



基于多目标支持向量机的ADHD分类

杜海鹏 邵立珍 张冬辉

ADHD classification based on a multi-objective support vector machine

DU Hai-peng, SHAO Li-zhen, ZHANG Dong-hui

引用本文:

杜海鹏, 邵立珍, 张冬辉. 基于多目标支持向量机的ADHD分类[J]. 工程科学学报, 2020, 42(4): 441-447. doi: 10.13374/j.issn2095-9389.2019.09.12.007

DU Hai-peng, SHAO Li-zhen, ZHANG Dong-hui. ADHD classification based on a multi-objective support vector machine[J]. *Chinese Journal of Engineering*, 2020, 42(4): 441-447. doi: 10.13374/j.issn2095-9389.2019.09.12.007

在线阅读 View online: <https://doi.org/10.13374/j.issn2095-9389.2019.09.12.007>

您可能感兴趣的其他文章

Articles you may be interested in

基于改进的支持向量回归机算法的磁记忆量化缺陷反演

Metal magnetic memory quantitative inversion of defects based on optimized support vector machine regression

工程科学学报. 2018, 40(9): 1123 <https://doi.org/10.13374/j.issn2095-9389.2018.09.014>

固溶时效工艺对6016铝合金力学性能的影响及多目标优化

Effect of solution and aging processes on the mechanical properties of 6016 aluminum alloy and multi-objective optimization

工程科学学报. 2017, 39(1): 75 <https://doi.org/10.13374/j.issn2095-9389.2017.01.010>

基于全局优化支持向量机的多类别高炉故障诊断

Multi-class fault diagnosis of BF based on global optimization LS-SVM

工程科学学报. 2017, 39(1): 39 <https://doi.org/10.13374/j.issn2095-9389.2017.01.005>

无数学模型的非线性约束单目标系统优化方法改进

Optimization method improvement for nonlinear constrained single objective system without mathematical models

工程科学学报. 2018, 40(11): 1402 <https://doi.org/10.13374/j.issn2095-9389.2018.11.014>

多目标多约束混合流水车间插单重调度问题研究

Research on rush order insertion rescheduling problem under hybrid flow shop with multi-objective and multi-constraint

工程科学学报. 2019, 41(11): 1450 <https://doi.org/10.13374/j.issn2095-9389.2018.11.27.002>

基于多目标支持向量机的 ADHD 分类

杜海鹏, 邵立珍[✉], 张冬辉

北京科技大学自动化学院工业过程知识自动化教育部重点实验室, 北京 100083

✉通信作者, E-mail: lishao@ustb.edu.cn

摘 要 注意力缺陷多动障碍(ADHD)是儿童期最常见的精神疾病之一,在大多数情况下持续到成年期。近年来,基于功能磁共振数据的 ADHD 分类成为了研究热点。文献中已有的大多数分类算法均假设样本是均衡的,然而事实上,ADHD 数据集通常是不平衡的。传统的学习算法会使得分类器倾向于多数类样本,从而导致性能下降。本文研究了基于不平衡神经影像数据的 ADHD 分类问题,即基于静息状态功能磁共振数据对 ADHD 进行分类。采用功能连接矩阵作为分类特征,提出了一种基于多目标支持向量机的 ADHD 数据分类方案。该方案将不平衡数据分类问题建模为具有三个目标的支持向量机模型,其中三个目标分别为最大化分类间隔、最小化正样本误差和最小化负样本误差,进而正负样本经验误差可以被分开处理。然后采用多目标优化的法向量边界交叉法对模型进行求解,并给出一组代表性的分类器供决策者进行选择。该方案在 ADHD-200 竞赛的五个数据集上进行测试评估,并与传统分类方法进行对比。实验结果表明本文提出的三个目标支持向量机分类方案比传统的分类方法效果好,可以有效的从算法层面解决数据不平衡问题。该方案不仅可用于辅助 ADHD 诊断,还可用于阿尔茨海默病和自闭症等疾病的辅助诊断。

关键词 多目标优化;功能磁共振数据;注意力缺陷多动障碍;支持向量机;不平衡数据集

分类号 TG181

ADHD classification based on a multi-objective support vector machine

DU Hai-peng, SHAO Li-zhen[✉], ZHANG Dong-hui

Key Laboratory of Knowledge Automation for Industrial Processes of Ministry of Education, School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China

