 月山西省手足口病的月度发病率数据; 以 2009 年 1 月至 2019 年 12 月山西省手足口病月度发病率数据作为样本建模数据分别在 MATLAB 7.6 软件构建相应的模型, 对比各个模型的预测效果, 根据误差值越小精度越高的原理, 选出最优模型。结果 使用 RMSE、MAE 误差指标对比不同模型在预测手足口病发病趋势中的预测效能, 结果显示, 改进的 CNN - BiLSTM 组合模型的预测效能优于单一的 LSTM 模型, 而基于智能算法优化的 CNN - BiLSTM 模型明显优于未优化的 CNN - BiLSTM 组合模型, 即 CNN - BiLSTM - PSO/GAPSO/SSA 的 RMSE、MAE 的值(1.943 3、1.309 7; 1.879 2、1.240 2; 1.419 5、1.169 1) 小于对应的 CNN - BiLSTM 模型(2.066 3、1.390 8); 其中, CNN - BiLSTM - SSA 组合模型表现最优。结论 基于单一时间序列预测模型(LSTM 模型)与 CNN - BiLSTM 组合模型相比, CNN - BiLSTM 组合模型的预测效果明显优于单一模型; 对其 CNN - BiLSTM 组合模型进行智能算法(PSO/GAPSO/SSA 算法)优化改进可发现, 基于智能算法优化的组合模型明显优于未优化前, 且 SSA 算法优化的 CNN - BiLSTM 模型效果更佳, 具有较好的预测性能和精度, 可用于山西省未来 HFMD 发病率的实时预测。

**关键词:**手足口病; LSTM 模型; CNN - LSTM 组合模型; 粒子群算法; 麻雀搜索算法; GAPSO 混合算法

中图分类号: R181.2; R183.4 文献标志码: A 文章编号: 1003 - 8507(2024)08 - 1364 - 07

DOI: 10.20043/j.cnki.MPM.202312044

## Applications of CNN - LSTM model based on intelligent algorithm optimization in the prediction of hand, foot and mouth disease

ZHOU Hao\*, DONG A-li, LI Hong, KANG Ya-nan, YANG Qi-yue, WANG Xing-yu, BAI Li-xia

\* Department of Epidemiology, School of Public Health, Shanxi Medical University, Taiyuan, Shanxi 030001, China

**Abstract:** Objective To analyze the application of CNN - BiLSTM combination model and intelligent algorithm optimization in the prediction and early warning of HFMD incidence and to discuss the optimization model for predicting the incidence of HFMD, so as to provide reference for relevant departments to formulate prevention and control measures. **Methods** The monthly incidence data of hand, foot and mouth disease in Shanxi Province from January 2009 to December 2019 and the year - end resident population data released by the Shanxi Statistical Yearbook 2008 - 2020 were collected from January 2009 to December 2019. The monthly incidence data of hand, foot and mouth disease in Shanxi Province from January 2009 to December 2019 were used as sample modeling data to construct the corresponding models in MATLAB 7.6 software, and the prediction effect of each model was compared, and the optimal model was selected according to the principle that the smaller the error value and the higher the accuracy. **Results** By comparing the root mean square error and mean absolute error obtained by predicting the monthly incidence of foot and mouth disease of hand with different models, it can be seen that the CNN - BiLSTM model optimized based on intelligent algorithm is significantly better than the unoptimized CNN - BiLSTM combination model, that is, the values of RMSE and MAE of CNN - BiLSTM - PSO/GAPSO/SSA (1.943 3, 1.309 7; 1.879 2, 1.240 2; 1.419 5, 1.169 1) is smaller than the corresponding CNN - BiLSTM model (2.066 3, 1.390 8), among which the CNN - BiLSTM - SSA combination model performs best. **Conclusion** The CNN - LSTM - SSA model has good predictive performance

基金项目: 山西省医学重点科技计划 (2021XM25)

作者简介: 周浩(1994—), 男, 硕士在读, 研究方向: 传染病流行病学、医院感染及疾病预防

通信作者: 白丽霞, E-mail: bailixia111@163.com

and accuracy in predicting the monthly incidence trend of HFMD, which can be used to predict the future incidence of HFMD in Shanxi Province.

