

带权重变化和决策融合的ELM在线故障检测

罗家祥^{1,2†}, 罗丹¹, 胡跃明^{1,2}

(1. 华南理工大学 自动化科学与工程学院, 广州 510640;

2. 华南理工大学 精密电子制造装备教育部工程研究中心, 广州 510640)

摘要: 在线极限学习机对样本数据分批或分块地学习,适用于分析生产过程的在线数据,进而检测生产过程的故障. 为了提高检测的准确性和快速性,提出一种权重变化和决策融合的极限学习机(ELM)在线故障检测方法. 该方法在学习过程中增加被当前数据监控模型错误预测的新样本权重,同时在数据监控模型中引入决策级融合的方法,提高模型的综合决策能力. 利用UCI数据集和TE过程进行仿真实验对比,对比结果表明所提出的方法在训练时间和检测准确率上都具有很好的性能.

关键词: 故障检测; 在线极限学习机; 权重变化; 决策级融合

中图分类号: TP277

文献标志码: A

A new online extreme learning machine with varying weights and decision level fusion for fault detection

LUO Jia-xiang^{1,2†}, LUO Dan¹, HU Yue-ming^{1,2}

(1. School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, China;

2. Engineering Research Center of Precision Manufacturing Equipments, Ministry of Education, South China University of Technology, Guangzhou 510640, China)

Abstract: Online sequential extreme learning machine learns samples one by one or chunk by chunk, so it is appropriate to analyze online process data and to detect system faults. To improve the detection accuracy and rapidity, a new online extreme learning machine algorithm with varying weights and decision level fusion has been proposed, which increases the weights of the new samples predicted wrongly by current data monitoring model in learning, and introduced decision level fusion to improve integrated decision-making ability of the model. Performance comparisons of the method are presented using UCI datasets and Tennessee Eastman process. The results show that the proposed algorithm produces comparable or better performance with higher accuracies and lower training time.

Keywords: fault detection; online extreme learning machine; varying weights; decision level fusion

0 引言

随着工业过程日益大型化、连续化,通过对过程数据的分析实现对过程故障的检测越来越成为企业确保安全和连续生产的重要手段. 在实际工业过程中,数据监控模型应根据数据的实时到达和动态变化及时更新,因此在线故障检测技术成为国内外自动化控制界的研究热点方向之一^[1],并在工业过程、电力系统、大型船舶、航空航天等领域得到了广泛应用^[1-4]. 传统的在线故障检测方法有神经网络(Neural network, NN)、支持向量机(Support vector machine, SVM)等^[5-6]. 这些方法分批或者分块地学习

在线过程数据样本的特性,都属于在线学习方法. 但这些传统方法的训练参数选择复杂,新到达的数据需要与旧数据一起学习,导致计算量大、训练时间长,影响了监控模型的实时性和灵活性.

在线极限学习机(Online sequential extreme learning machine, OS-ELM)^[7]具有训练速度快、参数选择简单、操作易实现等优点,因此在故障诊断、时间序列预测等领域具有很好的应用效果^[8-10]. 该方法在训练样本发生变化或有新样本加入时,根据上一步估计的网络模型递推地更新网络的训练参数,实现在线学习,基于OS-ELM的在线学习方法及其在故

收稿日期: 2017-03-15; 修回日期: 2017-06-27.

基金项目: 国家科技重大项目(2014ZX02503-3); 国家自然科学基金项目(06573146).

责任编委: 胡庆雷.

作者简介: 罗家祥(1979-),女,副教授,博士,从事工业过程监控与优化、智能计算等研究; 罗丹(1993-),女,硕士生,从事过程监控方法的研究.

†通讯作者. E-mail: luojx@scut.edu.cn

障检测中的应用已成为当前研究热点^[8-14]. 文献[9]提出了一种正则化OS-ELM,以提高学习机的泛化性能. 文献[10]为了更准确地建模时变非线性系统,提出了一种结合自适应正则因子的正则化OS-ELM,以降低模型结构风险. 文献[11]提出了一种在线核极限学习机,用核函数代替隐层激活函数,以提高在线学习能力. 文献[12]针对不稳定系统提出了一种具有低复杂度自适应遗忘因子的OS-ELM,以提高对不稳定系统的学习能力. 文献[13]提出了一种基于差分向量核主成分分析的动态循环OS-ELM故障检测方法,用来重建过程特征,检测实时非线性系统的过程故障. 文献[14]针对在线过程中数据类别不平衡的问题,提出了一种带权重的OS-ELM,以提高学习能力.

上述基于OS-ELM的在线学习方法或在线故障检测方法认为,在学习过程中,新到达的每个样本对当前模型的更新具有同等重要的作用. 然而,新到达的数据样本并不一定完全符合当前数据监测模型. 此外,由于传统OS-ELM对网络输入层参数一次随机赋值,导致模型决策单一、输出不稳定、在未知样本上的泛化性能弱. 本文提出一种带权重变化和决策融合的OS-ELM故障检测方法,该方法在学习过程中增加被当前模型错误分类样本的权重,提高错误分类新样本对模型更新的影响. 此外,为了提高模型在未知样本上的泛化能力以及改善单一模型的不稳定性,引入决策级融合方法^[15]进行集成决策. 为验证算法有效性,利用UCI数据集和TE过程故障数据进行仿真实验,结果表明,所提出的算法在训练时间和检测准确率上都表现出很好的性能.

