

控制与决策

Control and Decision

基于Frobenius和 $L_{2,1}$ 范数的多输出宽度学习系统

褚菲, 卢新宇, 苏嘉铭, 王雪松, 马小平

引用本文:

褚菲, 卢新宇, 苏嘉铭, 王雪松, 马小平. 基于Frobenius和 $L_{2,1}$ 范数的多输出宽度学习系统[J]. 控制与决策, 2023, 38(10): 2919–2924.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2021.1782>

您可能感兴趣的其他文章

Articles you may be interested in

基于数据分布特性的代价敏感宽度学习系统

Data distribution-based cost-sensitive broad learning system

控制与决策. 2021, 36(7): 1686–1692 <https://doi.org/10.13195/j.kzyjc.2019.1484>

一种基于池计算的宽度学习系统

A broad learning system based on reservoir computing

控制与决策. 2021, 36(9): 2203–2210 <https://doi.org/10.13195/j.kzyjc.2019.1729>

基于稀疏化神经网络的浮选泡沫图像特征选择

Selection method for froth image characters based on sparse neural network

控制与决策. 2021, 36(7): 1627–1636 <https://doi.org/10.13195/j.kzyjc.2019.1788>

行人重识别中度量学习方法研究进展

A survey on metric learning in person re-identification

控制与决策. 2021, 36(7): 1547–1557 <https://doi.org/10.13195/j.kzyjc.2020.0801>

输入约束不确定系统的点对点迭代学习控制与优化

Point-to-point iterative learning control and optimization for uncertain systems with constrained input

控制与决策. 2021, 36(6): 1435–1441 <https://doi.org/10.13195/j.kzyjc.2019.0908>

基于 Frobenius 和 $L_{2,1}$ 范数的多输出宽度学习系统

褚菲^{1,2†}, 卢新宇², 苏嘉铭², 王雪松², 马小平²

(1. 中国矿业大学人工智能研究院, 江苏徐州 221116;

2. 中国矿业大学信息与控制工程学院, 江苏徐州 221116)

摘要: 宽度学习系统 (broad learning system, BLS) 因其特征提取能力强、计算效率高而被广泛应用于众多领域. 然而, 目前 BLS 主要用于单输出回归, 当 BLS 存在多个输出时, BLS 无法有效发掘多个输出权重之间的相关性, 会导致模型预测性能的下降. 鉴于此, 通过 Frobenius 和 $L_{2,1}$ 矩阵范数的联合约束, 提出多输出宽度学习系统 (multi-output broad learning system, MOBLS). 首先, 在原有 BLS 的基础上构建新的目标函数, 将 L_2 损失函数替换为 $L_{2,1}$ 形式, L_2 正则化项替换为 Frobenius 和 $L_{2,1}$ 两项; 然后, 利用交替方向乘子法 (alternating direction method of multipliers, ADMM) 对新目标函数 BLS 的输出权重优化求解. 利用 11 个公共数据集和 1 个实际过程数据集验证了所提系统的有效性.

关键词: 宽度学习系统; 多输出回归; Frobenius 范数; $L_{2,1}$ 范数

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2021.1782

引用格式: 褚菲, 卢新宇, 苏嘉铭, 等. 基于 Frobenius 和 $L_{2,1}$ 范数的多输出宽度学习系统 [J]. 控制与决策, 2023, 38(10): 2919-2924.

Multi-output broad learning system based on Frobenius and $L_{2,1}$ norm

CHU Fei^{1,2†}, LU Xin-yu², SU Jia-ming², WANG Xue-song², MA Xiao-ping²

(1. Artificial Intelligence Research Institute, China University of Mining and Technology, Xuzhou 221116, China;
2. School of Information and Control Engineering, China University of Mining and Technology, Xuzhou 221116, China)

Abstract: Broad learning systems (BLS) have been widely used in many fields because of its strong feature extraction ability and high computational efficiency. However, the BLS is mainly used for single-output regression at present, When the BLS has multi-outputs, it cannot effectively explore the correlation between multi-output weights, which will lead to the degradation of model prediction performance. Therefore, by introducing the Frobenius and $L_{2,1}$ matrix norm, this paper proposes a multi-output broad learning system (MOBLS). Firstly, a new objective function is constructed on the basis of the original BLS. The L_2 -loss function is replaced by the $L_{2,1}$ form, and the L_2 regularization term is replaced by the Frobenius and $L_{2,1}$ two terms. Then, alternating direction method of multipliers (ADMM) is used to optimize the output weight of the BLS with the new objective function. Finally, 11 public datasets and 1 actual process dataset are used to verify its effectiveness.

