doi:10.6041/j.issn.1000-1298.2025.10.017

基于贝叶斯优化的鲃鱼投饲量多层感知机预测算法研究

朱烨均[1]　徐伟雄[1]　孙志成[1]　李东方[1]　吕铁力[1]　李宏然[2]　肖茂华[1]
(1.南京农业大学工学院,南京 210095;2.江苏海洋大学计算机工程学院,连云港 222005)

摘要:目前水产养殖中依赖经验判断的投饲方法缺乏长期跟踪预测投饲量的能力,易导致饲料浪费、水体污染并对水产动物生长产生负面影响。针对投饲量预测能力缺乏的问题,本文分析了影响鲃鱼投饲量的主要因素,选定水温、溶解氧含量、氨氮含量、鱼均质量、鱼总数及天气条件作为输入变量,建立了多层感知器(Multilayer perceptron,MLP)的投饲量计算模型,通过归一化处理的数据构建模型训练数据集,采用贝叶斯优化算法优化 MLP 模型的超参数,以提升模型性能。使用不同时期的数据进行测试,验证模型性能和泛化能力,在 11#塘口投饲高峰期 8、9 月平均绝对百分比误差稳定在 4% 以下,绝对误差控制在 700 kg 以下。研究结果对提高饲料利用率以及提升养殖效率具有重要意义。

关键词:水产养殖;贝叶斯优化;多层感知器;投饲量;预测算法
中图分类号:S964.3　　**文献标识码**:A　　**文章编号**:1000-1298(2025)10-0165-10　　**OSID**:

Bayesian Optimization-based Multilayer Perceptron Prediction Algorithm for Catfish Feeding Quantity

ZHU Yejun[1]　XU Weixiong[1]　SUN Zhicheng[1]　LI Dongfang[1]　LÜ Tieli[1]　LI Hongran[2]　XIAO Maohua[1]
(1. *College of Engineering*, *Nanjing Agricultural University*, *Nanjing* 210095, *China*
2. *School of Computer Engineering*, *Jiangsu Ocean University*, *Lianyungang* 222005, *China*)

Abstract: At present, the cost of aquaculture is continuously rising, with feed costs accounting for a significant proportion. Most feeding methods rely on empirical judgment and lack long-term tracking and prediction capabilities, which can easily lead to feed waste, water pollution, and negative impacts on the growth of aquatic animals. To address the lack of long-term prediction ability, the main factors affecting catfish feeding were first analyzed, water temperature, dissolved oxygen content, ammonia nitrogen content, average weight of fish, the total number, and meteorological conditions were selected as the input variables for the feeding calculation model, a multilayer perceptron (MLP) feeding calculation model was established, the data was normalized, and a model training dataset was constructed. After the model was established, a Bayesian optimization algorithm was used to optimize the hyperparameters of the MLP model to enhance its performance. Finally, a detailed comparative analysis between the actual feeding quantity and the model-predicted feeding quantity, using data from the three breeding ponds in the test set, verified the model's performance and generalization ability. The average absolute percentage error remained below 4%, and the absolute error was controlled at less than 700 kg during August and September, in the peak feeding period for the 11# ponds. However, during special periods such as fish disease, harvesting, and fishing in the breeding ponds, breeders would adjust the feeding amount accordingly, leading to a large error between the model-predicted feeding amount and the actual feeding amount. This also highlighted the model's limitations in coping with extraordinary events. Meanwhile, the traditional MLP model, random forest model, support vector machine model, and BO-MLP model were selected for comparative tests to verify the superior performance of the described model. Overall, the research results presented were of great significance as an important decision support tool for the aquaculture industry, improving feed utilization and enhancing aquaculture efficiency.

Key words: aquaculture; Bayesian optimization; multilayer perceptron; feeding quantity; prediction algorithm

收稿日期:2025-03-23　修回日期:2025-04-13
基金项目:江苏省重点研发计划项目(BE2022385)和江苏省现代农机装备与技术推广项目(NJ2024-11)
作者简介:朱烨均(1990—),男,讲师,博士,主要从事人工智能技术在农业中应用研究,E-mail:yjzhu@njau.edu.cn
通信作者:肖茂华(1981—),男,教授,博士,主要从事智能农机装备研究,E-mail:xiaomaohua@njau.edu.cn

0 引言

投饲是水产养殖过程中一个至关重要的环节，其效率直接影响到水产生物的生长发育。投饲率又称为投喂率或日粮水平，是指投饲量与养殖鱼类质量的比例[1]，投饲率过高会导致养殖成本和残饵量增加，诱发水体富营养化；而投饲率过低，会使鱼类生长速度减缓、养殖产量下降，两者都会影响水产养殖业的健康发展[2]。鲫鱼作为一种底栖鱼类，摄食行为并不活跃，当其被饲养在条件良好的鱼塘中时，其投饲率通常被设定为约 2%。目前大多数水产养殖的投饲量，还都是依靠养殖人员的生产经验或者是观察鱼群摄食行为来判断[3-4]，尚未有完整的模型能够精确确定最佳投饲量，在投饲智能化方面的应用也还停留在探索研究阶段，没有真正投入产业应用中[5]。因此，投饲量的精准把控就显得尤为重要。

关于精准控制和投饲量优化研究中，DAVISON 等[6]提出通过测量进食时间来预测单个动物的饲料摄入量，并据此估算饲料转化率。这种方法利用电子喂食系统收集进食时间数据，为养殖户提供了优化生产成本的指导；LI 等[7]开发了一种结合鱼类质量预测与需求分析的投饲量优化方法，通过模拟鱼类生长过程来实时决策投饲量，有效提升了鱼类生长性能和养殖产量，但该模型假设环境中的噪声输入是理想的，可能无法精确表示环境对鱼类生长的影响。目前针对智能化投饲的研究主要是利用计算机视觉技术分析鱼群摄食行为，或是通过 pH 值、温度与溶氧量含量等水质参数变化调控投喂量；还有一些研究是通过声学方法监测残留饵料和摄食声音对鱼的摄食行为进行反馈[8]，这些研究成果主要用于理论指导实际养殖生产，应用场景集中在工厂化和海水网箱养殖，同时传统模型并不能完全适应池塘养殖复杂的环境条件，需要综合考虑当地的环境因素和鱼类的真实生长状况。因此，需要通过引入更为先进的技术手段和理论，以更低的模型计算需求来实现投饲量的优化和精准控制，以提高养殖效率和资源利用率，促进水产养殖行业的持续健康发展。

本文提出一种更为科学高效的解决方案，以智能化手段优化投喂管理。综合考量影响投饲量的多种关键因素，包括水温、溶解氧含量、氨氮含量、鱼均质量、鱼数量及天气条件。为克服传统模型在处理此类多参数问题上的局限性，利用多层感知器(MLP)强大的非线性建模能力和灵活的网络结构来完成预测任务。但 MLP 容易陷入局部最优和超参数调整困难的局限，引入贝叶斯优化算法来调优 MLP 模型的超参数，建立基于贝叶斯优化的 MLP 模型来精准预测投饲量。

