

基于 BP 神经网络和遗传算法的养殖水域预警模型

徐云娟

(苏州托普信息职业技术学院 昆山 215311)

摘 要 随着我国经济的快速发展,环境保护工作面临前所未有的压力。为有效加强水产养殖水域环境的监管力度,应对突发性环境污染事故对社会生活和经济发展的影响,建立 BP 神经网络来拟合水产养殖水域饲料投放与总磷(TP)、总氮(TN)、透明度(SD)以及耗氧量(COD)等富营养指标变化情况的对应函数关系,并利用遗传算法来实现目标函数的优化方法,形成养殖水域预警模型,为水域环境治理和公共决策提供技术支撑。利用该模型对鄱阳湖新型水产养殖基地的样本进行分析,取得了很好的预测效果。

关键词 BP 神经网络,遗传算法,富营养化,预警模型

中图分类号 TP183,X502 文献标识码 A

Early Warning Model for Water Eutrophication Based on BP Artificial Neural Network and Genetic Algorithm

XU Yun-juan

(Suzhou Top Institute of Information Technology, Kunshan 215311, China)

Abstract With the economic development, the environmental protection work is facing unprecedented pressure. In order to enhance the aquatic environment control effectively and to deal with the impact of sudden environmental pollution accident on the social and economic development, the paper established BP neural network theory for fitting aquaculture feed, and total phosphorus, total nitrogen, transparency, as well as oxygen consumption, and other nutritious indicators of changes in the corresponding function. Furthermore, the paper used genetic algorithm to achieve optimization methods of the objective function, and formed a breeding waters early warning model. The model provides technical support for water environment governance and public decision-making. At the same time, the paper makes further analysis for the samples of Poyang Lake's new aquaculture base and forecasts good results by using the model.

Keywords BP neural network, Genetic algorithm, Eutrophication, Early warning model

随着我国经济的快速发展,环境保护问题日益突出。近年来,我国水产养殖业快速发展,养殖产业规模不断扩大,养殖方式由半集约化向高度集约化发展。为了加大水产品的产量,人们大量向养殖水体投放人工饵料,造成了严重的水体富营养化,水产养殖自身的污染与水域环境的矛盾也日益突出。在水产养殖与水域环境的相互关系上,人们大多只重视水域环境对养殖的影响,而对水产养殖所产生的负面影响关注不够,迫切需要对养殖水域的富营养化进行预警。

本文利用提出的模型对养殖水域的富营养化指标进行预警,并在对鄱阳湖新型水产养殖基地近 5 年的样本进行分析和拟合的基础上进行了有效的模拟预测。

1 BP 神经网络

BP(Back Propagation)神经网络是一种按误差逆向传播算法训练的多层前馈网络^[1]。BP 神经网络能学习和存储大量的输入-输出模式映射关系,而无需事前揭示描述这种映射关系的数学方程。它的学习规则使用最速下降法,通过反向传播来不断调整网络的权值和阈值,使网络的误差平方和最小。BP 神经网络模型拓扑结构包括输入层、隐层和输出层。

BP 算法由数据流的正向传播和误差信号的反向传播两个过程构成。正向传播时,传播顺序为输入层、隐层以及输出

层(见图 1),每层神经元的状态只影响下一层神经元。如果在输出层得不到期望的输出,则转向误差信号的反向传播过程。通过这两个过程的交替进行,在权向量空间执行误差函数梯度下降策略,动态迭代搜索一组权向量,使网络误差函数达到最小值,从而完成信息的提取和记忆。