**Keywords:** HFMD; LSTM; CNN - LSTM; PSO; SSA; GAPSO

手足口病(Hand-foot-mouth Disease, HFMD)是一种由多种肠道病毒感染的急性传染病,多发生于学龄前儿童。该病毒传染性强,在外界存活时间长,传播途径复杂,传播速度快,传播范围广,隐性感染比例大,易出现聚集性病例和暴发<sup>[1-3]</sup>。自 2008 年 5 月 2 日起,卫生部明确将手足口病列为丙类传染病管理。近年来,手足口病日趋严重,发病率和死亡率仍在上升,已然成为我国重点关注的公共卫生问题之一。因此,选择适当的方法开展手足口病疫情预测,对手足口病疫情研判和精准防控具有重大指导意义。目前,国内外许多学者在序列的特征提取、预测模型构建等关键技术领域开展了相关研究,其中基于参数的预测方法<sup>[4-6]</sup>和基于浅层机器学习的预测方法<sup>[7-8]</sup>在以往的手足口病发病率预测工作中已多次开展。近年来,深度学习模型已然成为当前研究的热点,广泛应用于时间序列预测领域<sup>[9-11]</sup>。长短时记忆网络模型在传染病领域已广泛应用,但 LSTM 网络本身无法捕捉数据的空间特征,必须人工将空间信息编码作为网络的输入,从而影响预测精度;其次, LSTM 在训练过程中,序列的时空特性会造成模型的自适应学习率误差较大,进而影响预测精度<sup>[12]</sup>。群智能优化算法的出现极大的丰富了模型优化问题的理论研究。群智能优化算法由于不受目标函数的可微,可导影响、连续性和其他性质的约束,本实用新型稳定性更好、具有高效性,收敛快的优点。为了充分提取时间序列的空间特征,本研究提出了 CNN - BiLSTM 组合模型对山西省 2009—2019 年 HFMD 发病情况进行拟合和预测,并结合智能优化算法机制进行该模型的优化训练,以提高模型的预测精度,为疾病防控提供可靠性依据。

## 1 材料与方法

**1.1 资料来源** 数据资料来源于公共卫生科学数据中心 (<https://www.phsciencedata.cn/Share/index.jsp>) 和山西省疾病预防控制中心,按发病日期检索山西省 2009 年 1 月至 2019 年 12 月的手足口病发病数据。山西省人口信息来源于《山西省统计年鉴》。

### 1.2 原理

**1.2.1 LSTM 模型** 长短时记忆网络(Long short-term memory, LSTM)是循环神经网络(Recurrent neural network, RNN)的一种特殊变体, Hochreiter 和 Schmidhuber 在 RNN 的基础上进行了改进<sup>[13]</sup>,引入门控单元系统,借助门的逻辑控制来决定数据单元是更

新还是丢弃,更好地解决了梯度爆炸和消失问题<sup>[14]</sup>。LSTM 不仅仅包含捕捉短期动态,还能够借助捕捉时间序列成分的长期动态(例如周期性、季节性和长期趋势)来增强 RNN。LSTM 采用 input Gates(输入门)、output Gates(输出门)和 forget Gates(遗忘门)对信息进行选择性控制,适当遗忘历史信息并依据新信息更新细胞状态。其结构如图 1 所示:

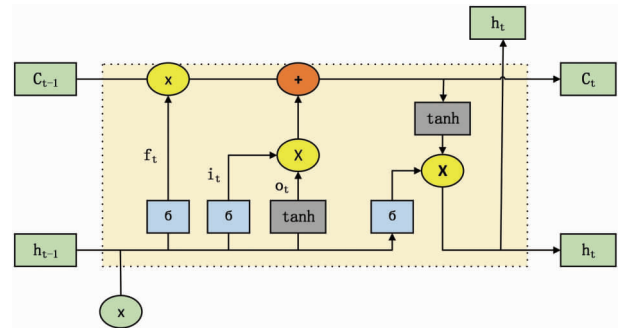


图 1 LSTM 模型

Fig.1 LSTM model

具体公式如下:

$$f_t = \sigma(W_f \cdot [h_{t-1}, X_t] + b_f)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, X_t] + b_i)$$

$$o_t = \sigma(W_o \cdot [h_{t-1}, X_t] + b_o)$$

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

其中  $i_t, f_t, o_t$  分别表示输入门、遗忘门和输出门;  $X_t, h_{t-1}, C_t$  分别表示  $t$  时刻的输入、上一时刻隐层状态和上一时刻记忆细胞状态;  $\sigma$  表示 Sigmoid 函数。

**1.2.2 双向长短期记忆神经网络(Bi-directional Long Short-Term Memory, BiLSTM)** BiLSTM 作为 LSTM 的改进模型,在标准的 LSTM 结构中增加了一层逆向的 LSTM,即由正向 LSTM 和逆向 LSTM 组成,如图 2 所示。BiLSTM 将时序数据分别传向正向 LSTM 和逆向 LSTM 从而可以得到两个不同的隐藏层特征,其训练过程同 LSTM 相同,分为信息的正向传播和误差的反向传递,之后通过线性融合得到最终结果。其优势在于同时考虑了过去与未来两个方向的信息,能更好的学习时间序列的特性和规律。

**1.2.3 CNN 模型** 卷积神经网络(convolutional neural network, CNN)是一类以卷积运算为核心的深度前馈神经网络,它拥有卷积层(convolutional layer)和池化层(pooling layer)构成的特征提取器。在卷积层中,每个神经元只连接到前一个输出层的一些神经

