

DOI:10.3969/j.issn.1005-202X.2024.08.017

医学信号处理与医学仪器

基于音频分析技术的甲状腺疾病诊断方法

刘彤, 张博维, 冯超, 顾毅, 高青山

四川省人民医院甲状腺外科, 四川 成都 610072

【摘要】通过结合基于SMOTE数据扩充方法对复杂音频数据的处理能力以及GA-SVM的准确预测能力,提出一种基于SMOTE数据扩充算法的GA-SVM模型,并将其应用于甲状腺疾病的诊断中。实验结果表明,本文方法在甲状腺疾病分类诊断方面表现出良好的性能。相较于其他方法,本文方法在诊断效果方面表现更佳,在甲状腺疾病的分类诊断中具有更高的适用性。

【关键词】甲状腺疾病;音频分析;GA-SVM模型;诊断方法

【中图分类号】R318

【文献标志码】A

【文章编号】1005-202X(2024)08-1031-05

Diagnostic method for thyroid disease using audio analysis technology

LIU Tong, ZHANG Bowei, FENG Chao, GU Yi, GAO Qingshan

Department of Thyroid Surgery, Sichuan Provincial People's Hospital, Chengdu 610072, China

Abstract: By combining the ability of SMOTE data augmentation method to process complex audio data with the accurate prediction ability of GA-SVM, a GA-SVM model based on SMOTE data augmentation algorithm is proposed and applied to the diagnosis of thyroid diseases. The results demonstrate the excellent performance of the proposed method in the classification and diagnosis of thyroid diseases. Compared with other methods, the proposed method has better diagnostic efficacy and higher applicability in the classification and diagnosis of thyroid diseases.

Keywords: thyroid disease; audio analysis; GA-SVM model; diagnostic method

前言

甲状腺是人体重要的内分泌腺,其功能包括产生关键的甲状腺激素,如血清总甲状腺素(T₄)和血清总三碘甲状腺原氨酸(T₃)。这些激素对身体的正常运转至关重要,其作用不仅局限于控制新陈代谢,还在多种生理过程中发挥作用,如调节能量代谢、维持体温、影响蛋白质合成等^[1]。甲状腺疾病可能受多种因素的影响,碘的不足或过量摄入、自身免疫性疾病、炎症、感染、遗传因素、肿瘤、环境和生活方式等,均可能导致甲状腺疾病的发生。这些因素的影响使得甲状腺可能产生过多或过少的甲状腺激素,或导致甲状腺功能异常,从而引发疾病^[2]。甲状腺疾病对人体健康带来不容忽视的危害。甲状腺功能亢进可

能导致心率增加、体质量减轻、焦虑和肌肉无力等不适症状。相反,甲状腺功能减退会导致疲劳、体质量增加、注意力不集中和情绪低落等问题^[3]。此外,甲状腺疾病还可能引起甲状腺结节、肿瘤,甚至是甲状腺炎等并发症。甲状腺疾病不仅会影响患者的生活质量,而且会增加患者出现心血管疾病、骨质疏松和其他健康问题的风险。如果不及时治疗和管理可能会导致症状的恶化,甚至导致严重并发症的发生^[4]。因此,对于甲状腺疾病的诊断、积极的治疗和有效的管理显得尤为重要,它可以有效地降低患者的病情风险,保障患者的健康和生活质量。

甲状腺疾病是一类常见疾病,而其诊断往往面临挑战,特别是在早期诊断和准确性方面。甲状腺疾病的早期诊断对患者的治疗和预后十分重要。然而,传统的诊断方法往往存在局限,例如在进行血液检测、影像学检查和临床症状观察等方面。虽然这些方法在一定程度上有效,但它们可能受到多种因素的影响,包括实验室的实验条件、医生的技术操作和主观判断。因此,寻求一种非侵入性、客观且准确的诊断方式显得尤为迫切。音频分析技术在医学诊