✉ Corresponding author, E-mail: lishao@ustb.edu.cn

ABSTRACT Attention deficit hyperactivity disorder (ADHD) is one of the most common mental disorders during childhood, which lasts until adulthood in most cases. In recent years, ADHD classification based on functional magnetic resonance imaging (fMRI) data has become a research hotspot. Most existing classification algorithms reported in the literature assume that samples are balanced; however, ADHD data sets are usually imbalanced. Imbalanced data sets can cause the performance degradation of a classifier by imbalanced learning, which tends to overfocus on the majority class. In this study, we considered an imbalanced neuroimaging classification problem: classification of ADHD using resting state fMRI. We used the functional connection matrix of fMRI as the classification feature and proposed a multi-objective data classification scheme based on a support vector machine (SVM) to aid the diagnosis of ADHD. In this scheme, the imbalanced data classification problem is formulated as an SVM model with three objectives: maximizing the margin, minimizing the sum of positive errors, and minimizing the sum of negative errors. Accordingly, the positive and negative sample empirical errors can be separately handled. Then, the model is solved by a multi-objective optimization method, i.e., normal boundary intersection method. A set of representative classifiers are computed for selection by decision makers. The proposed scheme was tested and evaluated on five data sets from the ADHD-200 consortium and compared with traditional classification methods. Experimental results show that the proposed three-objective SVM classification scheme is better than traditional classification methods

收稿日期: 2019-09-12

reported in the literature. It can effectively address the data imbalance problem from the algorithm level. This scheme can be used in the diagnosis of ADHD as well as other diseases, such as Alzheimer's and Autism.

KEY WORDS multi-objective optimization; functional magnetic resonance imaging; attention deficit hyperactivity disorder; support vector machine; imbalance data set

注意力缺陷多动障碍 (Attention deficit hyperactivity disorder, ADHD) 是儿童期最常见的精神疾病之一, 在大多数情况下持续到成年期. ADHD 在 DSM-5 中被定义为神经发育障碍, 主要表现为注意力缺陷, 过度活动和行为冲动等症状^[1]. 据报道, 全球儿童和青少年中 ADHD 的发病率为 3.4%. ADHD 的病因和发病机制尚不清楚, 目前 ADHD 的诊断主要依赖于医生的主观经验. 因此, ADHD 的客观诊断和有效治疗是神经科学领域的重要课题之一.

近年来, 脑电图^[2]、磁共振成像^[3]和功能性磁共振成像^[4]等技术已被用于 ADHD 的辅助诊断. 其中, 静息态功能磁共振成像 (Resting state functional magnetic resonance imaging, rs-fMRI) 在精神疾病的病理分析中显示出其特有的优势, 不仅可以用于诊断 ADHD, 还可以用于诊断精神分裂症^[4]和老年痴呆症^[5].

研究者们提出了各种特征提取、选择和分类方法用于基于 rs-fMRI 的 ADHD 分类中. Castellanos 等^[6]发现 fMRI 的功能连接信息可以成为 ADHD 诊断的一个突出特征. Du 等^[7]提出了一种判别子网络的方法来对 ADHD 进行分类, 该方法挖掘了来自全脑网络的判别子网络, 并使用基于图核的 PCA 来提取特征. Qureshi 等^[8]计算了 fMRI 的全局连通图, 并利用基于图谱的皮质分割的平均连通性度量作为分层极限学习机分类器的输入特征. Miao 和 Zhang^[9]提出了一种基于权重的 relief 算法来获得 rs-fMRI 中低频波动分数幅度的特征子集. Riaz 等^[10]集成了非影像数据和影像数据的机器学习框架, 研究 ADHD 和正常受试者之间功能连接的变化. 考虑到数据不平衡性, 合成少数类过采样技术 (Synthetic minority oversampling technique, SMOTE)^[11]用于生成少数类样本.

以上提到的大多数分类算法均假设样本是均衡的. 然而事实上, 基于 rs-fMRI 的 ADHD 数据分类问题中数据集是不平衡的. 若采用传统的分类方法, 通过不平衡学习会导致对多数类别样本的过度聚焦, 分类器性能下降. 已有的不平衡数据处理方法大体分为两大类: 数据层面的方法和算法层面的方法^[12]. 数据层面的方法通过数据采样来

处理数据不平衡问题, 而算法层面的方法通常在决策过程中对不同的错分样本引入不同的惩罚因子. 在 ADHD 数据分类中, SMOTE 方法已用于处理数据集不平衡问题. 但是, 通过对少数群体/多数群体进行随机过采样/欠采样, 这些创建平衡训练数据集的策略可能导致分类器性能欠佳^[13].

