

基于 ELM 的改进 CART 决策树回归算法^①



王宏, 张强, 王颖, 郭玉洁

(东北石油大学 计算机与信息技术学院, 大庆 163318)

通讯作者: 张强, E-mail: dqpi_zq@163.com

摘要: 为提高 CART (Classification And Regression Tree) 决策树回归算法的准确性, 提出一种基于 ELM (Extreme Learning Machine) 的改进 CART 决策树回归算法——ELM-CART 算法. 所提算法主要是在 CART 回归树创建过程中, 在每个叶节点使用极限学习机建模, 可以得到真正意义上的回归预测值, 提高泛化能力, 弥补 CART 决策树回归算法本身的容易过拟合以及预测输出为定值等缺点. 实验结果表明, 所提算法能够有效提高回归分析中目标数据的预测准确性, 其准确性优于所对比算法.

关键词: CART 决策树回归算法; 极限学习机; 叶节点; 预测输出; 回归分析

引用格式: 王宏, 张强, 王颖, 郭玉洁. 基于 ELM 的改进 CART 决策树回归算法. 计算机系统应用, 2021, 30(2): 201-206. <http://www.c-s-a.org.cn/1003-3254/7789.html>

Improved CART Decision Tree Regression Algorithm Based on ELM

WANG Hong, ZHANG Qiang, WANG Ying, GUO Yu-Jie

(School of Computer and Information Technology, Northeast Petroleum University, Daqing 163318, China)

Abstract: In order to increase the accuracy of the Classification And Regression Tree (CART) regression algorithm, we propose an improved CART regression algorithm based on Extreme Learning Machine (ELM-CART for short). The proposed algorithm mainly applies the ELM for modeling at each leaf node in the process of creating a CART, which can get the true regression prediction value, improve the generalization ability, and compensate for such disadvantages of the CART regression algorithm as easy overfitting and constant predictive output. The experimental results show that the proposed algorithm can effectively improve the prediction accuracy of target data in regression analysis, and its accuracy is higher than that of the counterparts.

Key words: CART decision tree regression algorithm; Extreme Learning Machine (ELM); leaf nodes; predictive output; regression analysis

在机器学习中, 回归问题同分类问题一样一直是一个兼具魅力和挑战的热门研究领域. 回归建模过程是通过算法学习寻找已有数据之间的关联, 求得一个复杂函数, 然后给出预测结果, 常用的回归算法有多元线性回归^[1]、支持向量机^[2]、决策树^[3]、极限学习机^[4]、人工神经网络^[5]等. 决策树算法由于具有抽取规则简单、对数据缺失值不敏感、模型易于构建、速度快、

效率高等特点, 深受相关学者专家重视. CART (Classification And Regression Tree) 决策树广泛应用于分类和回归问题, 本文主要针对研究回归问题, 目前 CART 回归树在医学, 农业, 环境, 石油, 煤炭, 电力, 气象等诸多领域都得到了大量的研究和应用, 效果显著, 对专业研究、生产实践都起到帮助和指导作用. 比如 Sut 等^[6]将回归树应用于颅脑损伤死亡率的预测; Müller 等^[7]

① 基金项目: 国家自然科学基金 (617020936); 大庆市指导性科技项目 (zd-2019-09)

Foundation item: National Natural Science Foundation of China (617020936); Science and Technology Guiding Program of Daqing Municipality (zd-2019-09)

收稿时间: 2020-06-23; 修改时间: 2020-07-14, 2020-07-29; 采用时间: 2020-07-31; csa 在线出版时间: 2021-01-27