1 材料与方法

1.1 数据集采集与构建

1.1.1 变量的选择

水质参数如水温和氨氮含量反映了鲫鱼的代谢和饲料投饲情况，溶解氧含量则会显著影响其食欲和消化能力[9]，进而对投饲率产生决定性影响。在实际养殖生产中，养殖技术人员会依据理想生长状态模型并结合水质因素来作出判断。同时，还需综合考虑天气因素，恶劣天气情况下可能会减少甚至暂停投喂，养殖技术人员将依据具体情况，采取人工方式灵活调节投饲量。这要求收集水温和其他水质指标数据，同时也需要准确测定鲫鱼的平均体质量、池塘内的鱼总数以及天气状况。因此，根据对数据采集要求的分析以及实际养殖关注的养殖参数，可以总结出在投饲量计算过程中，对投饲量有影响的有溶解氧含量、氨氮含量，有较大影响的有鱼均质量、鱼总数、水温、天气等。输入变量具体含意如下：

溶解氧含量 X_1：溶解氧是鲫鱼呼吸的氧源，溶解氧含量过低，鲫鱼缺氧会浮头，甚至死亡，进而影响投饲。

氨氮含量 X_2：氨氮不仅是水体中的关键营养物质，同时也具备一定的毒性。氨氮含量过高或过低都会影响投饲量。

鱼均质量 X_3：鲫鱼的平均体质量是决定投饲量的关键性指标之一。在当前阶段，该模型的训练数据来源于每隔大约 15 d 进行一次的人工打样采集，每次采集 100 尾样本的平均质量历史记录数据。随着监测平台的部署和启用，将会采用该平台提供的监测数据作为模型的输入量，以此来进一步优化和调整模型。

鱼总数 X_4：在鲫鱼池塘养殖的多变环境中，难免会出现一定的损耗，养殖工作人员通过定期巡视并捕捞池塘中的鱼来应对这一情况。因此，实际的鱼只总数会根据捕捞上来的死鱼数量进行相应调整。

天气条件 X_5：鲫鱼池塘养殖普遍采用风力投饲机进行饲料分配。然而，在遭遇恶劣的雨雪和大风天气时，投饲的效率可能大打折扣，同时鲫鱼的摄食意愿也会显著下降。这种情况下，饵料转化率会显著降低，不仅饲料利用率下降，还可能间接对水质造成不利影响。将天气划分为好、中、差 3 类，大风、雨、雪天为差，阴、雾、多云为中，晴天为好，并进行序号编码依次为 0，1，2。

水温 X_6：如图 1 显示，理化参数中的相对投饲率 k_t 与水温有强相关性，体现了水温对鲴鱼摄食行为的影响。在华东地区，水温变化较大，在适宜的水温条件下，可显著促进鲴鱼的食欲及其消化吸收能力的提升。

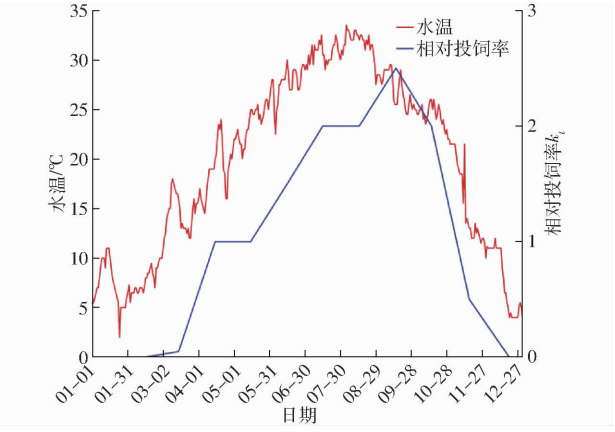


图 1　全年理论投饲率与池塘水温

Fig. 1　Theoretical feeding rate and pond water temperature throughout a year

本研究实施地点为南京通威水产科技有限公司，该水产养殖企业对于 k_t 的设定，提供了一个基于经验的初步估计值，图 1 为一整年相对投饲率与池塘水温变化图，说明 k_t 与池塘水温之间存在显著的相关性。

1.1.2　数据收集与预处理

使用的数据集来源于 2019 年 1 月 1 日至 2023 年 12 月 31 日南京通威水产养殖基地的生产记录。其中 X_1（溶解氧含量）、X_2（氨氮含量）、X_6（水温）水质参数为每日定时按养殖规程采集，X_3（鱼均质量）为养殖人员定期打样综合评价计算出，X_4（鱼总数）为初始总数减去损鱼数，X_5（天气条件）为养殖人员人工判定记录。所用数据皆源自通威水产养殖基地既有的生产监测体系，其采集过程属于员工工作的常规环节，通过内生性数据资源整合机制可实现零边际成本获取。

研究主要关注鲴鱼关键生长阶段——每年 5 月至 10 月，涵盖 8#、9#、11#、13#、26#、43# 6 个塘口的数据。这些池塘养殖密度基本一致，且水源均来自滁河，使用的养殖饲料和养殖方法均遵循公司的标准化规程。8#、13#、26# 3 个塘口在 2019 年至 2021 年期间，5 月 1 日至 10 月 31 日的1 656 d 的养殖数据被选作训练集。而 2022 年同一时间段内的 552 d 数据则用作验证集。此外，9#、11#、43# 3 个池塘在 2023 年 5 月 1 日至 10 月 31 日的 552 d 的养殖数据被用作测试样本。这样的样本分布设计，目的是确保模型在不同养殖池塘及不同年份能展现出良好的泛化能力。部分数据如表 1 所示。

表 1　部分数据

Tab. 1　Part of collection of information

序号	溶解氧含量/（mg·L^{-1}）	氨氮含量/（mg·L^{-1}）	鱼均质量/kg	鱼总数/万尾	天气条件	水温/℃	投饲量/kg
1	6.3	0.37	0.53	4.54	1	22.0	160
2	4.5	0.45	0.56	4.53	2	24.5	300
3	4.0	0.29	0.72	4.39	2	28.5	500
4	5.3	0.47	1.14	4.22	2	25.5	950
5	8.8	0.69	0.75	5.28	2	29.5	750
6	7.9	1.00	0.87	5.24	1	29.0	80
7	7.2	0.59	1.17	5.20	2	30.5	1 500
8	5.9	1.50	0.63	3.48	0	24.0	40
9	9.3	0.43	1.21	3.23	1	19.5	550
10	7.0	0.51	1.15	3.22	2	24.5	1 000

