

# 基于加权平均樽海鞘群算法和BP神经网络的 COVID-19预测\*

胡红萍<sup>1</sup>, 乔世昌<sup>1</sup>, 孔慧华<sup>1</sup>, 徐乔王<sup>2</sup>, 白艳萍<sup>1</sup>

(1. 中北大学 理学院, 山西 太原 030051; 2. 临汾市财政局 国际金融组织贷款服务中心, 山西 临汾 041000)

**摘要:** 新型冠状病毒肺炎以其高传染性和高致病性成为全球关注的问题之一. 有效预测COVID-19的累计确诊人数对COVID-19的防控具有重要价值. 本文提出加权平均樽海鞘群算法(AVSSA), 通过23个基准函数验证了AVSSA的有效性, 进而利用AVSSA优化BP神经网络建立预测模型AVSSA-BP, 实现COVID-19的预测. 实验结果表明预测模型AVSSA-BP有最小的误差和最高的确定性系数, 验证了AVSSA-BP的有效性.

**关键词:** 新型冠状病毒肺炎; 樽海鞘算法; BP神经网络; 函数优化; COVID-19预测

**DOI:** 10.13568/j.cnki.651094.651316.2021.03.30.0001

**中图分类号:** TP301.6 **文献标识码:** A **文章编号:** 2096-7675(2022)01-0019-07

**引文格式:** 胡红萍, 乔世昌, 孔慧华, 徐乔王, 白艳萍. 基于加权平均樽海鞘群算法和BP神经网络的COVID-19预测[J]. 新疆大学学报(自然科学版)(中英文), 2022, 39(1): 19-25.

**英文引文格式:** HU Hongping, QIAO Shichang, KONG Huihua, XU Qiaowang, BAI Yanping. Prediction of COVID-19 based on the weighted average salp swarm algorithm and BP neural network[J]. Journal of Xinjiang University(Natural Science Edition in Chinese and English), 2022, 39(1): 19-25.

## Prediction of COVID-19 Based on the Weighted Average Salp Swarm Algorithm and BP Neural Network

HU Hongping<sup>1</sup>, QIAO Shichang<sup>1</sup>, KONG Huihua<sup>1</sup>, XU Qiaowang<sup>2</sup>, BAI Yanping<sup>1</sup>

(1. School of Science North University of China, Taiyuan Shanxi 030051, China;

2. Linfen Finance Bureau, International Financial Organization Loan Service Center, Linfen Shanxi 041000, China)

**Abstract:** Corona Virus Disease 2019(COVID-19) is one of the global concerns due to its highly infectious and highly pathogenic coronavirus. It is of great value to effectively predict the cumulative number of confirmed cases of COVID-19 for the prevention and control of COVID-19. In this paper, the weighted average salp swarm algorithm is proposed, named by AVSSA, whose validation is performed by 23 benchmark functions. Then AVSSA is utilized to optimize the parameters of BP neural network to establish the predicted model AVSSA-BP for predicting the COVID-19. The experimental results show that the predicted model AVSSA-BP has the least errors and the highest coefficient of determination. Therefore, the proposed AVSSA is an effective algorithm.

**Key words:** Corona Virus Disease 2019; salp swarm algorithm; BP neural network; function optimization; prediction of COVID-19

## 0 引言

新型冠状病毒肺炎(Corona Virus Disease 2019, COVID-19)由于其高传染性和高致病性, 严重威胁到了人类的生命安全, 并已迅速蔓延到全球, 成为全球关注的问题之一. 自疫情暴发以来, 我国进行了全民总动员, 真正做到“早发现, 早隔离”. 但由于境外输入COVID-19和无症状感染的影响, 疫情防控仍然面临反弹压力. 因此有效地预测COVID-19的累计确诊人数对COVID-19的防控具有非常大的价值.

\* 收稿日期: 2021-03-30

基金项目: 山西省回国留学人员科研资助项目(2020-104); 国家自然科学基金(61774137; 61971381); 山西省重点研发计划项目(201903D121156).

作者简介: 胡红萍(1973-), 女, 博士, 教授, 从事智能算法与机器学习的研究, E-mail: hhp92@163.com.

目前,常用的预测方法有差分自回归移动平均乘积季节模型<sup>[1]</sup>,群智能算法优化神经网络模型<sup>[2-3]</sup>,多尺度组合模型<sup>[4]</sup>,随机森林算法和模糊信息粒化的混合算法<sup>[5]</sup>,经验模态分解和极限学习机的组合预测模型<sup>[6]</sup>,等等.随着群智能算法和机器学习的不断改进,预测模型也随之增加,更利于解决实际问题.如用萤火虫算法优化长短时记忆神经网络,并利用其预测生猪价格<sup>[2]</sup>;将粒子群算法和差分进化相结合优化自适应小波神经网络的初始化网络参数,进而对风电功率超短期多步预测<sup>[3]</sup>;利用改进的灰狼算法优化Elman神经网络,对基于基因表达谱的乳腺癌进行识别<sup>[7]</sup>;利用改进粒子群算法优化BP神经网络对甜菜产量进行预测<sup>[8]</sup>等.