将增强回归树用于比较阿尔巴尼亚和罗马尼亚的耕地放弃决定因素的研究; Park 等^[8] 使用回归树建立了预测水库系统中的叶绿素的应激响应模型来控制藻华; Larraondo 等^[9] 提出一种基于循环回归树的机场天气预报系统; 董红召等^[10] 将回归树用于城市交通道路氮氧化物浓度的预测研究; 郑向群等^[11] 提出一种空间数据挖掘算法, 使得空间关系变得简单, 以此提高了算法效率; 杜春蕾等^[12] 提出一种改进的 CART 算法决策树模型, 用来处理煤层底板突水数据预测, 缩短了算法运行的时间, 也提高了准确率; 刘云翔等^[13] 引入 Fayyad 边界判定定理, 使得挑选属性最优阈值的速度更快, 提高了运行效率; 毕云帆等^[14] 将梯度提升机与决策树相结合得到梯度提升决策树, 提高了电力短期负荷预测精度; 李正方等^[15] 调整决策树生成过程使其具有增量学习的能力, 提高了在气象信息系统中的实用性等等。很多专家学者们深入研究决策树相关算法, 并对决策树算法做出很多的改进优化, 并凭借这些改进算法更好的适应解决了相应研究领域的问题, 也获得了良好的效果, 但决策树本身仍存在问题有待研究改进, 如叶节点的输出值计算方式为该节点处全部样本的平均值, 即模型训练好后的预测输出值为有限个定值, 这种预测输出值计算方法并不合理, 并没有实现真正意义上的回归预测, 导致回归预测的准确性大大降低, 降低了泛化的能力, 可能在一些特定数据集上表现良好, 但不适合广泛的应用于一般数据集。因此, 为了提高准确性, 本文提出了该 ELM-CART 算法, 希望在回归树创建过程中, 通过在叶节点对样本子集分别采用 ELM (Extreme Learning Machine) 算法快速建模, 以此来构建合理的预测输出值, 提高预测准确性, 增强泛化能力, 实现真正意义上的回归预测。

1 相关算法介绍

1.1 极限学习机算法

二十一世纪初, 黄广斌等提出极限学习机算法^[16]。在此之前, 单隐层前馈神经网络 (Single Hidden Layer Feedforward Neural Network, SLFN) 是基于梯度的算法训练而成的, 而极限学习机并没有采用梯度算法, 而是随机生成输入层权重和隐藏层偏置。然后通过已知数据的输入, 在使得损失函数最小的情况下, 得到输出层权重, 到此整个极限学习机就训练完成了, 通过已知的输入层权重, 隐藏层偏置以及输出层权重就可以直

接计算测试数据的预测值了。与其他算法比, ELM 学习速度快、训练参数少、泛化能力强^[17]。其网络结构如图 1 所示。

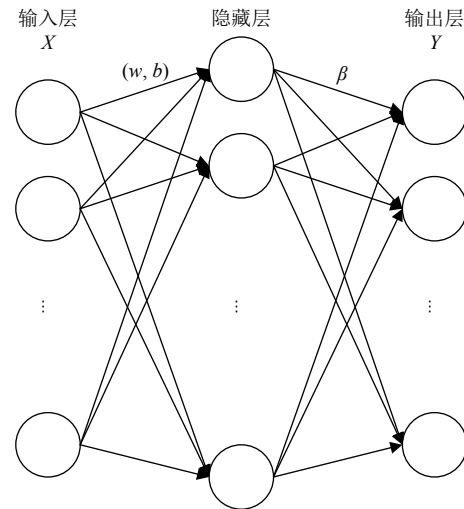


图 1 极限学习机网络

这里假设给定训练集 $\{x_i, t_i | x_i \in R^D, t_i \in R^m, i=1, 2, \dots, N\}$, 其中 x_i 表示第 i 个数据示例, t_i 表示第 i 个数据示例对应的标记, 所有的训练数据组成一个集合, 假设该网络中有 L 个隐层节点, 即:

$$\sum_{i=1}^L \beta_i g(W_i \cdot X_i + b_i) = o_j, \quad j=1, \dots, N \quad (1)$$