Keywords: broad learning system; multi-output regression; Frobenius norm; $L_{2,1}$ norm

0 引言

深度学习网络^[1-4]作为主流的机器学习方法在人脸识别、大规模数据处理、语义分割等领域应用广泛. 然而, 深度学习网络层数较多, 超参数的调整以及权重的迭代计算大大增加了网络的训练时间, 当新节点添加至网络后, 无法高效快速建模.

宽度学习系统 (broad learning system, BLS)^[5]起

源于随机向量函数链接神经网络^[6-8], 首先经特征层提取特征, 再经过增强层进行非线性转换, 最后共同连结到输出层, 通过伪逆计算网络的输出权值, 以保证 BLS 建模的效率. 近年来, 各个领域的众多学者为拓宽 BLS 的适用范围作出了重要贡献. 文献 [9-10] 在时序预测问题上对 BLS 进行改进; 文献 [11] 提出了将深度与宽度相结合的堆叠宽度学习系统, 并提出了

收稿日期: 2021-10-16; 录用日期: 2022-06-15.

基金项目: 国家自然科学基金项目 (61973304, 61873049, 62073060); 江苏省自然科学基金项目 (BK20191339); 江苏省第十六届“六大人才高峰”高层次人才选拔培养项目 (DZXX-045); 中国矿业大学研究生创新计划项目 (2022WLJCRCZL109); 中央高校基本科研业务费专项资金项目 (2019XKQYMS64).

责任编辑: 夏元清.

†通讯作者. E-mail: chufeizhufei@sina.com.

具有动态结构的堆叠宽度学习系统增量算法,该网络结构继承了宽度学习系统的快速性与准确率;文献[12]提出了一种图卷积宽度学习网络,利用BLS增强图卷积层特征,以提高模型的鲁棒性;文献[13]利用半监督自编码极限学习机获取映射特征,最后得到一个线性判别BLS模型,以增强BLS提取未标记数据特征的能力;文献[14-15]针对标签数据较少的问题,分别提出了半监督以及无监督BLS,用于进行分类任务;文献[16]针对BLS增量学习问题,提出一种具有语义特征提取层的BLS,以提高其快速建模能力;文献[17]针对噪声问题提出一种概率鲁棒BLS,用以处理带标签噪声的分类问题;文献[18]针对样本相关性以及标签错误分类问题,提出一种新颖的判别群稀疏宽度学习系统,增强了BLS在视觉识别方面的分类能力. 尽管BLS在很多领域取得了不错的进展,但针对多输出回归问题,目前还没有研究讨论.

标准BLS的损失函数与约束项均是 L_2 范数形式,BLS输出权重通常是一列矩阵,如果BLS有多个输出节点,则BLS的输出权重是多维行列矩阵,标准BLS的目标函数无法有效处理多个输出权重的相互关系. 因此,利用Frobenius和 $L_{2,1}$ 范数的联合约束,改写标准BLS的目标函数,此种联合约束方法类似于elastic net理论,将原目标函数改写为矩阵范数约束形式,以此来处理多个输出权重的相互关系. 通过Frobenius和 $L_{2,1}$ 范数的联合正则化,将BLS输入层至输出层的输出权重相互联合约束,探索不同输出权重之间的相关性,更加准确地预测多个输出,同时获得一个较为紧凑的网络结构. 此外,现实世界收集的数据具有不确定性,通常伴随采样误差、测量误差等情况,这会大大影响模型的建模能力. 因此,MOBLS利用 $L_{2,1}$ 范数的强鲁棒性建模,以降低离群值的影响. 针对改写后的目标函数,本文利用交替方向乘子法(alternating direction method of multipliers, ADMM)^[19-20]对新模型的输出权重进行优化求解.