最近很多学者将群智能算法和机器学习相结合对COVID-19进行预测.例如,文献[9]利用机器学习、自适应神经模糊推理系统和增强甲虫天线搜索群智能元启发式相结合的混合方法预测COVID-19确诊人数;文献[10]利用递归神经网络对COVID-19确诊人数进行预测;文献[11]利用集合经验模态分解和自回归移动平均模型对COVID-19全球预测系统预测结果进行改进.

樽海鞘群算法(Salp Swarm Algorithm, SSA)是受深海中樽海鞘的航行和觅食行为启发提出的<sup>[12]</sup>.自SSA提出以来,研究者已对SSA做了很多改进.例如,将SSA中樽海鞘链的领头者的搜索方式修改为贪婪的交叉选择方法,加强了SSA的性能且避免陷入局部最优<sup>[13]</sup>;将正余弦算法和干扰算子加到SSA得到了一种特征选择方法<sup>[14]</sup>;将混沌映射加入SSA中代替随机数建立了混沌SSA(CSSA),并且将CSSA与2维曲波变换和深度学习技术相结合用于X射线图像对COVID-19进行诊断<sup>[15]</sup>;将SSA的樽海鞘的位置更新修改为Singer的混沌映射,在开发阶段使用局部搜索算法,得到动态的SSA<sup>[16]</sup>.

本文针对其他樽海鞘对链中第*i*个跟随者均有影响的问题,提出加权平均樽海鞘群算法(Average Salp Swarm Algorithm, AVSSA),并通过23个基准函数验证AVSSA的有效性,进而用AVSSA优化BP神经网络的权值和偏差,建立预测境外输入COVID-19的预测模型AVSSA-BP.

## 1 樽海鞘群算法

樽海鞘群算法是受深海中樽海鞘的航行和觅食行为的启发提出的<sup>[12]</sup>.樽海鞘有透明的桶形身体,其组织非常类似于在海洋中发现的水母,它们像水母一样通过身体泵水前进.SSA是从樽海鞘的群体行为中获得的,它是以樽海鞘群体构成一个樽海鞘链,在该链中,第一个樽海鞘为领导者,其余的樽海鞘为跟随者.

设*F*是*n*维搜索空间中的食物源,将其作为樽海鞘链的目标.与*F*相关的领导者的更新位置的方式为:

$$x_i^1 = \begin{cases} F_i + c_1((ub_i - lb_i)c_2 + lb_i) & \text{若 } c_3 \geq 0.5 \\ F_i - c_1((ub_i - lb_i)c_2 + lb_i) & \text{若 } c_3 < 0.5 \end{cases} \quad (1)$$

其中: $x_i^1$ 和 $F_i$ 分别表示领导者 $x^1$ 和食物源*F*位置的第*i*个分量, $ub_i$ 和 $lb_i$ 分别表示上界*ub*和下界*lb*的第*i*个分量, $c_2$ 和 $c_3$ 都是服从区间[0,1]上的均匀分布的随机数, $c_1$ 用来平衡SSA中探索阶段和开发阶段,定义为:

$$c_1 = 2e^{-\left(\frac{4l}{L}\right)^2} \quad (2)$$

其中: $l$ 是当前迭代次数, $L$ 是迭代的最大次数.

根据牛顿运动定律,跟随者 $x^s$ 的位置更新方式为:

$$x_i^s = \frac{1}{2}at^2 + v_0t \quad (3)$$

其中: $s > 2$ , $x_i^s$ 是第*s*个跟随者位置的第*i*个分量, $v_0$ 是初速度, $a = \frac{v_{final}}{v_0}$ , $v = \frac{x-x_0}{t}$ .

由于优化的时间是迭代*t*,迭代之间的差异为1,在 $v_0=0$ 时,方程(3)变为:

$$x_i^s = \frac{1}{2}(x_i^s + x_i^{s-1}). \quad (4)$$

## 2 预测模型

### 2.1 加权平均樽海鞘群算法

在SSA中,第*i*个跟随者的位置更新方式定义为方程(4),为第*i*个跟随者和第*i*-1个跟随者的位置的平均值,这可能导致樽海鞘链中前*i*个跟随者对第*i*个跟随者的影响不大.为此,将樽海鞘链中第*i*个跟随者的位置更新改

进为前*i*个跟随者的加权平均, 即:

$$x_i^s = (1 - c_1) \frac{1}{s} \sum_{k=1}^s x_i^k \quad (5)$$

其中:  $c_1$ 与方程(2)中的定义相同. 我们将此改进称为加权平均SSA, 记为AVSSA.

