

基于改进遗传算法优化 BP 神经网络的土壤湿度预测模型^①



王佳楠¹, 王玉莹¹, 何淑林¹, 时龙闽¹, 张艳滴¹, 孙海洋¹, 刘 勇^{1,2}

¹(黑龙江大学 电子工程学院, 哈尔滨 150080)

²(黑龙江东部节水设备有限公司, 绥化 152001)

通信作者: 刘 勇, E-mail: liuyong@hlju.edu.cn

摘 要: 我国是农业大国, 在进行农业生产过程中, 对土壤的湿度进行精准预测具有非常重要的意义. 针对传统 BP (back propagation) 神经网络在预测过程中会出现局部最小化以及收敛速度慢的问题, 本文将改进的遗传算法 (genetic algorithm) 应用到传统 BP 神经网络模型当中, 提出了一种自适应遗传算法优化 BP 神经网络的土壤湿度预测方法. 通过 Matlab 仿真软件建立改进遗传算法优化 BP 神经网络的预测模型, 并且对哈尔滨地区玉米地的土壤湿度进行实验. 结果表明, 该模型的精度高于未优化的 BP 神经网络. 该模型能够大量减少湿度传感器的使用, 为农业生产减少了成本.

关键词: 神经网络; 遗传算法; 土壤湿度监测; 智慧农业

引用格式: 王佳楠, 王玉莹, 何淑林, 时龙闽, 张艳滴, 孙海洋, 刘勇. 基于改进遗传算法优化 BP 神经网络的土壤湿度预测模型. 计算机系统应用, 2022, 31(2): 273-278. <http://www.c-s-a.org.cn/1003-3254/8324.html>

Optimized BP Neural Network Model Based on Improved Genetic Algorithm for Soil Moisture Prediction

WANG Jia-Nan¹, WANG Yu-Ying¹, HE Shu-Lin¹, SHI Long-Min¹, ZHANG Yan-Di¹, SUN Hai-Yang¹, LIU Yong^{1,2}

¹(College of Electronic Engineering, Heilongjiang University, Harbin 150080, China)

²(Heilongjiang Water Saving in the East Co. Ltd., Suihua 152001, China)

Abstract: China is a large agricultural country. In the process of agricultural production, it is of great significance to accurately predict the soil moisture. In view of the local minimization and slow convergence in the prediction process of the traditional back propagation (BP) neural network, an improved genetic algorithm is applied to the traditional BP neural network model in this study. A soil moisture prediction method is proposed that optimizes the BP neural network by the adaptive genetic algorithm. A prediction model of the BP neural network optimized by the improved genetic algorithm is established by the Matlab simulation software and experimented on the soil moisture of corn fields in Harbin. The results show that the accuracy of the model is higher than that of the unoptimized BP neural network model. This model can greatly reduce the use of moisture sensor and thus reduce the agricultural production cost.

Key words: neural network; genetic algorithm; soil moisture monitoring; smart agriculture

随着我国农业与经济的发展和智慧农业对农田起到了至关重要的作用, 精准灌溉已经成为农业生产和生活的重要因素, 而实现对农作物精准灌溉, 对农作物

土壤湿度进行精准预测至关重要^[1]. 土壤湿度是旱涝评价的重要参考, 对地表能量、降雨再分配、水资源管理、农业灌溉等方面都有重要影响^[2]. 土壤湿度过低,

^① 基金项目: 中央引导地方科技发展项目 (SBZY2021E006)

收稿时间: 2021-04-20; 修改时间: 2021-05-19; 采用时间: 2021-06-02; csa 在线出版时间: 2022-01-17

会导致土壤干旱,对植物的光合作用产生严重影响,甚至导致植物干枯死亡;土壤湿度过高,会导致植物根部缺氧,严重时会导致植物根部腐烂^[3,4].通过对农作物土壤湿度的精准预测可以准确地反映农作物土壤的水分含量状况,从而为实现精准灌溉提供了科学决策的依据和方法.