需要说明的是,本文在更新模型时增加了在线学习过程中错误分类样本的权重,以提高模型对故障的识别性能. 该思想来源于文献[14],通过改变不平衡样本的权重来训练OS-ELM模型,能够有效地提高模型对不平衡样本的分类性能. 对于决策级融合方法,现有研究中的方法主要用于离线过程^[16-17],并表现出了良好的性能. 本文将该方法用于在线学习过程以提高在线监控模型的故障检测性能.

1 在线极限学习机基本原理

在线极限学习机是在极限学习机(Extreme learning machine, ELM)^[18]的基础上提出的在线学习方法. 不同于传统的神经网络方法,ELM是一种单层前馈神经网络,所有隐层参数均随机产生,且不需要繁琐的迭代过程,具有训练速度快、分类效果好等优点,已广泛应用于模型预测、故障诊断、语音识别、

图像识别等领域^[19-20].

设训练样本和标签对 $(x_i, t_i) \in D, D = \{(x_i, t_i) | i = 1, 2, \dots, N\}, t_i$ 为样本 x_i 的标签, $T = [t_1, t_2, \dots, t_N]^T$. ELM的结构如图1所示.

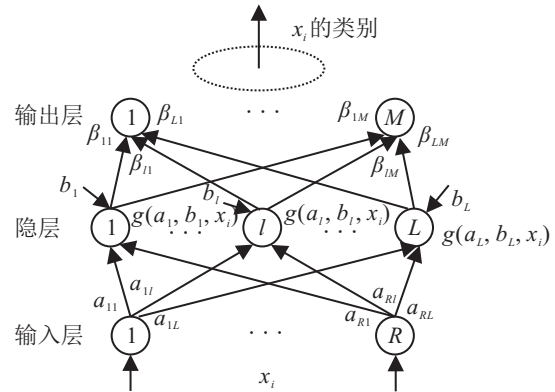


图1 ELM结构

R 为输入节点个数,即训练样本维数; L 为隐层节点个数; M 为输出节点个数,即样本标签类别数;输入权重矩阵为 $a_{R \times L}$;输出权重矩阵为 $\beta_{L \times M}$,第 l 个隐层节点激活函数为 $g(a_l, b_l, x_i)$, b_l 为隐层偏差. 在传统的ELM中,上述参数中矩阵 $a_{R \times L}$ 的元素和 $b_l = [b_1, b_2, \dots, b_l, \dots, b_L]$ 的元素是一次随机生成的,模型只需通过训练确定参数矩阵 $\beta_{L \times M}$. 上述网络模型所对应的输出为

$$T^* = H\beta. \quad (1)$$

其中

$$H = \begin{bmatrix} g(a_1, b_1, x_1) & \dots & g(a_L, b_L, x_1) \\ \vdots & \ddots & \vdots \\ g(a_1, b_1, x_N) & \dots & g(a_L, b_L, x_N) \end{bmatrix}, \quad (2)$$

$$\beta = [\beta_1, \dots, \beta_2, \dots, \beta_L]. \quad (3)$$

当模型零误差预测样本输出时,预测标签与实际标签一致,即 $T^* = T, T^* = [t_1^*, t_2^*, \dots, t_N^*]^T$. 因此,模型待定参数矩阵 $\beta_{L \times M}$ 可利用已有训练数据,通过最小二乘法同时最小化训练误差(经验风险)和输出权值矩阵的二次范数(结构风险)获得,即求解如下最小化问题:

$$\min F = \frac{C}{2} \|H\beta - T\|_2^2 + \frac{1}{2} \|\beta\|_2^2, \quad (4)$$

其中 C 为正则化参数. 利用求导得到式(4)的最优参数为

$$\beta = \left(H^T H + \frac{I}{C} \right)^{-1} H^T T. \quad (5)$$

根据 β ,输出层节点 m 对预测样本 x_i 的输出为

$$f_m(x_i) = \sum_{l=1}^L \beta_{lm} g(a_l, b_l, x_i). \quad (6)$$

将预测输出最大值的标签作为模型的输出,有

$$\text{class}(x_i) = \arg \max_{m=1,2,\dots,M} \{f_m(x_i)\}. \quad (7)$$

OS-ELM^[7]是在线ELM,其网络结构与ELM相同,只是在学习时既可分批学习也可逐样本学习,实现在线更新网络输出权重.OS-ELM包括初始化和在线学习两个阶段,设为训练样本数据块集合 $\{D_0, D_1, \dots, D_k, \dots, D_K\}$, $D_k = \{x_{ki}, t_{ki}\}_{i=1}^{N_k}$, N_k 为第 k 个样本块中样本的个数.

1) 初始化阶段. 随机生成输入权重矩阵 $a_{R \times L}$ 和隐层偏差向量 b_L ,根据初始样本数据块 D_0 和式(5)得初始训练参数为

$$\beta_0 = \left(H_0^T H_0 + \frac{I}{C}\right)^{-1} H_0^T T_0 = L_0 H_0^T T_0. \quad (8)$$

其中: $L_0 = K_0^{-1}$, $K_0 = H_0^T H_0 + \frac{I}{C}$, $T_0 = [t_1, t_2, \dots, t_{N_0}]^T$.