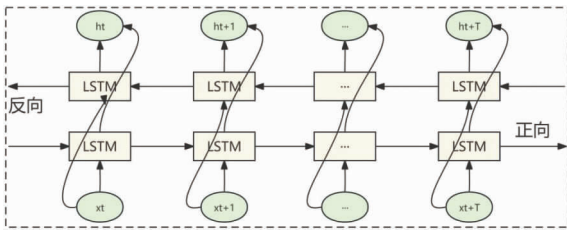


图 2 BiLSTM 网络单元结构

Fig. 2 Structure of the BiLSTM network unit

元,并生成多个特征图,卷积核大多数情况采取使用权重共享来减少卷积模型的参数量,同时保证卷积运算的平移不变性;池化层局部整合了卷积操作的激活区域,常采取使用最大池化或平均池化两种策略,减少中间隐藏层的维数,减少下一层的计算量并予以旋转不变性<sup>[15]</sup>。

**1.2.4 CNN - LSTM 组合模型** 尽管 CNN 模型能够对每个时间序列进行卷积操作,且对时间序列的局部特征具有较好的提取效果,但对时间序列特征并不敏感;而单独使用 LSTM 模型可以更好的提取出非线性数据信息,但无法得到时间序列的空间特征。因此采用 CNN 与 LSTM 相结合的方法,充分利用 CNN 的特征提取能力和 LSTM 对时间序列数据敏感性的特点,进一步改善模型的参数预测效果。CNN - LSTM 网络模型主要由两部分组成:首先,通过 CNN 网络模型的卷积和池化操作确定数据的输入,以实现数据特征的提取和降维;LSTM 网络模型的遗忘门、输入门和输出门通过大量数据的连续迭代训练调整自身参数,使它能从 CNN 网络中提取数据信息间的拟合关系,以便有效地动态输入和预测时间序列建模的输出数据;最后,通过 CNN - LSTM 网络拟合训练,数据输出预测值通过全连接神经网络连接,整个预测过程需要通过数据训练以确定网络模型参数<sup>[16-17]</sup>。CNN - LSTM 组合模型如图 3 所示。CNN - LSTM 网络模型的训练流程如图 4 所示。

**1.2.5 智能优化算法** 为了加快权重的拟合并提高网络输出的准确性和鲁棒性,为此引入智能优化算法。GA 算法和 PSO 算法都是经典的智能进化算法,PSO 算法简单,收敛速度快,但容易早熟,而 GA 算法全局搜索能力强,但收敛速度慢<sup>[18]</sup>。该方法不仅解决了单个 PSO 方法容易陷入局部最优化的缺陷,同时也解决了单个遗传优化方法在收敛速度上的不足,从而大大提高了 PSO 和 GA 的各阶段的综合性能,同时也大大提高了算法的整体稳定性。麻雀搜索算法 (SSA) 通过模拟麻雀的某些行为特征并将其应用于优化算法来解决全局优化问题,并对具有大量局部最优的实际问题提供一种全新的求解途径和方法,具有