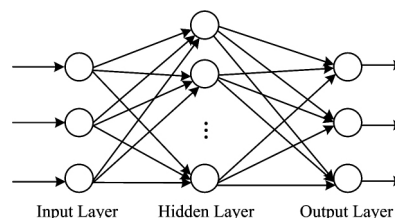


图 1 BP 神经网络

BP 神经网络学习的步骤一般如下。

- (1) 初始化神经网络和学习参数,一般设置初始权矩阵、学习率 η 以及修正参数 α 等;
- (2) 提供训练模式进行网络的训练,直到满足学习要求;
- (3) 前向传播过程:对给定的训练模式进行数据输入,计算网络的输出模式,并与期望模式比较,如果存在误差,且该误差超过规定的阈值,则执行第(4)步;否则,返回第(2)步;
- (4) 后向传播过程:
 - 1) 计算同一层单元的误差 δ_{pj} ;

2)修正权值: $W_{ji}(n+1)=W_{ji}(n)+\eta\delta_{pj}Z_{pi}+\alpha(W_{ji}(n)-W_{ji}(n-1))$,其中 Z_{pi} 为隐层神经元输出,阈值即 $i=0$ 时的连接权值;

3)返回第(2)步。

2 遗传算法

遗传算法(Genetic Algorithm)是一种全局优化算法^[2],它引用生物学的遗传观点,通过自然选择、遗传、变异等作用机制来提高生命体的适应性。

串行方式的遗传算法(Sequential Genetic Algorithm)的步骤描述如下:

(1)对待解决问题进行编码;

(2)设置进化代数计数器 $t=0$,利用随机函数初始化群体 $P(0)=(p_1, p_2, \dots, p_n)$;

(3)对当前群体 $P(t)$ 中每个个体 p_i 计算其适应度,适应度表示该个体性能的好坏;

(4)利用选择算子产生中间群体 $P_r(t)$;

(5)应用交叉算子以及变异算子对 $P_r(t)$ 产生新一代群体 $P(t+1)$,交叉算子和变异算子用于扩大有限个体的覆盖范围,避免陷入局部最优;

(6)进化代数计数器增加,即 $t=t+1$,并判断是否满足终止条件,如果不满足返回第(3)步。

3 遗传算法在优化神经网络结构中的应用

从结构上来说,神经网络包括网络的拓扑结构和节点转移函数。总是期望以简单的网络结构实现所需要的信号处理功能,并尽可能达到较高的性能指标。利用遗传算法设计神经网络可以根据性能评价的标准,如学习速度、泛化能力或结构复杂程度等,搜索结构空间中满足问题要求的最佳结构^[3]。

利用遗传算法优化神经网络结构的算法的步骤描述如下:

(1)随机产生若干个不同结构的神经网络,并对每个网络结构进行编码,使得一个码链对应一个网络结构,所有的码链构成一个种群;

(2)利用多种不同的初始连接权值分别对每个网络进行训练;

(3)计算与各个码链对应的神经网络的误差,利用误差或网络的泛化能力或结构复杂度等策略来确定各个个体的适应度;

(4)选择若干适应度最大的个体构成父本;

(5)利用交叉、变异等遗传操作算子对当前群体进行处理,由此产生新一代群体;

(6)重复上述步骤(2)一步骤(5),直到群体中的某个个体能满足要求为止,即该个体对应的神经网络最优。

4 养殖水域富营养指标

水体富营养化的根本原因是营养物质的增加。目前,一般是以总磷(TP)、总氮(TN)、透明度(SD)以及耗氧量(COD)等指标来衡量水体富营养化的程度^[6]。营养物质的增加与投放饲料有很大的关系。在水产养殖水域投放的饲料主要是各种肥料和复合饲料^[7]。表 1 列出了鄱阳湖地区的水体富营养化程度评价标准。

表 1 鄱阳湖地区富营养化程度的评价标准

指标等级	水质			
	COD/mg/L	TN/mg/L	TP/mg/L	透明度 SD/m
贫营养	0.48	0.08	0.005	8.00
贫-中营养	0.96	0.16	0.010	4.40
中营养	1.80	0.31	0.023	2.40
中-富营养	3.60	0.65	0.050	1.30
富营养	7.10	1.20	0.110	0.73
重富营养	14.00	2.30	0.250	0.40
严重富营养	27.00	4.60	0.560	0.22
异常富营养	54.00	9.10	1.230	0.12