1 基本理论

1.1 BP神经网络算法

BP (back propagation) 神经网络算法是一种典型的多层前馈神经网络,它通过误差逆向传播算法来进行学习训练^[5-7].BP神经网络是目前使用最为广泛,结构最为直观并且工作原理最为容易理解的神经网络,它的结构简单,可塑性强,且拥有较强的数据拟合能力^[8-10].BP神经网络主要包含输入层、隐含层和输出层,在训练过程中,神经网络不断调整输入层与隐含层以及隐含层和输出层之间的权值和阈值^[11,12],当神经网络输出值与目标值一致或者达到迭代次数时训练停止,这种神经网络拥有较强的泛化能力,神经网络的结构原理如图1所示.

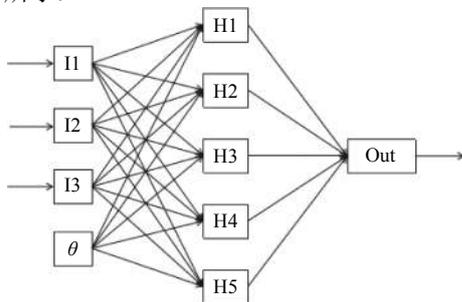


图1 BP神经网络结构原理

BP神经网络的计算公式如式(1)–式(3)所示:

$$f_s(u) = \frac{1}{1 + e^{-u}} \quad (1)$$

$$u_j = \sum_{i=1}^n w_{ij}x_i - o_j \quad (2)$$

$$p = f_s \left(\sum_{j=1}^n w_j o_j + \theta \right) \quad (3)$$

式(1)为神经元的激励函数,采用的是 Sigmoid 函数,式(2)为神经元的输入数据经过加权值求和之后与阈值相减, x_i 表示输入数据, w_{ij} 表示权值, o_j 表示阈值, p 表示局部输出,式(3)为将式(1)与式(2)联立所得的输出值.

常规的BP神经网络主要采用数据正向传播,误差反向传播的方式,根据误差值采用梯度下降的方法对网络的权值和阈值进行修正,从而不断使误差下降到预期数值,权值和阈值的修正如式(4)和式(5)所示:

$$w(k+1) = w(k) - \beta \frac{\partial E}{\partial w} \quad (4)$$

$$\theta(k+1) = \theta(k) - \beta \frac{\partial E}{\partial \theta} \quad (5)$$

式(4)和式(5)为权值和阈值修正的公式, E 表示神经网络的输出值与目标值的误差, β 表示神经网络的学习率.

1.2 改进遗传算法

1.2.1 遗传算法原理

遗传算法 (genetic algorithm, GA) 最早是由美国的 Holland 于 20 世纪 70 年代提出^[13],该算法是根据大自然中生物体进化规律而设计提出的^[14].它的工作原理是首先对输入数据进行编码,之后通过一定的概率进行选择、交叉和变异运算直到选择出适应度最大的个体作为目标值输出,之后停止运算.

1.2.2 适应度函数的选择

适应度函数是衡量种群中个体适应能力大小的标准,一般情况下,会将训练目标函数选作遗传算法的适应度函数,本算法中采用误差平方的倒数作为适应度函数,如式(6)和式(7)所示:

$$E = \frac{1}{2} \sum_{j=1}^n (P_j^{k-1}(w, x) - y_j)^2 \quad (6)$$

$$F(j) = \frac{1}{E(j)} \quad (7)$$

式(6)中, E 为误差函数, P 为整体输出, w 为权矢量, x 为输入矢量;式(7)中, F 为适应度, j 为迭代次数.