1 宽度学习系统

BLS^[5]作为一种新型的随机权值神经网络,具体构建过程描述如下.

假设 $X \in R^{S \times M}$ 为数据集的输入, $Y \in R^{S \times N}$ 为标签输出值. 其中: S 表示训练样本的数量, M 和 N 分别表示输入和输出数据的维度. 对于输入数据 X ,利用线性变换函数映射生成第 i 组映射特征

$$Z_i = \phi(XW_{e_i} + \beta_{e_i}). \quad (1)$$

其中: W_{e_i} 和 β_{e_i} 是由网络生成的随机权重, n 是特征

组数($i = 1, 2, \dots, n$). 每组的特征节点数为 k , $\phi(\cdot)$ 表示线性变换函数. 那么,将所有特征节点进行组合,记为

$$Z^n \equiv [Z_1, Z_2, \dots, Z_n]. \quad (2)$$

根据式(2)生成所有特征节点组 Z^n ,对其进行非线性函数变化,生成增强节点

$$H_j = \zeta_j(Z^n W_{h_j} + \beta_{h_j}). \quad (3)$$

其中: W_{h_j} 和 β_{h_j} 由网络随机生成, m 为增强组数($j = 1, 2, \dots, m$), $\zeta(\cdot)$ 用于非线性函数变换. 那么,所有增强节点可以表示为

$$H^m \equiv [H_1, H_2, \dots, H_m]. \quad (4)$$

根据式(2)和(4)可以得到输入层

$$A = [Z^n | H^m]. \quad (5)$$

因此,宽度网络的实际输出可以表示为

$$\hat{Y} = AW, \quad (6)$$

其中 W 是宽度网络的输出权重.

输出权重 W 可通过如下最小化目标公式得到:

$$\arg \min_W \|Y - AW\|_2^2 + \lambda \|W\|_2^2. \quad (7)$$

根据式(7),得到输出权重

$$W = (\lambda I + A^T A)^{-1} A^T Y. \quad (8)$$

其中: I 是单位矩阵, λ 是正则化参数.

2 多输出宽度学习系统

在本节中,将介绍本文所提出的多输出宽度学习系统,并利用交替方向乘子法以及增广拉格朗日乘子法给出其优化求解过程^[21-22].

探索各个输出目标之间的相关性,提高模型的预测精度^[23]是解决多输出回归问题的关键. 在回归任务中,标准BLS的目标函数利用了 L_2 范数,且通常只有一个输出节点,因此BLS只考虑了单个输出. 当BLS存在多个输出节点时,BLS的输出权重不再是单独一列矩阵,而是由多行列相互关联的矩阵,此时BLS具有多个输出. 传统的目标函数无法考虑多个输出之间的关联性,因此通过Frobenius和 $L_{2,1}$ 矩阵范数的联合约束,发掘多个输出之间的关联性,建立精确的数学模型. 下面给出这两种矩阵范数的定义:

$$\|A\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^2} = \sqrt{\text{tr}(A^T A)}, \quad (9)$$

$$\|A\|_{2,1} = \sum_{i=1}^n \|a_{i,\cdot}\|_2 = \sum_{i=1}^n \left(\sum_{j=1}^m |a_{ij}|^2 \right)^{1/2}. \quad (10)$$

其中: a 表示 A 的元素, $\text{tr}(A^T A)$ 表示 $A^T A$ 的迹.

