

加权子空间的随机向量函数链接网络集成方法^①

叶璇^{1,2}, 何玉林^{1,2}, 张曼静¹, 黄哲学^{1,2}

1. 人工智能与数字经济广东省实验室(深圳), 广东 深圳 518107;

2. 深圳大学 计算机与软件学院, 广东 深圳 518060

摘要: 随机向量函数链接网络(RVFL)是一种随机权网络模型,其基于非迭代权重更新方式,直接求解输出层权重完成模型训练,因此具有训练速度快的优点,已有的实验证明了其在分类和回归任务中均具有良好的泛化能力.目前在 RVFL 的改进工作中存在两点问题:改善网络结构会复杂化模型,容易造成过拟合现象;结合集成学习往往无法进一步通过增加集成多样性来提升模型性能.因此,本文基于子空间策略,提出了一种基于加权子空间的随机向量函数链接网络集成方法(WAB-RVFL).基于 RVFL 网络结构中输入层和输出层直连的线性特点,WAB-RVFL 引入属性优化的思想并提出属性加权矩阵的概念,对属性子空间进行加权转化获得更优质的加权子空间,使其更利于模型进行集成训练.通过在 8 个高维分类数据集上的实验测试,证实了 WAB-RVFL 的可行性、合理性和有效性,其能够获得比 6 种流行的 RVFL 网络模型更优的泛化能力.

关键词: 子空间策略;集成学习;泛化能力;粒子群优化;随机向量函数链接网络

中图分类号: TP391

文献标志码: A

文章编号: 1000-5471(2022)12-0001-10

Random Vector Function Link Network Ensemble Learning Method for Weighted Subspaces

YE Xuan^{1,2}, HE Yulin^{1,2}, ZHANG Manjing¹, HUANG Zhexue^{1,2}

1. Guangdong Provincial Laboratory of Artificial Intelligence and Digital Economy (Shenzhen), Shenzhen Guangdong 518107, China;

2. School of Computer and Software, Shenzhen University, Shenzhen Guangdong 518060, China

Abstract: Random vector function link network (RVFL) is a random weight network model, which is based on non-iterative weight update method and solves the output layer weights directly to complete the model training, so it has the advantage of fast training speed, and the existing experiments have proved its good generalization ability in both classification and regression tasks. There are two problems in the current work of improving RVFL: improving the network structure will complicate the model and easily cause overfitting phenomenon; combining with ensemble learning often cannot further improve the model per-

① 收稿日期: 2022-07-28

基金项目: 国家自然科学基金面上项目(61972261); 广东省自然科学基金面上项目(2314050006683); 深圳市基础研究重点项目(JCYJ20220818100205012); 深圳市基础研究面上项目(JCYJ20210324093609026).

作者简介: 叶璇, 主要从事概率密度估计技术、非线性优化方法和大数据系统计算技术的研究.

通信作者: 何玉林, 副研究员.

formance by increasing the integration diversity. Therefore, in this paper, a weighted subspace-based random vector function linked network ensemble learning method (WAB-RVFL) is proposed based on the subspace strategy. Based on the linear characteristics of the RVFL network structure in which the input and output layers are directly connected, WAB-RVFL introduces the idea of attribute optimization and proposes the concept of attribute weighting matrix to obtain a better quality weighted subspace by weighting transformation of the genus subspace, making it more conducive to the model for ensemble training. Through experimental tests on eight high-dimensional classification datasets, the feasibility, rationality and effectiveness of WAB-RVFL are confirmed, and it can obtain better generalization ability than six popular RVFL network models.

Key words: subspace strategy; ensemble learning; generalization ability; particle swarm optimization; random vector function linked network

随机向量函数链接网络(random vector functional link network, RVFL)作为最早的随机权网络模型^[1]于 1992 年在文献[2]里提出,文献[3]在此基础上进行了总结完善.作为随机权网络代表,RVFL 的输入层权重和隐含层偏置的初始值为随机生成,求解隐含层节点与输入层节点所构成的结合矩阵的伪逆并直接得到输出层权重,进而实现模型训练.因此,RVFL 具有极快的训练速度,且在文献[4]中的实验证明了其具有良好的泛化能力,能处理分类和回归任务,具有很多真实应用场景.如何有效提升 RVFL 的稳定性和泛化性,具有重要研究意义.

由于 RVFL 的结构简单化以及模型随机化,因此当前其改进方向主要集中在改善网络结构和结合集成学习两方面:

1) 改善网络结构的代表性工作.在网络参数选择方面,文献[5]提出了有效节点选择的 RVFL 改进模型,简记为 I-OI-RVFLNs.该模型首先引入正交化方法对隐含层输出矩阵进行正交化,显著提升模型的收敛速度;其次预先设置隐含层节点个数并迭代更新输出层权重,裁剪冗余节点以实现网络精简,提升模型的计算效率.文献[6]于 2017 年提出了一种基于贝叶斯法则的 RVFL 改进模型,其在训练阶段获取额外先验知识得到最优权重的概率分布,并在测试阶段线性搜索出合适的正则化因子,进而构造网络模型并使其获得更好的预测性能.在网络层数方面,文献[7]基于 RVFL 网络框架以及自编码器概念,首次提出了深度 RVFL 网络,简记为 sdRVFL. sdRVFL 引入了去噪准则,即不同的正则化法则,先对数据特征进行深层次的提取,然后将该部分特征输入到原始的 RVFL 进行预测,从而实现网络结构的深度化.在 sdRVFL 的基础上,文献[8]于 2021 年提出了另一种深度 RVFL 网络,简记为 dRVFL. dRVFL 将网络中每一层所提取到的数据特征,进行横向拼接构成拼接矩阵,并利用最小二乘直接计算输出层权重,从而完成 dRVFL 的训练. dRVFL 充分利用每层网络所提取的信息,有效提升了 RVFL 的泛化能力.此外,文献[9]提出了另一种新型的深度 RVFL 网络,简记为 FAF-DRVFL. FAF-DRVFL 结合迁移学习来进行数据预处理工作,并使用模糊激活函数代替常用的 Sigmoid 激活函数来计算隐含层输出,有效提高了模型的稳定性以及计算效率.

2) 结合集成学习的代表性工作.文献[10]于 2020 年提出了一种基于自适应机制的集成 RVFL 模型.该模型根据 RVFL 的预测准确率来分配不同权重,若样本被错误分类则在下一次训练中给予更大的权重,以此完成模型的集成优化.文献[11]基于残差网络机制,对数据特征进行有效提取,同时每个基模型 dRVFL 的输出层参数中添加噪声扰动,增加集成多样性,从而实现对集成 dRVFL(edRVFL)网络模型泛化能力的改进.

若在 RVFL 网络结构上进行改进,会改变其简单的模型结构,增加过拟合的风险;若结合集成方法对 RVFL 进行改进,又因无法增大基学习器差异性导致无法提升模型的稳定性和泛化性.为了在不改变 RV-

FL 简单结构的条件下有效提升 RVFL 的稳定性和泛化性, 本文提出了一种加权子空间的随机向量函数链接网络集成方法 (weighted attribute bagging-based random vector functional link network, WAB-RVFL), 为优化所选属性子空间使其更适应于 RVFL 的网络结构, 提出属性优化的思想, 引入属性加权矩阵概念, 提出加权子空间选择策略, 搜索得到更优的加权子空间, 使其能在 RVFL 的线性映射与非线性映射之间找到平衡点. 为寻找最优的属性加权矩阵, 本文利用训练集和验证集信息构建目标函数, 并采取粒子群优化算法迭代及最小化目标函数, 从而搜索得到最优属性加权矩阵. 本文在 8 个分类数据集上展开有说服力的实验, 分别验证 WAB-RVFL 的可行性、合理性和有效性. 实验结果表明, WAB-RVFL 具有更优的泛化能力.

本文的组织结构如下: 第 1 节对 RVFL 和贝叶斯子空间选择策略的原理进行介绍和分析; 第 2 节对本文提出的 WAB-RVFL 进行详细介绍; 第 3 节对 WAB-RVFL 的性能进行实验验证; 第 4 节进行总结并给出未来研究工作.

1 相关理论知识

1.1 随机向量函数链接网络

传统神经网络基于误差反向传播算法进行权重的迭代更新, 虽然能够较好地实现复杂的非线性映射关系并完成表征学习任务, 但却具有训练时间长、易陷入局部极小值等缺点. RVFL 是一种特殊的单隐含层前馈神经网络, 其结构特点是输入层同时与隐含层、输出层进行直接连接, 具体如图 1 所示. 由于避免了权重的迭代更新, 因此 RVFL 具有极快的训练速度、良好的函数拟合能力.

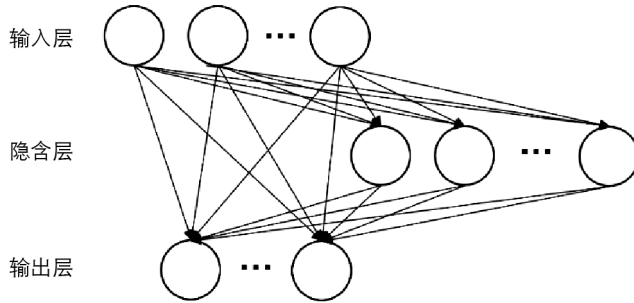


图 1 RVFL 的网络结构

RVFL 的训练步骤主要分为 3 个阶段: 随机初始化阶段、隐含层和输入层结合矩阵计算阶段、目标输出层权重计算阶段. 首先是随机初始化生成并固定输入层权重和隐含层偏置, 然后通过将数据样本进行非线性映射得到隐含层输出矩阵, 最后通过最小二乘法直接解析求得目标输出层权重.