考虑到分类问题的多目标性质, Shao 等^[14]提出了一种用于 ADHD 分类的双目标分类方法. 但是, 该方法并没有考虑数据集的不平衡性. 因此, 本文提出了采用基于 SVM 的多目标分类方案来解决 ADHD 数据不平衡问题, 该方案通过多目标优化单独惩罚错分的正负样本, 从而可以从算法层面有效地处理数据不平衡问题.

1 数据处理

本研究中使用的数据集来自于 ADHD-200 竞赛 (http://fcon_1000.projects.nitrc.org/)^[15]. 数据集主要从三个站点获取, 分别是 Kennedy Krieger Institute (KKI), New York University Medical Center (NYU) 和 Peking University (Peking). 实验采用了五个数据集, 分别为 KKI, NYU 和 Peking-1, Peking-2 和 Peking-joint, 其中 Peking-joint 由 Peking-1, Peking-2 和 Peking-3 三个数据集组成. 数据标签类型分为正常人群 (Normal control, NC) 和 ADHD 患者. 实验所用到的五个数据集的详细描述如表 1 所示.

表 1 ADHD-200 数据集描述

Table 1 Description of ADHD-200 data sets			
Data set	Total number of subjects	Number of ADHD subjects	Number of NC subjects
KKI	83	22	61
NYU	216	118	98
Peking-1	85	24	61
Peking-2	67	35	32
Peking-joint	194	78	116

数据的预处理过程采用了 DPARSF 工具箱 (<http://rfmri.org/DPARSF>). 预处理主要包括移除前十张不稳定图像, 时间层校正, 头动校正, 空间标准化, 带通滤波和平滑处理. 接下来对 90 个脑区分别计算其平均时间序列值, 进一步地计算 90 个

脑区两两之间的皮尔逊相关系数, 最终得到功能连接 (Functional connection, FC) 矩阵^[16], FC 矩阵采集的流程图如图 1 所示. 由于 FC 的对称性, 取下三角矩阵 $((90 \times 90 - 90) / 2 = 4005)$ 作为样本的特征.

由于实验所用数据集特征的个数远远大于样本个数, 采用 PCA (Principal component analysis)^[17] 对数据进行降维. 为了尽可能多地保留信息, 每个数据集的特征维度降至比其训练样本的数量少 1.

2 多目标分类方案

本文提出的基于多目标支持向量机的 ADHD 分类方案如图 2 所示. 首先建立三个目标分类模型, 其次通过多目标优化算法求解多目标优化问题以得到帕累托 (Pareto) 最优分类器, 然后评估分类器性能, 在交叉验证集中选出最佳分类器, 最后在测试集上进行测试.

2.1 三个目标 SVM 分类模型

在二分类问题中, 数据集中每个类别下的样本数目相差很大, 则该数据集被认为是不平衡的. 数目少的一类称作少数类样本, 数目较多的一类称为多数类样本. 文中所用数据集, 多数类为 NC 样本, 少数类为 ADHD 样本.

支持向量机 (Support vector machine, SVM) 是一种常用的有监督机器学习算法^[18]. 给定一个具有 m 个样本的训练集合 $S = \{(x^1, y^1), (x^2, y^2), \dots, (x^m, y^m)\}$, 其中 $x^i \in X \subset \mathbb{R}^n$ 表示第 i 个样本的特征, $y^i \in \{-1, +1\}$ 代表第 i 个样本的标签. SVM 的基本思想是在特征空间上找到最佳的分离超平面使得训练集上正负样本间隔最大. 分类超平面可以用 $y = w \cdot x + b$ 表示, 其中, $w \in \mathbb{R}^n$ 表示法向量, $b \in \mathbb{R}$ 表示截距.

传统的 SVM 对泛化能力和经验误差进行加权, 它在本质上是一个双目标优化问题^[19]. 文献 [14] 采用了基于 1 范数的双目标分类模型用于分类 ADHD 和 NC 受试者. 但是, 该模型没有考虑不平衡数据分布, 把正样本的误差和负样本的误差加和在了一起. 事实上, 当存在类不平衡时, 由于分类器的目的是最大化整体准确性, 传统的分类器将倾向于多数类. 换句话说, 分类器可能将所有样本分类为负值, 从而提高过高的准确度.

