

基于优化随机森林算法预测食品检验不合格指标

刘玉航¹, 曲媛^{1*}, 蒋嘉铭¹, 宗万里², 朱习军¹

(1. 青岛科技大学信息科学技术学院, 青岛 266061; 2. 威海市食品药品检验检测中心, 威海 264210)

摘要: **目的** 建立基于优化的随机森林算法模型实现对食品不合格指标的分类预测。**方法** 通过收集山东省食品药品监督管理局 2015—2019 年食品安全抽样检验产生的不合格数据, 并对其进行多项数据预处理操作, 采用超参数网格搜索和 10 折交叉验证方法建立基于随机森林的食品不合格指标的分类预测模型, 并通过对比传统随机森林模型的参数优化, 将其与决策树(decision tree, DT)、逻辑回归(logistic regression, LR)和梯度提升决策树(gradient boosting decision tree, GBDT)算法分类预测结果进行对比。**结果** 实验表明经过参数优化后的随机森林模型对食品中不合格指标的预测准确率能够达到 89.4%, 比 DT 算法提高了 11.0%, 比 LR 算法提高了 9.0%, 比 GBDT 算法提高了 8.1%。**结论** 基于优化的随机森林模型可以完成食品不合格指标分类预测任务, 有广阔的应用前景。

关键词: 食品安全数据; 决策树; 随机森林; 参数优化; 超参数网格搜索

Prediction of unqualified index of food inspection based on optimized random forest algorithm

LIU Yu-Hang¹, QU Yuan^{1*}, JIANG Jia-Ming¹, ZONG Wan-Li², ZHU Xi-Jun¹

(1. College of Information Science and Technology, Qingdao University of Science and Technology, Qingdao 266061, China; 2. Weihai Food and Drug Inspection and Testing Center, Weihai 264210, China)

ABSTRACT: Objective To establish a random forest algorithm model based on optimization and realize the classification and prediction of food unqualified indexes. **Methods** Through the collection of unqualified data generated by the food safety sampling inspection from 2015 to 2019 issued by the official website of Shandong Food and Drug Administration, and a number of data preprocessing operations, the hyper parameter grid search and 10-folds cross-validation method were used to establish a classification prediction model based on random forest-based food unqualified indicators. In addition, the parameters of the traditional random forest model was optimized, and compared with algorithm classification prediction results of decision tree (DT), logistic regression (LR) and gradient boosting decision tree (GBDT). **Results** Experiments showed that the random forest model after parameter optimization could achieve 89.4% prediction accuracy of unqualified indicators in food, which was 11.0% higher than the DT algorithm, 9.0% higher than the LR algorithm, and 8.1% higher than the GBDT algorithm. **Conclusion** The optimized random forest model can complete the classification and prediction task of food unqualified indicators, and has broad application prospects.

KEY WORDS: food safety data; decision tree; random forest; parameter optimization; hyper parametric grid search

基金项目: 山东省产教融合研究生联合培养示范基地项目(2020-19)

Fund: Supported by the Shandong Province Industry Education Integration Postgraduate Joint Training Demonstration Base Project (2020-19)

*通信作者: 曲媛, 硕士研究生, 主要研究方向为智慧医疗。E-mail: 2595989958@qq.com

*Corresponding author: QU Yuan, Master, College of Information Science and Technology, Qingdao University of Science and Technology, NO.99 Songling Road, Laoshan District, Qingdao 266061, China. E-mail: 2595989958@qq.com

0 引言

食品安全问题与人类健康发展息息相关,然而近年众多食品安全问题对人民生命健康构成极大的威胁^[1]。如今各食品安全监管监察部门在这一问题上所投入的精力、财力也在逐渐增加,因此也获得了大量的食品安全检测数据^[2],若能对其进行分析应用,并利用数据挖掘、机器学习等方式对食品安全数据进行建模,并由此展开后续食品安全质量检测工作,意义重大^[3]。