改进后的目标函数为

$$\arg \min_W \|Y - AW\|_{2,1} + \lambda_1 \|W\|_{2,1} + \lambda_2 \|W\|_F^2. \quad (11)$$

利用增广拉格朗日乘子法替换前两项约束,并写出相应的约束条件,有

$$\begin{aligned} & \arg \min_{D,E,W} \|D\|_{2,1} + \lambda_1 \|E\|_{2,1} + \lambda_2 \|W\|_F^2; \\ & \text{s.t. } D = Y - AW, \\ & E = W. \end{aligned} \quad (12)$$

设置迭代次数,利用ADMM进行梯度优化,对 E 、 B 、 D 的约束项执行迭代交替的最小化.式(12)的增广拉格朗日函数为

$$\begin{aligned} \mathcal{L}(D, E, W, U_1, U_2) = & \|D\|_{2,1} + \lambda_1 \|E\|_{2,1} + \lambda_2 \|W\|_F^2 + \\ & \frac{\mu}{2} \|Y - AW - D + U_1\|_F^2 + \\ & \frac{\mu}{2} \|E - W + U_2\|_F^2 - \frac{\mu}{2} \|U_1\|_F^2 - \frac{\mu}{2} \|U_2\|_F^2. \end{aligned} \quad (13)$$

其中: μ 表示惩罚因子, U_1 和 U_2 表示拉格朗日乘子.

在每次迭代中,ADMM由以下求解步骤组成.

1) W : 可以得到以下子优化问题:

$$\arg \min_W \lambda_2 \|W\|_F^2 + \frac{\mu}{2} \|Y - AW - D + U_1\|_F^2 + \frac{\mu}{2} \|E - W + U_2\|_F^2. \quad (14)$$

令式(14)对 W 求导等于0,可以得到

$$\begin{aligned} & \left[A^T A + \left(1 + \frac{2\lambda_2}{\mu} \right) I \right] W - \\ & A^T (Y - D + U_1) - (U_2 + E) = 0. \end{aligned} \quad (15)$$

因此,得到式(14)的解为

$$\begin{aligned} W = & \left[A^T A + \left(1 + \frac{2\lambda_2}{\mu} \right) I \right]^{-1} [A^T (Y - \\ & D + U_1) + U_2 + E], \end{aligned} \quad (16)$$

其中 I 表示单位矩阵.

2) D : 子目标优化函数如下:

$$\arg \min_D \frac{1}{\mu} \|D\|_{2,1} + \frac{1}{2} \|Y - AW - D + U_1\|_F^2. \quad (17)$$

利用vect-软阈值方法^[24]得到式(17)的解为

$$D = \text{vect} - \text{soft} \left(Y - AW + U_1, \frac{1}{\mu} \right), \quad (18)$$

其中

$$\text{vect} - \text{soft}(b, \tau) = b \left(\frac{\max\{\|b\|_2 - \tau, 0\}}{\max\{\|b\|_2 - \tau, 0\} + \tau} \right).$$

3) E : 子目标优化函数如下:

$$\arg \min_W \frac{\lambda_1}{\mu} \|E\|_{2,1} + \frac{1}{2} \|W - E - U_2\|_F^2, \quad (19)$$

其结构与式(17)相同.因此,可以得到解为

$$E = \text{vect} - \text{soft} \left(W - U_2, \frac{\lambda_1}{\mu} \right). \quad (20)$$

4) U_1, U_2 : 在每次迭代中更新两个参数

$$\begin{cases} U_1' = U_1 + (Y - AW - D), \\ U_2' = U_2 + (E - W). \end{cases} \quad (21)$$

3 实验与讨论

本节选取1个实际工业数据集和Mulan Java库中的11个公共数据集^[25]进行实验,训练集与测试集按照2:1的方式划分.表1给出了数据集的具体信息.实际工业数据集采用多级离心压缩机收集到的数据^[26].另外,向数据集的训练输出中加入不同比例的异常值($\alpha = 0\%, 20\%, 40\%$, 异常值的数量/样本的数量).

表1 数据集信息

数据集	训练样本	测试样本	特征	输出
andro	33	16	30	6
atp1d	225	112	411	6
atp7d	197	99	411	6
edm	103	51	16	2
enb	512	256	8	2
jura	239	120	15	3
oes10	269	134	298	16
oes97	223	111	263	16
scm20d	5977	2989	61	16
slump	69	34	7	3
wq	707	353	16	14
compressor	289	145	4	2

首先需要定义BLS和MOBLS的网络结构:特征节点数 $N_g = 20 \times 40$,增强节点数 $N_e = 200$.正则化参数通过网络搜索法获得,搜索范围为 $\{2^{-5}, 2^{-4}, 2^{-3}, \dots, 2^4, 2^5\}$.