5 数据材料来源

对鄱阳湖新型水产养殖基地的水域进行近 5 年的实地采样,在养殖水域设置 4 个基本采样点,现场测量透明度(SD),并通过水样的实验室分析来测定总磷(TP)、总氮(TN)以及耗氧量(COD)等数据。同时,养殖人员对这 5 年相应的投放饲料数量进行了严格的定量记录。这些数据构成了该养殖水域 BP 神经网络的输入数据矩阵和网络输出数据矩阵。

6 养殖水域 BP 神经网络的构建

利用上述投放的饲料数量作为 BP 神经网络的输入矩阵 P ,并将总磷、总氮、透明度以及耗氧量的采样值作为网络输出矩阵 T ,由此来构建该养殖水域的 BP 神经网络。

由于具有偏差和至少一个 S 型隐含层加上一个线性输出层的网络,能够逼近任何有理函数,本文利用三层 BP 网络建模逼近存在于训练数据中的对应关系。本文合理的隐层节点数为 2,隐层采用 Sigmoid 转化函数 $f(x)=1/(1+e^{-x})$,输出层采用线性转换函数。

以 MATLAB^[8-9]为例,建立如下神经网络算法:

```
net=newff(minmax(P),[20,1],{'logsig','purelin'},'trainlm');
net=train(net,P,T);
minmax_p=minmax(P);
minmax_t=minmax(T);
save('result','net','minmax_p','minmax_t');
```

使用 newff 函数来产生一个 BP 网络,并通过 train 函数对 newff 函数返回的具有权重和阈值的网络对象进行训练,计算网络输入矩阵数据中的最小值与最大值构成的矩阵 $\minmax_p$ 以及由网络输出矩阵数据中的最小值与最大值构成的矩阵 $\minmax_t$ 。

7 结果分析

经 6500 次训练后,对应的训练样本、检验样本和测试样本的平均绝对误差(AAE)分别为 0.1198、0.1257 和 0.1176;平均相对误差(ARE)分别为 3.14%、3.21%和 4.76%;均方根误差(RMSE)分别为 0.1542、0.1556 和 0.1514;相关系数分别为 0.9987、0.9973 和 0.9980。这些结果数据表明,以上建立的神经网络不仅对训练样本达到了很好的拟合效果,而且对检验样本和测试样本也具有很好的泛化和预测效果。

8 性能优化

在利用遗传算法对构建的神经网络进行寻优之前,需要将上述神经网络的输出数据转换成适用于遗传算法寻优的适应矩阵。转化算法描述如下:

```

function t_tmp=anntoga(p_tmp)
global net;global is_request_max;
global minmax_t;
min_t=minmax_t(1);
max_t=minmax_t(2);
t_tmp=sim(net,p_tmp);
if is_request_max
    t_tmp=t_tmp-min_t+(max_t-min_t);
else
    t_tmp=-t_tmp+max_t+(max_t-min_t);
end

```

anntoga 函数是利用刚刚建立好的神经网络来计算函数值,将在遗传算法程序中被调用,并且该函数值应该在 min_max_t 范围内。sim 函数实现神经网络训练结果的 $y=f(x)$ 关系。同时,根据 is_request_max 标识决定是求最大值还是最小值,将网络的输出值转换成适用于遗传算法寻优的适应值。