为验证所提算法AVSSA的有效性, 选取了5个比较的优化算法: 蚁狮优化器(ALO)、灰狼优化器(GWO)、飞蛾扑火优化(MFO)算法、鲸优化算法(WOA)和樽海鞘群算法(SSA). 本文所采用的23个基准函数 $F_1(x) \sim F_{23}(x)$ 来源于文献[17], 其中:  $F_1(x) \sim F_7(x)$ 为7个*n*维单峰函数,  $F_8(x) \sim F_{13}(x)$ 为6个*n*维多峰函数,  $F_{14}(x) \sim F_{23}(x)$ 为10个固定维数的基准函数. 在Win10 64bit专业平台和64 GB RAM下Matlab R2019a工具箱下运行这6个算法.

在这6个可比较的算法中, 种群大小为50, 最大迭代次数 $T = 500$ . 每个可比较的算法独立运行30次. 将这30次独立运行的每个算法所获得的全局最小值的平均值(Avg.)和标准差(Std.)作为评价算法的性能指标.

设*t*是当前迭代次数, *T*为最大迭代次数. 在ALO中, 比率 $I = 10^{\omega \frac{t}{T}}$ , 其中

$$\omega = \begin{cases} 2 & \text{若 } t > 0.1T \\ 3 & \text{若 } t > 0.5T \\ 4 & \text{若 } t > 0.75T \\ 5 & \text{若 } t > 0.9T \\ 6 & \text{若 } t > 0.95T \end{cases} \quad (6)$$

在GWO中, 收敛参数 $a = 2 - \frac{2t}{T}$ ; 在MFO中, 收敛参数 $a = -1 - \frac{t}{T}$ , 螺旋系数 $b = 1$ ; 在WOA中, 收敛参数 $a = 2 - \frac{2t}{T}$ ; 在SSA和AVSSA中, 控制参数 $c_1 = 2e^{-\left(\frac{4t}{T}\right)^2}$ .

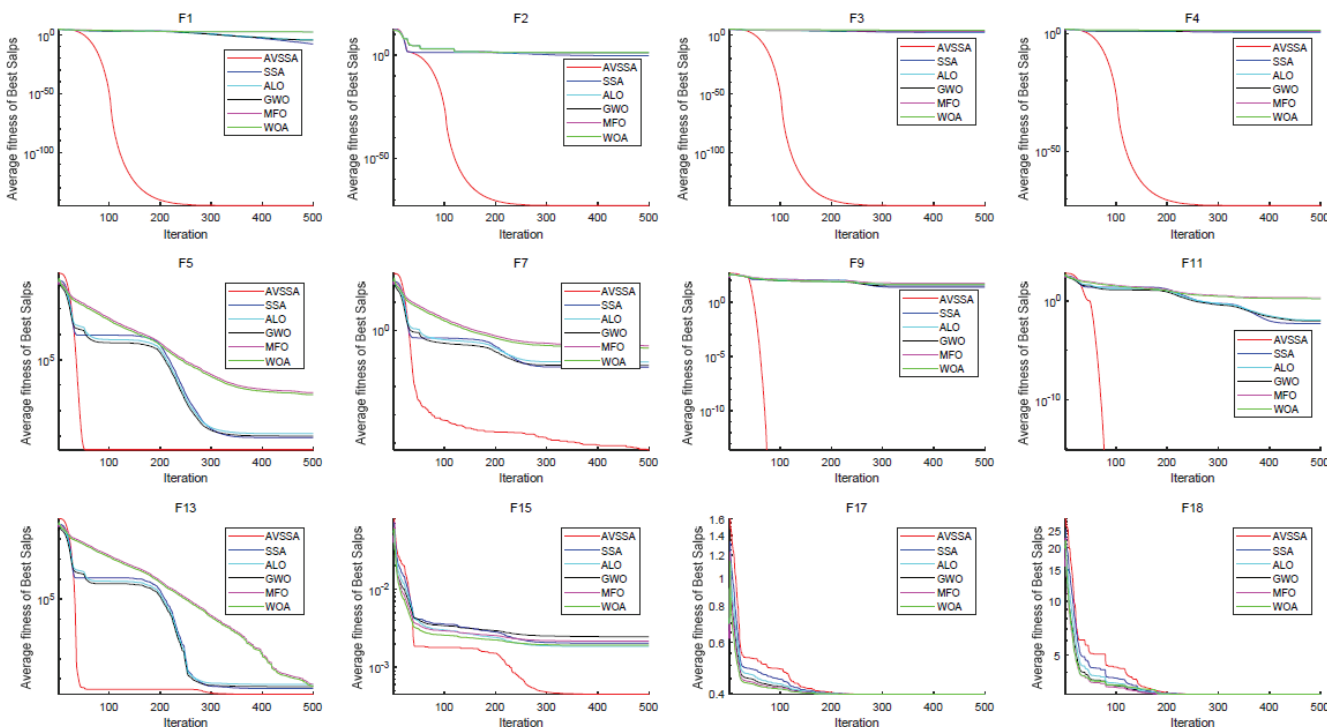


图 1 部分基准函数的收敛曲线

注: 横坐标表示迭代次数, 纵坐标表示最优樽海鞘的平均适应度值.