考虑到多个输出之间的关联性,本节采用以下指标评价模型性能:

平均均方根误差(average root mean squared error, aRMSE)

$$\text{aRMSE} = \frac{1}{m} \sum_{j=1}^m \sqrt{\frac{\sum_{i=1}^N (y_j^i - \hat{y}_j^i)^2}{N}}. \quad (22)$$

aRMSE值越小,模型性能越好.平均相对均方根误差(average relative root mean squared error, aRRMSE)

$$\text{aRRMSE} = \frac{1}{m} \sum_{j=1}^m \sqrt{\frac{\sum_{i=1}^N (y_j^i - \hat{y}_j^i)^2}{\sum_{i=1}^N (y_j^i - \bar{y}_j)^2}}. \quad (23)$$

aRRMSE值越小,模型性能越好.式中: m 表示输出

数, N 表示样本数, y_j^i 表示输出实际值, \hat{y}_j^i 表示模型预测值, \bar{y}_j 表示输出实际值的平均值.

本节采用上述参数在 $\alpha = 0\%$, 20% , 40% 的条件下通过网格搜索法进行模型训练. 另外, 针对宽度学习系统鲁棒性建模问题, 本文引入鲁棒宽度学习系统 (RBLS)^[27] 与 MOBLS 和 BLS 对比. 表 2 给出了 4 种算法在不同数据集上的 aRMSE. 在无异常值 ($\alpha = 0\%$) 的情况下, MOBLS 在 oes10、oes97、wq 数据集上的精度低于 BLS; 在 atp1d 数据集上, L1RBLS 与 MOBLS 持平, L1RBLS 与 ENRBLS 的预测精度相差不大; 在剩下的数据集上, MOBLS 精度高于其他 3 种模型. 在 $\alpha = 20\%$ 的情况下, 在 scm20d 数据集上, MOBLS 的预测精度与 ENBLS 持平, 低于 BLS; 在其余数据集上, L1RBLS 与 ENRBLS 的预测精度基本持平, 且优于 BLS, MOBLS 的预测精度最高. 在 $\alpha = 40\%$ 的情况下, MOBLS 的预测精度均高于其他 3 种模型, 且提升幅度较大; BLS 的预测精度最低, L1RBLS 与 ENRBLS 的预测精度均不如 MOBLS.

表 3 给出了 4 种算法在不同数据集上的 aRRMSE. 针对无离群点的情况, 在 compressor 数据集上, MOBLS、L1RBLS、ENRBLS 的 aRRMSE 值相近, 均高于 BLS; 在 scm20d 数据集上, MOBLS 表现不如 L1RBLS, 与 ENRBLS 持平; 在其他数据集上, 均有小幅的性能提升. 针对 $\alpha = 20\%$ 的情况, 整体上 MOBLS 的 aRRMSE 最低, L1RBLS 与 ENRBLS 的预测性能几乎持平, 均低于 BLS 的 aRRMSE, 预测性能提升幅度高于无离群值的情况. 针对 $\alpha = 40\%$ 的情况, 在 compressor 数据集上, MOBLS 的预测精度平均值提升程度最高; 在有噪声干扰的情况下, LRBLs、ENRBLS 的预测精度相较于 BLS 有一定程度的提升, 但仍低于 MOBLS. 因此, 相较于其他 3 种算法, MOBLS 的鲁棒性和泛化性都得到了很大的提升.

综合上述分析, 在无异常值情况下, MOBLS 在某些数据集上的性能不如 BLS, 但优于 L1RBLS 与 ENRBLS, 整体上略微提升了模型性能. 在有异常值的情况下, 随着数据中异常值的比例增加, BLS 模型性能下降较为明显, L1RBLS 与 ENRBLS 的性能下降程度低于 BLS, 而对 MOBLS 的性能影响最小, 这表明 MOBLS 对异常数据的建模能力最稳定. 综上所述, MOBLS 能够更有效地发掘多个输出之间的相关性, 提高模型的预测性能. 同时, 通过两种矩阵范数的联合约束, 能够获得更为紧凑的网络结构.