2) 在线学习阶段. 当第 k 个样本数据块 D_k 到来时更新 β_k ,有

$$\begin{aligned} \beta_k &= \left(\frac{I}{C} + \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}\right)^{-1} \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} T_{k-1} \\ T_k \end{bmatrix} = \\ &K_k^{-1} \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} T_{k-1} \\ T_k \end{bmatrix} = \\ &K_k^{-1} (H_{k-1}^T T_{k-1} + H_k^T T_k). \end{aligned} \quad (9)$$

令 $L_k = K_k^{-1}$,经过整理后满足

$$\begin{aligned} L_k &= \left(\frac{I}{C} + \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}\right)^{-1} = \\ &L_{k-1} - L_{k-1} H_k (I + H_k L_{k-1} H_k^T)^{-1} H_k L_{k-1}. \end{aligned} \quad (10)$$

将式(10)代入(9)可得

$$\beta_k = \beta_{k-1} + L_k H_k^T (T_k - H_k \beta_{k-1}).$$

综上,模型的更新公式可表示为

$$\begin{cases} L_k = L_{k-1} - L_{k-1} H_k^T (I + H_k L_{k-1} H_k^T)^{-1} L_{k-1}, \\ \beta_k = \beta_{k-1} + L_k H_k^T (T_k^T - H_k \beta_{k-1}). \end{cases} \quad (11)$$

上述OS-ELM是基于ELM的在线学习方法,继承了ELM结构简单、训练速度快等优点. 在线学习过程中,根据当前模型和新到达的样本不断更新模型. 如果当前模型能够正确预测新到达的样本,则式(11)中的 $T_k - H_k \beta_{k-1} = 0$,即当前模型无更新;否则模型被更新. 可见错误预测的样本对模型更新的影响较大,若增加当前模型错误预测的新样本对模型更新的重要性,则可能会提高模型预测精度. 此外,OS-ELM采用一次性随机初始化输入权重和隐层偏差,

往往会使得训练数据在OS-ELM特征空间中的分布具有一定的随机性,导致其输出稳定性和在未知样本上的泛化性能不太理想^[21-22].

2 WMOS-ELM算法

本文提出一种改进的带权重变化和决策融合的OS-ELM,简称WMOS-ELM. 在线学习过程中,利用当前模型对新样本进行测试,鉴于改变不平衡数据样本的权重的思想^[18],该方法在学习过程中增加当前模型分类错误的新到达样本权重,同时在模型中引入决策级融合方法,实现多个分类器融合,以提高模型的综合决策能力. 本文算法在一定程度上缓解了以往算法中所有新样本同等重要地更新模型而导致的检测精度不佳的问题,并通过决策融合提高了输出的稳定性,算法结构如图2所示.

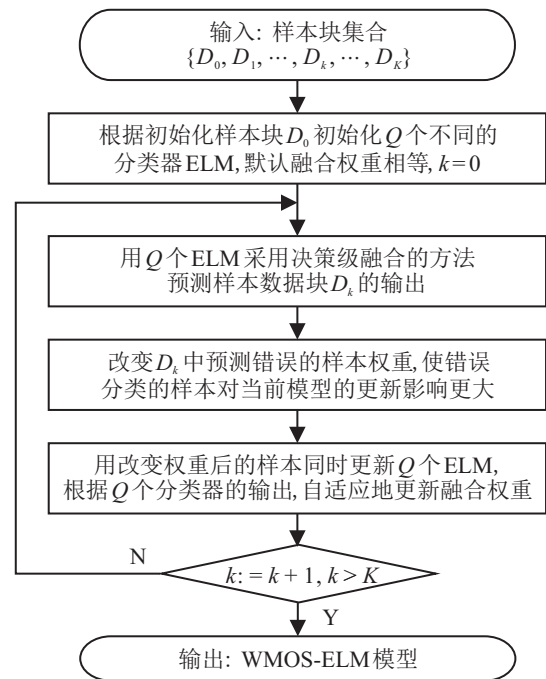


图2 WMOS-ELM算法流程

2.1 在线训练样本权重处理

引入样本权重矩阵

$$W = \text{diag}(w_1, w_2, \dots, w_n),$$

其中: n 为样本个数, w_i 为第 i 个样本的权重,初始化为1. 利用当前模型对新到达的样本进行预测,若第 i 个样本预测错误,则该样本权重变为 $w_i = 1 + \Delta$, Δ 为增加的权重值. 引入权重矩阵后,OS-ELM在初始化阶段的参数 β_0 可表示为

$$\beta_0 = K_0^{-1} H_0^T W_0 T_0. \quad (12)$$

其中: $K_0 = H_0^T W_0 H_0 + \frac{I}{C}$, $W = \text{diag}(w_{01}, w_{02}, \dots, w_{0N_0})$ 为初始样本块的权重矩阵.