其中, 输出层权重为 β_j , 激活函数为 $g(x)$, 输入层权重为 $W_i = [w_{i,1}, w_{i,2}, \dots, w_{i,n}]^T$, b_i 为第 i 个隐层单元偏置。算法学习的期望是最小化真实值与输出值的误差, 可以矩阵表示为:

$$H\beta = T \quad (2)$$

式 (2) 中, 输出层权重用 β 表示, 而 T 表示期望的输出, 隐藏层节点的输出用 H 表示, 其具体形式可表示为:

$$H(x_j) = g(W_i \cdot x_j + b_i) \quad (3)$$

网络中随机生成输入权重 W_i 和隐层偏置 b_i , 然后得到输出矩阵 H , 这时, 极限学习机的训练过程就可以看做是线性问题的求解过程, 因此, 通过以下公式就可以求解出输出权重 β , 可表示为:

$$\beta = H^+ T \quad (4)$$

其中, H^+ 是矩阵 H 的 Moore-Penrose 广义逆, 且可证明求得的解 β 的范数是最小并且唯一的。

1.2 CART 决策树回归算法

在经典机器学习算法中, 决策树相关算法一直是

研究热门,因其独特的树形结构有利于快速处理数据,表达数据之间的关联.分类回归树算法简称 CART 算法^[18,19].最初是由 Breiman^[20]于1984年提出的,可生成分类树或回归树.与 ID3^[21]不能处理连续型数据相比, CART 引入二元切分法即可处理连续型数据,且适合于对多特征变量的复杂数据进行建模,具有抽取规则简单、准确度高、可解释性强的优势,但算法稳定性差,容易过拟合.

假设 X 与 Y 分别为输入和输出变量,且 Y 为连续变量,给定训练数据集:

$$D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\} \quad (5)$$

对于输入空间的划分,子空间及其输出值与叶节点一一对应.这里通过依次遍历每个特征变量及其对应的每个特征值,假设当前切分变量为第 j 个变量,对应切分特征值为 s ,即可划分并定义两个区域:

$$\begin{cases} \alpha_1(j, s) = \{x | x^{(j)} \leq s\} \\ \alpha_2(j, s) = \{x | x^{(j)} > s\} \end{cases} \quad (6)$$

通过上述步骤,输入空间将被不停的分割成 L 个子空间 $\alpha_1, \alpha_2, \dots, \alpha_L$, 每个子空间 α_l 都包含部分的样本数据和输出值 β_l , 这时模型的解可表示为:

$$f(x) = \sum_{l=1}^L \beta_l I(x \in \alpha_l) \quad (7)$$

假设此时已经完成输入空间的划分,这时的损失函数的误差大小可以采用平方误差 $\sum (y_i - f(x_i))^2$ 来对比,通过平方误差最小化的准则,确定最优切分点和预测值.又根据最小二乘的原理可知,子空间 α_l 上的所有输出 y_i 的均值即为 β_l 的最优值 $\hat{\beta}_l$, 即可表示为:

$$\hat{\beta}_l = \text{ave}(y_i | x_i \in \alpha_l) \quad (8)$$

划分输入空间的关键是如何选择最优的切分属性 j 和属性值 s , 用公式表达为:

$$\min_{j,s} \left[\min_{\beta_1} \sum_{x_i \in \alpha_1(j,s)} (y_i - \beta_1)^2 + \min_{\beta_2} \sum_{x_i \in \alpha_2(j,s)} (y_i - \beta_2)^2 \right] \quad (9)$$

通过遍历全部的输入特征变量及其特征值,就可以找到当前最优的切分点 (j, s) , 然后根据切分点划分当前空间为两个子区域,此时若这两个子区域无法再划分时,即可得到相应的最优输出值,表示为:

$$\begin{cases} \hat{\beta}_1 = \text{ave}(y_i | x_i \in \alpha_1(j, s)) \\ \hat{\beta}_2 = \text{ave}(y_i | x_i \in \alpha_2(j, s)) \end{cases} \quad (10)$$

根据以上步骤,若该区域仍可继续划分,则对该区

域重复上述步骤,直到整个回归树完全不可划分停止.

2 ELM-CART 回归算法

相比于线性回归,树回归更适合处理复杂非线性的问题,即便如此,当输入数据特征维度较高,特征之间关系相对更加复杂时,构建全局模型往往导致准确性较低.考虑到提高回归预测的准确性,可以将输入数据集划分成若干个子集更易于建模.