本研究数据集来源于 2015—2019 年山东省食品药品监督管理局所记录的食品抽检时的不合格食品检测数据^[4],采用优化的随机森林模型对食品数据进行训练,利用现有数据集建立分类预测模型进行训练,将其用于食品抽样检测中质量不合格指标的预测。同时,对比分析各种分类算法的特点,并解释了模型分类的规则,对食品安全进行了监测和预警,为食品安全管理和控制决策提供科学依据。

1 材料与方法

1.1 随机森林基本原理

随机森林采用的是 Bootstrap 重采样技术,具体方法是通过有放回且随机地从原始训练样本 N 中抽取 k 个样本生成决策树,构成随机森林,最终分类由决策树的投票结果决定^[5]。随机森林算法可以看作是对原有决策树算法的整合和改进^[6],单棵决策树通过特征选择实现节点分裂,最终得到预测类别^[7],随机森林将 n 棵决策树系统地组合到一起,通过多数投票机制^[8]确定数据所属类别。随机森林流程图如图 1 所示。

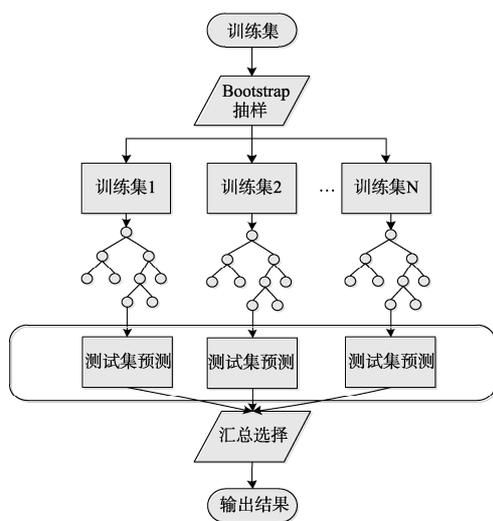


图 1 随机森林流程图

Fig.1 Random forest flow chart

训练随机森林的过程与训练每棵独立决策树相同^[9],

随机森林训练过程是并行的,这主要取决于每棵树独立的训练方式^[10],因此构建模型速度较快。

1.2 基于随机森林的食品不合格指标分类预测模型设计

随机森林模型在拟合能力强的基础上又可以较好地避免过拟合现象发生^[11]。随机森林算法能处理高维度数据,而食品安全检测数据中,单是同一大类的不合格指标就有多种具体不合格类项。在训练过程中,能够检测到 feature 间的相互影响及重要性,具有一定的参考意义^[12]。本研究所用食品不合格指标的分类预测模型结构如图 2 所示。

建立随机森林模型,首先需要构建随机森林中的基分类器—决策树。ID3 和 C4.5 决策树算法会尽可能多的对数据集挖掘信息,从而导致决策树模型复杂化,而分类回归树算法(classification and regression tree, CART)则可以通过剪枝处理简化决策树模型,优化决策树生成效率^[13],有利于每棵独立决策树对划分的食品检验样本集进行训练,所以本研究使用 CART 算法作为随机森林中每棵独立决策树的基生成构造算法。通过 bootstrap 抽样方法为每个已建立的 CART 决策树传递随机样本数据,决策树每次从随机抽取的特征子集中选择最优的分裂特征。CART 算法使用 Gini 系数作为特征选择与分裂的衡量标准并以此构建生成决策树,假设当前特征 X_i 包含 T 个类别,第 t 个类别的概率为 P_t ,则 Gini 系数可以通过下式计算得到^[14]:

$$\text{Gini} = \sum_{t=1}^T P_t(1 - P_t) = 1 - \sum_{t=1}^T P_t^2 \quad (1)$$

式中, Gini 代表基尼系数, T 代表特征类别数量, P_t 代表食品指标为当前特征的概率。

通过公式(1)可知,从数据中抽取两个样本,样本类别标志不一致的概率即为 Gini 系数。当 Gini 系数数值越小时,模型的分类纯度就越高,因此对数据的分类准确率也就越好^[15]。对决策树全部特征向量进行遍历操作并选择 Gini 系数最小的特征向量作为该节点的分裂特征,直到决策树构建完成,最后将叶子节点的类别作为该输入数据的类别。