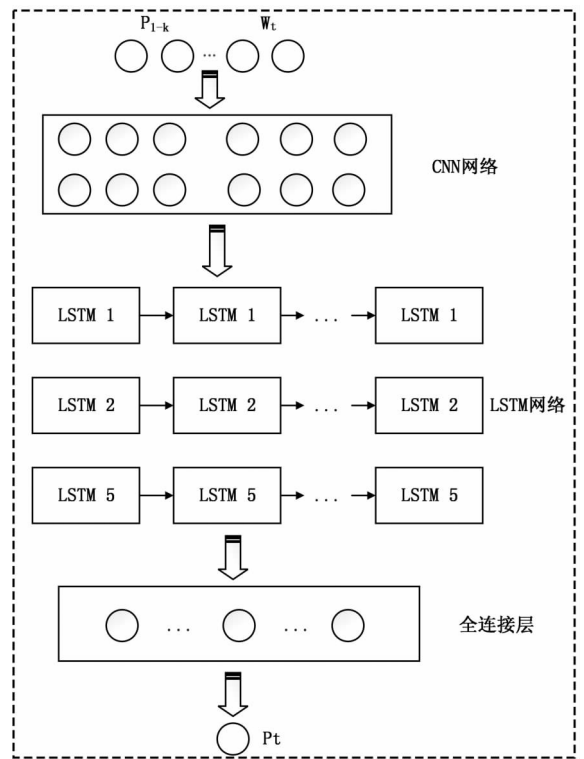


图 3 基于 CNN - LSTM 网络的手足口病发病率预测模型结构

Fig. 3 Structure of HFMD incidence prediction model based on CNN - LSTM network

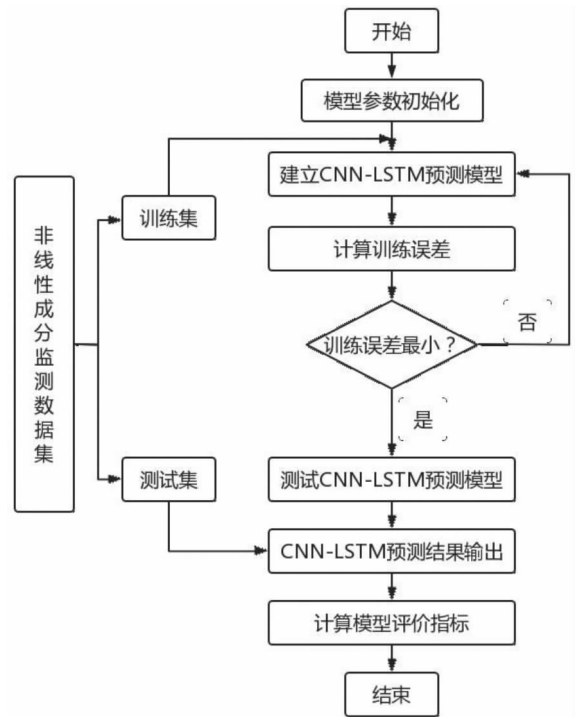


图 4 CNN - LSTM 网络的训练流程图

Fig. 4 Training flowchart of CNN - LSTM network

稳定性强、鲁棒性强、收敛速度快的特点<sup>[19]</sup>。

**1.3 模型评价指标** 本研究采用平均绝对误差

(Mean Absolute Error, MAE) 和均方根误差 (Root Mean Square Error, RMSE) 作为误差评估指标,来量化地分析该模型的预测结果,公式如下:

$$MAE = \frac{1}{n} \sum_{i=1}^n abs(dd(i) - tt(i))$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n abs(dd(i) - tt(i))^2}{n}}$$

**1.4 统计软件** 采用 Excel 2021 软件整理数据,双人录入双核验的方式。LSTM 网络模型、CNN - BiLSTM 组合模型及智能优化算法编程均采用 MATLAB 7.6 软件。

## 2 结果

### 2.1 发病率分析

**2.1.1 年发病率分析** 2009—2019 年山西省累积报告发病 283 903 例,年均发病率为 6.116 6/10 万,2009 年发病率最高(8.122 7/10 万),2016 年发病率最低(3.990 5/10 万)。

**2.1.2 月发病率分析** 山西省 HFMD 呈现明显的季节性流行特征,每年发病呈双峰变化趋势,主高峰在 2017、2018 年为 7 月,其余均为 6 月,而次高峰为 10—11 月,见图 5。

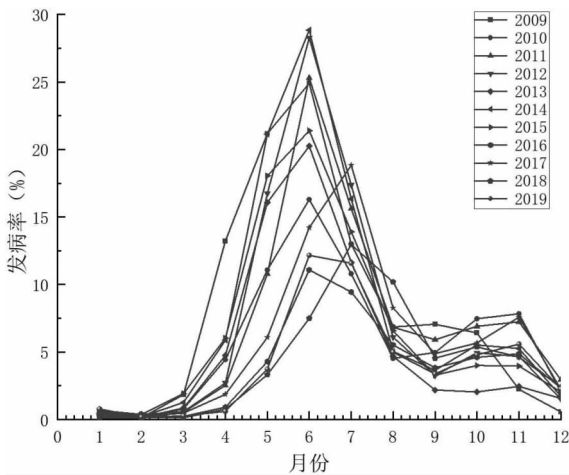


图 5 2009—2019 年山西省手足口病发病趋势图

Fig. 5 Incidence trend of hand, foot and mouth disease in Shanxi Province from 2009 to 2019

**2.2 模型构建** 历史研究中使用 SARIMA 通过若干次差分使其成为平稳序列,会造成信息损失的不足。而本文使用对数据无平稳性要求且考虑时间相关性的 LSTM 模型及其改进模型 CNN - BiLSTM 对山西省手足口病的发病情况进行预测。由于合适的模型参数对预测性能有较大影响,因此,首先对模型的时间步长(窗口长度)、隐藏层层数及节点数进行调参以获

得最优结果。鉴于不同的窗口长度对应不同的样本长度,在确定最佳窗口大小后,对山西省 2009 年 1 月至 2019 年 12 月的手足口病数据进行重构,并将重构数据分为训练集和测试集。本研究选取前 80% 的数据作为训练集,剩余数据作为测试集来验证模型性能,其中训练集的后 12 个数据作为内部验证集来判断模型是否合适。同时为了比较 LSTM、CNN - BiLSTM 模型对同一数据集的预测性能,本文将其设置为相同参数,并将 LSTM 作为调参模型。

**2.2.1 窗口大小设置** 因 HFMD 的数据存在明显的季节周期性,因此我们将窗口长度设置为 6 个月、12 个月。由于不同窗口长度会影响样本量,原始数据经归一化及数据重构后将其转换为 LSTM 模型输入要求的 3D 数据格式,其对应的输入数据格式如表 1 所示。结果显示,当窗口大小为 12 时, RMSE、MAE 在不同隐藏层下都基本到达了最低点,意味着当时间序列存在周期性变化时,将其一个周期的数据作为输入时,模型预测性能达到最优。因此最终窗口长度确定为 12。由此,可以确定输入层节点为 12,以预测下一个月手足口病的发病率,即输出层节点数为 1。

表 1 不同窗口长度下的模型预测误差比较

Table 1 Comparison of model prediction errors under different window lengths

不同窗口长度	6 个月	12 个月
训练样本	(99, 6, 1)	(93, 12, 1)
测试样本	(21, 6, 1)	(15, 12, 1)
指标	RMSE	2.707 7
	MAE	1.802 4
		2.308 3
		1.448 2