经研究发现一般回归树在建立过程中,就是将输入数据样本空间划分成若干个子样本空间,每个子样本空间都有唯一确定的输出值,这是因为在叶节点采用该节点处的样本平均值作为模型回归预测值,也就是说回归树拟合出来的是一个分段零阶函数,这些有限的离散定值极大程度上降低了决策树在回归问题上的预测准确率,因此,本文认为回归树叶节点处的样本子集还可以做进一步的处理,这里可以对每一个子集分别进行建模,充分利用叶节点处样本子集内部的特征关系,理论上是可以达到提高整个回归预测准确率的效果.结合 ELM 优点,本文提出一种对 CART 回归树算法的改进算法——基于 ELM 的改进 CART 决策树回归算法 (an improved regression algorithm for CART decision tree based on ELM, 简称 ELM-CART 算法).

2.1 算法原理

本算法在回归树创建过程中,采用二元切分法递归创建二叉树,每一次将原输入空间划分成两个独立的子区域,划分时希望划分的两个分支的误差越小越好,这使得每次的划分都是独立且最优的,越往树的底层深入,节点覆盖的样本越少,即随着树的生长,估计越来越不可靠,因此可以通过设置树停止生长的规则来提前结束树的生长,以此避免过拟合,达到最优树.在决策树生成结束后,在每一个叶节点处分别采用极限学习机对子集进行建模,以此来达到提高回归预测的准确性,如图 2 示例所示.

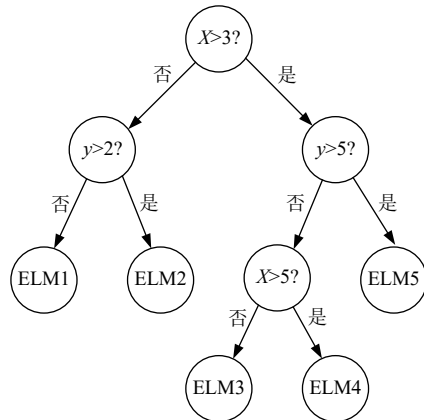
假设 X 与 Y 分别为输入和输出变量,且 Y 为连续变量,训练集为 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$, 其中 $x_i = (x_i^{(1)}, x_i^{(2)}, \dots, x_i^{(m)})$ 为输入特征数据, n 表示为特征维数, $i = 1, 2, \dots, N$, N 表示输入数据个数.

对输入样本空间的划分采用启发式方法,每次划分都循环遍历当前空间中的所有特征变量及其对应的所有特征值,再依据平方误差最小化准则唯一确定最优的切分特征和特征值组合.假设确定当前训练集 R 的第 j 个特征变量 $x^{(j)}$ 及其取值 s 为切分点,则有 $R_1(j, s) =$

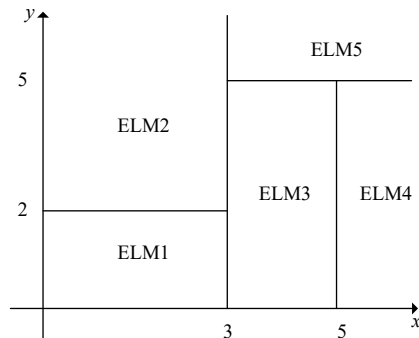
$\{x|x^{(j)} \leq s\}$ 和 $R_2(j, s) = \{x|x^{(j)} > s\}$ 两个划分的定义区域, 因此, 是否划分以及划分的最优切分点选取可由式 (11) 和式 (12) 平方误差和比较得出:

$$err = \sum_{x_i \in R(j,s)} (y_i - c_i)^2 \quad (11)$$

$$err_{1,2} = \min_{j,s} \left[\min_{c_1} \sum_{x_i \in R_1(j,s)} (y_i - c_{1i})^2 + \min_{c_2} \sum_{x_i \in R_2(j,s)} (y_i - c_{2i})^2 \right] \quad (12)$$



(a) 决策树生成结构图



(b) 对应该决策树的最终样本空间划分图

图2 特征空间划分

总之, 就是在小于 err 的 $err_{1,2}$ 中, 找出使得两个子区域平方误差之和 $err_{1,2}$ 最小的 j 和 s , 其中, 这里 c , c_1 和 c_2 的求解是通过 ELM 模型求得的估计值. 这里根据设置的隐含层节点数 L , 随机初始化生成服从任意的连续概率分布的输入权重 w_i 和隐含层阈值 b_i , 就可以通过下式计算得到输出矩阵 H .