数据归一化：由于所测数据及分类后的水质评价有不同的单位和量级，若直接进行模型的训练，会由于其中某个参数的量级太大而弱化其他量级较小的变量，因此为了避免影响结果的准确性[9]，需要对上述数据集进行归一化处理使得处理过后的数据均处于[0,1]之间，归一化公式为

$$C_i = \frac{c_i - c_{i,\min}}{c_{i,\max} - c_{i,\min}} \tag{1}$$

式中　C_i——归一化值　　c_i——当前值

$c_{i,\min}$——样本数据中最小值

$c_{i,\max}$——样本数据中最大值

表 1 数据经归一化处理后如表 2 所示。

表 2　归一化后部分数据

Tab. 2　Normalized partial data

序号	溶解氧含量	氨氮含量	鱼均质量	鱼总数	天气条件	水温	投饲量
1	0.412	0.231	0.138	0.513	0.5	0.618	0.080
2	0.136	0.281	0.154	0.510	1	0.693	0.150
3	0.091	0.181	0.262	0.463	1	0.858	0.250
4	0.227	0.294	0.754	0.407	1	0.731	0.475
5	0.773	0.431	0.200	0.760	1	0.895	0.375
6	0.636	0.625	0.292	0.747	0.5	0.880	0.040
7	0.500	0.369	0.789	0.733	1	0.925	0.750
8	0.318	0.938	0.169	0.160	0	0.663	0.020
9	0.864	0.269	0.819	0.077	0.5	0.475	0.275
10	0.545	0.319	0.797	0.073	1	0.693	0.500

1.2　投饲量预测方法

1.2.1　投饲量计算方法概述

在时间点 t，传统动物养殖每日投饲量 F_t 计算式为

$$F_t = k_t n_t \overline{w}_t \tag{2}$$

式中　n_t——t 时刻鱼的数量

$\overline{w}_t$——t 时刻鱼的平均质量

计算投饲量的过程涉及多参数和复杂的输入，这使得使用简单的函数模型来确定投饲量变得不合适[10-11]。本实验预测模型的目标是仿照养殖技术人员从实践中积累的经验法则来估算投饲量。采用了数据驱动的方法，应用人工神经网络回归模型来预测日投饲量，并特别选用了基于贝叶斯优化算法的 MLP 回归模型，以期达到预测精度的提高，满足实际应用需求。

1.2.2 多层感知器的应用

MLP 构成了最基础的前馈神经网络形式，其中神经元构成了网络的基本构建块[12]。在鱼群投饲量预测的应用中，MLP 能够处理和学习复杂的输入模式。如图 2 所示，神经元从 n 个不同的神经元接收输入信号，并将这些信号依据各自的权重进行汇总。累加后的信号若超过特定阈值，则通过激活函数进行处理后输出，其模型表示为

$$y_i = f\left(\sum_{i=1}^{n} w_{ij}x_i - \theta_j\right) \tag{3}$$

式中 x_i——输入值

w_{ij}——连接权重

θ_j——阈值

y_i——神经元输出

f——激活函数

激活函数 f 负责引入非线性到神经元输出中，常见的激活函数包括 Sigmoid 函数、双曲正切(Tanh)函数以及线性整流函数(ReLU)等。

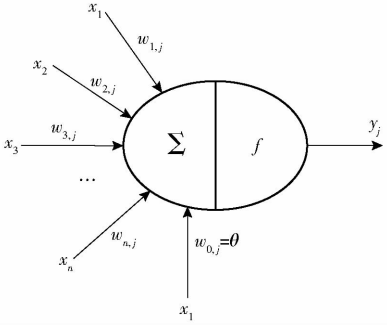


图 2 神经元模型

Fig. 2 Neuron model

MLP 由输入层、输出层以及隐藏层构成，数据处理过程为：输入层输入数据，经过隐藏层的逐层处理后，最终输出结果。如图 3 所示为本文建立的 MLP 网络结构，输入层接收 6 个特征变量，输出层则产生投饲预测量为单个输出值。隐藏层数量设定为 2，足以捕捉复杂数据问题中的非线性关系。每层神经元数量被设为待优化的超参数之一，各层间的神经元实现全连接模式。激活函数选用 ReLU，因为其有助于缓解梯度消失问题，并加快训练速度。同时，界定了需要优化的关键超参数及其取值范围：学习率设定在[0.001, 0.1]之间；隐藏层数量设定为[64, 1 024]，步长为 8，n 初始值设定为 64；迭代次数设定在[20, 100]之间。

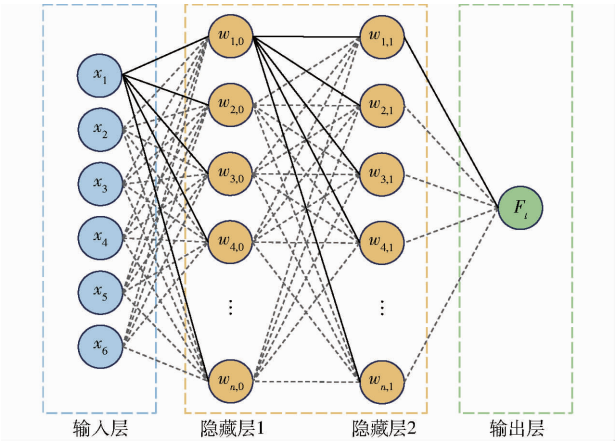


图 3 MLP 网络结构图

Fig. 3 Network structure of MLP

在运用 MLP 以执行回归任务之前，必须先通过训练数据样本来训练网络模型，优化其内部参数[13-15]。当 MLP 产生的输出与样本标签值之间的差异极小时，可视为该网络模型已训练完毕。神经网络的训练过程包括多个迭代周期，每个周期内包含前向传播和反向传播 2 个步骤，通过这一过程，并结合梯度下降方法对网络参数进行逐步调整。图 4 为 MLP 对鲫鱼投饲量进行预测的过程。

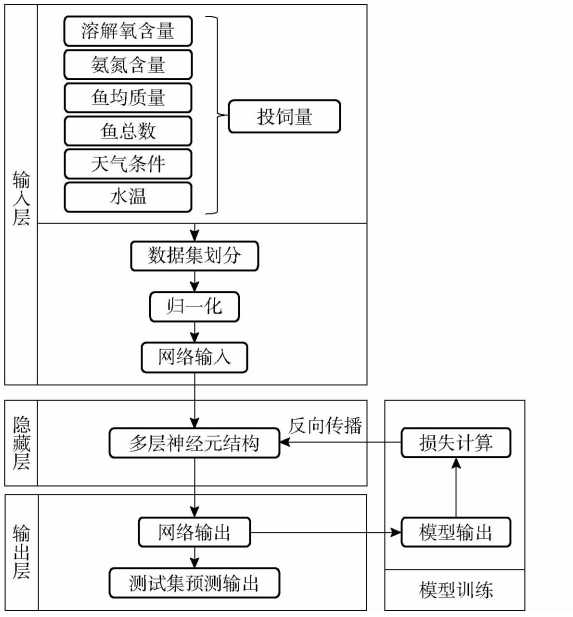


图 4 多层感知机网络训练流程图

Fig. 4 MLP network training

前向传播过程是将信息从输入层经过各隐藏层逐步向前传递，最终到达输出层的过程。在这一过程中，每层的输出都作为下一层的输入。令神经网络的总层数为 L，$a^{(l)}$ 表示第 l 层神经元输出，$\boldsymbol{W}^{(l)}$ 表示从第 $l-1$ 层到第 l 层的权重矩阵，$\boldsymbol{b}^{(l)}$ 为第 $l-1$