本文使用如下三个目标 SVM 模型 (T-SVM) 对正负样本的误差成本分别进行处理, 其中三个目标分别是最大化分类间隔, 最小化正样本经验误差之和与最小化负样本经验误差之和^[20].

$$\begin{aligned} (\text{T-SVM}) \min \quad & \|w\|_1, \sum_{\{i|y^i=+1\}} \xi_+^i, \sum_{\{i|y^i=-1\}} \xi_-^i \\ \text{s.t.} \quad & y^i(w \cdot x^i + b) + \xi_+^i \geq 1, i = 1, 2, \dots, m_+ \\ & y^i(w \cdot x^i + b) + \xi_-^i \geq 1, i = m_+ + 1, m_+ + 2, \dots, m \\ & \xi_+^i, \xi_-^i \geq 0 \end{aligned}$$

其中, m_+ 表示多数类样本个数, ξ_+^i, ξ_-^i 分别代表正负样本的经验误差, 目标函数反映了最大化分类间隔的同时最小化正负样本的经验误差. 为使优化问题 (T-SVM) 可解, 用两个正变量 w_+ 和 w_- 来表示 w , 即 $w = w_+ + w_-$, 因此第一个目标函数就表示为 $\|w\|_1 = e^T(w_+ + w_-)$, 其中 e 是全为 1 的列向量. 该模型是一个多目标线性优化问题 (Multi-objective linear programming, MOLP). 基于多目标优化最优解的概念, 本文给出如下 Pareto 最优分类器的定义.

对于一个训练集, 如果不存在任何一个可行解 $(\bar{w}_+, \bar{w}_-, \bar{b}, \bar{\xi}_+^i, \bar{\xi}_-^i)$ 支配 $(w_+, w_-, b, \xi_+^i, \xi_-^i)$, 即 $e^T(w_+ + w_-) \leq$

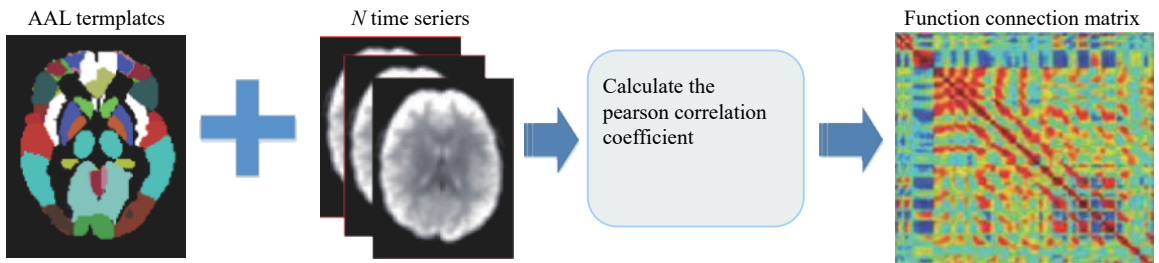


图 1 功能连接矩阵采集流程图

Fig.1 Flowchart of functional connection matrix acquisition

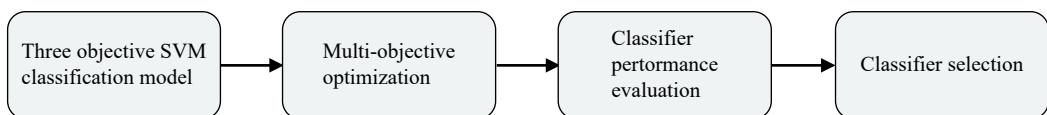


图 2 基于多目标支持向量机的 ADHD 分类方案

Fig.2 ADHD classification scheme based on multi-objective SVM

$e^T(\bar{w}_+ + \bar{w}_-)$, $\sum_{\{i|y^i=+1\}} \xi_+^i \leq \sum_{\{i|y^i=+1\}} \bar{\xi}_+^i$, $\sum_{\{i|y^i=-1\}} \xi_-^i \leq \sum_{\{i|y^i=-1\}} \bar{\xi}_-^i$, 则

称 $(\bar{w}_+, \bar{w}_-, \bar{b}, \bar{\xi}_+^i, \bar{\xi}_-^i)$ 为 Pareto 最优分类器. 所有 Pareto 最优分类器构成了 Pareto 前沿分类器. 由于所有 Pareto 最优分类器都是对分类间隔, 正负经验误差之间的权衡, 因此决策者可能会感兴趣. 此外, 由于具有最佳性能的分类器均在 Pareto 最优分类器集合中, 不需要考虑该集合之外的分类器.

