

doi:10.3969/j.issn.2095-0411.2021.04.009

基于交叉验证网格寻优的 GBDT-LSTM 水产养殖溶解氧预测

宦娟, 李慧, 李明宝, 陈波

(常州大学 计算机与人工智能学院, 江苏 常州 213164)

摘要: 为了进一步提高水产养殖溶解氧的预测精度, 提出了基于交叉验证网格寻优的 GBDT-LSTM 溶解氧预测模型。通过梯度提升决策树(GBDT)选取对溶解氧影响度较高的特征因子; 在 Keras 深度学习框架的基础上搭建长短时记忆神经网络(LSTM), 使用交叉验证网格寻优算法对 LSTM 参数进行优化; 将本模型运用到江苏省金坛市渔业基地一标准池塘进行溶解氧预测。试验表明: 该模型的评价指标均方根误差(R_{MSE})、平均绝对误差(M_{AE})、平均绝对百分误差(M_{APE})分别为 0.208, 0.158, 2.635, 其评价指标均优于其他对比预测模型。表明该模型具有较好的预测能力和泛化能力, 能够满足现代化水产养殖的实际需求。

关键词: 水产养殖; 溶解氧; GBDT; LSTM; 交叉验证网格寻优; 组合预测

中图分类号: TP 391; X 832

文献标志码: A

文章编号: 2095-0411(2021)04-0063-09

Prediction of Dissolved Oxygen in GBDT-LSTM Aquaculture Based on Cross-Validation Grid Optimization

HUAN Juan, LI Hui, LI Mingbao, CHEN Bo

(School of Computer & Artificial Intelligence, Changzhou University, Changzhou 213164, China)

Abstract: In order to further improve the prediction accuracy of dissolved oxygen in aquaculture, this paper proposes a dissolved oxygen prediction model based on cross-validation grid optimization GBDT-LSTM. First, the Gradient Boosting Decision Tree (GBDT) is used to select characteristic factors that have a high degree of influence on dissolved oxygen. Then the Long Short-Term Memory Network (LSTM) is built based on the Keras deep learning framework, and the cross-validation grid optimization algorithm is used to optimize the LSTM parameters. Finally, the model in this paper is applied to a standard pond in the fishery base of Jintan City, Jiangsu Province to predict dissolved oxygen. The experimental results show that the mean square root error (R_{MSE}), mean absolute error (M_{AE}) and

收稿日期: 2020-12-14。

基金项目: 国家自然科学基金资助项目(61803050)。

作者简介: 宦娟(1980—), 女, 江苏扬州人, 博士, 副教授。E-mail: huanjuan@cczu.edu.cn

引用本文: 宦娟, 李慧, 李明宝, 等. 基于交叉验证网格寻优的 GBDT-LSTM 水产养殖溶解氧预测[J]. 常州大学学报(自然科学版), 2021, 33(4): 63-71.

mean relative error mean (M_{APE}) of the model are 0.208, 0.158 and 2.635, respectively. The evaluation indexes are better than other comparison prediction models. It shows that the model has good predictive ability and generalization ability, which can meet the actual needs of modern aquaculture.

Key words: aquaculture; dissolved oxygen; GBDT; LSTM; cross-validation grid optimization; combined prediction

溶解氧是水产养殖中的一项重要指标^[1],溶解氧的含量影响着鱼类生长、水质环境。溶解氧含量容易受 pH、温度、氨氮含量等指标的影响^[2],具有非线性、大时滞、不稳定等特点。因此如何利用现有的数据对溶解氧进行预测,及时掌握溶解氧的变化趋势,做出准确、合理的决策已经成为国内外学者研究的热点问题^[3-4]。

为了对溶解氧进行及时准确的预测,近年来国内外很多学者进行了相关研究^[5-7]。如 CHEN 等^[8]提出了一种基于粒子群优化(PSO)的反向传播人工神经网络(BPANN)并结合 Kriging 方法的蟹类养殖池塘溶解氧的三维短期预测模型。徐龙琴等^[9]提出深度信念网络融合最小二乘支持向量机(DBN-LSSVR)对南美白对虾养殖溶解氧进行预测的模型。李明河等^[10]提出了针对污水溶解氧的基于 Levenberg-Marquardt 算法的溶解氧浓度神经网络预测控制。朱成云等^[11]提出了混沌变异的分布估计(CMEDA)算法优化最小二乘支持向量机模型(LSSVR)提高了溶解氧预测精度。上述涉及的预测方法都是利用现有的机器学习算法对溶解氧进行预测,未能结合溶解氧的影响因素有效地揭示其变化趋势,影响水产养殖预测预警系统的准确性。