层到第 l 层的偏置向量，f_l 为第 l 层所采用的激活函数。则前向传播计算过程可以表示为

$$a^{(l)} = f_l(\boldsymbol{W}^{(l)} a^{(l-1)} + \boldsymbol{b}^{(l)}) \tag{4}$$

权重 $\boldsymbol{W}^{(l)}$ 为神经网络需要学习的参数集合。通过迭代更新这些参数，神经网络寻找最小化损失函数的值，在回归任务中使用均方误差损失函数 L_m 来评估模型性能。该损失函数计算网络输出和实际饲料量之间的误差平方和，即

$$L_m = \frac{1}{2m} \sum_{i}^{m} (a_i - y_i)^2 \tag{5}$$

式中　a_i——饲料量预测值

y_i——饲料量实际值，由式(3)计算得到

m——样本数量

反向传播算法中，计算得到的误差会从输出层逆向传播至网络的每一层，利用梯度下降方法来逐步更新各层权重参数。这一过程持续进行，直至误差降至预定阈值以下，实现对网络性能的优化。

1.2.3　算法模型寻优过程

贝叶斯优化算法(BOA)是一种针对全局优化问题的概率模型方法[16]。适用于寻找计算成本高昂且无法直接通过求导等数学方法定位极值的函数的最优解。通过构建目标函数的概率模型来指导搜索过程，以此来有效地缩小搜索空间并集中在可能的全局最优解区域[17]。当面对高维度计算问题时，传统的遍历搜索方法由于维度爆炸而变得难以实现，而贝叶斯优化则通过学习目标函数的先验分布并逐步更新后验分布，实现对模型超参数的高效优化[18-19]。这一过程不仅考虑了参数间的相互依赖关系，而且通过平衡探索和利用的策略，有效地降低了重复率，优化计算资源的使用[20]；MLP 作为一种常用的深度神经网络结构，在处理复杂非线性问题上具有显著优势[21-22]。然而，MLP 的性能很大程度上依赖于准确的超参数设置，如层数、每层的神经元数量、激活函数类型等。在实践中，寻找最佳的超参数配置难度较大，需要大量的试验，这不仅耗时耗力，而且很难保证找到全局最优解。

将贝叶斯优化(BO)、MLP 的算法特性综合运用，在进行深度学习模型的超参数调优和特定问题解决上，呈现出一种较新的方法论。贝叶斯优化算法发挥其智能搜索的优势，通过对超参数空间的系统遍历，快速识别出可能最大化模型性能的超参数配置。设贝叶斯优化目标函数为

$$x_{\min} = \underset{x \in X}{\operatorname{argmin}} f(x) \tag{6}$$

式中　$x_{\min}$——超参数最终优化结果

$f(x)$——待优化目标函数

X——超参数 x 的配置搜索范围

首先定义超参数空间 $\mathscr{H}$(包括学习率、隐藏层数量和迭代次数)，选定均方误差(MSE)作为目标函数，然后选择超参数初始点来评估目标函数 MSE 并记录结果。由于 $f(x)$ 中每个特定的 x 都有唯一确定的 y 值对应，因此目标函数可以通过无偏差估计被有效观测。

利用先前轮次的采样数据结果与贝叶斯定理来推断目标函数的后验概率分布，基于这一后验分布，算法进一步执行参数优化迭代，最后发现目标函数的极值点。贝叶斯定理公式为

$$P(f|D) = P(f)\frac{P(D|f)}{P(D)} \tag{7}$$

式中，$D = \{(\boldsymbol{x}_1, y_1), (\boldsymbol{x}_2, y_2), \cdots, (\boldsymbol{x}_t, y_t)\}$ 为已观测的集合，即训练样本，$\boldsymbol{x}_t$ 为决策向量，y_t 为相应的观测结果，$P(f)$ 为 $f(x)$ 的先验概率分布，$P(f|D)$ 为给定观测数据集 D 后，函数 $f(x)$ 的后验概率分布，$P(D|f)$ 则是给定函数 f 下观测数据集 D 的似然函数。

采用高斯过程(GP)作为非参数化模型来实施贝叶斯优化过程。GP 模型因其在回归分析、超参数优化等领域的有效性而被广泛应用[23]，其具体公式为

$$f(x) \sim \mathrm{GP}(m(x), k(x, x')) \tag{8}$$

其中　$m(x) = E[f(x)]$，

$$k(x, x') = E[(f(x) - m(x))(f(x') - m(x'))]$$

式中　$m(x)$——均值函数

$k(x, x')$——核函数　　GP——高斯过程

为了简化运算过程，设 $m(x) = 0$，则该情况下 GP 取值只与 $k(x, x')$ 有关。

经过 GP 后，利用采集函数来决定下一个要评估的超参数组合，常用的采集函数有 3 种，分别是概率提升函数(POI)、最大化期望增量函数(EI)和置信边界函数(UCB)。由于上限置信区间函数具备计算简便且能够在探索(深度)与利用(宽度)之间实现平衡的特点，因此选用 UCB 函数作为优化过程中的采集函数。UCB 函数公式为

$$UCB(X) = \mu(x) + \sqrt{\beta}\sigma(x) \tag{9}$$

式中　$\mu(x)$——后验分布均值

β——用于平衡均值与协方差的系数

$\sigma(x)$——协方差

通过输入采集函数确定新的采样点，并生成一个新的高斯分布，并由性质可得联合高斯分布为

$$\begin{bmatrix} y \\ y_* \end{bmatrix} = N\left(0, \begin{bmatrix} \boldsymbol{K}(X, X) & \boldsymbol{K}(X, X_*) \\ \boldsymbol{K}(X_*, X) & \boldsymbol{K}(X_*, X_*) \end{bmatrix}\right) \tag{10}$$

其中　$\boldsymbol{K}(X_*, X) = \boldsymbol{K}^{\mathrm{T}}(X, X_*)$

式中　$\boldsymbol{K}(X, X_*)$——X 与 X_* 的协方差矩阵

可以得到预测分布函数为

$$P(y_*\mid D_{1:t},X_*)=N(\mu_t(x_*),\sigma_t^2(x_{t+1})) \tag{11}$$

其中

$$\begin{cases}\mu_t(x_*)=\boldsymbol{K}^{\mathrm{T}}(X,X_*)\boldsymbol{K}^{-1}(X_*,X)y\\ \sigma_t^2(x_{t+1})=\boldsymbol{K}(X_*,X_*)-\\ \qquad\boldsymbol{K}^{\mathrm{T}}(X,X_*)\boldsymbol{K}^{-1}(X_*,X)\boldsymbol{K}(X,X_*)\end{cases} \tag{12}$$

每次找到一个新的超参数组合后,都会将其加入到训练集中,并重新训练 GP 模型。随着更多数据的积累,模型对目标函数的逼近将更加准确,从而指导更有效的搜索方向。当满足预设的最大迭代次数或连续几次迭代中的改进幅度低于某个阈值时,停止搜索过程,输出最佳超参数配置。