假设有一个分类数据集容量大小为 N , 其具有 D 维属性以及 M 个类标:

$$S = \{(\mathbf{x}_n, \mathbf{t}_n) \mid \mathbf{x}_n = (x_{n1}, x_{n2}, \dots, x_{nD}), \mathbf{t}_n = (t_{n1}, t_{n2}, \dots, t_{nM})\} \quad (1)$$

1) 随机初始化阶段. 首先按照一定分布随机初始化 RVFL 的输入层权重和隐含层偏置, 分别表示为

$$\mathbf{W} = \begin{bmatrix} w_{11} & w_{12} & \dots & w_{1L} \\ w_{21} & w_{22} & \dots & w_{2L} \\ \vdots & \vdots & & \vdots \\ w_{D1} & w_{D2} & \dots & w_{DL} \end{bmatrix} \quad (2)$$

和

$$\mathbf{B} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_L \end{bmatrix} \quad (3)$$

2) 隐含层和输入层结合矩阵计算阶段. 数据样本经过输入层权重和隐含层偏置转换后, 进行非线性映射得到隐含层输出矩阵:

$$\mathbf{H}_{N \times L} = (\mathbf{h}_{nl})_{N \times L} = \begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1L} \\ h_{21} & h_{22} & \cdots & h_{2L} \\ \vdots & \vdots & & \vdots \\ h_{N1} & h_{N2} & \cdots & h_{NL} \end{bmatrix} \quad (4)$$

其中

$$h_{nl} = g\left(\sum_{d=1}^D x_{nd} w_{dl} + b_l\right), l = 1, 2, \dots, L \quad (5)$$

$g(\cdot)$ 表示激活函数, 通常使用 Sigmoid 函数:

$$g(u) = \text{Sigmoid}(u) = \frac{1}{1 + e^{-u}}, u \in \mathbb{R} \quad (6)$$

因此, 隐含层和输入层结合矩阵表示为

$$\mathbf{G}_{N \times (L+D)} = [\mathbf{H}_{N \times L}, \mathbf{X}_{N \times D}]_{N \times (L+D)} \quad (7)$$

其中 $\mathbf{X} = (x_{nd})_{N \times D}$ 是样本输入值. RVFL 的整个计算过程可以表示为

$$\mathbf{G}_{N \times (L+D)} \boldsymbol{\beta}_{(L+D) \times M} = \mathbf{Y}_{N \times M} \quad (8)$$

其中: $\boldsymbol{\beta}_{(L+D) \times M}$ 是目标输出层权重, $\mathbf{Y}_{N \times M}$ 是样本预测值.

3) 目标输出层权重计算阶段. 输出层权重是整个 RVFL 训练求解的核心部分. $\mathbf{T}$ 是样本真实值, 为使样本预测值和样本真实值尽可能接近, 需最小化以下目标函数:

$$\min_{\boldsymbol{\beta}} \|\mathbf{H}\boldsymbol{\beta} - \mathbf{T}\| \quad (9)$$

通过最小二乘法解析求解, 得到输出层权重为

$$\boldsymbol{\beta} = \begin{cases} (\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T \mathbf{T}, & \text{如果 } N \geq D \\ \mathbf{G}^T (\mathbf{G} \mathbf{G}^T)^{-1} \mathbf{T}, & \text{如果 } N < D \end{cases} \quad (10)$$

为了更好平衡经验风险和结构风险, 往往在求解过程中加入正则化因子进行计算, 此时输出层权重计算表示为

$$\boldsymbol{\beta} = \begin{cases} \left(\mathbf{G}^T \mathbf{G} + \frac{1}{\lambda}\right)^{-1} \mathbf{G}^T \mathbf{T}, & \text{如果 } N \geq D \\ \mathbf{G}^T \left(\mathbf{G} \mathbf{G}^T + \frac{1}{\lambda}\right)^{-1} \mathbf{T}, & \text{如果 } N < D \end{cases} \quad (11)$$

由 RVFL 训练过程可知, RVFL 不进行权重迭代调整, 从而获得了极快的训练速度. 已有实验结果表明, RVFL 具有良好的泛化性能^[4], 万能逼近能力理论也证明了 RVFL 的收敛性^[12].

1.2 贝叶斯子空间选择策略

随机子空间方法是一种常用的有效降低数据集维度的策略, 能够减小数据训练复杂度, 提升模型测试精度. 随机子空间方法基于集成学习策略实现, 通过获取多个随机属性子空间, 用于多个基学习器训练并对预测进行融合, 从而得到更稳定、更精确的最终模型.

贝叶斯子空间选择(bayesian attribute bagging, BAB)策略, 是文献[13]在贝叶斯法则的基础上所提出的一种有效的子空间选择策略. 基于贝叶斯法则计算条件属性与决策属性之间的属性决策度(decision-making degree, DMD); 根据最大 DMD 来生成最优的条件属性组(condition attribute group, CAG); 对得到的 CAG 按一定比例进行抽样, 生成多个 Bagging 属性组(bagging attribute group, BAG), 此时 BAG 即为得到的目标属性子空间.

不同于传统的属性 Bagging 策略对条件属性进行随机抽样, BAB 充分考虑条件属性与决策属性之间的影响程度, 得到的属性子空间更均衡、更稳定, 在后续的基学习器训练中会取得更佳的性能.

2 加权子空间的随机向量函数链接网络集成方法

BAB 策略能够有效获得更均衡、更稳定的属性子空间, 但由于 RVFL 网络结构具有输入层和输出层直连的特点, 因此在训练过程中, 输入属性具有线性映射和非线性映射两部分性质. 若简单将 BAB 和 RVFL 进行结合, 则会忽略掉该线性映射部分的特性. 为了更好结合 BAB 策略与 RVFL 的网络结构特点, 本文在 BAB 的基础上提出了加权子空间选择(weighted attribute bagging, WAB)策略, 用以构造加权子空间的随机向量函数链接网络集成模型(weighted attribute bagging-based RVFL, WAB-RVFL).