表 2 4 种算法在不同数据集上的 aRMSE

数据集	算法	α		
		0	20%	40%
andro	BLS	0.3547	0.4154	0.5067
	MOBLS	0.3266	0.3991	0.4934
	L1RBLS	0.3284	0.4012	0.4971
	ENRBLS	0.3266	0.3997	0.4955
atp1d	BLS	0.0807	0.1370	0.2200
	MOBLS	0.0761	0.1215	0.1511
	L1RBLS	0.0761	0.1284	0.1722
	ENRBLS	0.0762	0.1244	0.1512
atp7d	BLS	0.1664	0.1821	0.2690
	MOBLS	0.1581	0.1661	0.1930
	L1RBLS	0.1584	0.1763	0.2241
	ENRBLS	0.1585	0.1699	0.2114
edm	BLS	0.2132	0.2305	0.2977
	MOBLS	0.1851	0.2282	0.2779
	L1RBLS	0.1853	0.2286	0.2792
	ENRBLS	0.1853	0.2292	0.2788
enb	BLS	0.0752	0.1376	0.2339
	MOBLS	0.0738	0.0935	0.1498
	L1RBLS	0.0742	0.1107	0.1856
	ENRBLS	0.0741	0.1068	0.1796
jura	BLS	0.1041	0.1550	0.2325
	MOBLS	0.1041	0.1309	0.1417
	L1RBLS	0.1057	0.1447	0.1825
	ENRBLS	0.1052	0.1440	0.1757
oes10	BLS	0.0327	0.1044	0.1958
	MOBLS	0.0338	0.0554	0.0895
	L1RBLS	0.0340	0.0732	0.1457
	ENBLS	0.0339	0.0744	0.1445
oes97	BLS	0.0431	0.0970	0.1863
	MOBLS	0.0433	0.0602	0.0836
	L1RBLS	0.0445	0.0735	0.1322
	ENRBLS	0.0442	0.0721	0.1310
scm20d	BLS	0.1111	0.1676	0.2547
	MOBLS	0.1058	0.1708	0.2343
	L1RBLS	0.1088	0.1708	0.2447
	ENRBLS	0.1060	0.1733	0.2455
slump	BLS	0.1896	0.2099	0.2750
	MOBLS	0.1763	0.1992	0.2307
	L1RBLS	0.1789	0.2045	0.2512
	ENRBLS	0.1786	0.2036	0.2479
wq	BLS	0.2357	0.2671	0.3264
	MOBLS	0.2361	0.2441	0.2464
	L1RBLS	0.2366	0.2507	0.2955
	ENRBLS	0.2364	0.2513	0.2912
compressor	BLS	0.0037	0.1032	0.2232
	MOBLS	0.0029	0.0445	0.1107
	L1RBLS	0.0033	0.0672	0.1814
	ENRBLS	0.0029	0.0654	0.1715