**2.2.2 隐藏层层数及节点数选择** 本研究是在单隐层的结构下以确定隐藏层节点数。由于手足口病数据相对来说数据量不大,具体采用以隐藏层节点数为 2 的幂次方进行试验,见表 2。结果显示,当隐藏层节点为 128 时,模型的 RMSE、MAE 两个评价指标均最小。表明在其它参数固定的情况下,模型在单隐层、节点数为 128 时,即可达到最优预测性能。

综上,本研究最终选择时间步长为 12,即输入层节点数为 12,预测下个月 HFMD 发病率,输出层节点数为 1,隐藏层数为 1,隐藏层节点数为 128,迭代次数为 500 次,采用自适应学习率并将初始值设置为 0.001, Adam 优化器进行模型预测。实验过程中,为了防止训练过拟合,训练中采用 L2 正则化技术。

**2.3 拟合结果分析** 不同模型训练集和测试集下对应的山西省 HFMD 月发病率的预测结果与实际值对比如图 6 所示。可以看出,基于智能算法优化的 CNN - BiLSTM 模型明显优于未优化的 CNN - BiLSTM 组合模型。

表 2 单隐层下不同节点数的模型预测误差比较

Table 2 Comparison of model prediction errors with different numbers of nodes under a single hidden layer

指标	节点数(单隐层)						
	2	4	8	16	32	64	128
RMSE	4.873 7	2.303 6	3.023 1	4.313 3	5.210 9	2.198 2	1.637 2
MAE	4.538 9	1.727 7	1.984 9	3.177 9	3.018 4	1.484 1	1.157 0

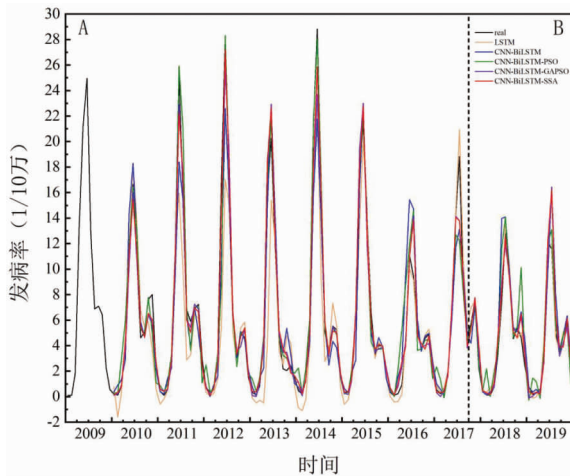


图 6 5 种模型训练集(A)和测试集(B)预测值与实际值的对比图

Fig. 6 Comparison of the predicted and actual values of the five model training sets (A) and test sets (B)

2.4 模型性能评估 为客观评价模型预测效果,使用 RMSE、MAE 进行不同模型间的对比(见表 3)。由表 3 可知,改进的 CNN - BiLSTM 组合模型的预测效能优于单一的 LSTM 模型,且相比 CNN - BiLSTM 组合模型而言,本研究提出的智能算法优化的 CNN - BiLSTM 组合模型预测性能均有所提高,其预测效果明显优于未优化前的 CNN - BiLSTM 组合模型,其中 CNN - BiLSTM - SSA 模型的 RMSE、MSE 分别为 1.4195、1.1691,其预测性能相比于 LSTM 分别提高了 31.3%、15.9%,CNN - BiLSTM - SSA 组合模型预测效果最优。结论证实对于具有非平稳、非线性的特征的序列,CNN - BiLSTM - SSA 模型预测 HFMD 月发病趋势具备较高的准确率及可行性,可用于山西省未来 HFMD 发病率的实时预测。

表 3 不同模型下预测误差比较

Table 3 Comparison of prediction errors under different models

	RMSE			MAE		
	RMSE	变化量	改善率(%)	RMSE	变化量	改善率(%)
LSTM	2.308 3	-	-	1.448 2	-	-
CNN - BiLSTM	2.066 3	-	-	1.390 8	-	-
CNN - BiLSTM - PSO	1.943 3	-0.123 0	5.900 0	1.309 7	-0.081 1	5.800 0
CNN - BiLSTM - GAPSO	1.879 2	-0.187 1	9.100 0	1.240 2	-0.150 6	10.800 0
CNN - BiLSTM - SSA	1.419 5	-0.646 8	31.300 0	1.169 1	-0.221 7	15.900 0

注: - 为参照值,即基于 PSO/GAPSO/SSA 算法优化的 CNN - BiLSTM 组合模型相比。

### 3 讨论

建立健全的传染病预测预警机制,根据病种对模型进行实时调整和优化,不断提高预测精度是当前疾病监测工作中的首要任务。对有季节性和趋势性的时间序列,ARIMA 模型能够有效地从数据中提取线性信息,很好地预测了时间序列的自相关性及季节性,但是对于无规律、波动较大的资料或序列长期预测结果并不理想<sup>[20-21]</sup>。而 LSTM 不仅能够对非线性数据进行建模,也可以有效利用序列的历史信息<sup>[22]</sup>。研究显示,LSTM 模型在手足口病的月发病趋势预测研究中已在部分地区广泛应用<sup>[23-24]</sup>。与 SARIMA 模型相比可知,LSTM 模型的拟合预测效果优于 SARIMA 模型<sup>[25-26]</sup>。后来,许多学者使用 LSTM 神经