2.1 加权子空间选择策略

2.1.1 属性加权矩阵

RVFL 的网络结构具有输入层到输出层、隐含层到输出层两部分目标输出权重, 因此其同时具有线性映射和非线性映射的特性. 对于 BAB 所得到的属性子空间, 并不完全适合 RVFL 模型训练, 为更贴合这两部分特性, 应该采取有效的策略进行属性子空间的优化转换. 对适合线性映射部分的属性进行加权放大增强, 对不适合线性映射部分的属性进行加权减小抑制, 从而实现更优模型的训练.

为了得到用于集成模型训练的更有效的属性子空间, 本文引入属性优化思想, 提出属性加权矩阵的概念, 用于属性子空间的加权转化, 其形式为

$$\mathbf{Q} = \begin{bmatrix} q_1 & 0 & \cdots & 0 \\ 0 & q_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & q_D \end{bmatrix} \quad (12)$$

其中 D 为子空间属性个数. 对于子空间输入样本 $\mathbf{X}$, 利用属性加权矩阵对其进行属性加权, 转化后的加权子空间可以表示为:

$$\mathbf{X} \cdot \mathbf{Q} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1D} \\ x_{21} & x_{22} & \cdots & x_{2D} \\ \vdots & \vdots & & \vdots \\ x_{N1} & x_{N2} & \cdots & x_{ND} \end{bmatrix} \begin{bmatrix} q_1 & 0 & \cdots & 0 \\ 0 & q_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & q_D \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{N1} \end{bmatrix} & \begin{bmatrix} x_{12} \\ x_{22} \\ \vdots \\ x_{N2} \end{bmatrix} & \cdots & \begin{bmatrix} x_{1D} \\ x_{2D} \\ \vdots \\ x_{ND} \end{bmatrix} \end{bmatrix} \quad (13)$$

属性加权矩阵是一个对角阵, 其对角线上的元素为子空间输入样本各个属性的加权. 在属性加权矩阵的作用下, 针对于 RVFL 特有的两部分结构, 若某些属性适合线性映射, 则其属性加权应该增大; 若某些属性适合非线性映射, 则其属性加权应该减小. 如何找到合适的属性加权矩阵是加权子空间选择策略的核心.

2.1.2 目标函数设计

在机器学习任务中, 目标函数是一个用来衡量模型拟合程度的函数, 通常如果函数值越小则表示模型的拟合效果越好. 如何有效最小化目标函数非常关键.

在属性加权矩阵的搜索过程中, 本文利用训练集和验证集信息来构建目标函数, 具体形式为:

$$E = [g(\tilde{\mathbf{X}}\mathbf{Q}\mathbf{W}), \tilde{\mathbf{X}}\mathbf{Q}][g(\mathbf{X}\mathbf{Q}\mathbf{W}), \mathbf{X}\mathbf{Q}]^+ \mathbf{Y} - \tilde{\mathbf{Y}} \quad (14)$$

其中: $\mathbf{X}$ 和 $\mathbf{Y}$ 是训练集样本的输入值和真实值, $\tilde{\mathbf{X}}$ 和 $\tilde{\mathbf{Y}}$ 是验证集样本的输入值和真实值, $g(\cdot)$ 表示激活函数. 为方便展示, (14) 式省略了节点偏置, 在实际计算中并不省略. 为更好进行目标函数优化, 本文使用 softmax 函数计算各类标的概率值并采用均方误差(mean squared error, MSE)作为衡量标准.

为搜索最优的属性加权矩阵, 本文选取粒子群优化(particle swarm optimization, PSO)算法^[14]进行迭代优化, 利用搜索得到属性加权矩阵 $\mathbf{Q}$, 使其对子空间的属性进行加权转化, 起到增强或抑制的作用, 从而获得加权后更优质、更适合 RVFL 网络结构的属性子空间.

2.2 WAB-RVFL 算法流程

WAB-RVFL 算法的核心是利用 PSO 迭代优化搜索得到最优的属性加权矩阵, 为更好地展示 WAB-

RVFL 算法，具体的流程如算法 1 所示.

算法 1: WAB-RVFL 算法

1 基于贝叶斯子空间选择(BAB)策略，得到 K 个属性子空间 BAGs;

2 基于加权子空间选择(WAB)策略，对 K 个 BAGs 进行加权优化:

3 for $k=1$ to K do

4 基于属性子空间 BAG_k ，进行 PSO 迭代更新搜索得到属性加权矩阵 Q_k ;

5 利用属性加权矩阵 Q_k ，对 BAG_k 进行加权得到优化后的 $\widetilde{BAG}_k$.

6 end for

7 基于优化后的加权子空间，进行 RVFL 集成训练和预测:

8 基于 K 个 $\widetilde{BAG}_k$ ，构建 K 个 RVFL 模型;

9 for $k=1$ to K do

10 利用训练好的 $RVFL_k$ ，对新样本进行预测得到 $y^{(k)}$.