$$H_i(x_j) = g(w_i \cdot x_j + b_i) \quad (13)$$

这里 x_i 为输入向量, $g(x)$ 表示激活函数, 这里使用 Sigmoid 函数. 接下来计算此时的输出权重 $\hat{\beta}$, 基于结构风险最小化的原则, 为了避免出现矩阵无法求逆以及防止模型的过拟合, 提高算法的泛化能力和鲁棒性, 引入 L_2 正则化项后, 此时输出权重 $\hat{\beta}$ 的计算公式为:

$$\hat{\beta} \left(H^T H + \frac{1}{C} \right)^{-1} H^T T \quad (14)$$

其中, C 为正则化系数. 由此根据式 (13), 式 (14) 即可分别求得当前训练集 R , R_1 和 R_2 的输出矩阵 H , H_1 , H_2 和输出权重 $\hat{\beta}$, $\hat{\beta}_1$, $\hat{\beta}_2$, 于是通过下式即可分别求得式 (11), 式 (12) 中的 c , c_1 和 c_2 .

$$\begin{cases} c = H\hat{\beta} \\ c_1 = H_1\hat{\beta}_1 \\ c_2 = H_2\hat{\beta}_2 \end{cases} \quad (15)$$

于是, 根据以上步骤, 若该区域仍可继续划分, 则对该区域重复上述步骤, 直到整个树完全不可划分停止.

对数据集进行训练过程中, 控制树的划分次数, 既可以减小时间开销; 又可以防止过度划分导致模型出现过拟合. 因此, 本算法设置 3 个超参数作为算法结束条件, 满足任意一个即训练结束: (1) 树的深度达到给定深度; (2) 误差小于容许的误差下降值; (3) 叶节点处的子区域数据量小于设定的数量.

3 实验分析

3.1 数据选取及处理

本文测试数据来自加州大学欧文分校的 UCI (University of California Irvine) 数据库, 这里面的数据集主要分为分类和回归两大类, 在很多机器学习算法的相关论文中都可以看到这些数据集的应用. 本次实验涉及的数据集分别有波士顿房价数据集, 鱼类毒性数据集, 翼型自噪声数据集, 联合循环电厂数据集, 混凝土抗压强度数据集, 水生毒性数据集以及加利福尼亚房价数据集等, 详细介绍如表 1 所示.

表 1 测试数据集

序号	数据集名称	实例数	特征数
D1	boston_housing DataSet	506	13
D2	QSARfishtoxicity DataSet	908	6
D3	airfoil_self_noise DataSet	1503	5
D4	CombinedCyclePowerPlant DataSet	9568	4
D5	ConcreteCompressiveStrength DataSet	1030	8
D6	QSARaquatictoxicity DataSet	546	8
D7	california_housing DataSet	20640	8

首先下载这些数据集文件, 然后读取不同类型文件并对格式进行处理统一, 再对数据归一化处理, 有利于提高模型的训练速度和准确性, 最后将数据集分为训练集和测试集, 为进一步的实验做好数据准备.

3.2 评价指标

本文采用均方误差 (Mean Squared Error, MSE),

平均绝对误差 (Mean Absolute Error, MAE) 以及 R^2 值 (R Squared) 3 种评价指标。

MSE 值越接近零, 表明模型越好, 其公式如下:

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2 \quad (16)$$

MAE 值越小, 模型也就越好, 其公式如下:

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - \hat{y}_i| \quad (17)$$

R^2 值不同于前两者, 它的范围为 0-1 之间, 并且随着 R^2 值增大, 训练模型的准确率也越来越高, 当 R^2 等于 0 时意味着不必进行预测, 常数模型直接取值的情况; 相反, 当 R^2 等于 1 时, 意味着所有预测值跟真实值是完全相等的, 也就是说模型是不犯错的, 这也是回归分析中追求的理想结果, 其公式如下:

$$R^2 = 1 - \frac{\sum_i (\hat{y}_i - y_i)^2}{\sum_i (\bar{y} - y_i)^2} \quad (18)$$