网络模型与其他模型组合的方法来预测传染病的发病率或流行情况,并取得了良好的效果,比如:LSTM 模型结合 SARIMA 模型对新冠<sup>[27]</sup>、肺结核<sup>[28]</sup>的发病情况进行预测,EEMD - LSTM 组合模型预测禽流感的发病情况<sup>[29]</sup>等。

目前,CNN 可较好的提取数据的空间特征<sup>[30]</sup>。GAPSO 混合优化算法在处理回归问题上具有较高的准确性,但是对于没有显著倾向的传染性疾病却没有很好的效果<sup>[18]</sup>。麻雀搜索算法(SSA)可解决全局优化问题,但对具有大量局部最优的实际问题具有稳定性强、鲁棒性强、收敛速度快的特点。CNN - LSTM 组合模型<sup>[31]</sup>及其结合 PSO 算法<sup>[32]</sup>及 GA 算法<sup>[33]</sup>虽然在某些领域已广泛应用,但 GAPSO 智能算法优化的 CNN - BiLSTM 组合模型在传染病预测中未见报道。

研究显示,单一预测模型及其优化算法<sup>[32]</sup>在对传染病预测分析时,其结构简单易于实现,而组合模型与智能优化算法结合时,可明显提升了传染病的预测精度。

由于受新冠肺炎的冲击,山西省 2020—2021 年度 HFMD 报告的 HFMD 报告病例数与往年比较,其数据的质量和稳定性都有所下降,从而使建立的模型存在很大的偏差,因此,为了减小误差,本研究选取 2009—2019 年山西省 HFMD 月发病率资料进行研究并构建相关模型,结果显示:基于智能算法优化的 CNN - BiLSTM 模型明显优于未优化的 CNN - BiLSTM 组合模型,而基于麻雀搜索(SSA)算法优化的 CNN - BiLSTM 组合模型预测效果明显优于 PSO、GAPSO 算法优化的 CNN - BiLSTM 组合模型,从而提升模型的预测效果。本研究在表明:智能优化算法通过获取神经网络的局部最优权重和阈值,可有效地提升 HFMD 月发病趋势预测结果的精度,且 SSA 比 PSO/GAPSO 算法对模型的优化效果更好,但是运算过程耗时更长。

本研究存在一定的局限性:不同病毒因流行周期不同,对 HFMD 的发生也会发生一定的影响;其次, HFMD 的传播、易感人群的保护等方面也会受所在地区经济状况等多种社会因素的影响。因此,未来对此进行深入分析并纳入模型,可能具有更好的预测效果,也有助于了解和监测 HFMD 的传播,减少感染的风险。

综上,本研究显示 CNN - BiLSTM - SSA 模型预测精度较高,具备较高的准确性及可行性,可用于山西省未来 HFMD 发病率的实时预测,为传染病的预测预警提供了新思路和有效方法,为制定 HFMD 防控策略提供科学的参考依据。随着深度学习算法操作能力的日益普及,越来越复杂、越来越精确的建模方法将在各个领域不断进行创新和发展,未来可进一步使用不同的数学模型,并结合全国或不同地区的传染病发病数据,建立预测性能更优的传染病监测系统。

**利益冲突声明** 本研究不存在任何利益冲突

## 参考文献

[1] Liu YX, Wang XJ, Pang CK, et al. Spatio-temporal analysis of the relationship between climate and hand, foot, and mouth disease in Shandong province, China, 2008 - 2012 [J]. BMC Infectious Diseases, 2015, 15: 146.

[2] 王雅婷,朋文佳,苏华林,等. 2011—2018 年中国手足口病发病的时空特征及影响因素研究[J]. 中华流行病学杂志,2022,43(10):1562 - 1567.

Wang YT, Peng WJ, Su HL, et al. Spatiotemporal characteristics of hand, foot and mouth disease and influencing factors in China from 2011 to 2018 [J]. Chinese Journal of Epidemiology, 2022, 43

(10): 1562 - 1567.

[3] 张静,李秀惠,李丽,等. 手足口病病原学和流行病学研究进展 [J]. 中华流行病学杂志,2022,43(5):771 - 783.

Zhang J, Li XH, Li L, et al. Etiology and epidemiology of hand, foot and mouth disease in China [J]. Chinese Journal of Epidemiology, 2022, 43(5): 771 - 783.

[4] Liu L, Luan RS, Yin F, et al. Predicting the incidence of hand, foot and mouth disease in Sichuan province, China using the ARIMA model [J]. Epidemiology and Infection, 2016, 144(1): 144 - 151.

[5] Meng D, Xu J, Zhao JJ, et al. Analysis and prediction of hand, foot and mouth disease incidence in China using Random Forest and XGBoost [J]. PLOS One, 2021, 16(12): e261629.