1.2.3 选择算子的改进

传统的遗传算法在工作过程中经常采用“轮盘赌”的方式,种群中个体被选中的概率是随机的,这种选择方式很有可能丢掉最优个体,在实际的运算过程中会产生较大的误差.因此,本文对选择算子进行改进,首先将种群个体利用排序法进行重新排列,重新排序之后个体被选择的概率如式(8)和式(9)所示:

$$s = \frac{p_0}{1 - (1 - p_0)^a} \quad (8)$$

$$p = s(1 - p_0)^{b-1} \quad (9)$$

式(8)和式(9)中, a 为遗传算法中种群的数量, p_0 为最

优个体可能被选择的概率, s 为将 p_0 进行标准化后的值, b 为在对种群重新排列后第 n 个个体所在的位置.

1.2.4 交叉算子的改进

传统的遗传算法在工作过程中一般会将交叉概率设为 0.3–0.8 之间的一个常数. 在运算过程中, 交叉概率设置过高会提高遗传算法的全局搜索能力, 但染色体的适应能力会有所下降, 而交叉概率设置得过低会降低遗传算法的全局搜索能力和收敛速度. 本文对交叉算子进行改进, 在算法迭代过程中会根据适应度的变化来调整交叉概率的变化, 改进的交叉概率如式 (10) 所示:

$$P_j = \begin{cases} P_{j\max}, & F < F_{\text{mean}} \\ P_{j\max} - \frac{P_{j\max} - P_{j\min}}{n_{\max}} \times n, & F \geq F_{\text{mean}} \end{cases} \quad (10)$$

其中, F 为种群中两个进行交叉的个体的最大适应度, F_{mean} 为整个种群的平均适应度, n 为遗传算法在当前工作过程中的迭代次数, n_{\max} 为遗传算子在工作过程中的最大迭代次数, 在遗传算法初始化时可以将最小交叉概率 $P_{j\min}$ 设置为 0.3, 最大交叉概率 $P_{j\max}$ 设置为 0.8.

1.2.5 变异算子的改进

传统的遗传算法在工作过程中一般会将变异概率设为 0.001–0.1 之间的一个常数. 在遗传算法运算初期, 种群个体的适应度相对平均适应度较低, 所以需要将变异的概率设为较小的值, 从而保留染色体中基因优良的个体. 在遗传算法运算的后期, 种群个体的适应度相对高于平均适应度, 因此需要将变异的概率设为较大的值来提高遗传算法的局部搜索能力. 本文对变异算子进行改进, 在遗传算法运算过程中会根据适应度的变化来调整变异概率的值, 改进的变异概率如式 (11) 所示:

$$P_b = \begin{cases} P_{b\min}, & F < F_{\text{mean}} \\ P_{b\max} - \frac{P_{b\max} - P_{b\min}}{n_{\max}} \times n, & F \geq F_{\text{mean}} \end{cases} \quad (11)$$

其中, F 为种群中两个进行变异的个体的最大适应度, F_{mean} 为整个种群的平均适应度, n 为遗传算法在当前工作过程中的迭代次数, n_{\max} 为遗传算子在工作过程中的最大迭代次数, 在遗传算法初始化时可以将最小变异概率 $P_{b\min}$ 设置为 0.001, 最大变异概率 $P_{b\max}$ 设置为 0.1.

2 改进遗传算法优化 BP 神经网络预测模型

2.1 预测模型的设计框架

BP 神经网络的初始权值和阈值对训练的结果有着很大的影响, 并且在训练过程中容易陷入局部最优,

因此为了提高 BP 神经网络的预测精度以及避免 BP 神经网络陷入局部最优的问题, 引入了改进遗传算法对 BP 神经网络的权值和阈值进行优化. 传统遗传算法的选择概率、交叉概率和变异概率使用的是固定的常数, 这样在训练的前期和后期容易产生局部最优解, 为了解决这个问题, 对选择算子、交叉算子和变异算子的概率进行了改进, 通过判断当前适应度值从而对选择概率、交叉概率和变异概率进行调整. 随着适应度值的不断变化, 预测模型的选择概率、交叉概率和变异概率会随之产生变化, 使预测模型能够将各个参数调整为模型最佳的状态, 形成了改进遗传算法优化 BP 神经网络预测模型.