2.2 多目标优化算法

本文采用法向边界交叉法 (Normal boundary intersection method, NBI)^[21] 来求解 MOLP 问题.

$$\begin{aligned} (\text{MOLP}) \min f(x) &= (f_1(x), f_2(x), \dots, f_p(x))^T \\ \text{s.t. } x \in \mathbb{R}^n : g(x) &= (g_1(x), g_2(x), \dots, g_p(x))^T \leq 0 \end{aligned}$$

其中, $X = \{x \in \mathbb{R}^n : g(x) = (g_1(x), g_2(x), \dots, g_p(x))^T \leq 0\}$ 代表决策变量可行域, 假设其非空, 则目标空间可行域为 $Y = \{f(x) : x \in X\}$.

对于 MOLP 问题, 如果不存在 $x \in X$ 使得 $f(x) \leq f(\hat{x})$, 则称 $\hat{x} \in X$ 为 MOLP 问题的一个有效解. 有效解组成的集合表示为 X_E , 称作决策空间中的有效集合. 相应地, $\hat{y} = f(\hat{x})$ 称作一个非支配点, $Y_N = \{f(x) : x \in X_E\}$ 称之为目标空间可行域中的非支配点解集. NBI 方法就是为了求得 MOLP 问题中的非支配点集 Y_N 的子集 R . 图 3 展示了 NBI 方法求解一个两个目标优化模型的示例. 该方法首先计算一个参考平面, 并在参考平面上放置均匀分布的参考点, 然后沿着法线方向将参考点投影到 Y 的边界, 最终得到多目标优化模型的代表性非支配点集合 R . 利用 NBI 解多个目标优化问题如算法 1 所示.

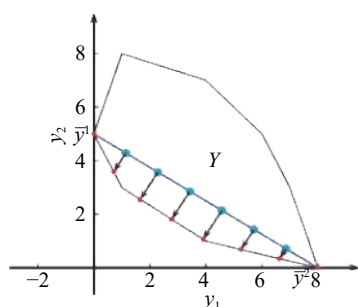


图 3 NBI 方法中获得的非支配点

Fig.3 Non-dominated points obtained using the NBI method

算法 1 NBI 算法求解 MOLP 问题

输入: 优化问题模型

(1) 求解三个目标模型中每个目标的最小值 $y_k^l = \min\{f_k(x) : x \in X\}$, $k = 1, 2, 3$, 假设最优解被表示为 $\bar{y}_1, \bar{y}_2, \bar{y}_3$, 构造 $\bar{y}_1, \bar{y}_2, \bar{y}_3$ 三个点组成的凸包作为参

考平面, 计算参考平面的法向量方向 $\hat{n}$;

(2) 在参考平面上布置均匀分布的参考点 $q_i, i = 1, \dots, k$;

(3) 遍历所有的 k 个参考点, 求解最优化问题:

$$\max t, \text{ s.t. } q_i + t\hat{n} = f(x), t \geq 0, x \in X$$

并把 $q_i + t\hat{n}$ 存储到集合 R .

输出: 代表性非支配点子集 R

通过 NBI 方法可以求得一组 MOLP 的代表性非支配点. 对于 (T-SVM) 问题, 每个非支配点都对应一个 Pareto 最优分类器. 决策者可以在交叉验证集上遍历所有的 Pareto 最优分类器, 选择出性能最优 Pareto 分类器作为最终的分器.

2.3 分类器性能评估

对于一组 Pareto 最优分类器, 决策者需要根据交叉验证集上的性能选择最终分类器. 常用的衡量分类器性能的评价指标有灵敏度 (Sensitivity)、特异性 (Specificity) 和准确性 (Accuracy). 敏感性表明少数群体的准确性, 特异性表明多数群体的准确性.

一般情况下, 准确率通常被视为评估标准之一. 但是, 在数据集不平衡的情况下准确率不能完全反映分类器的性能好坏. 例如, 对于正负样本比率为 1 : 9 的数据集, 即使判断所有少数类别的样本都错误, 准确率仍然可以达到 90%. 但是对于疾病诊断, 正确分类少数类样本 (患病) 是很重要的. 因此, 本文使用灵敏度和特异性的几何平均值 $g\text{-means}$ ($g\text{-means} = \sqrt{\text{sensitivity} \times \text{specificity}}$) 来评估分类器性能.

3 实验结果

本文使用提出的基于 1 范数 SVM 的三个目标分类方案对 ADHD-200 竞赛的五个数据集进行分类测试. 每个数据集都被随机分为三个数据集: 训练集、交叉验证集和测试集, 划分比例为 6 : 2 : 2, 即使用数据集的 60% 用于训练, 20% 用于模型交叉验证选取最终分类器, 剩余的 20% 用来测试衡量最终分类器的效果.