最终经过贝叶斯寻优后最佳超参数配置为:学习率为 0.037 2,隐藏层数量为 688,迭代次数 48,激活函数 ReLU,最低均方误差 3 311.47。

贝叶斯优化过程不仅显著提高了模型调优效率,还增强了模型准确性和泛化能力[24],能够有效解决传统 MLP 因影响投饲量的因素众多而预测精度较差的问题[25]。通过 BO-MLP 算法模型对这些复杂的参数进行综合分析和预测,可以为养殖者提供更为科学和精准的投饲建议,从而优化养殖管理,提高养殖效率,保障鲴鱼健康生长。BO-MLP 算法模型如图 5 所示。

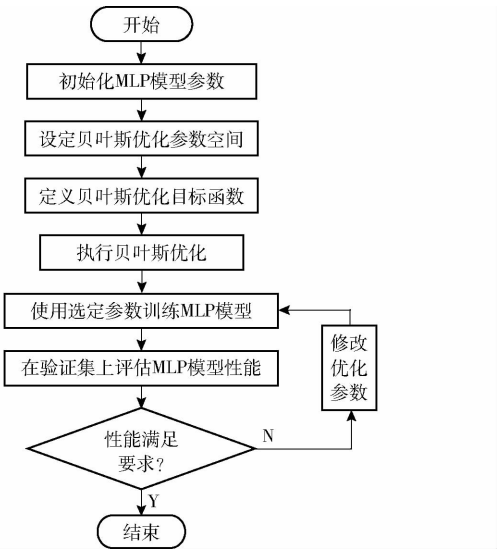


图 5 BO-MLP 算法模型流程图

Fig. 5 Process of BO-MLP model algorithm

贝叶斯优化应用于 MLP 模型的调优,首先建立 MLP 结构框架,确定网络层数、每层节点数以及激活函数[26]。模型权重和偏差初始值通常是随机分配的,打破对称性并开始训练[27]。设定学习率、批次大小和迭代次数。采用 GP 来对超参数空间进行建模,并利用获得的后验分布进行下一步超参数的选择。在每次迭代中,通过 GP 模型的推断来选择下一个最优的超参数组合(引导到目标函数最小化的超参数组合)。迭代过程中,评估每一组超参数对应的模型性能,并综合之前的评估结果动态调整搜索策略。MSE 直接反映模型预测值与真实值之间的偏差程度。贝叶斯优化通过最小化验证集上的 MSE,指导超参数选择,进一步推动模型性能向更优方向发展。最后,这个过程将在满足性能要求时停止。

2 试验

2.1 试验方案

为验证 BO-MLP 模型能够满足水产养殖中生长监测的实际应用需求,采用上述 2023 年 5 月 1 日至 10 月 31 日处理后 9#、11#、43#塘口的 552 组测试样本进行模型性能试验,将溶解氧含量、氨氮含量、鱼均质量、鱼总数、天气条件、水温输入至该模型获取的投饲量数据与公司记录的实际每天投饲量数据进行对照。

同时为评估 BO-MLP 模型在预测任务中的性能,选用 MLP、随机森林(RF)和支持向量机(SVM)几种常见模型进行对比试验。试验主要选用 9#塘口 8 月的数据进行投饲量预测分析,通过百分比误差评估各个模型在准确性方面的表现。

2.2 试验结果与分析

2.2.1 模型性能试验

为了验证本研究所构建模型在计算投饲量上的准确性和效果,将测试集 3 个塘口的 6 个月投饲量计算值与模型预测值进行对比分析,结果如图 6 所示。

采用测试样本对应输出与对照数据之间的绝对误差(AE)、平均绝对误差(MAE)、平均绝对百分比误差(MAPE)、均方误差(MSE)、均方根误差(RMSE)评价各模型性能。

9#塘口面积 3.6 hm^2,初始投苗 4.98 万尾,投苗规格为 0.195 kg/尾。模型预测效果如表 3 所示。

9#塘口 6 个月期间的均方根误差分别为 33.21、175.79、39.39、34.30、72.10、46.70 kg,证明了投饲预测模型在 7 月至 10 月期间具有较好的预测效果。而在 5 月,模型预测性能表现不佳,误差较大。通过审查当时的养殖日志,9#塘口的苗鱼投放时间相较正常投饲时间推迟,5 月初时池塘仍处于鲴鱼驯食阶段,因此投饲量不稳定。进入 6 月,6 月中旬模型预测与实际投饲量之间的差异显著增大,据生产日志记录,塘口发生了鱼病,导致暂停投饲 4 d,逐渐恢复正常投饲节奏后,6 月末模型

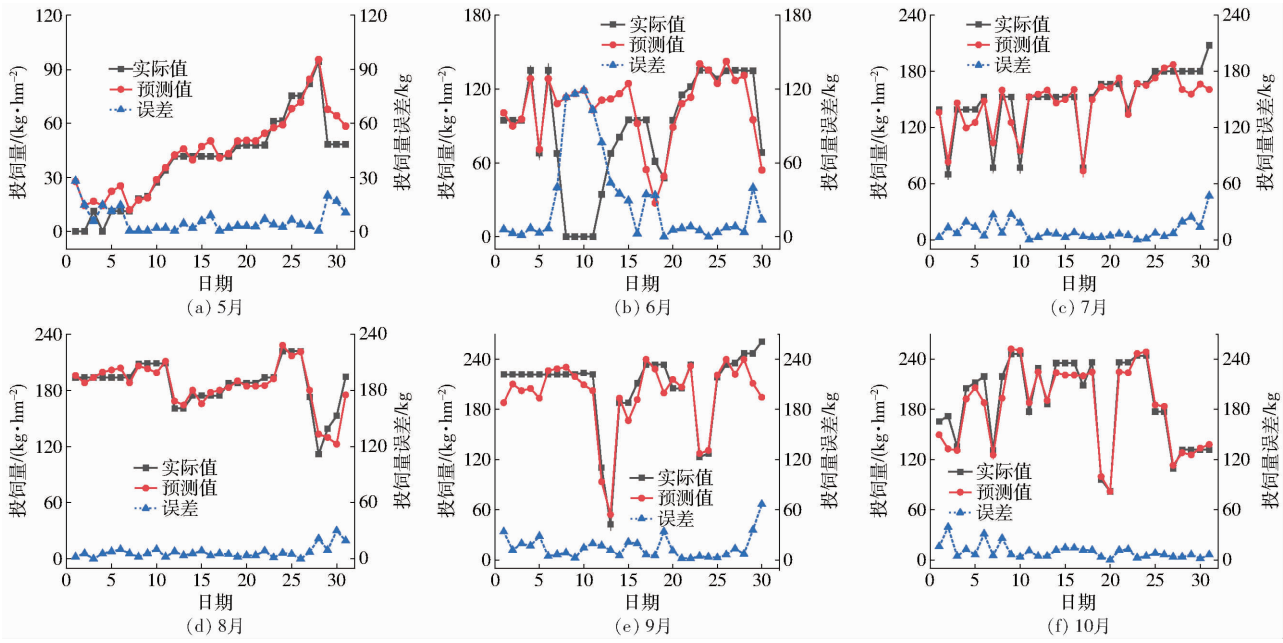


图 6　9#模型测量结果

Fig. 6　Model 9# measurement results

表 3　模型在 9#塘口预测性能

Tab. 3　Prediction performance of model in 9#

月份	AE/kg	MAE/kg	MAPE/%	MSE/kg²	RMSE/kg
5 月	714. 45	23. 05		1 102. 92	33. 21
6 月	3 270. 75	105. 51		30 901. 5	175. 79
7 月	926. 25	29. 87	5. 78	1 551. 38	39. 39
8 月	787. 35	25. 40	4. 12	1 176. 18	34. 30
9 月	1 572. 3	50. 72	7. 37	5 199. 02	72. 10
10 月	1 091. 65	35. 21	5. 09	2 181. 02	46. 70