式中, m 代表样本总数, \bar{y} 为真实平均值, y_i 和 \hat{y}_i 分别为第 i 个样本的真实值和预测值。

3.3 实验结果分析

本文基于以上标准数据集, 将 ELM-CART 回归算法与回归树, ELM, SVR, 人工神经网络算法等进行对比分析, 其中回归树的容许误差下降值默认设为 1, 切分的最小样本数取值在 10 以内; ELM 的 L_2 正则化参数 C 取值为 100 000, 隐层节点数设置有 20, 50, 1000 及 1200; SVR 采用的是使用最广泛的径向基核函数

(RBF); 人工神经网络采用的是两层及三层的神经网络; ELM-CART 算法可以设置树的深度, 一般默认为 2, 通过增加深度可以起到进一步划分数据集的效果, 原 CART 中保留的切分最小样本数基本同树的深度效果一样, 隐层节点数大于 20. 通过 MSE , MAE 以及 R^2 值 3 种评价指标得到以下结果, 如表 2 所示。

通过表 2 对比可知, 在这些标准数据集上, 本改进算法总体上在 MSE , MAE , R^2 值的对比中都明显优于 CART 回归树算法和 ELM 算法, 同时也总体略优于 SVR 和人工神经网络. 由于 MSE 评价指标采用平方计算使得对异常样本更加敏感, 通过对比 boston_housing、airfoil_self_noise、CombinedCyclePowerPlant、ConcreteCompressiveStrength 等数据集上的 MSE 值, 不难发现改进算法的 MSE 值相对明显较小, 这是因为在改进算法中引入了 L_2 正则化项, 添加了惩罚项使得改进算法在 MSE 评价指标下对比其他算法得到明显的下降, 使得模型更加适应异常样本存在的情形. 其次, 对比 MAE 值可以看出, 改进算法在这些数据集上的表现总体是优于所对比算法的. 接着从 R^2 值对比来看, 很明显改进算法的结果要优于其他算法, 特别是明显优于 CART 回归树算法和 ELM 算法, 甚至在 airfoil_self_noise 等数据集上有显著提高, 表明模型的准确率也更高. 利用 CART 算法对数据集的划分原理使得相关关系更紧密的样本数据划分到一起, 从而将全局模型的构建转化为多个局部模型的构建, 使得对比 CART 回归树和 ELM 明显提高了预测的准确性. 因此, 改进算法是将两个算法的优势结合在一起, 整体而言, 本文提出的改进算法达到了预期效果。

表 2 实验结果对比

数据集	回归树			ELM			SVR			人工神经网络			ELM-CART		
	MSE	MAE	R^2	MSE	MAE	R^2	MSE	MAE	R^2	MSE	MAE	R^2	MSE	MAE	R^2
D1	16.5343	2.9897	0.7952	14.9751	2.7632	0.8145	11.5267	2.2442	0.8572	16.8387	2.5556	0.7914	9.3306	2.3391	0.8844
D2	0.4028	0.4206	0.6953	0.3855	0.4079	0.7084	0.4231	0.4381	0.6799	0.3026	0.3693	0.7711	0.3174	0.3829	0.7599
D3	0.9839	0.7209	0.5661	0.9386	0.6679	0.5861	0.8694	0.6433	0.6166	0.8192	0.6322	0.6388	0.7722	0.6407	0.6595
D4	7.1029	2.0374	0.8521	6.6496	1.9480	0.8616	9.0965	2.1844	0.8106	12.4712	2.6396	0.7404	3.1372	1.2390	0.9347
D5	14.8490	2.8602	0.9495	15.2550	3.0766	0.9481	16.1809	3.1577	0.9450	17.6476	3.4008	0.9400	11.5832	2.5436	0.9606
D6	49.7077	5.0808	0.8128	37.0456	4.2151	0.8605	40.1256	4.5770	0.8489	38.0774	4.6774	0.8566	31.6776	4.0322	0.8807
D7	1.5959	0.9269	0.4028	1.0995	0.8211	0.5885	1.1139	0.7923	0.5832	1.0886	0.7893	0.5927	1.0859	0.8148	0.5937