独立运行这6个算法30次, 得到了23个基准函数的最小函数值的Avg.和Std., 如表1所示. 部分基准函数的收敛曲线如图1所示. 由图1可见, AVSSA有最好的收敛性能. 由表1可知, 利用AVSSA得到单峰函数 $F_1(x) \sim F_4(x)$ ,  $F_7(x)$ 的Avg.分别为 $1.18 \times 10^{-145}$ 、 $1.61 \times 10^{-73}$ 、 $2.66 \times 10^{-145}$ 、 $1.15 \times 10^{-73}$ 和 $5.81 \times 10^{-5}$ , 这些都最接近全局最小值0, 而AVSSA和SSA在函数 $F_6(x)$ 上有相近的Avg.. AVSSA在多峰函数 $F_9(x)$ 和 $F_{11}(x)$ 上的Avg.达到全局最

小值0, 在函数 $F_7(x)$ 、 $F_{10}(x)$ 和 $F_{12}(x)$ 的Avg.分别是 $5.81 \times 10^{-5}$ 、 $8.88 \times 10^{-16}$ 和 $4.40 \times 10^{-3}$ , 这些都最接近全局最小值0, 且在多峰函数 $F_{13}(x)$ 上的Avg.与其它可比较的算法具有相同的数量级. 对于固定维数的函数 $F_{16}(x) \sim F_{19}(x)$ , 所有可比较的算法AVSSA、SSA、ALO、GWO、MFO和WOA都达到了全局最小值-1.031 6、-3、-0.398和-3.86, 但AVSSA有最小的Std.. 对于固定维数的函数 $F_{22}(x)$ , AVSSA得到了最小的Avg. -8.326 6, 最接近于全局最小值-10.402 8.

表 1 23个基准函数的Avg.和Std.