在改进遗传算法优化 BP 神经网络预测模型中, 具体操作流程包括以下几个方面, 首先通过传感器对空气温度、空气湿度和光照强度 3 种气象数据进行采集; 其次, 对采集数据进行预处理, 由于采集后的数据可能存在错误数据, 将 3 组数据中的最大值和最小值去除, 并根据实际数据进行插值处理, 得到正确的气象数据; 再次, 将预处理的数据分为两组, 一组作为训练集, 另一组作为测试集. 最后将训练集数据导入预测模型中对预测模型进行训练, 再将测试集导入训练好的预测模型中进行预测评估. 改进遗传算法优化 BP 神经网络预测模型的设计框架如图 2 所示.

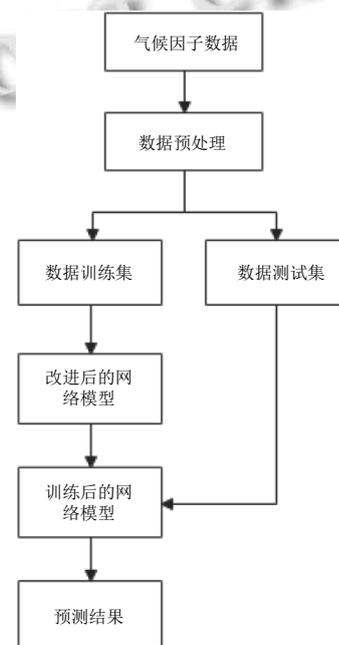


图 2 改进遗传算法优化 BP 神经网络框架图

2.2 预测模型的设计流程

实验所使用的软件平台为 Matlab R2018a, 计算机硬件为 Windows 7 系统, CPU 频率为 2.50 GHz, 内存为 8 GB, 硬盘为 500 GB. 将改进的遗传算法与 BP 神经网络结合, 建立了基于改进遗传算法优化 BP 神经网络, 提高了预测模型的预测精度. 改进遗传算法优化 BP 神经网络流程图如图 3 所示.

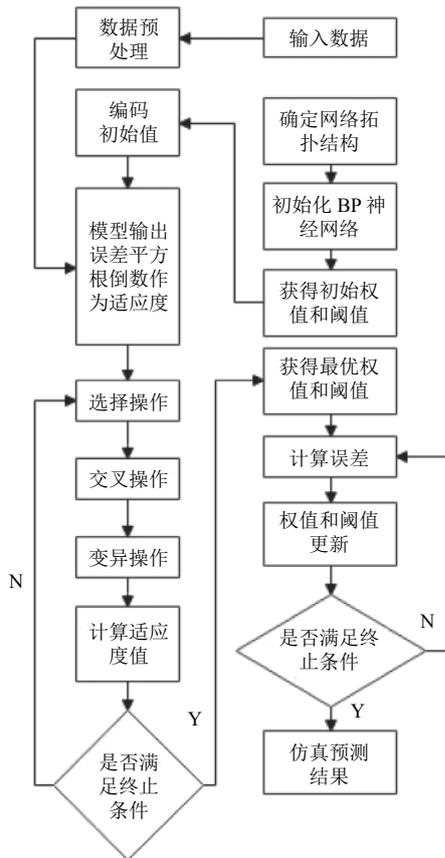


图 3 改进遗传算法优化 BP 神经网络流程图

3 实验仿真分析

3.1 预测模型的参数设置

进行预测模型的训练之前, 要对预测模型的参数进行设置, 在 Matlab R2018a 软件平台下构建基于改进遗传算法优化 BP 神经网络模型, 对改进遗传算法优化 BP 神经网络预测模型的相关参数进行设置, 输入层的节点数设置为 3 个, 隐含层的节点数设置为 9 个, 输出层的节点数设置为 1 个, 激活函数采用的是 Sigmoid 函数, 神经网络的迭代次数为 1 000.