[6] Yu CC, Xu CJ, Li YH, et al. Time series analysis and forecasting of the Hand - Foot - Mouth disease morbidity in China using an advanced exponential smoothing state space TBATS model [J]. Infection and Drug Resistance, 2021, 14: 2809 - 2821.

[7] Yang YS, Lu YL, Fang LY, et al. [Influence of meteorological factors on HFMD and construction of prediction model in Sichuan province] [J]. Sichuan Da Xue Xue Bao Yi Xue Ban, 2020, 51(5):685 - 690.

[8] 王牧雨,王妮,周阳,等. 基于时间序列相似性的患者结局预测模型 [J]. 北京生物医学工程,2022,41(3):249 - 254.

Wang MY, Wang N, Zhou Y, et al. Predictive model for patient outcomes based on time series similarity [J]. Beijing Biomedical Engineering, 2022, 41(3): 249 - 254.

[9] Gu JY, Liang LZ, Song HQ, et al. A method for hand - foot - mouth disease prediction using GeoDetector and LSTM model in Guangxi, China [J]. Scientific Reports, 2019, 9(1): 17928.

[10] Wang YB, Xu CJ, Zhang SK, et al. Development and evaluation of a deep learning approach for modeling seasonality and trends in hand - foot - mouth disease incidence in mainland China [J]. Scientific Reports, 2019, 9(1): 8046.

[11] Yoshida K, Fujimoto T, Muramatsu M, et al. Prediction of hand, foot, and mouth disease epidemics in Japan using a long short - term memory approach [J]. PLOS One, 2022, 17(7): e0271820.

[12] 李磊,张青苗,赵军辉,等. 基于改进 CNN - LSTM 组合模型的分时段短时交通流预测 [J]. 应用科学学报,2021,39(2):185 - 198.

Li L, Zhang QM, Zhao JH, et al. Short - Term traffic flow prediction method of different periods based on improved CNN - LSTM [J]. Journal of Applied Sciences, 2021, 39(2): 185 - 198.

[13] Hochreiter S, Schmidhuber J. Long short - term memory [J]. Neural Computation, 1997, 9(8): 1735 - 1780.

[14] 马亭亭,冀天娇,杨冠羽,等. 基于长短时记忆神经网络的手足口病发病趋势预测 [J]. 计算机应用,2021,41(1):265 - 269.

Ma TT, Ji TJ, Yang GY, et al. Incidence trend prediction of hand - foot - mouth disease based on long short - term memory neural network [J]. Journal of Computer Applications, 2021, 41(1): 265 - 269.

- Zhang XY, Zhuang SL, Yu XM, et al. Analysis of mortality level and disease burden of lung cancer among residents in Pukou District of Nanjing City from 2010 to 2017 [J]. Chinese Journal of Health Statistics, 2023, 40(2): 282-284.
- [22] 党静,马新颜,段宇帆. 2019 年河北省石家庄市肿瘤登记地区肺癌发病、死亡情况及 2014—2019 年变化趋势分析[J]. 中国肿瘤, 2023, 32(12): 891-898.
- Dang J, Ma XY, Duan YF. Incidence and mortality of lung cancer in Shijiazhuang cancer registration areas in 2019 and trend from 2014 to 2019 [J]. China Cancer, 2023, 32(12): 891-898.
- [23] Zheng RS, Zhang SW, Zeng HM, et al. Cancer incidence and mortality in China, 2016 [J]. Journal of the National Cancer Center, 2022, 2(1): 1-9.
- [24] 罗鹏飞,林萍,周金意. 肺癌与大气污染关系的流行病学研究进展[J]. 中国肿瘤, 2017, 26(10): 792-797.
- Luo PF, Lin P, Zhou JY. Progress on epidemiological studies of the relationship between lung cancer and ambient air pollution [J]. China Cancer, 2017, 26(10): 792-797.
- [25] Eckel SP, Cockburn M, Shu YH, et al. Air pollution affects lung cancer survival [J]. Thorax, 2016, 71(10): 891-898.
- [26] Godfrey KM, Barker DJ. Fetal nutrition and adult disease [J]. The American Journal of Clinical Nutrition, 2000, 71(5 Suppl): 1344S-1352S.
- [27] Gutiérrez OM, Muntner P, Rizk DV, et al. Dietary patterns and risk of death and progression to ESRD in individuals with CKD: a cohort study [J]. American Journal of Kidney Diseases, 2014, 64(2): 204-213.
- [28] 犹忆,李德俊,裴一霖,等. 1990—2019 年中国结直肠癌死亡趋势及年龄-时期-队列模型分析[J]. 现代预防医学, 2023, 50(18): 3281-3287.
- You Y, Li DJ, Pei YL, et al. Mortality trends and age-period-cohort model analysis of colorectal cancer in China, 1990-2019 [J]. Modern Preventive Medicine, 2023, 50(18): 3281-3287.