算法	$F_1(x)$		$F_2(x)$		$F_3(x)$		$F_4(x)$		$F_5(x)$	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
AVSSA	1.18	1.21	1.61	7.47	2.66	7.76	1.15	7.03	2.85	4.89
	$\times 10^{-145}$	$\times 10^{-146}$	$\times 10^{-73}$	$\times 10^{-75}$	$\times 10^{-145}$	$\times 10^{-146}$	$\times 10^{-73}$	$\times 10^{-75}$	$\times 10$	$\times 10^{-2}$
SSA	$9.34 \times 10^{-9}$	$9.09 \times 10^{-9}$	$5.05 \times 10^{-1}$	$7.35 \times 10^{-1}$	$2.27 \times 10^2$	$2.95 \times 10^2$	3.08	3.69	$1.04 \times 10^2$	$2.15 \times 10^2$
ALO	$3.20 \times 10^{-5}$	$5.41 \times 10^{-5}$	1.47	3.43	$7.80 \times 10^2$	$1.04 \times 10^3$	6.61	6.55	$1.88 \times 10^2$	$3.49 \times 10^2$
GWO	$2.40 \times 10^{-5}$	$4.88 \times 10^{-5}$	1.11	3.03	$5.85 \times 10^2$	$9.63 \times 10^2$	4.96	6.35	$1.48 \times 10^2$	$1.48 \times 10^2$
MFO	$2.67 \times 10^2$	$1.62 \times 10^3$	1.50	2.96	$4.16 \times 10^3$	$8.38 \times 10^3$	1.55	2.24	$1.07 \times 10^6$	$9.20 \times 10^6$
WOA	$2.22 \times 10^2$	$1.48 \times 10^3$	1.25	2.76	$7.94 \times 10^3$	$1.23 \times 10^4$	2.05	2.56	$8.90 \times 10^5$	$8.41 \times 10^6$
算法	$F_6(x)$		$F_7(x)$		$F_8(x)$		$F_9(x)$		$F_{10}(x)$	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
AVSSA	$2.15 \times 10^{-8}$	$5.80 \times 10^{-9}$	$5.81 \times 10^{-5}$	$4.77 \times 10^{-5}$	$-4.82 \times 10^4$	$7.73 \times 10^3$	0	0	$8.88 \times 10^{-16}$	0
SSA	$2.15 \times 10^{-8}$	$6.40 \times 10^{-9}$	$4.29 \times 10^{-2}$	$4.85 \times 10^{-2}$	$-4.75 \times 10^4$	$8.47 \times 10^3$	$2.10 \times 10$	$2.30 \times 10$	1.08	1.23
ALO	$2.15 \times 10^{-5}$	$7.93 \times 10^{-5}$	$6.89 \times 10^{-2}$	$5.85 \times 10^{-2}$	$-3.36 \times 10^4$	$2.10 \times 10^4$	$4.01 \times 10$	$3.82 \times 10$	1.64	1.74
GWO	$1.28 \times 10^{-1}$	$2.67 \times 10^{-1}$	$5.20 \times 10^{-2}$	$5.85 \times 10^{-2}$	$-2.67 \times 10^4$	$2.17 \times 10^4$	$3.06 \times 10$	$3.70 \times 10$	1.23	1.67
MFO	$3.33 \times 10^2$	$2.14 \times 10^3$	$9.97 \times 10^{-1}$	4.68	$-2.32 \times 10^4$	$2.07 \times 10^4$	$5.59 \times 10$	$6.22 \times 10$	3.78	6.35
WOA	$2.78 \times 10^2$	$1.96 \times 10^3$	$8.31 \times 10^{-1}$	4.29	$-2.11 \times 10^4$	$1.95 \times 10^4$	$4.66 \times 10$	$6.05 \times 10$	3.15	5.97
算法	$F_{11}(x)$		$F_{12}(x)$		$F_{13}(x)$		$F_{14}(x)$		$F_{15}(x)$	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
AVSSA	0	0	$4.40 \times 10^{-3}$	$1.02 \times 10^{-2}$	2.18	1.33	1.13	$4.31 \times 10^{-1}$	$4.63 \times 10^{-4}$	$2.48 \times 10^{-4}$
SSA	$4.53 \times 10^{-3}$	$9.35 \times 10^{-3}$	2.34	2.75	2.60	5.07	1.20	$5.99 \times 10^{-1}$	$1.23 \times 10^{-3}$	$2.07 \times 10^{-3}$
ALO	$1.04 \times 10^{-2}$	$1.39 \times 10^{-2}$	4.80	4.81	4.36	$1.03 \times 10$	1.19	1.09	$2.19 \times 10^{-3}$	$4.75 \times 10^{-3}$
GWO	$8.96 \times 10^{-3}$	$1.30 \times 10^{-2}$	3.61	4.65	3.38	9.06	2.15	2.47	$2.25 \times 10^{-3}$	$5.09 \times 10^{-3}$
MFO	1.38	$1.04 \times 10$	3.61	4.33	4.61	9.43	1.99	2.25	$2.00 \times 10^{-3}$	$4.58 \times 10^{-3}$
WOA	1.15	9.55	3.01	4.18	3.90	8.76	2.14	2.39	$1.78 \times 10^{-3}$	$4.21 \times 10^{-3}$
算法	$F_{16}(x)$		$F_{17}(x)$		$F_{18}(x)$		$F_{19}(x)$		$F_{20}(x)$	
	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.	Avg.	Std.
AVSSA	-1.036	$3.89 \times 10^{-14}$	$3.98 \times 10^{-1}$	$2.66 \times 10^{-14}$	3	$2.16 \times 10^{-13}$	-3.86	$1.64 \times 10^{-13}$	-3.25	$7.35 \times 10^{-2}$
SSA	-1.036	$3.98 \times 10^{-14}$	$3.98 \times 10^{-1}$	$3.21 \times 10^{-14}$	3	$2.34 \times 10^{-13}$	-3.86	$8.58 \times 10^{-13}$	-3.25	$6.99 \times 10^{-2}$
ALO	-1.036	$5.74 \times 10^{-14}$	$3.98 \times 10^{-1}$	$5.61 \times 10^{-14}$	3	$2.64 \times 10^{-13}$	-3.86	$7.02 \times 10^{-13}$	-3.25	$6.68 \times 10^{-2}$
GWO	-1.036	$1.20 \times 10^{-8}$	$3.98 \times 10^{-1}$	$1.86 \times 10^{-14}$	3	$9.69 \times 10^{-6}$	-3.86	$1.32 \times 10^{-3}$	-3.25	$6.87 \times 10^{-2}$
MFO	-1.036	$1.09 \times 10^{-8}$	$3.98 \times 10^{-1}$	$1.67 \times 10^{-5}$	3	$8.78 \times 10^{-6}$	-3.86	$1.18 \times 10^{-3}$	-3.24	$6.64 \times 10^{-2}$
WOA	-1.036	$1.01 \times 10^{-8}$	$3.98 \times 10^{-1}$	$1.53 \times 10^{-5}$	3	$8.21 \times 10^{-6}$	-3.86	$1.75 \times 10^{-3}$	-3.25	$6.81 \times 10^{-2}$
算法	$F_{21}(x)$		$F_{22}(x)$		$F_{23}(x)$					
	Avg.	Std.	Avg.	Std.	Avg.	Std.				
AVSSA	-7.05	3.27	-8.33	3.06	-8.31	3.29				
SSA	-7.77	3.14	-8.00	3.36	-8.48	3.12				
ALO	-7.47	3.05	-7.33	3.46	-7.74	3.37				
GWO	-7.89	2.92	-8.10	3.28	-8.39	3.17				
MFO	-7.60	3.10	-8.01	3.32	-8.43	3.20				
WOA	-7.87	3.02	-8.03	3.25	-8.31	3.22				

基于上述分析和讨论, 所提出的算法AVSSA优于其它可比较的算法SSA、ALO、GWO、MFO和WOA, 更

适于基准函数的极值寻优.