下面以 Peking-1 数据集为例, 给出分类器的选择过程. 首先, 用 NBI 方法来求解由训练集构成的 (T-SVM) 问题, 从而获得一组非支配点, 共 11 个, 如图 4(a) 所示, 每个对应一个 Pareto 最优分类器, 图中三个坐标轴分别代表分类间隔 $\|w\|_1$, 正样本经验误差之和 $\sum \xi_+$ 和与负样本经验误差之和 $\sum \xi_-$. 进一步地, 计算出 11 个分类器在训练集和交叉验证集上的准确率和 $g\text{-means}$ 值, 见表 2.

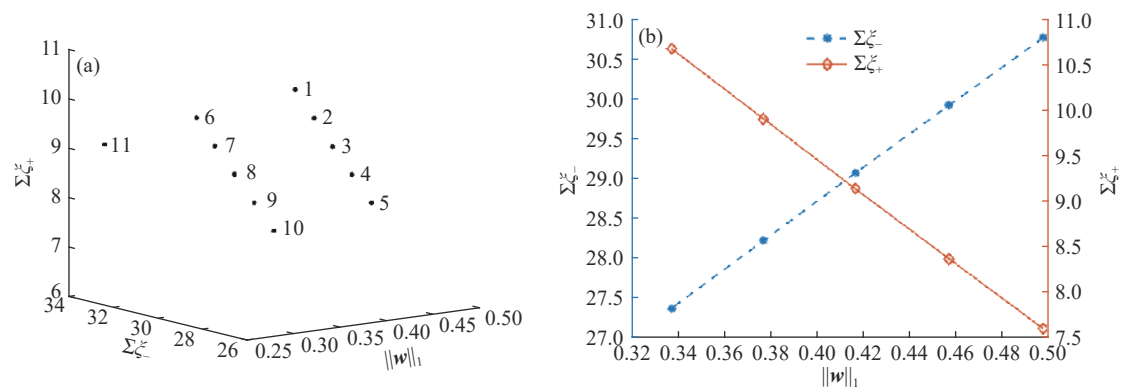


图 4 Peking-1 数据集上非支配点集. (a)非支配点集; (b)非支配点 1-5 的权衡关系

Fig.4 Non-dominated points on Peking-1 data set: (a) non-dominated points; (b) trade-off information of non-dominated points 1-5

表 2 训练集/交叉验证集上的性能评价

Table 2 Evaluation of the training/cross-validation data set

Classifier	1	2	3	4	5	6
Accuracy	0.6600/0.6842	0.6400/0.6842	0.6600/0.7368	0.6800/0.3842	0.6600/0.7368	0.6000/0.6316
G-means	0.6547/0.5311	0.6607/0.6202	0.6929/0.7161	0.7237/0.6794	0.7182/0.7596	0.6299/0.5883

Classefier	7	8	9	10	11
Accuracy	0.6000/0.6842	0.6200/0.6842	0.6000/0.6842	0.6200/0.6842	0.5800/0.5789
G-means	0.6299/0.6794	0.6726/0.6794	0.6547/0.7161	0.6841/0.7161	0.6268/0.5991

为了清楚地展示非支配点三个目标之间的权衡关系,选取了前五个非支配点,在图 4(b)和图 5 中展示其所对应的分类器的性能. 图 4(b)展示了三个目标($\|\mathbf{w}\|_1, \Sigma \xi_{+}, \Sigma \xi_{-}$)之间的权衡关系,可以看出随着 $\|\mathbf{w}\|_1$ 的增加,正样本经验误差之和 $\Sigma \xi_{+}$ 随之上升,而负样本经验误差之和 $\Sigma \xi_{-}$ 随之下降. 图 5 进一步展示了这五个分类的性能,其中 5(a)展示了分类误差, 5(b)展示了 g-means 值. 两个子图中以菱形点绘成的曲线代表正样本的分类误差之和,图中具有相同横坐标值的其他点代表分类器在训练集与交叉验证集上的性能. 从图中可以看出,正样本(少数类)经验误差的减小是以牺牲多