在线学习阶段的训练参数 β_k 为

$$\beta_k = \left(\frac{I}{C} + \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} W_{k-1} & 0 \\ 0 & W_k \end{bmatrix} \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix} \right)^{-1} \times \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} W_{k-1} & 0 \\ 0 & W_k \end{bmatrix} \begin{bmatrix} T_{k-1} \\ T_k \end{bmatrix}, \quad (13)$$

其中 W_k 为第 k 个样本数据块的权重矩阵. 由式(13), 式(9)中 K_k 的表达式转换为

$$K_k = \frac{I}{C} + \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} W_{k-1} & 0 \\ 0 & W_k \end{bmatrix} \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix} = \frac{I}{C} + H_{k-1}^T W_{k-1} H_{k-1} + H_k^T W_k H_k = K_{k-1} + H_k^T W_k H_k. \quad (14)$$

则式(13)中 β_k 的后半部分可进一步表示为

$$\begin{aligned} & \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} W_{k-1} & 0 \\ 0 & W_k \end{bmatrix} \begin{bmatrix} T_{k-1} \\ T_k \end{bmatrix} = \\ & H_{k-1}^T W_{k-1} T_{k-1} + H_k^T W_k T_k = \\ & K_{k-1} K_{k-1}^{-1} H_{k-1}^T W_{k-1} T_{k-1} + H_k^T W_k T_k = \\ & K_{k-1} \beta_{k-1} + H_k^T W_k T_k = \\ & (K_k - H_k^T W_k H_k) \beta_{k-1} + H_k^T W_k T_k = \\ & K_k \beta_{k-1} - H_k^T W_k H_k \beta_{k-1} + H_k^T W_k T_k. \end{aligned} \quad (15)$$

将式(14)和(15)代入(13), 得到 β_k 的更新公式为

$$\begin{aligned} \beta_k &= K_k^{-1} \begin{bmatrix} H_{k-1} \\ H_k \end{bmatrix}^T \begin{bmatrix} W_{k-1} & 0 \\ 0 & W_k \end{bmatrix} \begin{bmatrix} T_{k-1} \\ T_k \end{bmatrix} = \\ & K_k^{-1} (K_k \beta_{k-1} - H_k^T W_k H_k \beta_{k-1} + H_k^T W_k T_k) = \\ & \beta_{k-1} + K_k^{-1} H_k^T W_k (T_k - H_k^T \beta_{k-1}). \end{aligned} \quad (16)$$

令 $L_k = K_k^{-1}$, 经类似于式(10)中对 L_k 的整理, 并结合式(16), 得到模型的更新公式

$$\begin{cases} L_k = \\ L_{k-1} - L_{k-1} (W_k^{-1} + H_k L_{k-1} H_k^T)^{-1} H_k L_{k-1}, \\ \beta_k = \beta_{k-1} + L_k H_k^T W_k (T_k^T - H_k \beta_{k-1}). \end{cases} \quad (17)$$

2.2 决策级融合

决策级融合^[15-17]按照一定的准则和每个分类器的可信度对各基分类的信息进行融合, 作出全局最优决策. 决策级融合与集成学习^[23]的融合策略相似, 但后者迭代训练多个分类器, 并从多个分类结果中选择一个结果, 而决策级融合独立一次训练多个分类器, 再将分类器的信息综合产生一个结果, 因此具有更好的实时性和容错性^[24-25]. 常用的决策级融合策略有基于投票和基于贝叶斯的方法, 后者具有更高的准确

性和适用性而被广泛应用^[24].

为提高模型输出的稳定性和对未知数据的泛化性能, 本节引入数据融合思想, 根据第2.1节方法同时训练多个 OS-ELM 模型, 并采用基于贝叶斯法的决策级融合法集成多个分类器信息产生分类结果. 参考文献[17], 给定预测样本 x_i , 训练好的 Q 个学习机的预测输出为 $f_1(x_i), \dots, f_q(x_i), \dots, f_Q(x_i)$, 其中 $f_q(x_i) = [f_{1q}(x_i), f_{2q}(x_i), \dots, f_{Mq}(x_i)]^T$. 所采用的决策融合步骤如下.

Step 1: 计算模型的概率输出矩阵.

Step 1.1: 对每个学习机的各输出节点的输出进行概率映射, 得到

$$p_{mq}(f_{mq}(x_i)) = \frac{1}{1 + \exp(-f_{mq}(x_i))}, \quad m = 1, 2, \dots, M, \quad q = 1, 2, \dots, Q; \quad (18)$$

Step 1.2: $P(x_i)$ 中的元素 $p_{mq}(x_i)$ 即为 $p_{mq}(f_{mq}(x_i))$ 归一化的值, 有

$$p_{mq}(x_i) = \frac{p(f_{mq}(x_i))}{\sum_{t=1}^M p(f_{tq}(x_i))}, \quad m = 1, 2, \dots, M, \quad q = 1, 2, \dots, Q. \quad (19)$$

Step 2: 根据概率输出矩阵 $P(x_i)$ 进行集成决策.