1.3 随机森林模型的参数优化

为了对随机森林算法中超参数进行优化,提高随机森林模型的性能和效果,本研究使用了超参数网格搜索算法^[16],给随机森林选择一组最优超参数,避免手动依次重复测试。其中,设定随机森林中决策树的数量为 $n_estimators$, $n_estimators$ 太小容易欠拟合,太大不能显著的提升模型,所以设置 $n_estimators$ 的取值范围为[50, 100, 200, 300, 500, 800];设定随机森林中每棵决策树的最大深度为 max_depth ,取值范围为[2, 3, 4, 5]。通过对上述 2 项参数的调优以实现整个模型参数优化。利用 sklearn 库中的 Grid Search CV 函数,系统地遍历多种参数,通过 10 折交叉验证确定最优的模型参数组合^[17]。通过 10 折交叉验证将数据集分成 10 分,依次取其中的 9 份作为训练数据,剩余的 1

份作为测试数据, 进行实验, 每次实验都会得到一个相应的正确率^[18], 取 10 次结果的正确率的平均值作为对算法精度的估计, 返回正确率最高时对应的最优参数组合。

1.4 数据采集

本研究将山东省食品药品监督管理局官网公布的 2015—2019 年公布的不合格食品抽检信息作为基础数据, 将其整合为 3877 条结构化记录, 作为后续模型所需的数据集。网站提供的食品检测信息中每条数据含有 14 个特征属性及 1 个不合格项的目标属性, 其部分数据及特征信息如表 1 所示。

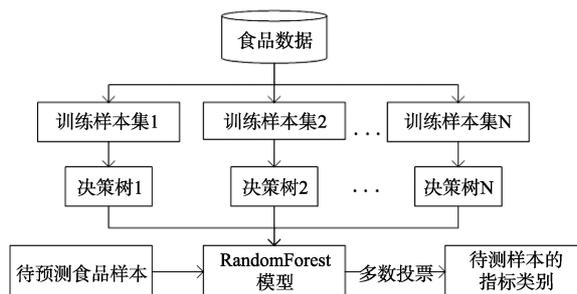


图 2 基于随机森林的食品安全检测模型
Fig.2 Food safety detection model based on random forest

表 1 食品抽检不合格产品部分信息

Table 1 Partial information of unqualified products in food sampling inspection

标称食品名	标称规格	生产日期	检验结果	标准值	分类	不合格项目
小碗	/	2019.05.21	0.058 mg/100 cm ²	不得检出	餐饮食品	阴离子合成洗涤剂
餐盘	/	2019.06.05	0.11 mg/100 cm ²	不得检出	餐饮食品	阴离子合成洗涤剂
仲明香瓜子	1 kg/袋	2019.05.05	1.2 g/100 g	≤0.8	炒货食品及坚果制品	过氧化值(以脂肪计)
水果味蘸酱饼干	散装称重	2019.05.01	0.21 g/kg	0.1 g/kg	饼干	二氧化硫残留量
塑料餐盘	散装	2019.05.10	0.048 mg/100 cm ²	不得检出	餐饮食品	阴离子合成洗涤剂

1.5 数据预处理

1.5.1 数据清洗

数据清洗依据数据需要通过缺失值填补、数据降噪、归一化属性等操作来去除源数据中的噪声数据和无关数据^[19]。对于数据中由于个人或者其他人为原因导致出现属性值空缺的情况进行填补^[20], 对于食品日期的空缺值填补为与同一批多数食品日期相同的值。食品规格类型空缺值全部填补为预包装类型。在来源地字段首先考虑生产地址、当生产地址为空时选择抽样地址。

1.5.2 数据集成与归约

数据集成通常是对用于数据分析的分散数据源的集成。由于在不同时期发布的食物抽样信息在特定属性字段中具有某些差异, 因此在数据清理后合并数据源时, 主要解决属性名称不一致的问题。经过清理、集成的食品安全检验数据集中包含 3651 条记录食品安全检验数据。

在不影响数据特征的情况下, 采取更为有效的数据挖掘技巧十分重要^[21]。本次预处理中, 针对不合格食品安全检测数据, 保留了食品名称、食品种类、食品规格、生产日期 4 个特征属性以及一个不合格指标的目标属性。