【收稿日期】2024-01-23

【基金项目】四川省国际科技创新项目(21GJHZ0066)

【作者简介】刘彤,主治医师,研究方向:甲状腺、甲状旁腺疾病,E-mail: L838516t888@163.com

【通信作者】张博维,博士,主治医师,研究方向:甲状腺、甲状旁腺疾病,E-mail: zhangbowei@med.uestc.edu.cn

断领域中拥有广泛的应用,尤其是在甲状腺疾病的诊断中,它作为一种潜在的新兴方法,提供一个全新的视角来探究甲状腺疾病的诊断。声音信号中携带大量有关声带振动、声音特征以及生理状况的信息,这些声音特征可能会受到甲状腺疾病的影响,例如声带功能异常或生理变化可能在声音信号中留下特定的痕迹^[5]。然而,收集的音频数据包含大量的复杂信息,由于每个人的声音特征都不尽相同,且受到个体差异、环境影响以及录音设备的影响,导致采集的音频数据具有高度的个体差异性和多样性;同时,音频数据往往会受到环境噪音和其他干扰因素的影响,例如背景噪音、录音设备的噪声,以及说话者的情绪、情感等^[6]。这些噪音和干扰会干扰声音信号的纯净度,对声音信号的解释造成困难。传统模型可能无法精确捕捉这些复杂的多因素关系,从而导致预测不准确,而机器学习方法能够处理大规模和复杂的数据,可以有效地识别隐藏在声音信号中的模式。因此,本文考虑通过机器学习方法结合音频分析技术,为甲状腺疾病提供早期诊断。

学者们在研究甲状腺疾病诊断方面进行大量的研究,Al-muwaffaq等^[7]提出一种针对甲状腺疾病预测的决策树算法,并开发了名为MLTDD的机器学习工具,用于智能诊断甲状腺疾病。Razia等^[8]将多种机器学习方法引入甲状腺疾病的分类诊断,通过对比实验验证了其方法的有效性,进一步证实了机器学习在疾病分类中的优势。Tyagi等^[9]利用UCI机器学习库的数据集,提出多种不同的机器学习和诊断方法,用于预测患者罹患甲状腺疾病的风险。Begum等^[10]将基于分类的数据挖掘方法应用于甲状腺疾病诊断,使用不同的分类方法预测甲状腺疾病,并计算不同激素与甲状腺功能亢进和功能减退的相关性。其他相关研究包括Aversano等^[11](2021年)、Alyas等^[12](2022年)、Islam等^[13](2022年)、Sultana等^[14](2023年)、Naeem等^[15](2023年)以及Pal等^[16](2022年)。尽管当前的研究在甲状腺疾病的分类诊断方面取得显著的进展,但仍然存在一些问题。数据不平衡是这些研究面临的主要问题之一。当数据中患者和健康受试者的比例失衡时,传统的机器学习算法倾向于产生不准确的预测结果,因为它们更倾向于较大类别的预测,而忽略了较小类别的特征。这种问题可能导致过度拟合,即模型在训练数据上表现良好,但对新数据的泛化能力较差。同时,针对机器学习算法的超参数设置也是一项具有挑战性的任务。超参数的选择对模型的性能有着重要的影响,但手动调整这些参数通常是一项繁琐而耗时的任务,因此需要更智能和自适应的方法来解决这个问题。为了

克服这些障碍,本文提出一种创新的方法,即基于SMOTE数据扩充算法的GA-SVM算法。该方法通过过采样技术平衡了不平衡的数据集,同时利用遗传算法对SVM的参数进行优化,以获得更加精确和可靠的预测模型。这种方法的引入不仅提高模型的性能和准确性,还增强其对数据不平衡性的鲁棒性。