3.2 预测模型的误差评估

为了测试预测模型的精确度, 采用 *MAE* (平均绝对误差)、*MAPE* (平均绝对百分率误差) 以及 *RMSE* (均方根误差) 3 种误差分析方法对预测模型的精确度进行评估, 从而验证预测模型的准确性. 3 种误差公式如式 (12)、式 (13)、式 (14) 所示.

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - m_i| \quad (12)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^N \frac{|y_i - m_i|}{y_i} \quad (13)$$

$$RMSE = \sqrt{\sum_{i=1}^N (y_i - m_i)^2} \quad (14)$$

3.3 仿真数据采集

实验数据来自黑龙江省哈尔滨市呼兰区玉米种植基地, 基地采用农业物联网技术, 通过物联网技术对种植基地的各种气象数据进行采集, 将采集的数据进行处理后, 选取 1 000 组环境气象数据作为本文的试验数据, 试验数据包括空气温度、空气湿度和光照强度. 试验将 1 000 组数据分为两部分, 随机选取 980 组数据作为预测模型的训练数据, 剩余 20 组数据作为测试数据, 表 1 为部分试验数据.

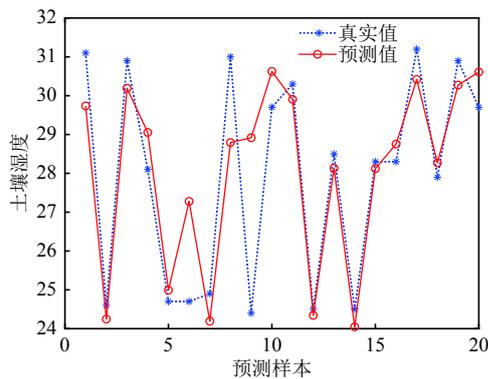
表 1 部分采集数据

序号	空气温度 (°C)	空气湿度 (%RH)	光照强度 (lx)	土壤湿度 (%RH)
1	14.5	85.6	18 215	28.4
2	14.7	81.7	11 046	28.2
3	15.0	78.3	11 384	28.1
4	15.1	77.3	10 477	28.7
5	15.2	76.8	12 892	28.5
...
999	18.8	78.0	3 275	24.7
1000	19.1	76.1	3 840	24.5

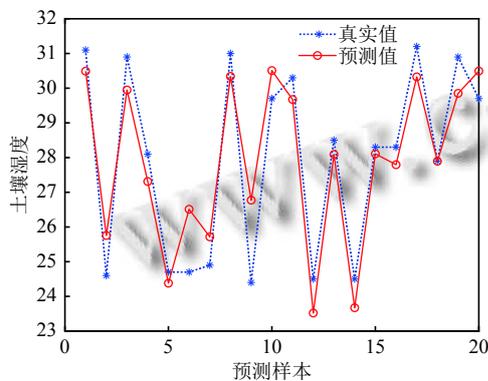
3.4 实验仿真结果

使用处理后的训练数据对模型进行训练, 将训练后的预测模型通过测试集数据测试, 并将 BP 神经网络、遗传算法优化 BP 神经网络和改进遗传算法优化 BP 神经网络的测试结果进行对比, 数据预测结果如图 4 所示.

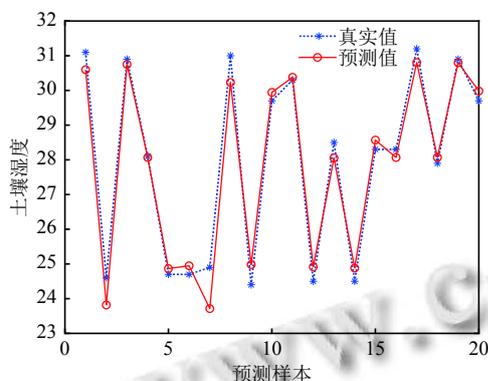
图 4(a)、图 4(b) 和图 4(c) 分别是 BP 神经网络、遗传算法优化 BP 神经网络和改进遗传算法优化 BP 神经网络的预测结果对比图, 由图可知改进遗传算法优化的 BP 神经网络预测效果最好, BP 神经网络预测效果最差.