Step 2.1: 确定每个学习机中的最大输出概率值

$$p_q = \max_{m=1,2,\dots,M} \{p_{mq}(x_i)\}; \quad (20)$$

Step 2.2: 利用每个学习机的最大输出概率值, 计算每个学习机 q 在集成决策中的权重

$$u_q = \frac{p_q}{\sum_{q'=1}^Q (p_{q'})}; \quad (21)$$

Step 2.3: 计算模型对样本 x_i 的预测标签输出为

$$\text{class}(x_i) = \arg \max_{m=1,2,\dots,M} \left\{ \sum_{q=1}^Q u_q p_{mq}(x_i) \right\}. \quad (22)$$

在上述过程中, 首先对每个学习机的输出值进行概率映射并归一化, 再根据每个学习机的最大概率输出值确定该学习机在集成决策中的权重, 最后对每个学习机采用权重的方式进行决策融合. 这样充分考虑了各分类器针对不同输入样本识别性能不同的特性, 确定较为合理的集成决策权值, 解决了传统 OS-ELM 对样本预测输出出现过于接近的值而造成的误分类问题.

3 仿真实验

本文采用 3.2 GHz CPU, 8 G RAM, 64 bit 主机, 在 Matlab 2014 环境中测试算法. 由于故障检测问题本

质上是一个分类问题,首先用UCI分类数据集分别验证增加错误分类样本权重策略和引入决策级融合策略的有效性以及算法的鲁棒性,然后利用TE过程故障数据进行仿真实验.将所提出的WMOS-ELM与传统在线故障检测方法(如PNN、SVM)以及传统的正则化在线极限学习机ROS-ELM进行对比.选择测试准确率作为实验结果,计算过程如下:

$$Acc = \left(\frac{1}{N} \sum_{i=1}^N I(t(i) = t^*(i)) \right) \times 100. \quad (23)$$

其中

$$I(t(i) = t^*(i)) = \begin{cases} 1, & t(i) = t^*(i); \\ 0, & t(i) \neq t^*(i); \end{cases}$$

$t(i)$ 为样本实际标签, $t^*(i)$ 为模型预测输出.

以下实验中,每个测试结果均是算法重复50次的平均结果.

3.1 参数设置

算法具体参数设置如下.

1) 输入权重矩阵 $a_{R \times L}$ 和隐层偏差向量 $b_L: a_{R \times L}$ 的元素 $a_{r,l}$ 和 b_L 的元素 b_i 取 $[-1,1]$ 区间的随机数^[10].

2) 隐层节点个数 L :参考文献[7],对每个数据集使用交叉验证的方法进行预实验,在1~500范围内优选隐层节点个数 L 的值.

3) 正则化因子 C :用有效留一交叉检验的方法在 $[2^{-30}, 2^{30}]$ 范围内优选^[9].

4) 分类错误的样本权重增加量 Δ :将 Δ 依次设置为 $[0,10]$ 范围内以0.1为步长的数值,使用交叉验证方法对代表性数据集(Haberman, magic, statlog)进行预实验.实验结果表明,最好的 Δ 值集中在 $\{0.1, 0.2, 0.3, 0.4\}$,因此,对于每个数据集,使用交叉验证法在集合 $\{0.1, 0.2, 0.3, 0.4\}$ 中优选权重变化量 Δ .

5) 分类器个数 Q :在多个学习机融合时,学习机个数过多会导致训练时间增长,个数过少会使融合策略效果不明显,通过预实验选择3个学习机进行融合.

6) 激活函数:分别采用sigmoid、sin、radbas三种激活函数的算法进行测试.图3给出了当 $L = 1 \sim 500$ 时,算法在Segment数据集上的测试结果.结果表明,采用Sigmoid函数的算法取得了更好的结果.为此,所有算法均选择Sigmoid作为隐层激活函数.

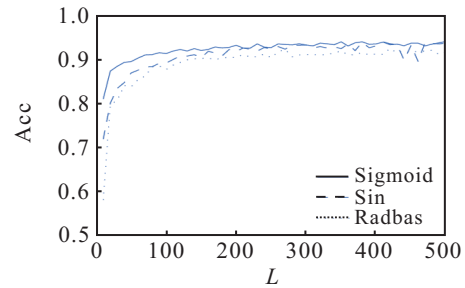


图3 不同激活函数预测准确率

另外,根据文献[7],初始阶段训练样本个数 $N_0 = L + 50$,由于个别训练样本数太少,初始阶段样本个数为 $N_0 = L + 10$.

表1 离线测试实验单个故障检测结果

数据集	样本个数	ROS-ELM		WROS-ELM		MROS-ELM		WMOS-ELM	
		Acc / %	Rate / %	Acc / %	Rate / %	Acc / %	Rate / %	Acc / %	Rate / %
Segment	2 100	93.05	93.62	0.61	93.80	0.81	94.12	1.15	
Satellite	4 435	96.46	96.58	0.12	96.74	0.29	96.94	0.50	
Gesture	1 747	91.52	91.99	0.51	92.00	0.52	92.08	0.61	
Seeds	210	94.43	94.60	0.18	94.43	0.00	95.06	0.67	
Wine	1 599	60.42	60.13	-0.48	61.28	1.42	60.94	0.86	
Haberman	306	73.94	74.86	1.24	73.76	-0.24	75.12	1.60	
letter	20 000	82.48	83.74	1.53	84.33	2.24	86.48	4.85	
Contraceptive	1 437	52.42	52.88	0.88	51.94	-0.92	53.20	1.49	
Fertility	100	85.59	87.06	1.72	88.65	3.58	87.18	1.86	
magic	19 020	84.07	84.4	0.39	84.72	0.77	84.88	0.96	
Hayes	132	53.91	55.36	2.69	53.64	-0.50	56.68	5.14	
Pendigits	10 992	98.49	98.62	0.13	98.75	0.26	98.81	0.32	
Climate	540	91.76	92.13	0.40	91.24	-0.57	91.83	0.08	
Statlog	6 435	88.12	88.37	0.28	88.46	0.39	88.54	0.48	
SPECTF	267	78.58	79.53	1.21	79.44	1.09	79.62	1.32	
Vowel	990	82.35	87.76	6.57	85.24	3.51	88.78	7.81	
Wilt	4 839	95.65	96.01	0.38	95.79	0.15	96.18	0.55	
Yeast	682	58.37	60.30	3.31	58.71	0.58	60.37	3.43	
平均值	-	-	-	1.2	-	0.74	-	1.87	