11 end for

12 对于 K 个 RVFL 的预测 $y^{(1)}, y^{(2)}, \dots, y^{(K)}$ ，计算信息量权重 $v_1, v_2, \dots, v_K$ 并进行加权融合输出^[13]:

$$y = \sum_{k=1}^K v_k y^{(k)}$$

3 实验验证与分析

本次实验在 8 个高维分类数据集中开展，数据集信息如表 1 所示^[15]. 本文对数据集进行清洗，并为减少训练时间采取随机抽样方式得到小规模数据集，用于模型的训练和测试. 数据被划分成了 80% 训练集和 20% 测试集，并采用最大最小归一化方法，将属性区间缩放到[0, 1]. 本实验采取准确率(Accuracy)作为模型的衡量标准，且所有实验均是重复 5 次试验的平均结果.

PSO 算法的实验参数设置如下：种群迭代次数为 200，种群大小为 50，个体值区间为[0, 2]，速度值区间为[-0.2, 0.2]，惯性因子为 0.729 8，学习因子均为 1.494 45. 将训练集以 0.5 的比例划分得到验证集用于目标函数的迭代优化. 通过 10 次重复试验遍历 RVFL 得到最优精度对应的节点个数. 在 WAB-RVFL 实验中将 BAB 策略的特征抽样比例参数设置为 0.8.

表 1 数据集信息

数据集	来源	属性/个	样本/个	类标/个
Arrhythmia	UCI	279	420	12
* LVR	UCI	309	126	2
Parkinson's Disease	UCI	754	756	2
P53(2% Sampling)	UCI	5 408	302	2
Gisette(10% Sampling)	* SFSR	5 000	700	2
Urban Land Cover	UCI	147	675	9
Musk(version 1)	UCI	168	476	2
Musk(version 2)(10% Sampling)	UCI	168	660	2

注：* LVR 为 LSVT Voice Rehabilitation 的简写；* SFSR 为 Scikit-Feature Selection Repository 的简写.

3.1 WAB-RVFL 可行性验证

在 WAB-RVFL 中，如何有效最小化该目标函数非常关键. 本文采取 PSO 优化算法对目标函数进行迭代最小化，从而实现最优属性加权矩阵的搜索，优化属性子空间而进一步提升模型性能. 本文选取了 4 个数据集进行迭代优化展示，具体如图 2 所示.

从图 2 可以看出，随着迭代次数的增加，目标函数值逐渐优化减小；在迭代次数为 150 之后，图像基本上趋于收敛状态. 这表明 WAB-RVFL 所构造的目标函数是有效的，其随着 PSO 的优化迭代而不断减小，验证了 WAB-RVFL 算法的可行性.

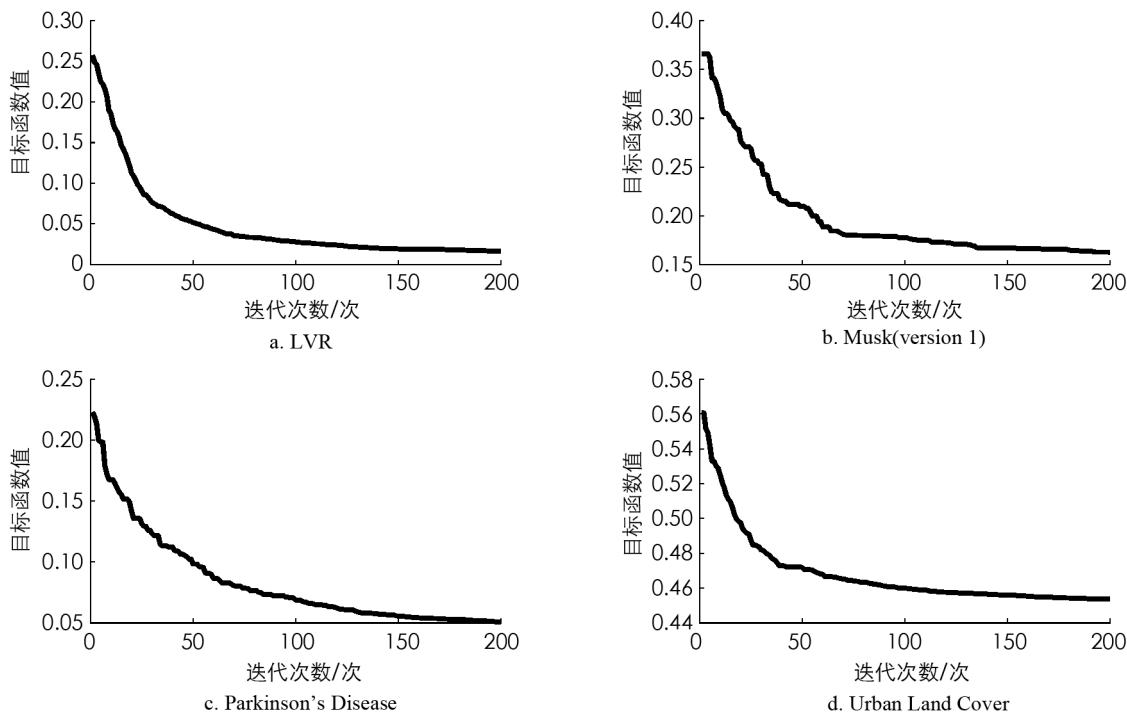


图 2 在 PSO 优化下目标函数值的迭代展示

3.2 WAB-RVFL 合理性验证

为了验证 WAB-RVFL 的合理性, 本节对 3 种基模型进行了比较, 即 Single RVFL 模型、Single BAB-RVFL 模型和 Single WAB-RVFL 模型, 其分别是基于投票策略的 RVFL(Voting-RVFL)、基于贝叶斯子空间的 RVFL(BAB-RVFL)和基于加权子空间的 RVFL(WAB-RVFL)3 种集成模型的基学习器. 具体的实验结果如图 3 所示.