数类样本的准确性为代价的. 根据所有非支配点对应的 Pareto 最优分类器在交叉验证集的表现,对于 Peking-1 数据集,选择了具有最高 g-means(值为 0.7596)的分类器 5 作为该数据集的最终分类器. 其他几个数据集的处理过程与 Peking-1 数据集一样,由于篇幅限制这里不再给出. 为了进一步地展示本文提出的三个目标分类方案的性能,将本文提出的方法与 1 范数 SVM (L_1 SVM)、2 范数 SVM(L_2 SVM)、随机森林(random forest, RF)^[22]、极限学习机(extreme learning machine, ELM)^[23]以及两个目标分类方案(B-SVM)进行了

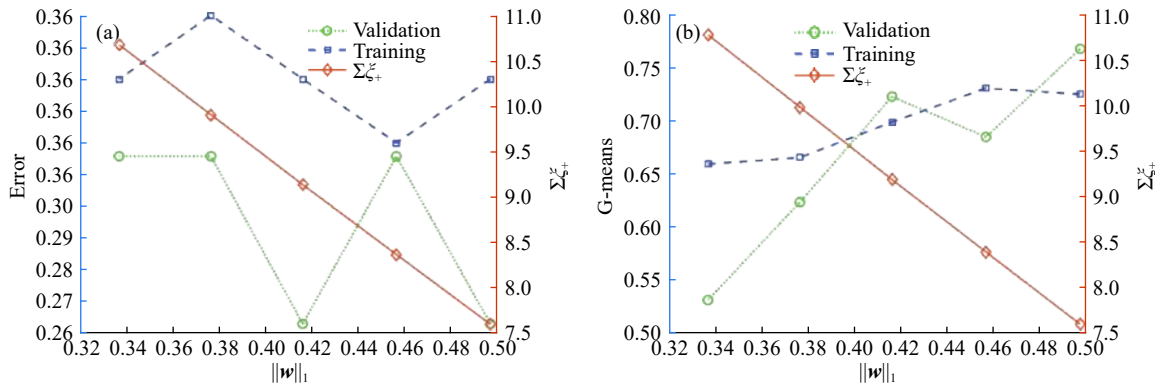


图 5 Peking-1 数据集上 1-5 Pareto 最优分类器的性能. (a)范数与经验误差的关系; (b)范数与 g-means 的关系

Fig.5 Performance of Pareto optimal classifiers 1-5 for Peking-1: (a) norm versus empirical error; (b) norm versus g-means

对比分析. 其中, L_1 SVM 和 L_2 SVM 中的超参数 $C=0.8$; RF 中树的个数为 50, 每棵树的最大深度为 5; ELM 中隐层节点个数为 30. 实验运行了十次, 最终给出了每个数据集下十次的平均结果, 如表 3 所示. 表中每个数据集最高的准确度以及最高的 g-means 值均以黑体加粗的形式给出. 对比结果表明, 本文提出的三个目标的方案在所有 ADHD 数据集上的表现都优于其他对比方法.

除 ADHD 数据集外, 本文也选取了 University of California Irvine(UCI) Machine Learning Repository 上的 MNIST 数据集来测试提出的方法的有效性. MNIST 数据集有 10 类, 我们对其进行下采样并随机选取标签为“9”的 50 个样本作为正样本, 其他类均为负样本构造分类问题. 分类结果如表 3 所示, 结果表明本文提出的方法对一般的不平衡数据集也具有较好的分类效果.

表 3 不同方法的平均准确度/g-means 值

Table 3 Average accuracy/g-means value for different methods

Data set	L_1 SVM	L_2 SVM	B-SVM	RF	ELM	T-SVM
KKI	0.635/0.421	0.634/0.515	0.732/0.527	0.725/0.530	0.696/0.622	0.753/0.606
NYU	0.545/0.543	0.556/0.542	0.643/0.624	0.608/0.610	0.588/0.594	0.703/0.698
Peking-1	0.725/0.683	0.714/0.664	0.801/0.677	0.770/0.688	0.677/0.647	0.813/0.711
Peking-2	0.636/0.637	0.665/0.683	0.807/0.776	0.635/0.649	0.564/0.601	0.845/0.851
Peking-joint	0.630/0.615	0.624/0.611	0.742/0.764	0.665/0.686	0.625/0.613	0.751/0.743
MNIST	0.977/0.783	0.978/0.797	0.979/0.800	0.975/0.790	0.969/0.00	0.984/0.849

4 结论

本文提出了一种基于多目标支持向量机的 ADHD 数据分类方案. 该方案使用基于 1 范数 SVM 的三个目标优化模型, 分别考虑了正负样本的经验误差, 从而可以从算法层面有效地处理类不平衡问题. 通过求解多目标优化问题, 可以得到一组代表性的 Pareto 最优分类器以供决策者进行选择. 该分类方案在 ADHD-200 数据集上进行了测试并和文献中的方法进行了对比分析. 实验结果表明, 本文提出的三个目标 SVM 分类方案在所有测试数据集上的表现优于 1 范数 SVM, 2 范数 SVM, 随机森林、极限学习机和双目标 SVM 方法.