3.2 实际分类数据集仿真实验和结果

3.2.1 算法策略的有效性分析

参考文献[7],用18个UCI实际分类数据集验证本文所提策略的有效性,将本文算法WMOS-ELM与ROS-ELM、采用增加错误分类样本权重策略的OS-ELM算法(WROS-ELM)和采用决策级融合策略的OS-ELM算法(MROS-ELM)进行比较.实验中所有属性均标准化在 $[-1,1]$ 范围内,2/3的数据集用于训练,1/3的数据集用于测试.测试结果如表1所示.

表1中,Rate列中的数值为各算法结果相对于传统ROS-ELM结果的变化值,有

$$\text{Rate} = \frac{\text{Acc} - \text{Acc}_{\text{ROS-ELM}}}{\text{Acc}_{\text{ROS-ELM}}} \times 100\%.$$

由表1可以得出以下结论:WROS-ELM和MROS-ELM的平均准确率相对于ROS-ELM分别提高了1.20%、0.74%,其中在Vowel数据集上,WROS-ELM提高6.75%,MROS-ELM提高3.51%.同时加入两种策略的WMOS-ELM的平均准确率相对于ROS-ELM提高了1.87%,其中在Vowel数据集上比传统ROS-ELM的准确率提高幅度达到7.81%.表明了两种策略均能提高算法分类准确率,而两者的综合应用能得到更好的分类性能.

3.2.2 算法鲁棒性分析

为验证WMOS-ELM对噪声数据的鲁棒性,在表1中随机选择2个数据集Segment和Gesture进行实验,按照噪声级别0,0.1,⋯,0.5的顺序给两个数据集的所有样本增加噪声,噪声添加方法如下^[26]:1)计算样本相应属性的标准差 σ ;2)产生满足 $N(0, \sigma)$ 分布的随机数叠加到样本数据对应属性值上,同时保持样本数据的类别标记值不变,实验结果如图4所示.由

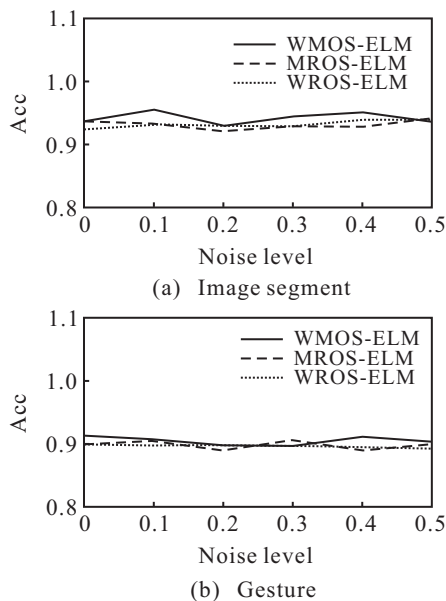


图4 噪声数据实验结果

图4可见,随着噪声级别的增加,MROS-ELM、WROS-ELM和WMOS-ELM的分类准确率均保持平稳,算法具有较好的鲁棒性.

3.3 TE过程实验仿真和结果

田纳西-伊斯曼过程(Tennessee-Eastman process, TEP)是实际化工过程的模拟程序模型,已广泛应用于多变量控制、自适应控制、非线性控制和故障检测等领域的测试平台^[27-28].TE过程中共含有52个测量变量,其中41个是过程变量,11个是操作变量^[27-28].

基于文献[27-29],本文的TE过程测试数据选取其中的故障1、2、4、6、7、8、13、14、17、18、19和正常模式数据进行仿真实验.实验前将所有数据标准化到 $[-1,1]$.

1) 离线测试实验.

将WMOS-ELM算法与传统的故障检测方法SVM、PNN、ROS-ELM进行对比.实验中数据样本分为训练集和测试集,各算法在测试集上的测试结果如表2和表3所示.表2为单个故障检测结果,表3为所有故障检测结果的平均值和时间消耗情况.用偏差(Bias)评估分类器性能,Bias越小说明该分类器性能越好,即

$$\text{Bias} = \frac{\text{Acc}_{\text{max}} - \text{Acc}}{\text{Acc}_{\text{max}}} \times 100\%. \quad (24)$$

其中: Acc_{max} 为比较方法中准确率的最大值,Acc为比较算法得到的准确率.