4 结论与展望

本文提出了基于 ELM 的改进 CART 决策树回归算法, 将复杂的全局模型分解为多个相对容易的局部模型, 从而降低回归分析的难度, 提高模型准确性, 并

结合 ELM 具有的优点, 将 CART 回归树与 ELM 算法融合, 弥补了 CART 回归树算法本身的缺点与不足, 达到提高回归预测的准确率, 并通过上述实验结果对比表现出了改进算法的优化效果. 在接下来的工作中, 将

会在更多数据集上测试算法的准确性,同时将会继续研究相关算法,以期进一步提高回归预测精度及稳定性。

参考文献

- 1 冷建飞,高旭,朱嘉平.多元线性回归统计预测模型的应用.统计与决策,2016,(7):82-85.
- 2 谢永华,张鸣敏,杨乐,等.基于支持向量机回归的城市PM_{2.5}浓度预测.计算机工程与设计,2015,36(11):3106-3111.
- 3 张棧,曹健.面向大数据分析的决策树算法.计算机科学,2016,43(S1):374-379,383.
- 4 徐睿,梁循,齐金山,等.极限学习机前沿进展与趋势.计算机学报,2019,42(7):1640-1670.
- 5 侯大力,孙雷,潘毅,等.人工神经网络预测高含CO₂天然气的含水量.西南石油大学学报(自然科学版),2013,35(4):121-125. [doi: 10.3863/j.issn.1674-5086.2013.04.017]
- 6 Sut N, Simsek O. Comparison of regression tree data mining methods for prediction of mortality in head injury. Expert Systems with Applications, 2011, 38(12): 15534-15539. [doi: 10.1016/j.eswa.2011.06.006]
- 7 Müller D, Leitão P J, Sikor T. Comparing the determinants of cropland abandonment in Albania and Romania using boosted regression trees. Agricultural Systems, 2013, 117: 66-77. [doi: 10.1016/j.agsy.2012.12.010]
- 8 Park Y, Pachepsky YA, Cho KH, et al. Stressor-response modeling using the 2D water quality model and regression trees to predict *chlorophyll-a* in a reservoir system. Journal of Hydrology, 2015, 529: 805-815. [doi: 10.1016/j.jhydrol.2015.09.002]
- 9 Larraondo PR, Inza I, Lozano JA. A system for airport weather forecasting based on circular regression trees. Environmental Modelling & Software, 2018, 100: 24-32.
- 10 董红召,许慧鹏,卢滨,等.城市交通道路氮氧化物浓度的CART回归树预测研究.环境科学学报,2019,39(4):1086-1094.
- 11 郑向群,赵政.基于S-CART决策树的多关系空间数据挖掘方法.计算机应用,2008,28(3):749-752.
- 12 杜春蕾,张雪英,李凤莲.改进的CART算法在煤层底板突水预测中的应用.工矿自动化,2014,40(12):52-56.
- 13 刘飞翔,吴浩.基于改进CART决策树建立水华预警模型.中国农村水利水电,2018,(1):26-28. [doi: 10.3969/j.issn.1007-2284.2018.01.007]
- 14 毕云帆,张健,胥晓晖,等.基于梯度提升决策树的电力短期负荷预测模型.青岛大学学报(工程技术版),2018,33(3):70-75.
- 15 李正方,杜景林,周芸.基于改进CART算法的降雨量预测模型.现代电子技术,2020,43(2):133-137,141.
- 16 Huang GB, Zhu QY, Siew CK. Extreme learning machine: Theory and applications. Neurocomputing, 2006, 70(1-3): 489-501.
- 17 甘露.极限学习机的研究与应用[硕士学位论文].西安:西安电子科技大学,2014.
- 18 Loh WY. Classification and regression trees. Wires Data Mining and Knowledge Discovery, 2011, 1(1): 14-23. [doi: 10.1002/widm.8]
- 19 苏锑,杨明,王春香,等.一种基于分类回归树的无人车汇流决策方法.自动化学报,2018,44(1):35-43.
- 20 Breiman L, Friedman J, Olshen R, et al. Classification and Regression Trees. New York: Chapman and Hall, 1984.
- 21 苗红星,余建坤.基于决策树的ID3算法和C4.5算法的比较.现代计算机(专业版),2014,(15):7-10,14.