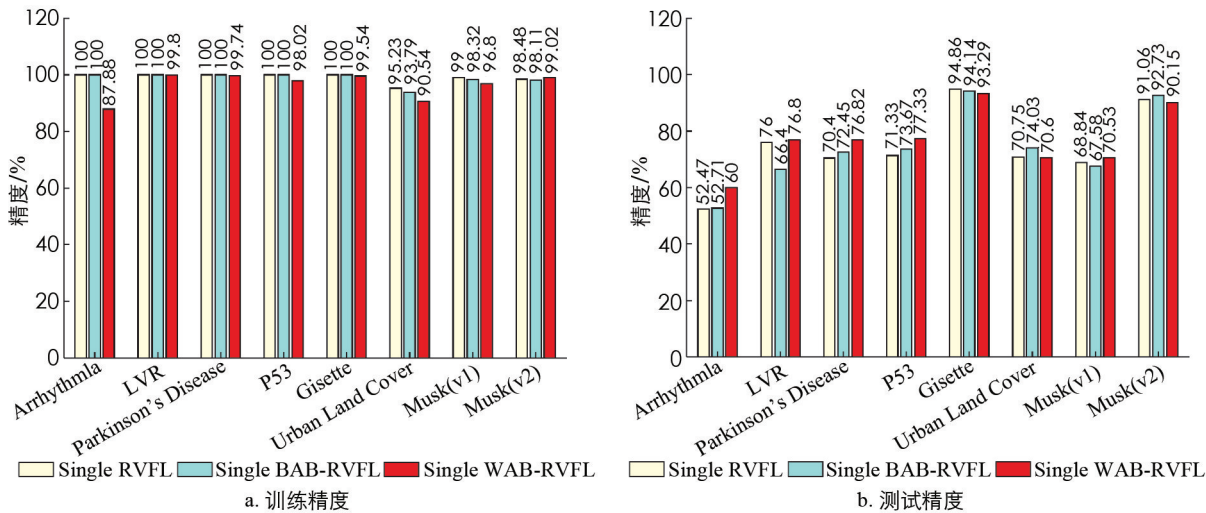


图 3 3 种基模型在 8 个分类数据集下的比较

由图 3 可知, 在训练部分, Single RVFL 和 Single BAB-RVFL 在 5 个数据集上的精度达到 100%, Single WAB-RVFL 的训练精度普遍较低, 只有在 Musk(version 2)数据集中表现比其他两个模型高; 在测试部分, Single WAB-RVFL 则在多个数据集下普遍表现比其他两个基模型的要好, 即精度更高. 在 Musk(version 2), Urban Land Cover 和 Gisette 3 个数据集中, 虽然 Single WAB-RVFL 测试精度不是最高, 但与 Single RVFL 相比较为相近.

总结分析得, WAB-RVFL 的基模型 Single WAB-RVFL 经过属性加权矩阵的优化之后, 在 8 个数据集

的测试集上有 5 个表现要比 Single RVFL 和 Single BAB-RVFL 好, 其中 3 个数据集上表现均与 Single RVFL 相差不大. 这说明本文提出的属性加权矩阵和 WAB-RVFL 模型具有合理性, 属性加权矩阵能够较好地子空间进行加权优化, 从而提升模型的测试精度. 这为下一小节中验证最终集成模型的性能提供了合理性依据.

3.3 WAB-RVFL 有效性验证

在上一小节中的 WAB-RVFL 合理性实验, 分析得到 Single WAB-RVFL 在大多数数据集上比 Single RVFL 和 Single BAB-RVFL 性能更佳. 本小节将通过丰富的模型对比实验, 验证 WAB-RVFL 的有效性.

本文搭建实验平台, 将 WAB-RVFL 与以下 6 个模型: RVFL, Voting-RVFL, Bagging-RVFL, dRVFL, edRVFL 以及 BAB-RVFL 进行比较, 其中 dRVFL 和 edRVFL 的隐含层数设置为 10, 具体的实验结果如图 4 所示.