本文旨在探讨和利用音频分析技术在甲状腺疾病诊断中的潜在价值,通过记录的声音信号,探索声带功能和相关神经活动的特征,并建立声音特征与甲状腺疾病之间的联系。通过结合机器学习算法,能够更好地识别声音信号中的生物标记和异常模式,并最终实现甲状腺疾病的智能化诊断。通过本文的研究,发掘音频分析技术在甲状腺疾病诊断中的潜在优势,为未来发展出更加客观、准确的诊断工具提供理论支持,从而为甲状腺疾病的早期诊断和治疗提供新的视角和解决方案。

1 基于音频分析技术的甲状腺疾病诊断方法

1.1 SMOTE数据扩充算法

目前,音频分析技术产生的数据通常呈现不平衡分布。在解决类别不平衡问题时,常用的方法包括过采样和欠采样。这两种方法旨在处理数据集中类别数量不均衡的情况,以提高模型训练的效果和性能。过采样方法通过增加少数类样本的数量,而欠采样方法则通过减少多数类样本的数量。然而值得注意的是欠采样方法可能会导致信息丢失,因此在某些情况下,尤其是在医学研究中更倾向于使用过采样方法。在过采样方法中,最常用的算法之一是2002年由Chawla团队提出的SMOTE算法。SMOTE是一种数据增强技术,专门针对数据集中的少数类样本,通过合成新的少数类样本来平衡数据集,以提高模型性能^[17]。它的基本思想是通过将少数类样本进行插值来生成新的合成样本。具体步骤如下:

(1) 选择一个少数类样本。对于每个少数类样本 V_i ,计算其与所有相邻样本 V_{si} (最近邻)的距离。定义 D 为两个样本之间的距离:

$$D = \sqrt{\sum_{i=1}^n (V_i - V_{si})^2}, i = 1, \dots, n \quad (1)$$

(2) 选择少数类样本的邻居。选择该少数类样本的若干个最近邻样本。

(3) 生成新的合成样本。对于少数类中的每个样本,搜索其最近邻样本,并从这些最近邻数据集中随机选择 N 个样本,然后在原始样本和其最近邻样本之间进行插值操作,即将差值乘以一个随机数,然后

加上被选中的少数类样本,生成新的合成样本。

上述过程将在少数类样本中重复进行,直到数据集中的各个类别样本的数量达到所设定的平衡水平。下面是SMOTE算法的基本公式:假设 V_i 是少数类样本, V_{si} 是 V_i 的一个近邻, $0 < \gamma < 1$ 是随机选取的值, V_{new} 是生成的新样本:

$$V_{new} = V_i + (V_{si} - V_i)\gamma \quad (2)$$

其中,新合成样本 V_{new} 是由 V_i 和 V_{si} 之间的差值乘以一个随机数 γ 后加到 V_i 上。

SMOTE方法的运用旨在通过生成合成样本来弥补数据集中的不平衡,以改善模型的泛化能力和性能。通过这种方式,模型能够更好地学习和理解少数类样本的特征,从而提高对少数类的预测准确性。这在解决分类问题中,尤其是在少数类数据较为重要的情况下,发挥重要作用。

1.2 改进SVM诊断方法

SVM是一种监督学习算法,用于分类和回归分析,其主要目标是找到一个能够划分数据集并且最大化分类间隔的超平面^[18-19]。假设一个诊断是否患有甲状腺疾病的二分类问题,模型中输入的特征向量由甲状腺疾病的主要影响因素构成,即设输入的变量为 $\mathbf{z}_i = (z_{i1}, z_{i2}, \dots, z_{ip})$, p 为输入变量的个数,模型以是否患甲状腺疾病为输出,设为 $y_i \in \{-1, 1\}$,因此,构成回归数据集 $\{\mathbf{z}_i, y_i\}, i = 1, \dots, n$,对于一个二分类问题,线性SVM通过以下优化问题寻找一个超平面 $\mathbf{w}^T \mathbf{z} + c = 0$,其中 \mathbf{w} 为法向量, c 为截距^[20]。这个优化问题的目标是找到一个最大间隔的超平面,当 $y_i = 1$ 时, $\mathbf{w}^T \mathbf{z} + c > 0$;反之, $\mathbf{w}^T \mathbf{z} + c < 0$ 来分割数据集,其对应表达式如下所示:

$$\min_{\mathbf{w}, c} \frac{1}{2} \|\mathbf{w}\|^2 \quad (3)$$

$$\text{s.t.} (\mathbf{w}^T \mathbf{z}_i + c) y_i \geq 1, i = 1, \dots, n$$

其中, $\|\mathbf{w}\|$ 是 \mathbf{w} 的范数。因此,对应的分类函数为 $g(\mathbf{z}) = \text{sign}(\mathbf{w}^T \mathbf{z} + b)$ 。然而,在医学研究中,通常会出现数据线性不可分的情况,为解决这个问题,引入松弛变量 ζ_i 以及惩罚参数 K ,因此式(3)可以转换为:

$$\min_{\mathbf{w}, c, \zeta} \frac{1}{2} \|\mathbf{w}\|^2 + K \sum_{i=1}^n \zeta_i \quad (4)$$

$$\text{s.t.} (\mathbf{w}^T \mathbf{z}_i + c) y_i \geq 1 - \zeta_i, \zeta_i \geq 0$$

此时需要通过求解对偶问题的方法求解上述参数,将多个拉格朗日乘子 $b_{1i} \geq 0, b_{2i} \geq 0$ 引入到式(4)中,可得:

$$F(\mathbf{w}, c, \zeta, b_{1i}, b_{2i}) = \frac{1}{2} \|\mathbf{w}\|^2 + K \sum_{i=1}^N \zeta_i - \sum_{i=1}^N b_{1i} (y_i (\mathbf{w}^T \mathbf{z}_i + c) - 1 + \zeta_i) - \sum_{i=1}^N b_{2i} \zeta_i \quad (5)$$

原始问题的对偶问题:

$$\begin{aligned} \min_b & \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n b_{ij} b_{ij} y_i y_j (\mathbf{z}_i^T \mathbf{z}_j) - \sum_{i=1}^n b_{1i} \\ \text{s.t.} & \sum_{i=1}^n b_{1i} y_i = 0, 0 \leq \alpha_i \leq C \end{aligned} \quad (6)$$

设 $\mathbf{b}_1^* = (b_{11}^*, b_{12}^*, \dots, b_{1n}^*)^T$ 是式(6)的一个解,若存在 \mathbf{b}_1^* 的一个分量 $b_{ij}^*, 0 < b_{ij}^* < K$,则对应的分类决策函数为:

$$g(\mathbf{z}) = \text{sign} \left(\sum_{i=1}^n b_{1i}^* y_i (\mathbf{z} \mathbf{z}_i) + c^* \right) \quad (7)$$

目标函数式(6)中的内积 $\mathbf{z}_i^T \mathbf{z}_j$ 可以通过核函数来取代,这个核函数使得数据在特征空间中经过一个映射,将其转换为一个新的特征空间。当映射函数为非线性函数时,采用含有核函数的SVM便成为非线性分类模型。常见的核函数包括线性核函数、多项式核函数以及径向基函数,它们被应用于SVM中以处理非线性问题^[21]。

SVM模型参数对性能表现至关重要,然而调整这些参数通常依赖于人工经验和迭代试错,尤其由于SVM具有众多参数且意义多样,因此优化这些参数显得异常繁琐,容易陷入局部最优,导致精度提升有限^[22-23]。参数的最佳组合极大程度上取决于训练的样本数据。噪声的存在会影响模型学习和演算进程,降低训练模型的泛化能力和稳定性。然而,利用遗传算法可以有效地改善这一不足^[24]。GA基于生物进化理论,通过模拟自然选择和遗传机制,利用遗传复制、交叉和变异思想,演化出适应特定环境的最优解^[25]。它具有随机性、并行性和全局性,能够自动获取全局范围内的空间信息,并自适应地逼近最优解。遗传算法在各领域如目标优化、调度方案、模式识别和机器学习等得到广泛应用,因其鲁棒性强、效率高,成为一种重要的优化方法。算法优化的具体流程如图1所示。