1.5.3 数据编码

对数据集进行预处理后, 数据集中只保留了食品名称、食品种类、食品规格、生产日期和不合格指标 5 个字段, 其类型均为类别变量。随机森林算法建模时需要进行属性转换, 本研究利用数字标号对数据集中的类别变量进行编码。其中食品种类有 26 个类别, 食品类型有 74 个类别, 食品规格有 3 个类别, 食品日期有 4 个类别, 食品指标有 10 个类别。食品规格、食品日期和食品指标编码格式如

表 2 所示, 食品种类和食品类型编码同上, 由于类别较多, 未具体列出。

表 2 类别变量转换表

Table 2 Category variable conversion table

食品规格	食品日期	食品指标
包装=1 散装=2 自制=3	保质期四分之一=1 保质期四分之一=2 保质期四分之三=3 保质期四分之四=4	化学物质=1, 添加剂=2, 金属元素=3, 微生物=4, 农药残留=5, 非法添加=6, 兽药残留=7, 生物毒素=8, 污染物=9, 碘化物=10

2 结果与分析

2.1 实验仿真

数据预处理完成后, 得到 3651 条可用实验数据, 为了使模型能够得到更多的训练数据, 通过交叉验证的方法重复对数据集进行划分, 以此来获取不同的训练集与测试集的组合。本实验的数据量及实验数据的特征数量, 基本满足采用随机森林建模要求。最终按照 8:2 的比例将数据集划分为训练集和测试集, 并通过该数据集进行建模。

实验环境为 Anaconda 3, 并在此环境下手动构建随机森林模型。首先, 通过对数据集进行清洗、打乱及数据编码等预处理操作。对随机森林模型进行训练时, 将食品大类、食品品种、食品规格、食品日期作为随机森林模型的输入变量, 食品指标作为分类变量, 随机抽取训练样本, 对特征进行节点分裂, 生成决策树, 从而生成随机森林,

利用超参数网格搜索算法和交叉验证遍历超参数组合, 返回最优参数组合。通过训练得到最优参数组合为: $n_estimators: 100; max_depth: 5$ 。为验证使用随机森林算法对该数据集进行建模的有效性, 同时使用 LR (logistic regression) 算法、决策树算法及梯度提升决策树 (gradient boosting decision tree, GBDT) 算法与之进行比较。

2.2 评价标准

为了检验基于随机森林算法所构建的食品安全检验模型的有效性及模型的泛化能力, 通过模型的预测准确率、召回率及特异度情况进行分析, 观察模型效果。

可以将准确率 (accuracy)、召回率 (recall)、特异度 (specificity) 通过公式 (2)、(3)、(4) 表示。

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$recall = \frac{TP}{TP + FN} \quad (3)$$

$$specificity = \frac{TN}{TN + FP} \quad (4)$$

其中, P 代表的是正类样本的数量, N 是负类样本的数量。因此, TP (true positive) 代表正例预测正确的个数, FP (false positive) 代表负例预测错误的个数; TN (true negative) 代表负例预测正确的个数, FN (false negative) 代表正例预测错误的个数^[22]。

准确率是该模型预测的指标与实际指标一致的不合格食品检验数据与全部不合格食品检验数据总量的比例, 也就是该模型针对食品不合格指标分类是正确的概率。根据上述公式 (2), 可以得到该模型的模型准确率为 95.2%, 即输入一条指标未知的不合格食品检验数据, 模型对其正确分类的概率为 95.2%, 具体表示为式 (5) 所示:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} = \frac{3264}{3651} \approx 89.4\% \quad (5)$$

召回率是该模型正确预测的不合格指标的食品数据的数量和真实不合格指标的食品数据的数量的比例, 也就是有多少不合格食品数据能被该模型预测出来。根据上述公式 (3), 可以得到该模型的召回率为 92.7%, 具体表示为式 (6) 所示。

$$recall = \frac{TP}{TP + FN} = \frac{3264}{3264 + 399} \approx 89.1\% \quad (6)$$