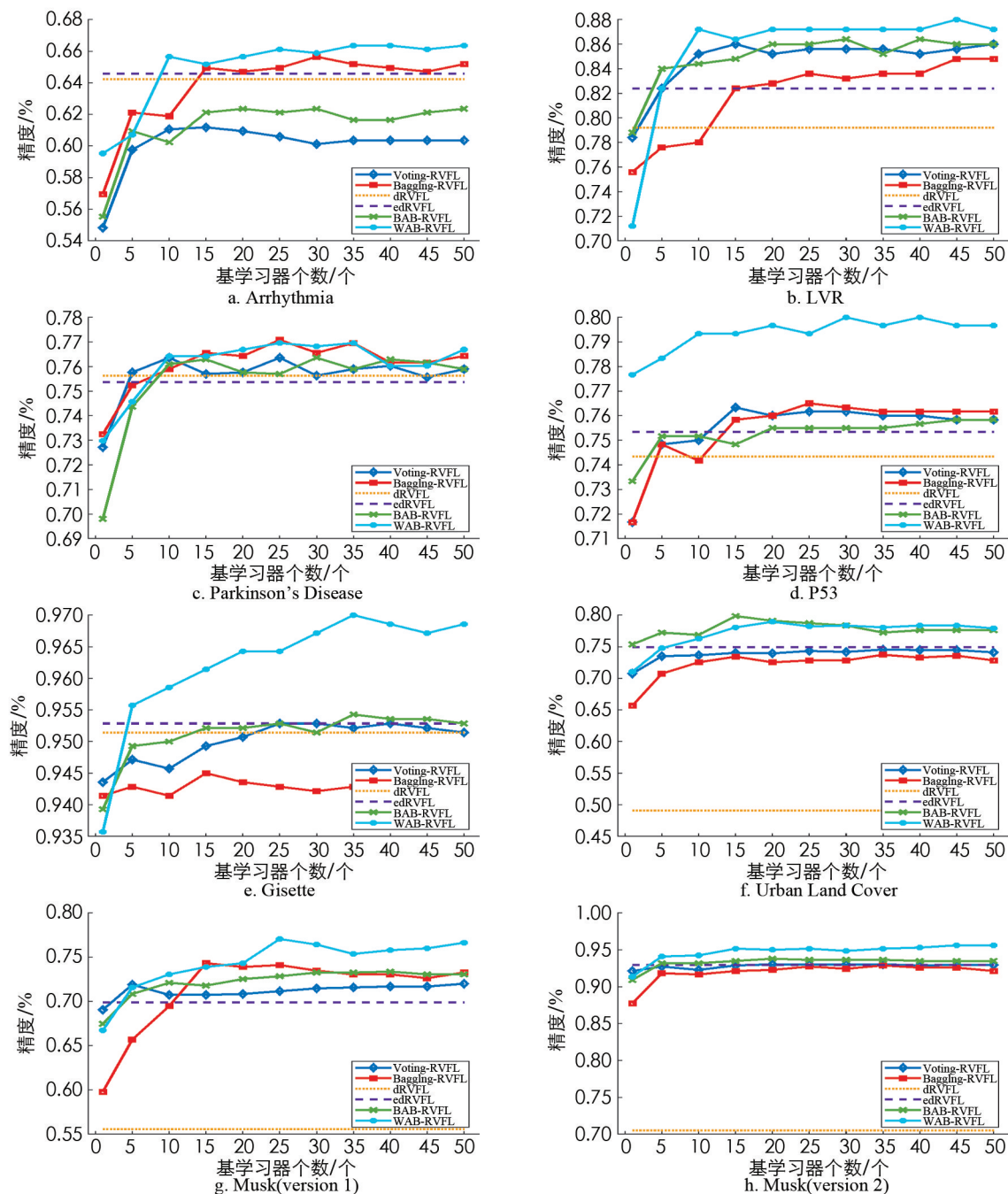


图 4 在分类任务中 WAB-RVFL 与其他 RVFL 模型的比较

由图 4 可知, 随着基学习器的增加, 所有 RVFL 集成算法均趋于收敛, 表明 RVFL 结合集成方法的应用是可行的, 模型精度能够不断提升并趋于稳定状态. 本文提出的 WAB-RVFL 在 8 个高维分类数据集下, 与其他 RVFL 模型相比, 均能取得更高的预测精度. 实验结果表明, 加权子空间选择策略能有效缓解模型过拟合现象, WAB-RVFL 具有更佳的泛化性能, 验证了 WAB-RVFL 的有效性.

本文还使用 Nemenyi 检验方法^[16]来进行多个模型之间的统计检验分析, 显著性水平设置为 0.1, CD 值表示为临界差值区间, 固定基学习器为 50, 取测试集精度进行模型比较. 由图 5 可知, 在测试过程中, WAB-RVFL 和 BAB-RVFL 共处同一临界差值区间, 且 WAB-RVFL 的平均秩值更小, 其他模型则分别与 WAB-RVFL 处于不同的临界差值区间, 这表明 WAB-RVFL 的预测性能优于 BAB-RVFL; 与其他 5 种模型的差异较大, 且均优于这 5 种模型. 因此, 本文提出的 WAB-RVFL 具有更稳定且更优的预测性能.