收稿日期: 2023-01-24

## (上接第 1369 页)

- [15] 张荣,李伟平,莫同. 深度学习研究综述[J]. 信息与控制, 2018, 47(4): 385-397, 410.
- Zhang R, Li WP, Mo T. Review of deep learning [J]. Information and Control, 2018, 47(4): 385-397, 410.
- [16] Muhammad LJ, Haruna AA, Sharif US, et al. CNN-LSTM deep learning based forecasting model for COVID-19 infection cases in Nigeria, South Africa and Botswana [J]. Health and Technology, 2022, 12(6): 1259-1276.
- [17] 胡海文. 基于机器学习的新冠肺炎疫情影响因素及预测分析[D]. 兰州: 兰州交通大学, 2022.
- Hu HW. Influencing factors and predictive analysis of novel coronavirus pneumonia epidemic based on machine learning [D]. Lanzhou: Lanzhou Jiaotong University, 2022.
- [18] 李红亚,彭忠显,邓楚燕,等. GA 与 PSO 的混合研究综述[J]. 计算机工程与应用, 2018, 54(2): 20-28, 39.
- Li HY, Peng YZ, Deng CY, et al. A review of mixed studies on GA and PSO [J]. Computer Engineering and Applications, 2018, 54(2): 20-28, 39.
- [19] 薛建凯. 一种新型的群智能优化技术的研究与应用: 麻雀搜索算法[D]. 上海: 东华大学, 2019.
- Xue JK. Research and application of a new type of swarm intelligence optimization technology: Sparrow search algorithm [D]. Shanghai: Donghua University, 2019.
- [20] 付之鸥,周扬,陈诚,等. 时间序列分析与机器学习方法在预测肺结核发病趋势中的应用[J]. 中国卫生统计, 2020, 37(2): 190-195.
- Fu ZO, Zhou Y, Chen C, et al. Application of time series analysis and machine learning methods in predicting the incidence of tuberculosis [J]. Chinese Journal of Health Statistics, 2020, 37(2): 190-195.
- [21] Zou JJ, Jiang GF, Xie XX, et al. Application of a combined model with seasonal autoregressive integrated moving average and support vector regression in forecasting hand-foot-mouth disease incidence in Wuhan, China [J]. Medicine, 2019, 98(6): e14195.
- [22] 高秋菊,周宇畅,赵树青,等. ARIMA 乘积季节模型和 LSTM 深度神经网络对石家庄市手足口病疫情预测效果的比较[J]. 中华疾病控制杂志, 2020, 24(1): 73-78.
- Gao QJ, Zhou YC, Zhao SQ, et al. Comparison on predictive capacity of ARIMA model and LSTM model for incidence of hand, foot and mouth disease in Shijiazhuang [J]. Chinese Journal of Disease Control & Prevention, 2020, 24(1): 73-78.
- [23] Zhang R, Guo Z, Meng YJ, et al. Comparison of ARIMA and LSTM in forecasting the incidence of HFMD combined and uncombined with exogenous meteorological variables in Ningbo, China [J]. International Journal of Environmental Research and Public Health, 2021, 18(11): 6174.
- [24] 王瑞,李瑞沂,曹沛根,等. 基于 ARIMA-LSTM 混合模型对传染病的预测分析[J]. 现代信息科技, 2024, 8(1): 116-120.
- Wang R, Li RY, Cao PG, et al. Prediction analysis of infectious diseases based on ARIMA-LSTM mixed model [J]. Modern Information Technology, 2024, 8(1): 116-120.
- [25] 赖晓莹,钱俊. ARIMA-LSTM-XGBoost 加权组合模型在肺结核发病趋势预测的研究[J]. 现代预防医学, 2021, 48(1): 5-9.
- Lai XY, Qian J. ARIMA-LSTM-XGBoost weighted combination model in prediction of tuberculosis incidence trend [J]. Modern Preventive Medicine, 2021, 48(1): 5-9.
- [26] 何振欢,肖建华. 基于 EEMD-LSTM 模型的禽流感预测研究[J]. 动物医学进展, 2022, 43(11): 34-38.
- He ZH, Xiao JH. Research on prediction of fowl cholera based on EEMD-LSTM model [J]. Progress in Veterinary Medicine, 2022, 43(11): 34-38.
- [27] Manoharan SN, Kumar KMVM, Vadivelan N. A novel CNN-TLSTM approach for dengue disease identification and prevention using IoT-Fog cloud architecture [J]. Neural Processing Letters, 2023, 55(2): 1951-1973.
- [28] Zhou W, Li Y, Xiao YJ, et al. The application of automatic identification system information and PSO-LSTM neural network in CRI prediction [J]. Computational Intelligence and Neuroscience, 2022, 2022: 8699322.
- [29] 吴文博,李虹艾,万鹏程,等. 利用遗传算法优化的 ARIMA-BP 组合模型预测手足口病发病趋势[J]. 中南医学科学杂志, 2014, 42(6): 572-576.
- Wu WB, Li HA, Wan PC, et al. Predicting the incidence trend of Hand-Foot-Mouth disease by using the GA optimized ARIMA-BP combination modeling [J]. Medical Science Journal of Central South China, 2014, 42(6): 572-576.

收稿日期: 2023-12-04