溶解氧预测数据是一组时间序列数据集^[12],时间序列具有高维度、动态性、易受噪声干扰等特点^[13]。在对时间序列进行预测时,考虑到时间序列的特殊性,应首先对时间序列进行处理,如降维、相似性或相异性处理等,主要分为两类:特征抽取和特征选择^[14]。特征选择以最小化冗余和最大相关性选择特征子集,对原始特征不进行任何改变,比特征抽取具有更好的可读性和可理解性^[15]。GBDT 是一种经典集成模式的特征选择方法^[16],由多棵决策树组成,可解释性强,且对数据字段缺失不敏感,在考虑多组特征相互影响的情况下,具有很好的非线性。目前田嫦丽等^[17]针对互联网广告点击率提出基于 GBDT 的多维特征提取,节省了人工成本,提高了模型精度;王洪伟等^[18]使用 GBDT 模型对评论质量进行特征降维,效果较好。

循环神经网络(Recurrent Neural Network, RNN)是近几年来发展起来的深度学习算法,以序列数据为输入的一种递归神经网络算法^[19-20]。由于 RNN 具有记忆性、参数共享、图灵完备等特点,因此 RNN 被广泛运用于对序列的非线性特征的学习,如语音识别、时间序列预测等^[21]。LSTM 是目前最流行的 RNN 神经网络变种模型, LSTM 在 RNN 的基础上解决了 RNN 的梯度消失和梯度爆炸等问题,很大程度地提升了预测的精度^[22]。李泽龙等^[23]提出 LSTM-RNN 模型进行硅含量预测,预测误差稳定,精度较高;殷兆凯等^[24]构建了基于 LSTM 针对不同预见期的径流预报的流域降雨径流模型,实现了对流域降雨径流的预测。目前, LSTM 神经网络在农业方面的应用还相对较少。

在前人研究的基础上,本文提出了基于交叉验证网格寻优的 GBDT-LSTM 预测模型,拟提高溶解氧预测精度。以中国江苏常州金坛区渔业基地采集到的数据为数据源,对模型进行训练和预测,展开试验。

1 材料与方法

1.1 试验区域

试验数据样本采集自江苏省常州市金坛区渔业基地,该基地位于江苏省金坛市(31°75'N, 119°57'E),

紧邻长荡湖,水产资源丰富,占地面积约 1.53 km²。区域内水质清澈,河流众多,建有池塘循环水系统,配备 DO 传感器、增氧泵、无线监控系统等现代化渔业装备。该基地充分利用综合能源服务平台,对鱼塘实时监测,另增加光储一体化供电方式保障供电可靠性,实现鱼塘的电气化和智能化。本文选取场内一个面积为 667 m² 的标准池塘作为试验区域,养殖品种为河蟹,塘内配有一个增氧泵。

1.2 数据采集

试验数据通过水产养殖物联网在线监测终端数据管理系统采集,其结构如图 1 所示。该水产养殖基地配有 pH 传感器、水温传感器、溶解氧传感器、农业气象站等,监测中心将传感器采集到的水质信息通过 GPRS 传输到服务器,用户可以通过手机或电脑访问服务器,获取实时水质信息。水产养殖物联网在线监测终端数据管理系统每 10 min 对溶解氧、水温、pH、空气温度、空气湿度、风速、气压等生态环境数据在线采样一次,本文选用 2020 年 5 月 24 号至 6 月 23 号共 31 d 的 4 464 个数据作为试验数据。选取前 4 176 个数据为训练集,后 288 个为测试集。

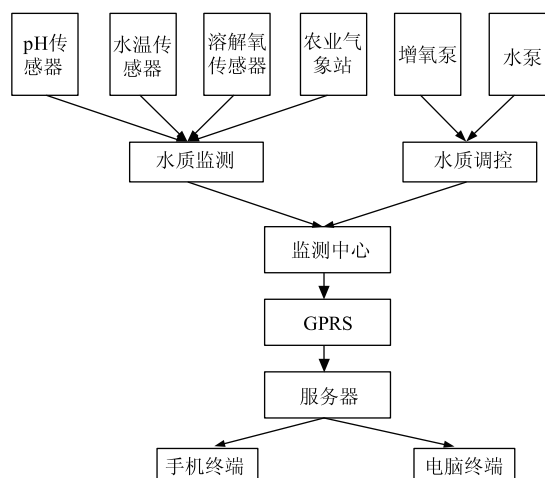


图 1 物联网在线监测系统结构图

Fig.1 Structure diagram IoT online monitoring system

1.3 研究方法

1.3.1 GBDT 基本理论

GBDT 全称梯度提升决策树,由 Jerome Friedman 在 1999 年提出的一种迭代决策树算法的集成模型,是集成法的代表模型。GBDT 以回归树为基学习器,集成梯度、boosting 算法、回归树。成熟的 GBDT 算法步骤为:

- 1) 初始化所有样本在原始数据集的 K 个类别上的估计值,估计使得损失函数极小化的常数值;
- 2) 在当前模型中,对 K 个样本计算损失函数的负梯度值,将其作为残差的估计;
- 3) 将 2) 中得到的残差值作为样本新的真实值,得到新的回归树,估计回归树叶节点的区域,拟合残差近似值;
- 4) 对得到的叶子区域利用线性搜索,计算最佳拟合值,使损失函数极小化;
- 5) 更新回归树;
- 6) 将 2) 到 5) 迭代计算 M 次;
- 7) 得到最终的模型。

GBDT 的特征选择是对各个特征在决策树节点分裂后模型的结构增益,累计求和后对所有的树进行评价,选择合适的特征,在不改变原始特征的前提下,以最小冗余性和最大相关性选择特征子集,实现将原始高维特征映射到新的低维特征空间。在 GBDT 算法中,某个特征 j 的全局重要度是通过特征 j 的重要度平均值来衡量的,公式为

$$\hat{J}_j^2 = \frac{1}{M} \sum_{m=1}^M \hat{J}_j^2(T_m) \quad (1)$$

式中 M 是树的个数。特征 j 在一颗树中的重要度公式为

$$\hat{J}_j^2(T) = \sum_{t=1}^{L-1} \hat{i}_t^2 1(v_t = j) \quad (2)$$

式中: L 是决策树叶子节点的数量, $L-1$ 是非叶子节点的数量; v_t 是与 t 节点相关联的特征; i_t^A 是节点 t 分裂产生的平方损失。

将原始数据样本带入式(1)和式(2),即可得到各个特征的贡献度,根据贡献度,对原始数据样本进行数据选择。

1.3.2 LSTM 基本理论

LSTM 神经网络是一种循环神经网络,基于传统 RNN 神经网络的基础上提出,并在标准 RNN 的基础上,增加了更多的神经网络层,每层以特殊的方式进行交互,同时引入了门的概念,实现选择性地让信息通过。目前,使用最广泛的 LSTM 神经网络基本架构如图 2 所示。其中包括遗忘门、输入门、输出门。LSTM 的第一步要通过遗忘门,读取上一个细胞的输出 h_{t-1} 和当前细胞的输入 x_t ,结合式(3),输出一个 0~1 之间的

数值,决定从细胞状态中丢弃的信息;再通过输入门,首先使用式(4),由 σ 层决定哪些信息需要更新,然后利用式(5), \tanh 层生成备选更新内容;最后利用式(6)确定让多少新的信息加入到当前细胞中,即将上个细胞的丢弃信息和当前细胞的更新信息相结合;最后通过输出门,第一步运用 σ 层来决定输出哪部分的细胞状态,即式(7);第二步通过 \tanh 层处理,得到 -1~1 之间的数值,将其与式(7)中数值利用式(8)相结合,最终决定细胞状态的哪些内容输出。整个 LSTM 的基本架构的主要公式表示为:

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (3)$$

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (4)$$

$$\tilde{C}_t = \tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (5)$$

$$C_t = f_t * C_{t-1} + i_t * \tilde{C}_t \quad (6)$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (7)$$

$$h_t = O_t * \tanh(C_t) \quad (8)$$

式中: f_t, i_t, O_t 分别为 t 时刻遗忘门、输入门、输出门的计算方式; \tilde{C}_t 为 t 时刻的备选更新内容的计算方法; C_t 为 t 时刻记忆细胞的计算方法; σ, \tanh 为 sigmoid 和双曲正切激活函数; W_f, W_i, W_c, W_o 为权重矩阵变量; b_f, b_i, b_c, b_o 为偏置量; h_{t-1} 为上一个细胞的输出; x_t 为当前细胞的输入; b 为神经网络的偏置; w 为神经网络的权重。