(a) BP神经网络



(b) 遗传算法优化BP神经网络



(c) 改进遗传算法优化BP神经网络

图4 各个模型预测误差

通过测试集数据对预测模型进行预测后,分别计算各个模型的相对误差,误差对比情况如表2所示。

表3为模型预测误差,通过各个预测模型的结果计算它们的MAE、MAPE以及RMSE三种误差,从表3能够看出,IGA-BP模型的预测精度最高,MAE、MAPE以及RMSE分别是0.014 2、0.000 4和0.063 4,与BP预测模型相比分别提高了3.13%、0.11%和14.02%;与GA-BP模型相比分别提高了2.55%、

0.09%和11.4%,所以IGA-BP预测模型的预测质量更高。

表2 各个模型预测结果

序号	预测输出	BP模型输出	相对误差	GA-BP模型输出	相对误差	IGA-BP模型输出	相对误差
1	31.1	29.735 8	0.043 9	30.490 8	0.019 6	30.597 9	0.016 1
2	24.6	24.247 9	0.014 3	25.755 3	0.047 0	23.817 2	0.031 8
3	30.9	30.192 5	0.022 9	29.947 5	0.030 8	30.750 3	0.004 8
4	28.1	29.053 3	0.033 9	27.311 5	0.028 1	28.067 5	0.001 2
5	24.7	24.991 1	0.011 8	24.381 3	0.012 9	24.867 5	0.006 8
6	24.7	27.278 7	0.104 4	26.511 0	0.073 3	24.947 4	0.010 0
7	24.9	24.191 0	0.028 5	25.715 7	0.032 8	23.713 9	0.047 6
8	31.0	28.793 8	0.071 2	30.340 8	0.021 3	30.231 1	0.024 8
9	24.4	28.916 7	0.185 1	26.772 5	0.097 2	24.981 9	0.023 8
10	29.7	30.624 9	0.031 1	30.508 1	0.027 2	29.944 0	0.008 2
11	30.3	29.905 9	0.013 0	29.673 4	0.020 7	30.378 4	0.002 6
12	24.5	24.342 1	0.006 4	23.521 3	0.039 9	24.916 1	0.017 0
13	28.5	28.134 5	0.012 8	28.091 0	0.014 4	28.055 8	0.015 6
14	24.5	24.046 0	0.018 5	23.673 1	0.033 7	24.882 0	0.015 6
15	28.3	28.120 5	0.006 3	28.107 8	0.006 8	28.569 1	0.009 5
16	28.3	28.754 1	0.016 0	27.795 4	0.017 8	28.068 3	0.008 2
17	31.2	30.421 3	0.025 0	30.326 1	0.028 0	30.804 7	0.012 7
18	27.9	28.270 4	0.013 3	27.894 5	0.000 2	28.071 0	0.006 1
19	30.9	30.273 3	0.020 3	29.851 9	0.033 9	30.799 7	0.003 2
20	29.7	30.610 6	0.030 7	30.493 2	0.026 7	29.983 5	0.009 5

表3 模型预测误差比较

预测模型	MAE	MAPE	RMSE
BP	0.045 5	0.001 5	0.203 6
GA-BP	0.039 7	0.001 3	0.177 4
IGA-BP	0.014 2	0.000 4	0.063 4

通过对训练结果的分析,改进的遗传算法优化的BP神经网络具有更好的稳定性和精确性,对BP模型、GA-BP模型和IGA-BP模型输出的相对误差可以得出,IGA-BP模型的相对误差普遍较小;对3种预测模型的MAE、MAPE以及RMSE进行分析可以得出IGA-BP模型的预测效果更精准。