2 实验

2.1 数据集简介

本文采用的数据来自四川省人民医院甲状腺外科,其中包含673例患者的临床诊断数据以及与患者声音相关的特征,包括声音特征、声带振动特征和共振特征,这些声音特征是利用音频分析技术收集的。音频分析技术将患者发出的声音信号进行录音并提取声音数据中的基频、声音强度和其他声学参数。声带振动特征包括颤音、规则长短声带震动等,这些信息可以给出声带的震动情况。同时,共振特征考察声音信号在喉部和口腔中的共振情况,如共振谱峰值和频谱宽度,这些数据能够反映声音在不同区

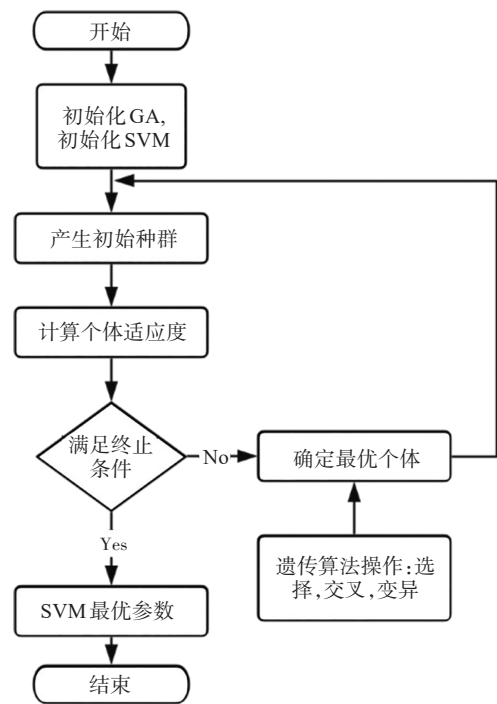


图1 遗传算法优化SVM参数的过程

Figure 1 Optimizing SVM parameters with genetic algorithm

域的变化。通过结合音频分析技术,将得到的声音特征用于评估甲状腺疾病的诊断和监测。所有患者数据的收集和使用都严格遵循相关的数据隐私法规,以确保未侵犯任何患者的权利。为了统一不同特征的数据量纲,采用归一化处理,以便对数据进行规范化和比较。表1详细列出了主要影响特征及其取值范围。根据上述音频信号,将建立的模型应用于上述音频信号研究中,用以识别音频信号中与甲状腺疾病相关的特定模式或特征。通过对数据集的分析,分析出音频特征与甲状腺疾病之间的关联,从而实现对患者疾病状态的预测和分类。

表1 主要影响特征和取值范围

Table 1 Major influence characteristics and the value ranges

| 特征名称 | 参数取值范围 |
|----------|-----------|
| 基频 | 51~271 Hz |
| 声音强度 | 37~89 dB |
| 共振峰 | 0, 1 |
| 颤音 | 15~221 Hz |
| 共振谱峰值 | 30~97 dB |
| 规则长短声带震动 | 0, 1 |
| 正常峰谷比 | 0, 1 |

2.2 数据处理及参数选择

该数据集涉及甲状腺疾病临床诊断以及与患者声音相关的特征数据,呈现了19.39:1的典型数据不平衡情况。鉴于此,本文采用SMOTE方法对原始数据集进行预处理,表2中展示了处理不平衡数据集所采用的超参数和取值结果。

表2 SMOTE算法的超参数与取值

Table 2 SMOTE algorithm hyperparameters and their respective values

| 超参数 | 含义 | 参数取值 |
|-------------------|-------|-------|
| K_neighbors | 最近邻数量 | 6 |
| sampling_strategy | 采样比例 | 1:1 |
| random_state | 随机种子 | 2 023 |