LSTM 中有众多的参数,其中主要的参数有批尺寸、迭代次数、优化算法、学习率、网络权值初始值、神经元激活函数、正则化、隐藏层神经元数量等。每个参数对整个 LSTM 神经网络的有着不同的作用。例如学习率的设定过高或者过低,整个网络的学习效果将会变得非常差、学习速度非常慢,甚至没有任何的进展;激活函数的选择不当容易出现梯度消失等问题,进而增加深度学习网络的难度;迭代次数的数量过少,整个网络就没有足够时间学习适合的参数,过多又会导致网络对训练集过拟合。因此,整个参数的选择对于 LSTM 神经网络预测精度至关重要。

1.3.3 基于交叉验证网格寻优的 GBDT-LSTM 的预测模型

对于溶解氧的预测,由于影响溶解氧的众多因素之间存在不同程度的相似性,导致溶解氧的预测存在信息冗余,增加了 LSTM 对其预测的时间,降低了算法的效率。因此,本文结合以上算法提出了基于交叉验证网格寻优的 GBDT-LSTM 溶解氧预测模型克服这一不足。算法在使用 LSTM 对溶解氧进行

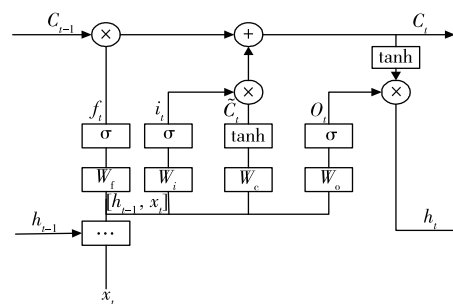


图 2 LSTM 基本架构

Fig.2 Basic architecture of LSTM

预测前,先利用 GBDT 对水质数据进行特征选择^[25-26],去除冗余信息,选择对溶解氧影响较大的特征,解决 LSTM 在预测时存在的信息的冗余问题。

在使用 LSTM 对溶解氧进行预测时,选取不同的参数值对预测结果有着较大的影响,因此有必要对 LSTM 的参数进行优化,即选择合适的 LSTM 参数。网格搜索法是目前比较直观,能够穷尽搜索直到得到 LSTM 最优化参数的参数优化方法^[27]。网格搜索就是将所有可能的参数进行组合,然后对各组参数进行训练,并验证结果是否是最优的,即找到最优的参数组合。交叉验证是一种统计学方法,可以消除由取样的随机性带来的训练误差^[28]。基于此,本文将交叉验证与网格搜索相结合,采用交叉验证网格寻优调整 LSTM 参数,提高参数选优的准确度和效率,极大地降低了样本抽样的随机性对整个算法性能的影响。

为了提高溶解氧的预测精度,本文将 3 种算法进行组合,构建了基于交叉验证网格寻优的 GBDT-LSTM 预测模型,主要包括获取水质数据、溶解氧预测的特征选择、交叉验证网格寻优优化 LSTM 参数、LSTM 对溶解氧进行预测、模型评价、模型应用,具体的溶解氧预测过程如图 3 所示。

1)数据采集:通过在线监测终端数据管理系统实时获取试验基地传感器和小型气象站采集到的水质和天气数据。

2)特征选择:利用 GBDT 选择出对溶解氧有关键影响的特征,降低了输入数据的维度,解决了信息冗余的问题。

3)参数调优:将数据集划分为训练集和测试集,利用交叉验证网格寻优对 LSTM 神经网络参数进行调整,提高了 LSTM 的预测精度。

4)数据预测:将训练集输入模型,构建基于交叉验证网格寻优的 GBDT-LSTM 溶解氧预测模型。

5)对比分析:运用测试集对整个模型的性能进行测试,与其他预测模型进行对比,实现对溶解氧的精准预测。

1.4 模型评价指标

对整个预测模型使用平均绝对误差(M_{AE}),平均绝对百分误差(M_{APE})和均方根误差(R_{MSE})对预测性能进行评价,这 3 个指标都能从不同的数学角度反映出实际预测误差的大小。式(9)为平均绝对误差;式(10)为平均绝对百分误差;式(11)为均方根误差。