能较为准确预测投饲量。这一情况表明，当前模型未能有效适应突发鱼病这类非常态事件，需要养殖人员主动介入系统，停止饲料供给。但从全年角度来看，鲴鱼恢复健康后，模型整体上能够满足正常养殖条件下投饲量预测需求。

11#塘口面积 3. 7 hm^2，初始投苗 4. 90 万尾，投苗规格为 0. 206 kg/尾。预测效果如表 4 所示。

表 4　模型在 11#塘口预测性能

Tab. 4　Prediction performance of model in 11#

月份	AE/kg	MAE/kg	MAPE/%	MSE/kg²	RMSE/kg
5 月	333. 09	10. 74	5. 89	199. 58	14. 13
6 月	459. 65	15. 32	5. 09	301. 83	17. 37
7 月	678. 62	21. 89	5. 39	947. 11	21. 89
8 月	615. 45	19. 85	3. 67	641. 89	25. 34
9 月	619. 75	19. 99	3. 73	713. 55	26. 71
10 月	733. 55	23. 66	4. 34	950. 19	30. 83

在整个主要养殖周期中，平均绝对百分比误差的最高值出现在 5 月，为 5. 89%，而最低值则在 8 月，降至 3. 67%。在投饲高峰期 8 月和 9 月平均绝对百分比误差稳定在 4% 以下，绝对误差控制在每月 700 kg 以下，满足鲴鱼的正常生长。证明模型的投饲量预测效果优良，能够与人工经验计算高度一致。

43#塘口 9. 7 hm^2，初始投苗 12. 56 万尾，投苗规格 0. 205 kg/尾。预测效果如表 5 所示。

表 5　模型在 43#塘口预测性能

Tab. 5　Prediction performance of model in 43#

月份	AE/kg	MAE/kg	MAPE/%	MSE/kg²	RMSE/kg
5 月	897. 35	28. 95	11. 72	1 479. 48	38. 46
6 月	1 230. 50	39. 69	5. 29	3 337. 72	57. 77
7 月	1 353. 82	43. 67	3. 80	2 743. 08	52. 37
8 月	1 511. 25	48. 75	3. 68	4 703. 52	68. 58
9 月	1 701. 05	54. 87	4. 23	5 593. 56	74. 79
10 月	3 326. 25	107. 30	9. 45	31 614. 5	177. 80

该塘口绝对百分比误差最高值出现在 5、10 月，分别达到 11. 72%、9. 45%，误差较大。据生产月报分析，5 月上旬也同上述 9#塘口情况，10 月初驯食阶段，该池塘则共销售 1. 95 万尾鱼，每尾鱼的质量为 0. 875 kg，总质量达到 17 062. 5 kg。由于模型主要服务于投苗至捕捞准备期的日常养殖决策，不考虑此大规模捕捞对养殖造成的影响，需要养殖人员主动对投饲策略进行一定调整，因此造成预测模型误差。而在 7 ~ 9 月模型依旧表现良好。

2. 2. 2　不同模型性能对比

为验证 BO - MLP 模型对鲴鱼投饲量的预测能力，选择 MLP、RF、SVM 和 BO - MLP 模型进行对比

试验,使用 9#塘口 8 月数据进行预测,其中 BO - MLP 模型与 MLP 性能对比如图 7 所示,BO - MLP 模型在当月预测值 MAPE 为 4.12%,相较 MLP 模型 4.83% 的平均绝对百分比误差,优化后的超参数配置使得模型性能表现有明显提升。

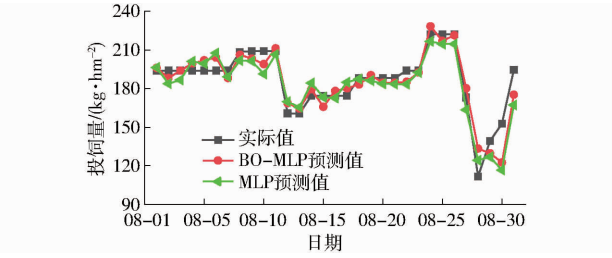


图 7 BO - MLP 模型与 MLP 模型预测值对比

Fig. 7 Comparison of prediction results between BO - MLP and MLP models

RF 模型是一种基于决策树算法构建的集成学习方法,它通过构建多棵决策树并进行集成,从而提高预测的准确性与稳定性。而 SVM 是二分类监督学习算法,广泛应用于分类和回归问题,SVM 通过将输入数据映射到高维空间中并找到一个最优超平面来实现预测[28]。二者都能做出良好的预测任务,图 8 为 RF 模型和 SVM 模型对投饲量预测效果对比。

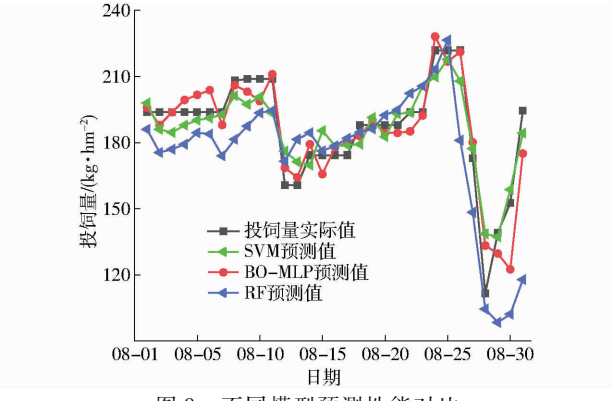


图 8 不同模型预测性能对比

Fig. 8 Comparison of predictive performance of different models

RF 模型和 SVM 模型预测值的 MAPE 分别为 9.24% 和 5.17%,RF 模型虽然能够处理特征之间的非线性关系,但在处理高度复杂的数据模式时,其性能仍有所局限;SVM 通常适用于数据较为平衡且特征维度较低的情况。相比之下,BO - MLP(MAPE 为 4.12%)通过多层神经网络结构,可以更好地捕捉复杂的特征关系和高维数据的潜在模式,其性能显著优于传统 MLP(MAPE 为 4.83%)、RF 和 SVM 模型。贝叶斯优化的引入使得 BO - MLP 能够自动调整超参数,从而提升了模型的泛化能力和精确度。

通过对不同模型的分析和评估,BO - MLP 模型在模拟不同年份以及不同养殖池塘中专业养殖人员的投饲行为上展现了良好的适应性和准确性。不仅能够捕捉到各个养殖周期内饲料需求的变化规律,还能根据不同养殖环境的具体条件,如塘口的尺寸、水质、天气条件以及鱼类的生长情况等因素,进行精准的投饲量预测。高度的灵活性和准确度表明,该模型能够作为养殖行业中重要的决策支持工具,帮助养殖人员优化饲料使用,提高养殖效率,增加经济效益。