利用遗传算法进行神经网络寻优的算法描述如下:

```

global net;global is_request_max;
global minmax_t;
load('result');
min_p=minmax_p(1);
max_p=minmax_p(2);
[ori_gen,bits,min_p,max_p]=encoding(min_p,max_p,scale,popsize); %编码函数求初始种群
[p_gen,fitness]=decoding('anntoga',ori_gen,bits,min_p,max_p);
%利用解码函数解码个体
gen=1; %代数初始化
while gen < generation %generation 为总的代数
    [tmp_gen,body_best,max_fitness]=selection(ori_gen,fitness);
    %选择
    tmp_gen=crossover(tmp_gen,pc); %交叉,pc 为交叉概率
    tmp_gen=mutation(tmp_gen,pmut); %变异,pmut 为变异概率
    new_gen=[tmp_gen;body_best;body_best];
    [p_gen,fitness]=decoding('anntoga',new_gen,bits,min_p,max_p);
    gen=gen+1;%下一代
    ori_gen=new_gen;
end
[max,index]=max(fitness);
individual=p_gen(index,:);
if is_request_max
    maximal_t_value=max+min_t-(max_t-min_t)
else
    minmal_t_value=-max+max_t+(max_t-min_t)
end
itsvariable=individual

```

上述遗传算法包括编码、选择、交叉、变异等过程。利用 encoding 函数对输入的数据进行编码,对应的 decoding 函数为解码函数。通过选择算子函数 selection、交叉算子函数 crossover 以及变异算子函数 mutation,在有限次 (generation

限定)循环后,实现神经网络的寻优。

结束语 本文根据鄱阳湖新型水产养殖基地近 5 年的数据建立 BP 神经网络模型,通过大量数据的训练对水产养殖水域饲料投放量与多种富营养指标之间的对应关系进行了有效的拟合,并利用遗传算法实现对神经网络模型的寻优方法,形成养殖水域富营养化预警模型。该模型的建立方法可以避免其他如模糊综合评价和灰色聚类等方法在权值赋值和隶属函数确定中人为因素的影响,使得建立的神经网络模型具有较好的客观性、通用性和实用性。

该模型有助于对养殖水域的富营养化指标进行科学的测评认定,有利于公共决策层及时发现养殖过程中的富营养化倾向,对环境的治理能起到很好的指导工作。

参 考 文 献

- [1] RAFIQ M Y,BUGMANN G,EASTERBROOK D J. Neural network design for engineering applications[J]. Computers and Structures,2001,79(17):1541-1552.
- [2] JAVADI A A,FARMANI R,TAN T P. A hybrid intelligent genetic algorithm[J]. Advanced Engineering Informatics,2005,19(4):255-262.
- [3] DEB K,PRATAP A,AGARWAL S,et al. A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. IEEE Transactions on Evolutionary Computation,2002,6(2):182-197.
- [4] RAMIREZ R,HAZAN A. Inducing a Generative Expressive Performance Model using a Sequential-Covering Genetic Algorithm[C]// Genetic And Evolutionary Computation Conference (GECCO 2007). 2007.
- [5] WANG P,JIANG H L. Neural Network Optimization Algorithm Based on Improvement Genetic Algorithm[C]// International Symposium on Intelligent Information Systems And Applications. 2009.
- [6] CLOERN J E. Our evolving conceptual model of the coastal eutrophication problem[J]. Marine Ecology-Progress Series,2001,210:223-253.
- [7] KARAMOUZ M,TAHERIYOUN M,BAGHVAND A,et al. Optimization of Watershed Control Strategies for Reservoir Eutrophication Management [J]. Journal of Irrigation and Drainage Engineering-Asce,2010,136(12):847-861.
- [8] LIU L H,CHEN J,XU L X. Realization and application research of BP neural network based on MATLAB[C]// International Seminar on Future Biomedical Information Engineering (FBIE 2008). 2008.
- [9] QUN M,HUI H,YANG H,et al. Assessment of water quality of Nansi Lake using BP neural network based on MATLAB[C]// Proceedings of Information Technology And Environmental System Sciences(ITESS 2008). 2008.
- [10] ZENG Q W,XU Z H,WU J A. Prediction of Electricity Consumption Based on Genetic Algorithm-RBF Neural Network[C]// 2ND IEEE International Conference on Advanced Computer Control (ICACC 2010). 2010.