### 2.2 基于AVSSA和BP神经网络的预测模型

利用所提出的AVSSA优化有一个隐含层的BP神经网络的参数: 输入层与隐含层的连接权值 $W$ , 隐含层与输出层的连接层的连接权值 $W'$ , 隐含层的偏差 $b$ 和输出层的偏差 $b'$ . 将此预测模型记为AVSSA-BP, 具体步骤为:

步骤1: 初始化. 初始化的参数包括樽海鞘链中樽海鞘的大小、最大迭代次数、隐含层中神经元节点的个数、终止条件.

步骤2: 将链中的每个樽海鞘映射为连接权值 $W$ 和 $W'$ , 偏差 $b$ 和 $b'$ . 选取函数

$$o = \frac{1}{Q} \sum_{s=1}^Q \sum_{i=1}^t (y_{is} - \bar{y}_{is})^2 \tag{7}$$

作为AVSSA的适应度函数, 其中 $Q$ 是样本数,  $y_{is}$ 和 $\bar{y}_{is}$ 分别是第 $s$ 个样本的预测输出和目标输出.

步骤3: 执行AVSSA算法, 利用公式(1)和(5)分别更新链中领导者和跟随者的位置.

步骤4: 若满足终止条件, 则返回AVSSA算法的食物源 $F$ , 转向步骤5. 否则执行AVSSA, 转向步骤2.

步骤5: 将食物源 $F$ 映射为连接权值 $W$ 和 $W'$ , 偏差 $b$ 和 $b'$ , 训练和测试BP神经网络.

预测模型AVSSA-BP的流程图如图2所示.

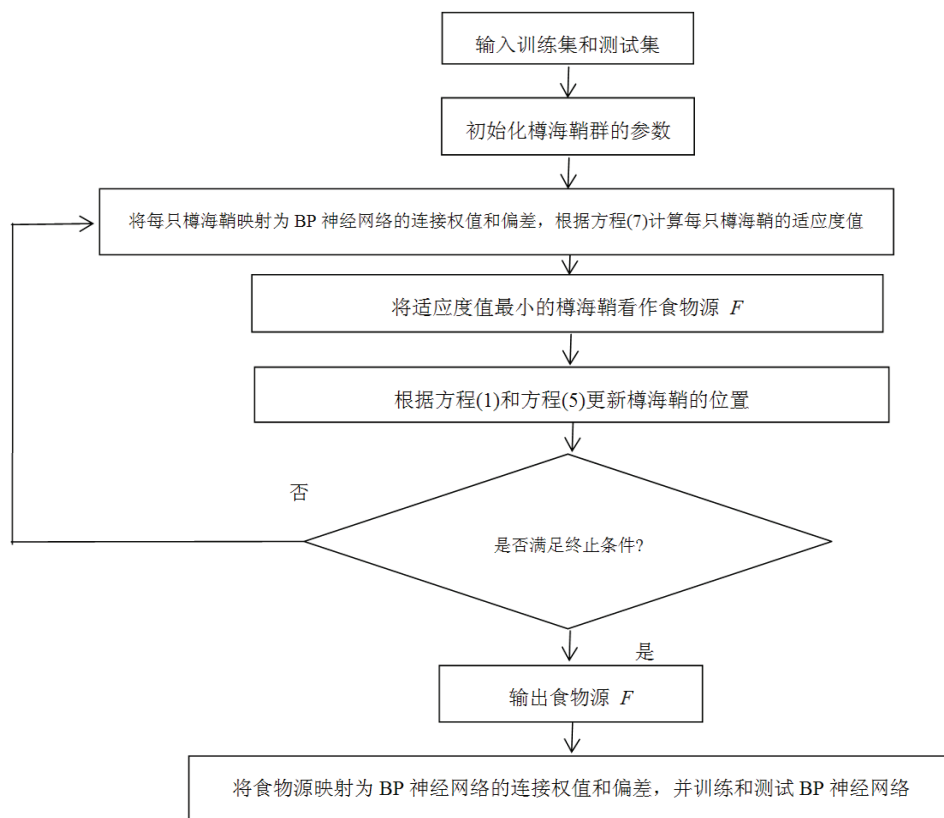


图 2 AVSSA-BP模型的流程图

## 3 实验

### 3.1 数据源

本文所采用的COVID-19数据为2020年3月5日至2020年10月6日国家卫健委公布的境外输入COVID-19累计确诊病例数, 网站为[http://www.nhc.gov.cn/xcs/xxgzbd/gzbd\\_index.shtml](http://www.nhc.gov.cn/xcs/xxgzbd/gzbd_index.shtml). 所采用的COVID-19数据利用前3天的COVID-19累计确诊病例数预测第4天的COVID-19累计确诊病例数. 得到了214组3维输入数据, 再删除最后一组3维数据, 这样就得到了213组输入数据, 将前192组数据作为训练数据, 剩下的21组数据作为测试数据.