$$M_{AE} = \frac{1}{N} \sum_{i=1}^N |x_i - x'_i| \quad (9)$$

$$M_{APE} = \frac{100}{N} \sum_{i=1}^N \frac{|x_i - x'_i|}{x_i} \quad (10)$$

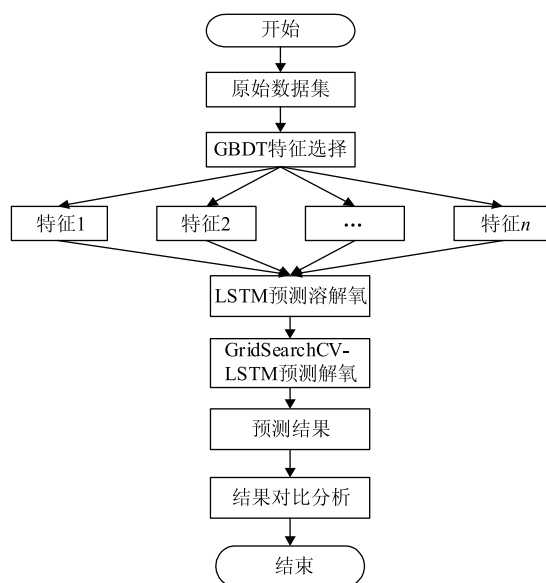


图3 基于交叉验证网格寻优的 GBDT-LSTM 溶解氧预测流程图

Fig.3 GBDT-LSTM dissolved oxygen prediction flow chart based on cross-validation grid optimization

$$R_{\text{MSE}} = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - x'_i)^2} \quad (11)$$

式中: N 为测试集样本总数; x_i 为溶解氧真实值; x'_i 为溶解氧预测值。

2 结果与讨论

2.1 仿真分析

2.1.1 基于 GBDT 的特征选择

水产养殖在线监测系统监测数量较多,类型不一,目前包括时间、溶解氧、水温、pH、湿度、气温、压强、风速、风级等,这些因子对溶解氧的影响程度各不相同,若全部直接作为模型的输入会使整个算法过于庞大,计算量过大,运算时间过长。这些因子中除去溶解氧自身作为研究对象外,时间因素是必不可少的,对于剩余因子,本文使用 GBDT 算法来进行特征选择。图 4 是 GBDT 算法计算出的水温、湿度、气温、pH、风速、风级、压强对于溶解氧的特征重要度。

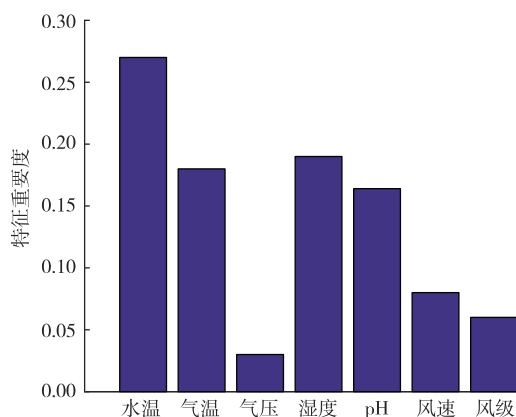


图 4 各因子特征重要度

Fig.4 Characteristic importance of each factor

从图中可以较容易地观察到各因子的特征重要度大小依次排序为水温、湿度、气温、pH、风速、风力、压强,且水温对溶解氧的影响较大,湿度、气温、pH 对溶解氧的影响相差不大,但与风速、风级、气压的特征重要度相比,湿度、气温、pH 更为重要。因此本文选取水温、湿度、气温、pH、溶解氧作为主要特征,降低了数据的维度和计算的复杂度,以此构建溶解氧预测模型的输入样本。

2.1.2 基于 LSTM 的溶解氧预测

根据选择出的特征,选择合适的预测算法对溶解氧进行预测,溶解氧的预测是依赖于时间序列的,即依赖于过去一段时间内溶解氧的时间特征来预测未来一段时间内溶解氧的变化,这样的时间序列预测模型是依赖于事件发生的顺序的,同样的值输入模型的顺序改变了产生的结果也会不同。所以本文选取对时间序列预测能力较强的 LSTM 神经网络算法进行预测。但是 LSTM 中有众多的参数,在不通过算法的情况下对参数进行取值时,一般通过损失函数 train loss 和 val loss 的变化情况手动调整参数。最终,得到较合适的 LSTM 参数值的组合。

表 1 为 LSTM 神经网络模型在参数调整的过程中精度较高的 4 个结果精度分析,从表中可以清晰地看出:当 epochs 为 1 000, batch_size 为 20, 学习率为 0.000 1, 时间步为 10, 隐藏层节点数为 30, 优化算法为 SGD 时整个模型精度较高。图 5 为利用 LSTM 进行溶解氧预测测试集数据的变化趋势。从图 5 和表 1 中可以看出通过手动调整参数的值, LSTM 的预测模型的精度和拟合性能已经有所提高,但是调节过程过于繁琐。