3 结论

(1)设计了一种基于贝叶斯优化的 MLP 模型,用于预测水产养殖中的投饲量。该模型综合考虑了影响投饲量的关键因素,包括水温、溶解氧含量、氨氮含量、鱼均质量、鱼总数及天气条件。模型所需的各项数据完全来源于渔业生产中的常规监测指标,其采集过程已嵌入标准化作业流程,在零边际成本条件下即可获取完备数据支持,这一特性显著降低了技术应用门槛,为产业化推广奠定了可复制的数据基础。

(2)利用水产养殖基地不同年份的实际养殖数据对模型效果进行试验,将不同年份数据分为训练集、验证集以及测试集 3 部分,对数据进行归一化处理后导入模型进行计算分析。

(3)BO - MLP 模型在不同塘口的表现显示出了出色的预测性能和泛化能力。尽管在某些特殊情况下,如驯食阶段和疾病发生时,模型的预测性能有所下降,但从全年来看,模型整体上能够满足正常养殖条件下的投饲量预测需求。在 11#塘口 5 月至 10 月期间 MAPE 最高仅为 5.89%,在 8 月和 9 月 MAPE 稳定在 4% 以下,同时经过不同模型性能对比,验证了 BO - MLP 模型优良的预测效果。

参 考 文 献

[1] 卫育良, 王建学, 徐后国, 等. 红鳍东方鲀幼鱼适宜投喂频率和投喂水平的研究[J]. 动物营养学报, 2021, 33(3): 1755 - 1765.
WEI Yuliang, WANG Jianxue, XU Houguo, et al. Study on optimum feeding frequency and feeding level of juvenile takifugu rubripes[J]. Chinese Journal of Animal Nutrition, 2021, 33(3): 1755 - 1765. (in Chinese)

[2] 潘伟平, 王伊凡, 应李民, 等. 投饲频率对青鱼生长、生理生化指标及摄食相关基因表达水平的影响[J]. 中国粮油学报, 2020, 35(4): 83 - 91.

PAN Weiping, WANG Yifan, YING Limin, et al. Effects of feeding frequency on growth, physiological and biochemical indexes and expression levels of ingestion-related genes of black carp[J]. Journal of the Chinese Cereals and Oils Association, 2020, 35(4): 83 – 91. (in Chinese)

[3] 赵思琪,丁为民. 水产养殖精准投喂关键技术研究进展[J]. 智能化农业装备学报(中英文), 2023, 4(1): 42 – 53.
ZHAO Siqi, DING Weimin. Research progress on key technology of precision feeding in aquaculture[J]. Journal of Intelligent Agricultural Mechanization, 2023, 4(1): 42 – 53. (in Chinese)

[4] 俞国燕, 张宏亮, 刘皞春, 等. 水产养殖中鱼类投喂策略研究综述[J]. 渔业现代化, 2020, 47(1): 1 – 6.
YU Guoyan, ZHANG Hongliang, LIU Haochun, et al. Research summary on fish feeding strategies in aquaculture[J]. Fishery Modernization, 2020, 47(1): 1 – 6. (in Chinese)

[5] 高倩倩, 张宇雷, 黄达. 我国智能投饲技术研究进展[J]. 南方农业, 2022, 16(5): 183 – 186.

[6] DAVISON C, MICHIE C, TACHTATZIS C, et al. Feed conversion ratio (FCR) and performance group estimation based on predicted feed intake for the optimisation of beef production[J]. Sensors, 2023, 23(10): 4621.

[7] LI H, CHATZIFOTIS S, LIAN G, et al. Mechanistic model based optimization of feeding practices in aquaculture[J]. Aquacultural Engineering, 2022, 97: 102245.

[8] 李道亮, 李万超, 杜壮壮. 基于音视频信息融合与 Self – Attention – DSC – CNN6 网络的鲈鱼摄食强度分类方法[J]. 农业机械学报, 2025, 56(1): 16 – 24.
LI Daoliang, LI Wanchao, DU Zhuangzhuang. Classification method of feeding intensity of sea bass based on self-attention – DSC – CNN6 and multi-modal fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2025, 56(1): 16 – 24. (in Chinese)

[9] 冯婧然. 基于 ZigBee 的鲫鱼养殖远程监测系统研究[D]. 保定:河北农业大学, 2019.

[10] 王颢然, 肖建华, 欧阳茂霖, 等. 基于贝叶斯累加回归树评估中国非洲猪瘟发生风险区域[J]. 农业工程学报, 2022, 38(11): 180 – 187.
WANG Haoran, XIAO Jianhua, OUYANG Maolin, et al. Assessing the risk areas for African swine fever in China using Bayesian additive regression trees[J]. Transactions of the CSAE, 2022, 38(11): 180 – 187. (in Chinese)

[11] 赵秀艳, 王彬, 都晓娜, 等. 基于时域卷积网络与 Transformer 的茶园蒸散量预测模型[J]. 农业机械学报, 2024, 55(9): 337 – 346.
ZHAO Xiuyan, WANG Bin, DU Xiaona, et al. Evapotranspiration prediction model of tea garden based on temporal convolutional network and Transformer[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(9): 337 – 346. (in Chinese)

[12] 黎亮亮, 张著洪, 张永丹. 多输入傅里叶神经网络及其麻雀搜索优化[J]. 北京航空航天大学学报, 2024, 50(2): 623 – 633.
LI Liangliang, ZHANG Zhuhong, ZHANG Yongdan. Multi-input Fourier neural network and its sparrow search optimization[J]. Jouranl of Beijing University of Aeronautics and Astronautics, 2024, 50(2): 623 – 633. (in Chinese)

[13] 王现磊, 王义江, 陈春江, 等. 基于循环神经网络的弹药着速预测[J]. 火力与指挥控制, 2024, 49(3): 119 – 123.
WANG Xianlei, WANG Yijiang, CHEN Chunjiang, et al. Prediction of ammunition target velocity basedon recurrent neural network[J]. Fire Control & Command Control, 2024, 49(3): 119 – 123. (in Chinese)

[14] 李阳, 彭彦昆, 李永玉. 基于距离校正和数据融合的苹果可溶性固形物含量预测模型优化[J]. 农业机械学报, 2024, 55(增刊 1): 336 – 345.
LI Yang, PENG Yankun, LI Yongyu. Optimization of apple soluble solids content prediction models based on distance correction and data fusion[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(Supp. 1): 336 – 345. (in Chinese)

[15] ZHONG J, ZHAN H, XU F, et al. Research on a multidimensional digital printing image quality evaluation method based on MLP neural network regression[J]. Applied Sciences, 2024, 14(14): 5986.