2.3 结果分析

利用参数优化的随机森林算法构建模型并训练后, 将其各项评价指标分别与 LR 算法、决策树算法及 GBDT 算法所构建的模型进行比较。通过比较验证本实验所采取参数优化随机森林算法所构建的模型是否优于其他算法模型, 最终得出的结果如表 3 所示。

通过表 3 可以得出, 经过参数优化后的随机森林算法在准确率上比其他算法具有一定的优势, 准确率达到 89.4%, 较 LR 算法准确率提高了 11.0%, 较决策树算法准确率提高了 9.0%, 较 GBDT 算法准确率提高了 8.1%。同时, 为了更好地比较各个算法在食品不合格指标分类预测的性能, 将不同算法所绘制的 ROC 曲线 (receiver operating characteristic curve, ROC) 绘制在同一坐标系, 得出的 ROC 曲线如图 3 所示。

表 3 不同算法所构建模型的评价指标

Table 3 Evaluation indexes of models constructed by different algorithms

分类算法	准确率/%	召回率/%	耗时/s
随机森林	89.4	89.1	0.74
LR	78.4	77.3	0.06
决策树	80.4	79.7	0.11
GBDT	81.3	80.9	0.27

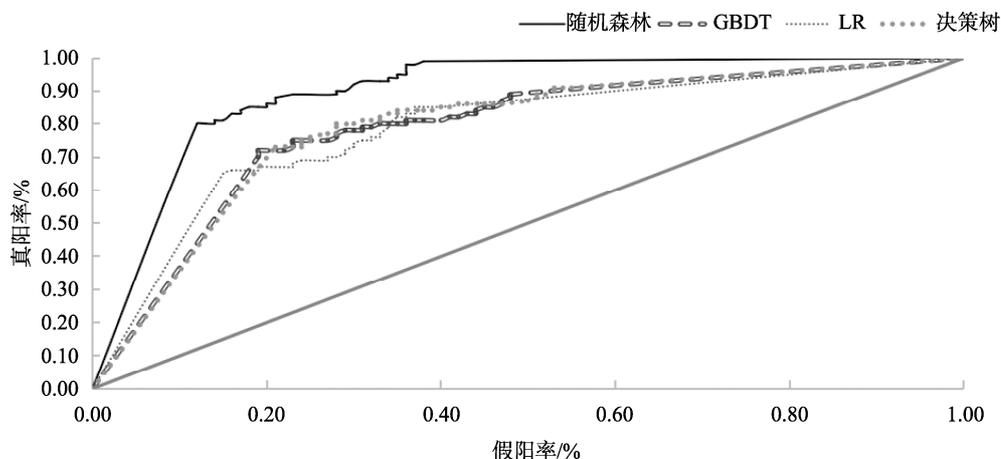


图 3 不同模型的 ROC 结果

Fig.3 ROC results of different models

通过表 3 和图 3, 可以直观地看到本研究应用于食品安全检验不合格数据的参数优化随机森林模型在准确率、召回率等各项评价标准上均优于其他模型, 其 ROC 曲线最接近坐标系左上角, 即其曲线下面积(area under the curve, AUC)值是上述各模型中最高的, 这充分说明了随机森林算法进行参数优化后作为食品安全数据不合格检验模型相较于其他算法所构建的模型表现更为出色, 能够更好地实现分类任务, 预测结果较之其他模型也更为准确。

2.4 结果解读

随机森林模型中所生成的决策树的部分数据挖掘规则如下:

If“食品种类”==“饼干制品”, Then 不合格指标为“微生物”;

If“食品种类”==“茶产品”, Then 不合格指标为“农药残留”;

If“食品种类”==“再制蛋”, Then 不合格指标为“重金属元素”;

If“食品种类”==“发酵豆制品”, Then 不合格指标为“添加剂”;

.....