表2 离线测试实验单个故障检测结果

故障类型	SVM		PNN		ROS-ELM		WMOS-ELM	
	Acc / %	Bias / %	Acc / %	Bias / %	Acc / %	Bias / %	Acc / %	Bias / %
1	90.69	1.28	91.39	0.52	91.87	0.00	91.77	0.11
2	92.73	0.00	87.96	5.14	92.27	0.50	91.70	1.11
4	94.60	0.00	92.78	1.92	93.15	1.53	93.73	0.92
6	72.64	2.76	66.04	11.59	74.41	0.39	74.70	0.00
7	92.23	2.85	93.13	1.91	94.08	0.91	94.94	0.00
8	88.44	0.00	79.04	10.63	73.85	16.50	83.89	5.14
13	90.06	3.27	93.10	0.00	83.74	10.05	86.27	7.34
14	58.76	16.28	49.64	29.28	60.23	14.19	70.19	0.00
17	79.14	8.79	49.10	43.41	82.24	5.22	86.77	0.00
18	63.54	18.09	44.95	42.05	73.85	4.80	77.57	0.00
19	79.58	7.45	80.00	6.97	84.31	1.95	85.99	0.00
平均	-	5.52	-	13.95	-	5.09	-	1.33

表3 离线测试实验所有故障结果的平均值

	SVM	PNN	ROS-ELM	WMOS-ELM
测试准确率 / %	79.02	45.32	79.32	81.48
训练时间 / s	3.87	12.14	0.01	0.07
测试时间 / s	0.97	5.55	0.17	0.19

由表2可见,WMOS-ELM的偏差平均值只有1.33%,远小于SVM、PNN、ROS-ELM的平均值,分别为5.52%、13.95%和5.09%.由表3可见,WMOS-

ELM的平均准确率为81.48%,高于SVM、PNN和ROS-ELM的结果,说明WMOS-ELM在单个故障和所有故障的检测上表现出了更好的性能.虽然WMOS-ELM的训练时间和测试时间略长于ROS-ELM,但都少于0.2s.

2) 在线故障检测实验.

为了对比在线过程算法的性能,选择故障1、2、4、7的故障数据进行实验.参考文献[18],实验结果为当前模型对每一批新来的数据进行测试,如表4、图5和图6所示.表4中的平均正确率是各批数据测试准确率的平均值,训练时间是所有批数据训练完所需要的时间.图5和图6显示了各批次的测试准确率和时间消耗情况.

表4 在线故障检测实验结果

	SVM	PNN	ROS-ELM	WMOS-ELM
平均准确率/%	93.73	92.00	93.83	93.98
训练时间/s	5.67	15.57	0.88	2.44

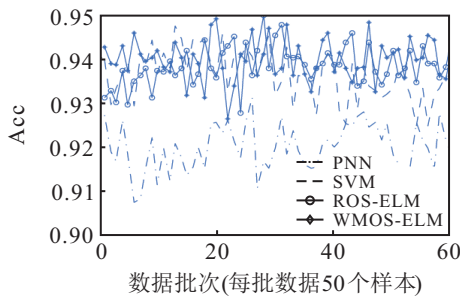


图5 在线故障检测准确率对比

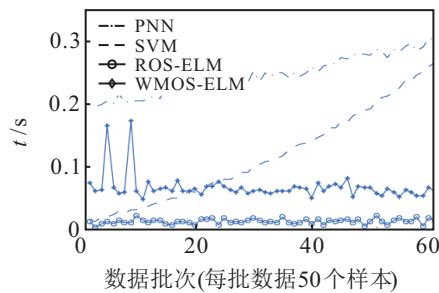


图6 在线过程训练时间对比

由表4可知,WMOS-ELM在测试集上的平均准确率为93.98%,略优于SVM、PNN、ROS-ELM.由图5可知,WMOS-ELM各批次测试准确率的总体情况略优于SVM、PNN、ROS-ELM.因为SVM和PNN的训练过程是新旧样本一起更新模型,因此训练时间随批次的增加而增加.如图6所示,其训练时间随样本批次的增加呈线性增长,其中PNN的增加梯度较大,而ROS-ELM和WMOS-ELM的训练时间基本保持不变.总体上,WMOS-ELM在故障检测准确率和训练时间上均表现出良好的性能.

综上所述,文中提出的WMOS-ELM算法用于

TE过程在线故障检测时,在准确率和训练时间上都表现出良好的性能.

4 结论

为了能够准确及时地检测生产过程故障,本文提出了一种带权重变化和决策融合的ELM在线故障检测方法.该方法在学习过程中,先用当前模型对新加入的样本进行检测,再在模型更新中增加错误分类样本的权重,并引入决策级融合方法进行集成决策.利用UCI数据集分别验证了两种策略的有效性,利用TE过程故障数据进行仿真实验,并与传统的用于在线故障检测的分类算法PNN和SVM进行对比,结果表明,WMOS-ELM在训练时间和检测准确率上都表现出较好的性能.