参 考 文 献

[1] American Psychiatric Association. Diagnostic and statistical manual of mental disorders. *BMC Med*, 2013, 17: 133

[2] Saad J F, Kohn M R, Clarke S, et al. Is the theta/beta EEG marker for ADHD inherently flawed? *J Attention Disord*, 2018, 22(9): 815

[3] Chang C W, Ho C C, Chen J H. ADHD classification by a texture analysis of anatomical brain MRI data. *Front Syst Neurosci*, 2012, 6: 66

[4] Kuang L D, Lin Q H, Gong X F, et al. Model order effects on ICA of resting-state complex-valued fMRI data: application to schizophrenia. *J Neurosci Methods*, 2018, 304: 24

[5] Hojjati S H, Ebrahimzadeh A, Khazaei A, et al. Predicting conversion from MCI to AD using resting-state fMRI, graph theoretical approach and SVM. *J Neurosci Methods*, 2017, 282: 69

[6] Castellanos F X, Margulies D S, Kelly C, et al. Cingulate-precuneus interactions: a new locus of dysfunction in adult attention-deficit/hyperactivity disorder. *Biol Psychiat*, 2008, 63(3): 332

[7] Du J Q, Wang L P, Jie B, et al. Network-based classification of ADHD patients using discriminative subnetwork selection and graph kernel PCA. *Comput Med Imag Graph*, 2016, 52: 82

[8] Qureshi M N I, Jo H J, Lee B. ADHD subgroup discrimination with global connectivity features using hierarchical extreme learning machine: resting-state FMRI study // 2017 *IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*. Melbourne, 2017: 529

[9] Miao B, Zhang Y L. A feature selection method for classification of ADHD // *Proceedings of 4th International Conference on Information, Cybernetics and Computational Social Systems (ICCSS)*. Dalian, 2017: 21

[10] Riaz A, Asad M, Alonso E, et al. Fusion of fMRI and non-imaging data for ADHD classification. *Comput Med Imag Graph*, 2018, 65: 115

[11] Chawla N V, Bowyer K W, Hall L O, et al. SMOTE: synthetic minority over-sampling technique. *J Artif Intell Res*, 2002, 16: 321

[12] Krawczyk B. Learning from imbalanced data: open challenges and future directions. *Prog Artif Intell*, 2016, 5(4): 221

[13] He H B, Garcia E A. Learning from imbalanced data. *IEEE Trans Knowl Data Eng*, 2009, 21(9): 1263

[14] Shao L Z, Xu Y D, Fu D M. Classification of ADHD with bi-objective optimization. *J Biomed Inf*, 2018, 84: 164

[15] Bellec P, Chu C, Chouinard-Decorte F, et al. The neuro bureau ADHD-200 preprocessed repository. *Neuroimage*, 2017, 144: 275

[16] Friston K J. Functional and effective connectivity: a review. *Brain*

- Connect*, 2011, 1(1): 13
- [17] Reris R, Brooks J P. Principal component analysis and optimization: a tutorial // *Proceedings of 14th INFORMS Computing Society Conference*, Richmond, Virginia, US, 2015: 212
- [18] Cortes C, Vapnik V. Support-vector networks. *Mach Learn*, 1995, 20(3): 273
- [19] Aytug H, Sayin S. Exploring the trade-off between generalization and empirical errors in a one-norm SVM. *Eur J Oper Res*, 2012, 218(3): 667
- [20] Aşkan A, Sayin S. SVM classification for imbalanced data sets using a multiobjective optimization framework. *Ann Oper Res*, 2014, 216(1): 191
- [21] Das I, Dennis J E. Normal-boundary intersection: a new method for generating the Pareto surface in nonlinear multicriteria optimization problems. *SIAM J Optim*, 1998, 8(3): 631
- [22] Breiman L. Random forests. *Mach Learn*, 2001, 45(1): 5
- [23] Peng X L, Lin P, Zhang T S, et al. Extreme learning machine-based classification of ADHD using brain structural MRI data. *PloS One*, 2013, 8(11): e79476