为验证AVSSA-BP的预测性能, 我们仍然选取SSA、ALO、GWO、MFO和WOA优化BP神经网络的参数, 建立预测模型SSA-BP、ALO-BP、GWO-BP、MFO-BP和WOA-BP进行比较, 并与BP神经网络进行比较.

在本文中,平均绝对误差(Mean Absolute Error, MAE),根均方差误差(Root Mean Square Error, RMSE),平均绝对百分比误差(Mean Absolute Percentage Error, MAPE)和确定性系数 $R^{2[16]}$ 作为预测模型AVSSA-BP的评价指标.

### 3.2 实验结果

在预测模型AVSSA-BP的BP神经网络中,输入层神经元的个数为3,输出层的节点个数为1. 经过大量的实验,选取预测模型AVSSA-BP中BP神经网络的隐含层的神经元节点数为6. 这样我们将AVSSA-BP中樽海鞘链的每个樽海鞘映射为输入层与隐含层之间的连接权值 $W$ ,隐含层与输出层之间的连接权值 $W'$ ,隐含层的偏差 $b$ 和输出层的偏差 $b'$ ,这样预测模型AVSSA-BP中樽海鞘链的每个樽海鞘的维数为 $3 \times 6 + 6 \times 1 + 6 + 1 = 31$ ,并选择方程(7)作为AVSSA的适应度函数.

在实验中,AVSSA-BP、SSA-BP、ALO-BP、GWO-BP、MFO-BP和WOA-BP的种群大小为50,最大迭代次数为100. 独立运行BP神经网络,AVSSA-BP、SSA-BP、ALO-BP、GWO-BP、MFO-BP、WOA-BP和BP各30次,得到7个预测模型AVSSA-BP、SSA-BP、ALO-BP、GWO-BP、MFO-BP、WOA-BP和BP的21个测试样本的预测值的平均值和评价指标,分别见表2和表3.

表 2 测试数据的预测值

日期	实际值	预测值						
		AVSSA-BP	SSA-BP	ALO-BP	GWO-BP	MFO-BP	WOA-BP	BP
2020-09-16	2 674	2 687.64	2 688.59	2 687.36	2 687.73	2 684.44	2 674.41	2670.62
2020-09-17	2 706	2 699.07	2 698.88	2 696.1	2 698.5	2 693.36	2 685.52	2 678.07
2020-09-18	2 720	2 716.39	2 716.66	2 715.36	2 716.03	2 714.93	2 720.45	2 707.16
2020-09-19	2 730	2 731.63	2 732.66	2 731.31	2 733.13	2 729.2	2 734.17	2 715.22
2020-09-20	2 742	2 741.21	2 740.51	2 739.97	2 741.98	2 738.94	2 741.56	2 728.04
2020-09-21	2 748	2 749.6	2 749.85	2 749.88	2 750.6	2 748.58	2 753.13	2 739.45
2020-09-22	2 758	2 756.93	2 757.78	2 757.43	2 758.79	2 756.33	2 758.35	2 744.48
2020-09-23	2 765	2 766.24	2 766.78	2 767.44	2 767.91	2 765.74	2 769.63	2 754.41
2020-09-24	2 773	2 775.5	2 776.1	2 776.42	2 777.97	2 774.95	2 776.2	2 759.77
2020-09-25	2 788	2 786.08	2 785.6	2 786.38	2 788.6	2 785.07	2 784.73	2 767.34
2020-09-26	2 802	2 801.34	2 799.94	2 802.05	2 803.66	2 800.39	2 801.1	2 779.87
2020-09-27	2 823	2 818.05	2 816.78	2 818.33	2 821.55	2 816.94	2 813.39	2 790.41
2020-09-28	2 835	2 837.28	2 836.19	2 838.76	2 840.09	2 836	2 833.89	2 808.18
2020-09-29	2 854	2 850.26	2 851.18	2 850.06	2 853.49	2 848.31	2 842.16	2 817.11

表 3 测试数据的评价指标

评价指标	AVSSA-BP	SSA-BP	ALO-BP	GWO-BP	MFO-BP	WOA-BP	BP
MAE	4.311 0	5.180 7	6.491 3	5.190 2	6.060 6	5.557 3	2.964 0 $\times 10$
RMSE	1.413 4	1.652 7	2.135 3	1.707 0	1.990 1	1.645 0	7.272 0
MAPE	$1.547 0 \times 10^{-1}$	$1.853 4 \times 10^{-1}$	$2.309 6 \times 10^{-1}$	$1.856 1 \times 10^{-1}$	$2.173 7 \times 10^{-1}$	$1.984 0 \times 10^{-1}$	1.042 2
$R^2$	$9.938 2 \times 10^{-1}$	$9.912 9 \times 10^{-1}$	$9.889 4 \times 10^{-1}$	$9.921 1 \times 10^{-1}$	$9.915 5 \times 10^{-1}$	$9.920 9 \times 10^{-1}$	$9.899 7 \times 10^{-1}$

由表3可知,利用AVSSA-BP得到的境外输入病例确诊的COVID-19累计病例数预测的平均MAE、RMSE和MAPE最小,分别为4.311 0、1.413 4和0.154 7,且具有最高的平均确定性系数. 因此,预测模型AVSSA-BP优于其它预测模型SSA-BP、ALO-BP、GWO-BP、MFO-BP、WOA-BP和BP. 实验结果表明AVSSA算法能有效地优化神经网络的参数实现COVID-19的预测.