2.3 实验结果与分析

为了全面评估本文提出算法的效果,研究采用多个分类算法进行评估。经过数据预处理后,将数据按5:1的比例划分为训练集与测试集,并将决策树、Logistic回归、BP神经网络、K最近邻(KNN)、SVM、基于SMOTE的SVM(SMOTE-SVM)的算法以及本文提出的基于SMOTE数据扩充算法的GA-SVM模型(SMOTE-GA-SVM)的算法应用于数据集,本文选取评价各模型的准则包括召回率、准确率以及F1值。实验结果见表3。

表3 不同分类算法的对比结果

Table 3 Comparison of different classification algorithms

| 分类算法 | 准确率 | 召回率 | F1值 |
|--------------|------|------|------|
| 决策树 | 0.79 | 0.14 | 0.12 |
| Logistic回归 | 0.86 | 0.00 | 0.00 |
| BP神经网络 | 0.86 | 0.17 | 0.13 |
| KNN | 0.85 | 0.00 | 0.00 |
| SVM | 0.86 | 0.00 | 0.00 |
| SMOTE-SVM | 0.86 | 0.30 | 0.23 |
| SMOTE-GA-SVM | 0.87 | 0.40 | 0.37 |

根据表3的结果可知,未经过SMOTE数据扩充的算法,除了决策树和BP神经网络之外,其他算法的召回率和F1值均为0,说明这些算法在处理不平衡数据集时表现不佳。然而,通过SMOTE数据扩充后,SVM算法对于少数类样本的识别有了显著改善,召

回率和F1值均显著增加。为了更好地使模型适应数据特点,本研究利用遗传算法优化了SVM模型的参数,提出SMOTE-GA-SVM方法。该方法在处理数据不平衡问题上表现出色,不仅保持了高准确率,还显著提高了对少数类别的识别能力,使模型更全面地适应数据集的特征。这表明该算法在解决不平衡数据集问题上具有很大潜力,并且在选择算法时可以综合考虑多个参数,以找到最适合特定数据的最优模型。因此,本文提出的方法更适用于甲状腺疾病的分类诊断。

3 结 论

本文通过音频分析技术与SMOTE数据扩充方法相结合,提出一种基于GA-SVM的模型,用于甲状腺疾病的分类诊断。实验结果表明,这一方法在处理甲状腺疾病的分类诊断方面表现良好。特别是经过SMOTE数据扩充后,SVM算法对于识别少数类样本有了显著改善,从而提高召回率和F1值。通过遗传算法优化SVM模型参数,提出的SMOTE-GA-SVM算法在解决数据不平衡问题上表现优越。该算法不仅保持了高准确率,还显著提高了对少数类别的识别能力,使得模型更全面地适应数据集特征。研究结果体现了该算法处理不平衡数据集的潜力,但同时也体现了综合考虑多个参数对于算法选择的重要性。因此,本文提出的SMOTE-GA-SVM算法在甲状腺疾病分类诊断中展现出较高的实用价值和前景。