表 1 不同参数精度对比

Table 1 Comparison of different parameters

迭代次数	批尺寸	学习率	时间步	隐藏层节点数	优化算法	均方根误差
1 000	20	0.000 1	10	30	SGD	0.235 6
1 400	25	0.000 1	20	13	Adam	0.241 2
1 200	50	0.001 0	30	30	RMSprop	0.246 7
1 300	10	0.001 0	10	30	SGD	0.253 4

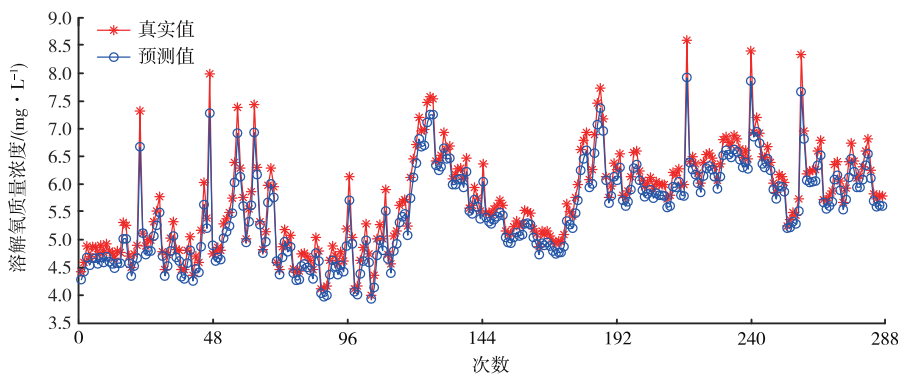


图 5 LSTM 溶解氧预测

Fig.5 LSTM dissolved oxygen prediction

2.1.3 交叉验证网格寻优改进 LSTM 参数

由于通过观察损失函数图对参数进行调整存在人为选择参数的盲目性,同时调整各个参数的值,找到最优参数的组合耗时较长。故本文选用交叉验证网格寻优对参数进行自动寻优,很大程度地避免了盲目地人为调参,极大地缩短了计算时间。

表 2 为交叉验证网格寻优对 LSTM 进行自动调优输出得分最高的 5 个结果,从表中可以看出当 epochs 为 1 300, batch_size 为 40, 学习率为 0.001, 时间步为 20, 隐藏层节点数为 25, 优化算法为 Adam 时整个组合得分最高,即当前组合是自动寻优后得到的最佳参数组合。图 6 为基于交叉验证网格寻优的 GBDT-LSTM 溶解氧预测结果,对比图 5 和图 6,可以看出本文模型的拟合效果更好,同时计算时间更短。

表 2 交叉验证网格寻优调优结果

Table 2 Cross-validation of grid optimization

迭代次数	批尺寸	学习率	时间步	隐藏层节点数/个	优化算法	得分
1 300	40	0.001	20	25	Adam	0.872
1 600	60	0.010	10	20	Adam	0.867
1 150	30	0.001	15	25	SGD	0.862
900	10	0.010	20	30	SGD	0.857
1 000	20	0.010	30	16	Adam	0.831

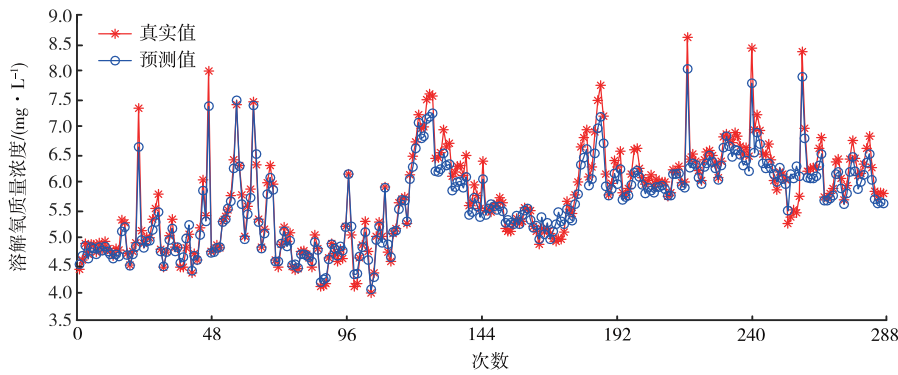


图 6 GBDT-LSTM 溶解氧预测

Fig.6 GBDT-LSTM dissolved oxygen prediction

2.2 讨 论

为了验证基于交叉验证网格寻优的 GBDT-LSTM 模型的预测效果,采用不同的预测模型进行对比,分别为极限学习机(ELM)模型、BP(BP)神经网络模型、粒子群(PSO)改进的最小二乘支持向量机