饼干制品的微生物指标的不合格率较高, 提示我们虽然通常情况饼干的水分含量较低, 微生物含量较少, 微生物指标不易超标, 但是由于饼干的类型多种多样, 有些饼干以散装的形式出售, 造成了饼干中微生物出现不合格的概率较高。

茶叶的不合格指标主要为农药残留, 茶叶中农药残留超标的原因主要为茶树从土壤中吸附持久性农药、茶农大量使用农药。

再制蛋在制作过程中为了使蛋白质沉淀, 常将蛋制品浸泡在含有铅、铜的溶液中, 因此常会造成蛋制品中金属元素超标。

豆制品的添加剂不合格可能性较大, 豆制品在生产过程中由于防腐、调味、着色等工艺的需要往往使用添加剂来达到这些需求, 但在使用添加剂的过程中会出现超量和超范围的情况。

3 结论与讨论

本研究通过收集食品不合格抽检数据并通过数据清洗、集成规约、数据编码的方式对其进行处理, 同时建立优化的随机森林模型, 实现对食品抽检不合格指标的分类预测, 并和其他算法模型进行对照实验, 实验表明基于优化的随机森林模型可以较好地预测食品不合格指标类别。

构建优化的随机森林模型, 可以更好地服务于当前相关部分, 进行食品不合格指标检验任务, 为食品安全检

测环节提供了有力的帮助。

由于本研究是以山东省食品药品监督管理局 5 年内的食品安全检测数据作为基础数据集, 因此如何解决将数据关联规则扩充至更多省份、更高时间跨度乃至实时数据成为将来需要研究的重点问题。同时, 受研究环境限制, 本研究所使用的随机森林模型在解决高并发数据问题时表现不够优秀, 当数据量过于庞大时可以考虑使用大数据架构 Spark、Hadoop 等服务器并行计算技术, 充分发挥模型的应用价值。

参考文献

- [1] 旭日干, 庞国芳. 中国食品安全现状、问题及对策战略研究[M]. 北京: 科学出版社, 2015.
XU RQ, PANG GF. Research on current situation, problems and countermeasures of food safety in China [M]. Beijing: Science Press, 2015.
- [2] 徐向前, 徐华琳. 我国食品安全现状和食品检测的发展方向[J]. 现代食品, 2020, (7): 153-154.
XU XQ, XU HL. The present situation of food safety in China and the development direction of food testing [J]. Mod Food, 2020, (7): 153-154.
- [3] 李燕. 食品安全现状及食品安全检测技术应用浅析[J]. 医学食疗与健康, 2020, 18(4): 207-208.
LI Y. The present situation of food safety and the application of food safety detection technology [J]. Med Diet Health, 2020, 18(4): 207-208.
- [4] 山东省食品药品监督管理局官网[Z]. <http://www.sdfda.gov.cn/>
Shandong food and drug administration website [Z]. <http://www.sdfda.gov.cn/>
- [5] PAUL A, MUKHERJEE DP, DAS P, *et al.* Improved random forest for classification [J]. IEEE T Image Process, 2018; 1. DOI: 10.1109/TIP.2018.2834830
- [6] 吕红燕, 冯倩. 随机森林算法研究综述[J]. 河北省科学院学报, 2019, 36(3): 37-41.
LU HY, FENG Q. Review of random forest algorithm [J]. J Hebei Acad Sci, 2019, 36(3): 37-41.
- [7] 王奕森, 夏树涛. 集成学习之随机森林算法综述[J]. 信息技术, 2018, 12(01): 49-55.
WANG YS, XIA ST. Survey of random forest algorithm for ensemble learning [J]. Inform Commun Technol, 2018, 12(1): 49-55.
- [8] 凡建. 基于近红外光谱传感器的便携式食品检测器的设计与实现[D]. 南京: 南京邮电大学, 2019.
FAN J. Design and implementation of portable food detector based on near infrared spectroscopy sensor [D]. Nanjing: Nanjing University of Posts and Telecommunications, 2019.
- [9] 赵庆明. 基于邻传播算法和随机森林回归模型的台区线损率计算方法[J]. 电力系统及其自动化学报, 2020, 32(9): 94-98.
ZHAO QM. Line loss rate calculation method based on nearest neighbor propagation algorithm and random forest regression model [J]. Proc CSU-EPSSA, 2020, 32(9): 94-98.
- [10] HAN Y, LI XD, PANG HJ, *et al.* Application of random forest algorithm in hail forecasting over Shandong peninsula [J]. Atmos Res, 2020, 244: 105093.