【参考文献】

- [1] Farling PA. Thyroid disease[J]. Br J Anaesth, 2000, 85(1): 15-28.
- [2] Klein I, Danzi S. Thyroid disease and the heart[J]. Circulation, 2007, 116(15): 1725-1735.
- [3] Wu P. Thyroid disease and diabetes[J]. Clinical diabetes, 2000, 18(1): 38-38.
- [4] Vanderpump MP. The epidemiology of thyroid disease[J]. Brit Med Bull, 2011, 99(1): 39-51.
- [5] Hemdan EE, El-Shafai W, Sayed A. CR19: a framework for preliminary detection of COVID-19 in cough audio signals using machine learning algorithms for automated medical diagnosis applications[J]. J Ambient Intell Humaniz Comput, 2023, 14(9): 11715-11727.
- [6] Peruzzi G, Pozzebon A, Van Der Meer M. Fight fire with fire: detecting forest fires with embedded machine learning models dealing with audio and images on low power IoT devices[J]. Sensors, 2023, 23(2): 783.
- [7] Al-muwaffaq I, Bozkus Z. MLTDD: use of machine learning techniques for diagnosis of thyroid gland disorder[J]. Comput Sci Inf Technol, 2016, 2016: 63-67.
- [8] Razia S, Swathi Prathyusha P, Vamsi Krishna N, et al. A comparative study of machine learning algorithms on thyroid disease prediction[J]. Int J Eng Technol, 2018, 7(2.8): 315-319.
- [9] Tyagi A, Mehra R, Saxena A. Interactive thyroid disease prediction system using machine learning technique[C]//2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC). Piscataway, NJ, USA: IEEE, 2018: 689-693.
- [10] Begum A, Parkavi A. Prediction of thyroid disease using data mining techniques [C]//2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS). Piscataway, NJ, USA: IEEE, 2019: 342-345.
- [11] Aversano L, Bernardi ML, Cimitile M, et al. Thyroid disease treatment prediction with machine learning approaches[J]. Procedia Comput Sci, 2021, 192: 1031-1040.
- [12] Alyas T, Hamid M, Alissa K, et al. Empirical method for thyroid disease classification using a machine learning approach[J]. Biomed Res Int, 2022, 2022: 9809932.
- [13] Islam SS, Haque MS, Miah MS, et al. Application of machine learning algorithms to predict the thyroid disease risk: an experimental comparative study[J]. PeerJ Comput Sci, 2022, 8: e898.
- [14] Sultana A, Islam R. Machine learning framework with feature selection approaches for thyroid disease classification and associated risk factors identification[J]. J Electr Syst Inf Technol, 2023, 10(1): 32.
- [15] Naeem AB, Senapati B, Chauhan AS, et al. Hypothyroidism disease diagnosis by using machine learning algorithms[J]. Int J Intell Syst Appl Eng, 2023, 11(3): 368-373.
- [16] Pal M, Parija S, Panda G. Enhanced prediction of thyroid disease using machine learning method[C]//2022 IEEE VLSI Device Circuit and System (VLSI DCS). Piscataway, NJ, USA: IEEE, 2022: 199-204.
- [17] Chawla NV, Bowyer KW, Hall LO, et al. SMOTE: synthetic minority over-sampling technique[J]. J Artif Intell Res, 2002, 16(1): 321-357.
- [18] Cortes C, Vapnik V. Support-vector networks[J]. Mach Learn, 1995, 20(3): 273-297.
- [19] Wang HF, Hu DJ. Comparison of SVM and LS-SVM for regression [C]//2005 International Conference on Neural Networks and Brain. Piscataway, NJ, USA: IEEE, 2005: 279-283.
- [20] Cherkassky V, Ma YQ. Practical selection of SVM parameters and noise estimation for SVM regression[J]. Neural Netw, 2004, 17(1): 113-126.
- [21] Yu H, Kim S. SVM tutorial - classification, regression and ranking[M]// Rozenberg G, Bäck T, Kok JN. Handbook of Natural Computing. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012: 479-506.
- [22] Zhou T, Lu HL, Wang WW, et al. GA-SVM based feature selection and parameter optimization in hospitalization expense modeling[J]. Appl Soft Comput, 2019, 75: 323-332.
- [23] Huang WC, Liu HY, Zhang Y, et al. Railway dangerous goods transportation system risk identification: comparisons among SVM, PSO-SVM, GA-SVM and GS-SVM[J]. Appl Soft Comput, 2021, 109: 107541.
- [24] Shi XH, Liang YC, Lee HP, et al. An improved GA and a novel PSO-GA-based hybrid algorithm[J]. Inf Process Lett, 2005, 93(5): 255-261.
- [25] Scrucca L. GA: a package for genetic algorithms in R[J]. J Stat Softw, 2013, 53(4): 1-37.

(编辑:陈丽霞)