## 4 讨论与结论

本文对SSA中樽海鞘链中第 $i$ 个跟随者的位置更新进行了改进,得到加权平均的SSA,记为AVSSA. 将AVSSA与BP神经网络相结合,对境外输入确诊COVID-19累计病例建立了预测模型AVSSA-BP,且AVSSA-BP具有最小

的预测误差MSE、MAE、RMSE和MAPE,最高的确定性系数 $R^2$ 。实验结果表明AVSSA-BP优于SSA-BP、ALO-BP、GWO-BP、MFO-BP、WOA-BP和BP模型。

在以后的研究中,我们还可以对SSA算法进行改进,并将其与其他一种或多种群智能算法相结合,建立新的混合算法来提高其性能。这些新算法可以用来优化某些机器学习算法的参数,以解决预测和分类问题。

### 参考文献:

- [1] 刘伟,刘远,胡文穗,等.差分自回归移动平均乘积季节模型预测广州市肺结核发病趋势[J].中华疾病控制杂志,2021,25(2): 240-243+248.
- [2] 刘怡然,王东杰,邓雪峰,等.基于长短时记忆神经网络的生猪价格预测模型[J].江苏大学学报(自然科学版),2021,42(2): 190-197.
- [3] 卢继平,曾燕婷,喻华,等.基于改进AWNN的风电功率超短期多步预测[J].太阳能学报,2021,42(1): 166-173.
- [4] 曹建飞.基于多尺度组合模型的我国棉花价格预测[J].统计与决策,2021,37(2): 176-180.
- [5] 熊景华,茹璟.基于随机森林算法和模糊信息粒化的汇率预测组合模型研究[J].数量经济技术经济研究,2021,38(1): 135-156.
- [6] 刘合兵,韩晶晶,马新明,等.基于EMD-ELM模型的农产品价格预测研究[J].农业大数据学报,2020,2(3): 68-74.
- [7] HU H, WANG H, BAI Y, et al. Determination of endometrial carcinoma with gene expression based on optimized Elman neural network[J]. Applied Mathematics and Computation, 2019, 341: 204-214.
- [8] 顾丽丽,刘勇,甄佳奇.基于改进粒子群算法优化BP神经网络的甜菜产量预测方法[J].新疆大学学报(自然科学版)(中英文),2021,38(2): 191-196.
- [9] ZIVKOVIC M, BACANIN N, VENKATACHALAM K, et al. COVID-19 cases prediction by using hybrid machine learning and beetle antennae search approach[J]. Sustainable Cities and Society, 2021, 66: 102669.
- [10] REDDY K S S, REDDY YCAP, RAO C M. Recurrent neural network based prediction of number of COVID-19 cases in India[J]. Materials Today: Proceedings, 2020, <https://doi.org/10.1016/j.matpr.2020.11.117>.
- [11] LIU C, HUANG J, JI F, et al. Improvement of the global prediction system of the COVID-19 pandemic based on the ensemble empirical mode decomposition(EEMD) and autoregressive moving average(ARMA) model in a hybrid approach[J]. Atmospheric and Oceanic Science Letters, 2020, <https://doi.org/10.1016/j.aosl.2020.100019>.
- [12] MIRJALILI S, GANDOMI A H, MIRJALILI S Z, et al. Salp swarm algorithm: a bio-inspired optimizer for engineering design problems[J]. Advances in Engineering Software, 2017, 114: 163-191.
- [13] SULTAN H M, MENESY A S, KAMEL S, et al. Parameter identification of proton exchange membrane fuel cells using an improved salp swarm algorithm[J]. Energy Conversion and Management, 2020, 224: 113341.
- [14] NEGGAZ N, EWEES A A, ELAZIZ M A, et al. Boosting salp swarm algorithm by sine cosine algorithm and disrupt operator for feature selection[J]. Expert Systems with Applications, 2020, 145: 113103.
- [15] ALTAN A, KARASU S. Recognition of COVID-19 disease from X-ray images by hybrid model consisting of 2D curvelet transform, chaotic salp swarm algorithm and deep learning technique[J]. Chaos, Solitons and Fractals, 2020, 140: 110071.
- [16] TUBISHAT M, JA'AFAR S, ALSWAITTI M, et al. Dynamic salp swarm algorithm for feature selection[J]. Expert Systems with Applications, 2021, 164: 113873.
- [17] HU H, AO Y, BAI Y, et al. An improved harris's hawks optimization for sar target recognition and stock market index prediction[J]. IEEE Access, 2020, 8: 65891-65910.

责任编辑:赵新科