表 3 4 种算法在不同数据集上的 aRRMSE

数据集	算法	α		
		0	20 %	40 %
andro	BLS	1.6120	1.8357	2.3679
	MOBLS	1.5145	1.7716	2.3034
	L1RBLS	1.5157	1.8125	2.3544
	ENRBLS	1.5164	1.7776	2.3064
atp1d	BLS	0.2657	0.3679	0.5092
	MOBLS	0.2482	0.3143	0.4348
	L1RBLS	0.2587	0.3335	0.4566
	ENRBLS	0.2486	0.3315	0.4512
atp7d	BLS	0.6026	0.5834	0.5872
	MOBLS	0.5935	0.5554	0.5464
	L1RBLS	0.5951	0.5752	0.5623
	ENRBLS	0.5954	0.5731	0.5601
edm	BLS	0.6192	0.7264	0.8985
	MOBLS	0.6029	0.7132	0.7260
	L1RBLS	0.6124	0.7196	0.7766
	ENRBLS	0.6107	0.7182	0.7754
enb	BLS	0.2697	0.3622	0.6321
	MOBLS	0.2502	0.2961	0.4029
	L1RBLS	0.2577	0.2987	0.4726
	ENRBLS	0.2546	0.2987	0.4623
jura	BLS	0.5254	0.6775	0.9438
	MOBLS	0.5201	0.6007	0.7152
	L1RBLS	0.5219	0.6273	0.8148
	ENRBLS	0.5218	0.6211	0.8121
oes10	BLS	0.2998	0.5199	0.6242
	MOBLS	0.3003	0.4851	0.5679
	L1RBLS	0.3007	0.5004	0.5891
	ENBLS	0.3012	0.5007	0.5844
oes97	BLS	0.2656	0.2840	0.4232
	MOBLS	0.2614	0.2834	0.3966
	L1RBLS	0.2634	0.2840	0.4122
	ENRBLS	0.2626	0.2844	0.4100
scm20d	BLS	0.1111	0.1676	0.2547
	MOBLS	0.1058	0.1708	0.2343
	L1RBLS	0.1054	0.1632	0.2431
	ENRBLS	0.1058	0.1664	0.2461
slump	BLS	0.6858	0.9362	1.3581
	MOBLS	0.6637	0.9712	1.1999
	L1RBLS	0.6654	0.9814	1.2014
	ENRBLS	0.6638	0.9755	1.2215
wq	BLS	0.9066	0.9265	1.1185
	MOBLS	0.9034	0.9249	0.9480
	L1RBLS	0.9043	0.9281	1.0016
	ENRBLS	0.9049	0.9258	0.9504
compressor	BLS	0.0143	0.3315	0.7255
	MOBLS	0.0128	0.1719	0.3751
	L1RBLS	0.0132	0.1882	0.6512
	ENRBLS	0.0129	0.1884	0.6414

4 结 论

本文通过 Frobenius 和 $L_{2,1}$ 矩阵范数的联合约束提出一种多输出宽度学习系统, 考虑不同列输出权重的互信息, 旨在发掘多个输出之间的相关性, 并精确预测多个目标, 同时 MOBLS 能够有效处理数据中的异常值. 通过在多个数据集上进行的仿真实验, 验证了所提方法的有效性.

参考文献(References)

- [1] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets[J]. *Neural Computation*, 2006, 18(7): 1527-1554.
- [2] Hinton G E, Salakhutdinov R R. Reducing the dimensionality of data with neural networks[J]. *Science*, 2006, 313(5786): 504-507.
- [3] Salakhutdinov R, Hinton G. Deep boltzmann machines[C]. *Artificial Intelligence and Statistics*. PMLR, 2009: 448-455.
- [4] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J/OL]. 2014, arXiv: 1409.1556.
- [5] Chen C L P, Liu Z L. Broad learning system: An effective and efficient incremental learning system without the need for deep architecture[J]. *IEEE Transactions on Neural Networks and Learning Systems*, 2018, 29(1): 10-24.
- [6] Pao Y H, Park G H, Sobajic D J. Learning and generalization characteristics of the random vector functional-link net[J]. *Neurocomputing*, 1994, 6(2): 163-180.
- [7] Tyukin I Y, Prokhorov D V. Feasibility of random basis function approximators for modeling and control[C]. *IEEE Control Applications(CCA) & Intelligent Control (ISIC)*. St. Petersburg, 2009: 1391-1396.
- [8] Pao Y H, Takefuji Y. Functional-link net computing: Theory, system architecture, and functionalities[J]. *Computer*, 1992, 25(5): 76-79.
- [9] Han M, Feng S B, Chen C L P, et al. Structured manifold broad learning system: A manifold perspective for large-scale chaotic time series analysis and prediction[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2019, 31(9): 1809-1821.
- [10] Xu M L, Han M, Chen C L P, et al. Recurrent broad learning systems for time series prediction[J]. *IEEE Transactions on Cybernetics*, 2020, 50(4): 1405-1417.
- [11] Liu Z L, Chen C L P, Feng S, et al. Stacked broad learning system: From incremental flatted structure to deep model[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2021, 51(1): 209-222.
- [12] Zhang T, Wang X H, Xu X M, et al. GCB-net:

- Graph convolutional broad network and its application in emotion recognition[J]. *IEEE Transactions on Affective Computing*, 2022, 13(1): 379-388.
- [13] Liu Z, Huang S L, Jin W, et al. Broad learning system for semi-supervised learning[J]. *Neurocomputing*, 2021, 444: 38-47.
- [14] Zhao H M, Zheng J J, Deng W, et al. Semi-supervised broad learning system based on manifold regularization and broad network[J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2020, 67(3): 983-994.
- [15] Kong Y, Cheng Y H, Chen C L P, et al. Hyperspectral image clustering based on unsupervised broad learning[J]. *IEEE Geoscience and Remote Sensing Letters*, 2019, 16(11): 1741-1745.
- [16] Zhang L, Li J H, Lu G Q, et al. Analysis and variants of broad learning system[J]. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 2022, 52(1): 334-344.
- [17] Jin J W, Li Y T, Chen C L P. Pattern classification with corrupted labeling via robust broad learning system[J]. *IEEE Transactions on Knowledge and Data Engineering*, 2022, 34(10): 4959-4971.
- [18] Jin J W, Li Y T, Yang T J, et al. Discriminative group-sparsity constrained broad learning system for visual recognition[J]. *Information Sciences*, 2021, 576: 800-818.
- [19] Wei E M, Ozdaglar A. Distributed alternating direction method of multipliers[C]. *The 51st IEEE Conference on Decision and Control*. Maui, 2012: 5445-5450.
- [20] Chen C H, He B S, Ye Y Y, et al. The direct extension of ADMM for multi-block convex minimization problems is not necessarily convergent[J]. *Mathematical Programming*, 2016, 155(1/2): 57-79.
- [21] Kannan B K, Kramer S N. An augmented Lagrange multiplier based method for mixed integer discrete continuous optimization and its applications to mechanical design[J]. *Journal of Mechanical Design*, 1994, 116(2): 405-411.
- [22] Rockafellar R T. Augmented Lagrange multiplier functions and duality in nonconvex programming[J]. *SIAM Journal on Control*, 1974, 12(2): 268-285.
- [23] Zhen X T, Yu M Y, He X F, et al. Multi-target regression via robust low-rank learning[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018, 40(2): 497-504.
- [24] Wright S J, Nowak R D, Figueiredo M A T. Sparse reconstruction by separable approximation[J]. *IEEE Transactions on Signal Processing*, 2009, 57(7): 2479-2493.
- [25] Tsoumakas G, Katakis I, Vlahavas I. Mining multi-label data[M]. *Data Mining and Knowledge Discovery Handbook*. Boston: Springer US, 2009: 667-685.
- [26] Chu F, Wang F L, Wang X G, et al. Performance modeling of centrifugal compressor using kernel partial least squares[J]. *Applied Thermal Engineering*, 2012, 44: 90-99.
- [27] Jin J W, Philip Chen C L. Regularized robust broad learning system for uncertain data modeling[J]. *Neurocomputing*, 2018, 322: 58-69.

作者简介

褚菲(1984—), 男, 副教授, 博士生导师, 从事复杂工业过程建模、控制与优化、机器学习、运行状态评价等研究, E-mail: chufeizhufei@sina.com;

卢新宇(1996—), 男, 硕士生, 从事机器学习的研究, E-mail: 541528235@qq.com;

苏嘉铭(1996—), 男, 硕士生, 从事机器学习的研究, E-mail: sujiaming@cumt.edu.cn;

王雪松(1974—), 女, 教授, 博士生导师, 从事机器学习、生物信息学、人工智能等研究, E-mail: wangxuesongcumt@163.com;

马小平(1961—), 男, 教授, 博士生导师, 从事控制理论与应用、网络控制、计算机技术及应用等研究, E-mail: xpma@cumt.edu.cn.