4 结论与展望

本文提出了一种土壤湿度预测模型,该模型具有良好的预测性能,能够提高土壤湿度的预测精度,为土壤湿度的预测提供了新的模型和方法。该模型的具体实施步骤是:(1)对数据进行预处理,初始化BP神经网络,从而获得初始权值和阈值;(2)对初始权值和阈值进行编码,导入改进的遗传算法模型中并且初始化遗传算法;(3)通过改进的遗传算法能够得到最优的权值和阈值,从而提高了预测模型的精确性。本文通过对黑

龙江省哈尔滨市呼兰区玉米种植基地采集的数据进行训练,之后取 20 组数据对模型进行测试,结果显示本文提出的土壤预测模型具有更高了预测精度.该模型还有很多不足之处,比如该模型的适应性有待提升,对环境气候不同地区的预测可能会产生误差,因此未来的改进方向是如何更好地适应不同地区土壤环境预测,对于湿度不同地区的土壤湿度进行准确预测.

参考文献

- 1 孙浩然,孙琳,毕春光,等.基于粒子群与模拟退火协同优化的农田物联网混合多跳路由算法.智慧农业(中英文),2020,2(3):98-107.
- 2 王冰玉.基于神经网络灰色模型的农田土壤湿度预测研究.节水灌溉,2020,(7):32-35.[doi:10.3969/j.issn.1007-4929.2020.07.007]
- 3 Ignace DD. Determinants of temperature sensitivity of soil respiration with the decline of a foundation species. PLoS One, 2019, 14(10): e0223566. [doi: 10.1371/journal.pone.0223566]
- 4 张绪成,马一凡,王红丽,等.不同耕作方式对旱地全膜双垄沟播玉米产量和水分利用的影响.生态学报,2021,41(9):3603-3611.
- 5 张峰峰,张欣,陈龙,等.采用改进遗传算法优化神经网络的双目相机标定.中国机械工程,2021,32(12):1423-1431. [doi: 10.3969/j.issn.1004-132X.2021.12.005]
- 6 闫春,厉美璇,周潇.基于改进的遗传算法优化 BP 神经网络的车险欺诈识别模型.山东科技大学学报(自然科学版),2019,38(5):72-80.
- 7 周婧,白云龙,张小宝,等.基于 MEA-BP 神经网络的土壤养分评价模型.中国农机化学报,2020,41(2):231-236.
- 8 彭彬彬,闫献国,杜娟.基于 BP 和 RBF 神经网络的表面质量预测研究.表面技术,2020,49(10):324-328,337.
- 9 魏良良,丁祥,蔡甜,等.基于 BP 神经网络与改进遗传算法的泵站优化调度.水电能源科学,2019,37(5):168-171.
- 10 张然,刘敏,张启坤,等.一种基于 SA-SOA-BP 神经网络的网络安全态势预测算法.小型微型计算机系统,2020,41(10):2157-2163. [doi: 10.3969/j.issn.1000-1220.2020.10.022]
- 11 Zhao Z, Chen WH, Wu XM, *et al.* LSTM network: A deep learning approach for short-term traffic forecast. IET Intelligent Transport Systems, 2017, 11(2): 68-75. [doi: 10.1049/iet-its.2016.0208]
- 12 Kuremoto T, Kimura S, Kobayashi K, *et al.* Time series forecasting using a deep belief network with restricted Boltzmann machines. Neurocomputing, 2014, 137: 47-56. [doi: 10.1016/j.neucom.2013.03.047]
- 13 Holland J. Adaptation in natural and artificial systems: An introductory analysis with application to biology. Control & Artificial Intelligence, 1975.
- 14 宋尧,仰燕兰,叶桦.改进遗传算法在 MSPSP 问题中的验证.计算机系统应用,2020,29(10):235-241. [doi: 10.15888/j.cnki.csa.007591]