(LSSVM)模型和单一的 LSTM 模型,5 种模型采用相同的测试集,对未来 12 h 内的溶解氧进行预测,图 7 为 5 种模型的预测曲线。

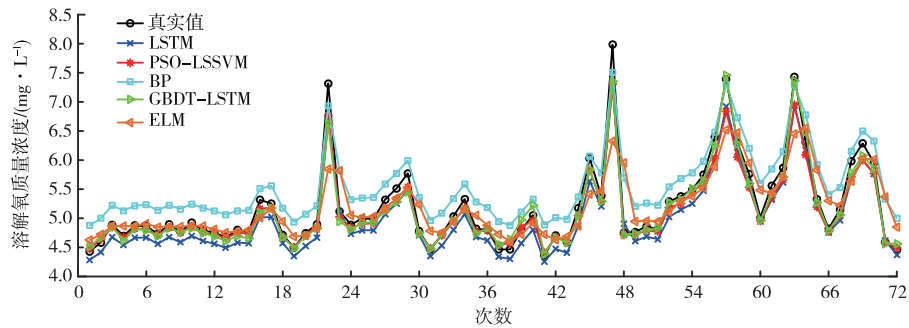


图 7 5 种模型的预测曲线结果

Fig.7 Prediction of curve results for 5 models

从图 7 可以看出,基于交叉验证网格寻优的 GBDT-LSTM 模型的预测结果曲线比其他 4 种模型的预测曲线更加接近真实值,有更好的预测精度,能更好地拟合溶解氧时间序列。为了更准确地比较 5 种模型的预测情况,本文计算出 5 种模型的均方根误差(R_{MSE})、平均绝对误差(M_{AE})、平均绝对百分误差(M_{APE}),其结果精度见表 3。从表 3 中可以看出,建立的 GBDT-LSTM 模型的 $R_{MSE} = 0.208$, $M_{AE} = 0.158$, $M_{APE} = 2.635$,与单一的 LSTM 模型比较 3 项指标都有所提高,说明降低冗余度对模型提高计算精度起着重要的作用,预测精度也略优于其他模型。所以本文提出的预测模型的预测精度高、拟合能力强,能够有效的预测水产养殖中的溶解氧含量,可以为水产养殖的预测预警提供帮助。

表 3 5 种模型预测结果精度分析

Table 3 Accuracy analysis of prediction results for 5 models

模型	平均绝对 误差	平均绝对百分 误差/%	均方根 误差
BP	0.358	6.543	0.379
ELM	0.211	3.608	0.322
PSO-LSSVM	0.227	3.652	0.278
LSTM	0.211	3.659	0.231
GBDT-LSTM	0.158	2.635	0.208

3 结 论

为了进一步提高溶解氧的预测精度,实现池塘现代化管理,本文提出了基于交叉验证网格寻优的 GBDT-LSTM 溶解氧预测模型。首先对采集到的原始数据集进行 GBDT 特征选择,选出对溶解氧有关键影响的特征因子,再利用交叉验证网格寻优算法对 LSTM 神经网络进行优化,自动寻找 LSTM 最优参数对其进行调整,随后使用参数优化过的 LSTM 神经网络模型对溶解氧进行预测,最终得到预测结果。试验结果表明,与 BP,ELM,PSO-LSSVM 以及单一 LSTM 相比,本文提出的模型 $R_{MSE} = 0.208$, $M_{AE} = 0.158$, $M_{APE} = 2.635$,预测效果较好,精度较高,能有效地对水质溶解氧进行预测,对水质智能化管理具有重要的意义。未来计划在溶解氧预测中增加在线预测技术,帮助水产养殖工作者及时做好调控。

参考文献:

- [1]胡金有,王靖杰,张小栓,等. 水产养殖信息化关键技术研究现状与趋势[J]. 农业机械学报, 2015, 46(7): 251-263.
- [2]段青玲,刘怡然,张璐,等. 水产养殖大数据技术研究进展与发展趋势分析[J]. 农业机械学报, 2018, 49(6): 1-16.
- [3] SUN Y N, LI D L, DU S F, et al. WSN-based intelligent detection and control of dissolved oxygen in crab culture[J]. Sensor Letters, 2013, 11(6): 1050-1054.