[16] 高媛媛, 余振华, 杜方, 等. 基于贝叶斯优化的无标签网络剪枝算法[J]. 计算机应用, 2023, 43(1): 30 – 36.
GAO Yuanyuan, YU Zhenhua, DU Fang, et al. Unlabeled network pruning algorithm based on Bayesian optimization[J]. Journal of Computer Applications, 2023, 43(1): 30 – 36. (in Chinese)

[17] 张晶莹, 陈滇豫, 马永胜, 等. 基于贝叶斯分析的干热河谷区橙子林冠层蒸腾耗水模拟[J]. 农业机械学报, 2024, 55(1): 305 – 315.
ZHANG Jingying, CHEN Dianyu, MA Yongsheng, et al. Canopy transpiration water consumption simulation of orange forest in dry and hot valley area based on Bayesian analysis[J]. Transactions of the Chinese Society for Agricultural Machinery, 2024, 55(1): 305 – 315. (in Chinese)

[18] TEMYANOV K B, NIGMATULLIN R. Concentration of measure and global optimization of Bayesian multilayer perceptron. part I[J]. Lobachevskii Journal of Mathematics, 2024, 45 (3): 1275 – 1286.

[19] 崔佳旭, 杨博. 贝叶斯优化方法和应用综述[J]. 软件学报, 2018, 29(10): 3068 – 3090.
CUI Jiaxu, YANG Bo. Survey on Bayesian optimization methodology and applications[J]. Journal of Software, 2018, 29(10): 3068 – 3090. (in Chinese)

[20] 孟祥俊, 陈进东, 张健. 融合非结构化文本信息的中小企业信用风险预测研究[J]. 系统科学与数学, 2024, 44(6):

1608 – 1629.
MENG Xiangjun, CHEN Jindong, ZHANG Jian. Research on credit risk prediction for small and medium-sized enterprises by integrating unstructured textual information[J]. Journal of Systems Science and Mathematics, 2024, 44(6): 1608 – 1629. (in Chinese)
[21] MPOUZIOTAS D, BESHARAT J, TSOULOS G I, et al. AliAmvra-enhancing customer experience through the application of machine learning techniques for survey data assessment and analysis[J]. Information, 2024, 15 (2): 83.
[22] HULIN J, YONG-GUK K, ZHIRAN J, et al. Optimization and analysis of bioenergy production using machine learning modeling: multi-layer perceptron, Gaussian processes regression, K-nearest neighbors, and artificial neural network models [J]. Energy Reports, 2022, 8:13979 – 13996.
[23] 赵帅, 李妍君, 熊伟丽. 基于 KPCA – Bagging 的高斯过程回归建模方法及应用[J]. 控制工程, 2019, 26(1): 131 – 136.
ZHAO Shuai, LI Yanjun, XIONG Weili. Gaussian process regression modeling based on KPCA-bagging algorithm and its application[J]. Control Engineering of China, 2019, 26(1): 131 – 136. (in Chinese)
[24] ZHANG Xinyi, DAI Chengyuan, LI Weiyu, et al. Prediction of compressive strength of recycled aggregate concrete using machine learning and Bayesian optimization methods[J]. Frontiers in Earth Science, 2023, 11: 1112105.
[25] ZHANG Bin, HE Jiawen, LIU Peishun, et al. An automated multi-layer perceptron discriminative neural network based on Bayesian optimization achieves high-precision one-source single-snapshot direction-of-arrival estimation[J]. Scientific Reports, 2024, 14(1): 10300.
[26] LEE H, CHO W, PARK H J, et al. Predicting liquid natural gas consumption via the multilayer perceptron algorithm using Bayesian hyperparameter autotuning[J]. Energies, 2024, 17(10): 2290.
[27] 孙龙清, 吴雨寒, 孙希蓓, 等. 基于 IBAS 和 LSTM 网络的池塘水溶解氧含量预测[J]. 农业机械学报, 2021, 52(增刊): 252 – 260.
SUN Longqing, WU Yuhan, SUN Xibei, et al. Dissolved oxygen prediction model in ponds based on improved beetle antennae search and LSTM network[J]. Transactions of the Chinese Society for Agricultural Machinery, 2021, 52(Supp.): 252 – 260. (in Chinese)
[28] 高振, 卢彩云, 李洪文, 等. 基于计算机视觉的种子分布信息检测关键技术研究现状与趋势[J]. 智能化农业装备学报(中英文), 2023, 4(3): 50 – 60.
GAO Zhen, LU Caiyun, LI Hongwen, et al. Research progress and the prospect of crucial technology of seed spacing information detection based on computer vision[J]. Journal of Intelligent Agricultural Mechanization, 2023, 4(3): 50 – 60. (in Chinese)

(上接第 146 页)

[25] 袁华荣. 南海北部三种典型放流鱼类幼鱼驯化技术初步研究[D]. 上海:上海海洋大学,2012.
YUAN Huarong. Preliminary research on the taming technology of 3 typical releasing juvenile fishes in northern south China sea [D]. Shanghai:Shanghai Ocean University,2012. (in Chinese)
[26] HY/T 0341—2022 人为水下噪声对海洋生物影响评价指南[S]. 2022.
[27] TOUGAARD J, HENRIKSEN O D, MILLER L A. Underwater noise from three types of offshore wind turbines: estimation of impact zones for harbor porpoises and harbor seals[J]. Journal of the Acoustical Society of America, 2009, 125(6):3766 – 3773.
[28] 牛富强,杨燕明,许肖梅,等. 海上风电场运营期水下噪声测量及特性初步分析[J]. 振动与冲击, 2016, 35(12):215 – 220.
NIU Fuqiang, YANG Yanming, XU Xiaomei, et al. Measurements and analysis of underwater noise from the operation of offshore wind farms[J]. Journal of Vibration and Shock, 2016, 35(12):215 – 220. (in Chinese)
[29] 苏文,吴霓,章柳立,等. 海上风电工程对海洋生物影响的研究进展[J]. 海洋通报, 2020, 39(3):291 – 299.
SU Wen, WU Ni, ZHANG Liuli, et al. A review of research on the effect of offshore wind power project on marine organisms [J]. Marine Science Bulletin, 2020, 39(3):291 – 299. (in Chinese)
[30] 黄六一,王羿宁,黄桂芳,等. 海上风电场对鱼类福利的影响研究进展[J]. 水产学报, 2022,46(11):2226 – 2240.
HUANG Liuyi, WANG Yining, HUANG Guifang, et al. Advances in research on the effects of offshore wind farm on fish welfare[J]. Journal of Fisheries of China, 2022,46(11):2226 – 2240. (in Chinese)
[31] 阳杰,张建华,马兆荣,等. 海上风电与海洋牧场融合发展趋势与技术挑战[J]. 南方能源建设,2024,11(2):1 – 16.
YANG Jie, ZHANG Jianhua, MA Zhaorong, et al. Development trend and technical challenges of the integration of offshore wind turbine with marine ranch[J]. Southern Energy Construction,2024,11(2):1 – 16. (in Chinese)
[32] POPPER A N, HAWKINS A D, FAY R R, et al. ASA S3/SC1.4 TR—2014 Sound Exposure Guidelines for Fishes and Sea Turtles: a technical report prepared by ANSI-accredited standards committee S3/SC1 and registered with ANSI[M]. Springer International Publishing,2014.