- [11] 谢桦, 陈俊星, 郭志星, 等. 基于随机森林算法的架空输电线路状态评价方法[J]. 现代电力, 2020, 37(6): 559–565.
XIE H, CHEN JX, GUO ZX, *et al.* State evaluation method of overhead transmission lines based on stochastic forest algorithm [J]. Mod Electr Power, 2020, 37(6): 559–565.
- [12] 闫佰忠, 孙剑, 安娜. 基于随机森林模型的地下水水质评价方法[J]. 水电能源科学, 2019, 37(11): 66–69.
YAN BZ, SUN J, AN N. Groundwater quality evaluation method based on random forest model [J]. Water Resour Pow, 2019, 37 (11): 66–69.
- [13] KANG B, SEOK C, LEE J. Prediction of molecular electronic transitions using random forests [J]. J Chem Inf Model, 2020, 60(12). DOI: 10.26434/chemrxiv.12482840
- [14] MAO YX, HE YH, LIU LM, *et al.* Disease classification based on eye movement features with decision tree and random forest [J]. Front Neurosci-Switz, 2020, 14: 798.
- [15] 王思培, 韩涛. 基于随机森林算法的潜在高价值专利预测方法研究[J]. 情报科学, 2020, 38(5): 120–125.
WANG SP, HAN T. Research on the prediction method of potential high value patent based on stochastic forest algorithm [J]. Inf Sci, 2020, 38(5): 120–125.
- [16] 温博文, 董文瀚, 解武杰, 等. 基于改进网格搜索算法的随机森林参数优化[J]. 计算机工程与应用, 2018, 54(10): 154–157.
WEN BW, DONG WH, XIE WJ, *et al.* Optimization of random forest parameters based on improved grid search algorithm [J]. Comput Eng Appl, 2018, 54(10): 154–157.
- [17] 魏勤, 陈仕军, 黄炜斌, 等. 利用随机森林回归的现货市场出清价格预测方法[J]. 中国电机工程学报, 2021, 41(4): 1360–1367, 1542.
WEI Q, CHEN SJ, HUANG WB, *et al.* The method of forecasting the price of clearing out of spot market by random forest regression [J]. Proc CSEE, 2021, 41(4): 1360–1367, 1542.
- [18] 徐辉煌, 张海宇, 林勇. 基于多粒度级联森林的骨质疏松性骨折预测研究[J]. 北京生物医学工程, 2019, 38(4): 384–391.
XU HH, ZHANG HY, LIN Y. Prediction of osteoporotic fracture based on multi granularity cascade forest [J]. Beijing Biomed Eng, 2019, 38(4): 384–391.
- [19] 刘振鹏, 苏楠, 秦益文, 等. FS-CRF: 基于特征切分与级联随机森林的异常点检测模型[J]. 计算机科学, 2020, 47(8): 185–188.
LIU ZP, SU N, QIN YW, *et al.* FS-CRF: outlier detection model based on feature segmentation and cascade random forest [J]. Comput Sci, 2020, 47(8): 185–188.
- [20] 姚雨虹, 杨小兵. 基于多指标加权分析的航空客户价值预测模型[J]. 中国计量大学学报, 2020, 31(1): 97–103.
YAO YH, YANG XB. Aviation customer value prediction model based on multi index weighted analysis [J]. J China Univ Metrol, 2020, 31(1): 97–103.
- [21] 李宇琪. 基于随机森林的房价预测模型[J]. 通讯世界, 2018, (9): 306–308.
LI YQ. House price forecasting model based on random forest [J]. Telecom World, 2018, (9): 306–308.
- [22] NUGENT CM, ELLIOTT TA, RATNASINGHAM S. Coil: An R package for cytochrome oxidase I (COI) DNA barcode data cleaning, translation, and error evaluation [J]. Genome, 2020, 63(6): 291–305.

(责任编辑: 郑丽于梦娇)

作者简介



刘玉航, 硕士研究生, 主要研究方向为数据分析。
E-mail: 564275986@qq.com



曲媛, 硕士研究生, 主要研究方向为智慧医疗。
E-mail: 2595989958@qq.com