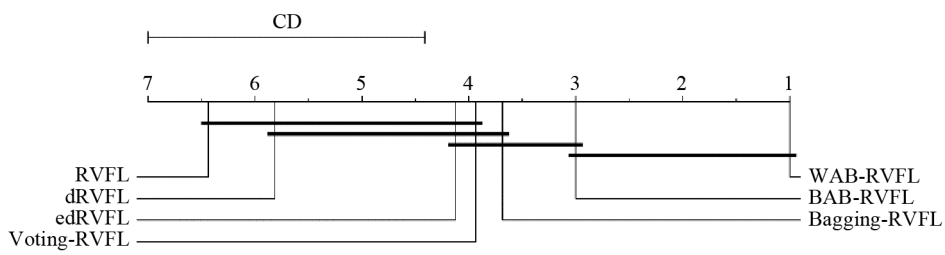


图 5 模型测试精度的统计检验结果

4 结论与未来工作

本文基于 RVFL 输入层和输出层直连的结构特点, 提出了一种有效的、鲁棒性更强的加权子空间的随机向量函数链接网络集成方法 (weighted attribute bagging-based random vector functional link network, WAB-RVFL). 首先, 基于贝叶斯子空间选择策略, 对 RVFL 结构进行分析并提出合适的加权子空间选择策略, 引入属性加权矩阵概念, 构造有效的目标函数并利用 PSO 算法迭代优化; 其次, 设置 3 个实验验证了 WAB-RVFL 的可行性、合理性和有效性. 实验结果表明, PSO 对目标函数迭代优化收敛证明了 WAB-RVFL 的可行性; 在多个数据集下, 3 种基学习器的对比实验中 Single WAB-RVFL 预测性能更好, 体现了 WAB-RVFL 的合理性; 在 7 种 RVFL 模型的对比实验中, WAB-RVFL 具有更高测试精度, 验证了 WAB-RVFL 的有效性.

- 在本文中仍有部分内容值得进一步地探究, 未来的工作展望主要有以下两点:
- 1) 其他智能算法的使用. 针对 WAB-RVFL 模型中使用到的经典 PSO 优化算法, 探究是否能使用其他更优的智能算法进行模型的迭代优化, 减少实验的开销, 提升模型的性能;
 - 2) 大规模数据的处理. WAB-RVFL 模型具有并行运算的潜力, 而目前在大数据时代下对模型精度以及运算效率要求都特别高, 因此接下来将考虑是否能将本文提出的两种模型, 在大数据框架下进行分布式计算, 处理真实的大数据问题.

参考文献:

[1] CAO W P. A Review on Neural Networks with Random Weights [J]. Neurocomputing, 2018, 275: 278-287.

[2] SCHMIDT W F, KRAAIJVELD M A, DUIN R P W. Feedforward Neural Networks with Random Weights [C]//11th IAPR International Conference on Pattern Recognition. Vol. II. Conference B: Pattern Recognition Methodology and Systems. New York: IEEE Press, 1992: 1-4.

[3] PAO Y H. Learning and Generalization Characteristics of the Random Vector Functional-Link Net [J]. Neurocomputing, 1994, 6(2): 163-180.

[4] ZHANG L. A Comprehensive Evaluation of Random Vector Functional Link Networks [J]. Information Sciences, 2016,

367-368; 1094-1105.

- [5] ZHOU P. Data Modeling for Quality Prediction Using Improved Orthogonal Incremental Random Vector Functional-Link Networks [J]. *Neurocomputing*, 2019, 365: 1-9.
- [6] SCARDAPANE S, WANG D H, UNCINI A. Bayesian Random Vector Functional-Link Networks for Robust Data Modeling [J]. *IEEE Transactions on Cybernetics*, 2018, 48(7): 2049-2059.
- [7] KATUWAL R. Stacked Autoencoder Based Deep Random Vector Functional Link Neural Network for Classification [J]. *Applied Soft Computing*, 2019, 85: 105854.
- [8] SHI Q S. Random Vector Functional Link Neural Network Based Ensemble Deep Learning [J]. *Pattern Recognition*, 2021, 117: 107978.
- [9] SHARMA R. FAF-DRVFL: Fuzzy Activation Function Based Deep Random Vector Functional Links Network for Early Diagnosis of Alzheimer Disease [J]. *Applied Soft Computing*, 2021, 106: 107371.
- [10] HU M H, SHI Q S, SUGANTHAN P N, et al. Adaptive Ensemble Variants of Random Vector Functional Link Networks [C]//The 27th International Conference on Neural Information Processing. New York: IEEE Press, 2020: 30-37.
- [11] CHENG W X. Time Series Classification Using Diversified Ensemble Deep Random Vector Functional Link and Resnet Features [J]. *Applied Soft Computing*, 2021, 112: 107826.
- [12] IGELNIK B, PAO Y H. Stochastic Choice of Basis Functions in Adaptive Function Approximation and the Functional-Link Net [J]. *IEEE Transactions on Neural Networks*, 1995, 6(6): 1320-1329.
- [13] HE Y L, YE X, HUANG J Z, et al. Bayesian Attribute Bagging-Based Extreme Learning Machine for High-Dimensional Classification and Regression [J]. *ACM Transactions on Intelligent Systems and Technology*, 2022, 13(2): 1-26.
- [14] KENNEDY J, EBERHART R. Particle Swarm Optimization [C]//Proceedings of ICNN'95-International Conference on Neural Networks. New York: IEEE Press, 1995: 1942-1948.
- [15] DUA D, GRAFF C. UCI Machine Learning Repository [EB/OL]. (2019-04-03)[2021-09-05]. <http://archive.ics.uci.edu/ml>.
- [16] DEMŠAR J. Statistical Comparisons of Classifiers over Multiple Data Sets [J]. *J Mach Learn Res*, 2006, 7: 1-30.

责任编辑 张枸