- [4]郑思宁,刘强,郑逸芳. 规模化水产养殖技术效率及其影响因素分析[J]. 农业工程学报, 2016, 32(20): 229-235.
- [5]TAN G H, YAN J Z, GAO C, et al. Prediction of water quality time series data based on least squares support vector machine[J]. Procedia Engineering, 2012, 31: 1194-1199.
- [6]陈彦,殷建军,项祖丰,等. 基于时间序列模型的海洋溶解氧分析与预测[J]. 轻工机械, 2012, 30(3): 83-87, 96.
- [7]徐敏,曾光明,谢更新,等. 混沌理论在河流溶解氧预测中的应用初探[J]. 环境科学学报, 2003, 23(6): 776-780.
- [8]CHEN Y Y, XU J, YU H H, et al. Three-dimensional short-term prediction model of dissolved oxygen content based on PSO-BPANN algorithm coupled with kriging interpolation[J]. Mathematical Problems in Engineering Theory, 2016, 2016: 1-10.
- [9]徐龙琴,刘双印,张垒,等. 基于 DBN-LSSVR 的南美白对虾养殖溶解氧预测[J]. 仲恺农业工程学院学报, 2017, 30(4): 1-7.
- [10]李明河,周磊,王健. 基于 LM 算法的溶解氧神经网络预测控制[J]. 农业机械学报, 2016, 47(6): 297-302.
- [11]朱成云,刘星桥,李慧,等. 工厂化水产养殖溶解氧预测模型优化[J]. 农业机械学报, 2016, 47(1): 273-278.
- [12]蒋艳,彭期冬,骆辉煌,等. 淮河流域水质污染时空变异特征分析[J]. 水利学报, 2011, 42(11): 1283-1288.
- [13]田中大,李树江,王艳红,等. 经验模式分解与时间序列分析在网络流量预测中的应用[J]. 控制与决策, 2015, 30(5): 905-910.
- [14]姚旭,王晓丹,张玉玺,等. 特征选择方法综述[J]. 控制与决策, 2012, 27(2): 161-166, 192.
- [15]张丽新,王家钦,赵雁南,等. 机器学习中的特征选择[J]. 计算机科学, 2004, 31(11): 180-184.
- [16]GUPTA A, GUSAIN K, POPLI B. Verifying the value and veracity of extreme gradient boosted decision trees on a variety of datasets[C]//2016 11th International Conference on Industrial and Information Systems (ICIIS). Roorkee: IEEE, 2016: 457-462.
- [17]田嫦丽,张珣,潘博,等. 互联网广告点击率预估模型中特征提取方法的研究与实现[J]. 计算机应用研究, 2017, 34(2): 334-338.
- [18]王洪伟,孟园. 在线评论质量有用特征识别: 基于 GBDT 特征贡献度方法[J]. 中文信息学报, 2017, 31(3): 109-117.
- [19]LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.
- [20]BENGIO Y. Learning deep architectures for AI[J]. Foundations and Trends in Machine Learning, 2009, 2(1): 1-127.
- [21]叶佳音,郑文明,李阳,等. 基于深度神经网络的多模态情感识别[J]. Journal of Southeast University (English Edition), 2017, 33(4): 444-447.
- [22]HOCHREITER S, SCHMIDHUBER J. Long short-term memory[J]. Neural computation, 1997, 9(8): 1735-1780.
- [23]李泽龙,杨春节,刘文辉,等. 基于 LSTM-RNN 模型的铁水硅含量预测[J]. 化工学报, 2018, 69(3): 992-997.
- [24]殷兆凯,廖卫红,王若佳,等. 基于长短时记忆神经网络(LSTM)的降雨径流模拟及预报[J]. 南水北调与水利科技, 2019, 17(6): 1-9, 27.
- [25]江萧君,李为民,肖金科. 基于参数优化的 SVM 联合作战战场建设方案优选[J]. 火力与指挥控制, 2017, 42(7): 110-113, 123.
- [26]韩晓慧,杜松怀,苏娟,等. 基于参数优化的最小二乘支持向量机触电电流检测方法[J]. 农业工程学报, 2014, 30(23): 238-245.
- [27]吴峰,陈后金,姚畅,等. 基于网格搜索的 PCA-SVM 道路交通标志识别[J]. 铁道学报, 2014, 36(11): 60-64.
- [28]孙美,张晓琳,冯绍元,等. 基于交叉验证的农田土壤饱和导水率传递函数研究[J]. 农业机械学报, 2014, 45(10): 147-152.

(责任编辑:李艳)