参考文献(References)

- [1] Peng K X, Zhang K, Li G. Online contribution rate based fault diagnosis for nonlinear industrial processes[J]. Acta Automatic Sinica, 2014, 40(3): 423-430.
- [2] 童晓阳,王睿晗,王洪彬,等.基于有限PMU的电网故障在线识别算法研究[J].电力系统保护与控制,2016,44(19):49-55.
(Tong X Y, Wang R H, Wang H B, et al. Online detection algorithm of power system fault using limited PMUs[J]. Power System Protection and Control, 2016, 44(19): 49-55.)
- [3] 徐巧宁,喻峰,周华,等.基于鲁棒观测器的船舶舵机伺服系统故障检测[J].华中科技大学学报:自然科学版,2015,43(1):101-106.
(Xu Q N, Yu F, Zhou H, et al. Robust observer based fault detection for a ship rudder servo system[J]. Journal of Huazhong University of Science and Technology: Natural Science Edition, 2015, 43(1): 101-106.)
- [4] Zhao X, Wang S C. Real-time fault detection method based on belief rule base for aircraft navigation system[J]. Chinese J of Aeronautics, 2013, 26(3): 717-729.
- [5] Messai A, Mellit A. On-line fault detection of a fuel rod temperature measurement sensor in a nuclear reactor core using ANNs[J]. Progress in Nuclear Energy, 2015, 79: 8-21.
- [6] Li Y F, Wang Z F, Yuan J Q. On-line fault detection using SVM-based dynamic MPLS for batch processes[J]. Chinese J of Chemical Engineering, 2006, 14(6): 754-758.
- [7] Liang N Y, Huang G B, Sarathandran P, et al. A fast and accurate online sequential learning algorithm for feedforward networks[J]. IEEE Trans on Neural Networks, 2006, 17(6): 1411-1423.
- [8] 许玉格,邓文凯,陈立定.基于核函数的加权极限学习

- 机污水处理在线故障诊断[J]. 化工学报, 2016, 67(9): 3817-3825.
(Xu Y G, Deng W K, Chen L D. Online fault diagnosis in wastewater treatment process by kernel-based weighted extreme learning machine[J]. CIESC Journal, 2016, 67(9): 3817-3825.)
- [9] Shao Z F, Meng J E. An online sequential learning algorithm for regularized extreme learning machine[J]. Neurocomputing, 2016, 173(3): 778-788.
- [10] Lu X J, Zhou C, Huang M H, et al. Regularized online sequential extreme learning machine with adaptive regulation factor for time-varying nonlinear system[J]. Neurocomputing, 2016, 174(PB): 617-626.
- [11] Wan Y D, Ong Y S, Tan P S, et al. Online sequential reduced kernel extreme learning machine[J]. Neurocomputing, 2016, 174(PA): 72-84.
- [12] Li M J, Lee S. Low complexity adaptive forgetting factor for online sequential extreme learning machine for application to nonstationary system estimations[J]. Neural Comput and Application, 2013, 22(3/4): 569-576.
- [13] Xu Y, Ye L L, Zhu Q X. A new DROS-extreme learning machine with differential vector-KPCA approach for real-time fault recognition of nonlinear processes[J]. J of Dynamic Systems Measurement and Control, 2015, 137(5): 1-10.
- [14] Mirza B, Lin Z Q, Toh K A. Weighted online sequential extreme learning machine for class imbalance learning[J]. Neural Process Letter, 2013, 38(3): 465-486.
- [15] Jiang N, Qiu M, Dai W. SROC: A speaker recognition with data decision level fusion method in cloud environment[J]. J of Signal Processing Systems, 2016, 86(2/3): 123-133.
- [16] Niu Q, Lee S S, Yang B S, et al. Decision fusion system for fault diagnosis of elevator traction machine[J]. J of Mechanical Science and Technology, 2008, 22(1): 85-95.
- [17] 张文博, 姬红兵. 融合极限学习机[J]. 电子与信息学报, 2013, 35(11): 2728-2732.
(Zhang W B, Ji H B. Fusion of extreme learning machines[J]. J of Electronics and Information Technology, 2013, 35(11): 2728-2732.)
- [18] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: A new learning scheme of feedforward neural networks[C]. Proc of Int Joint Conf on Neural network. Budapest: IEEE, 2004: 985-990.
- [19] 陈晓, 王新民, 黄誉. 倾转旋翼机飞控系统的变精度粗糙集-OMELM故障诊断方法[J]. 控制与决策, 2015, 30(3): 433-440.
(Chen X, Wang X M, Huang Y. Fault diagnosis for tilt-rotor aircraft flight control system based on variable precision rough set-OMELM[J]. Control and Decision, 2015, 30(3): 433-440.)
- [20] 徐嘉明, 张卫强, 杨登舟, 等. 基于流形正则化极限学习机的语种识别系统[J]. 自动化学报, 2015, 41(9): 1680-1685.
(Xu J M, Zhang W Q, Yang D Z, et al. Manifold regularized extreme learning machine for language recognition[J]. Acta Automatic Sinica, 2015, 41(9): 1680-1685.)
- [21] 徐勇, 王东, 张慧. 基于自适应在线极限学习机模型的预测方法[J]. 统计研究, 2016, 33(7): 103-109.
(Xu Y, Wang D, Zhang H. Study on prediction method based on adaptive ensemble online sequential extreme learning machine[J]. Statistical Research, 2016, 33(7): 103-109.)
- [22] 林梅金, 罗飞, 苏彩红, 等. 一种新的混合智能学习机[J]. 控制与决策, 2015, 30(6): 1078-1084.
(Lin M J, Luo F, Su C H, et al. A improved hybrid intelligent extreme learning machine[J]. Control and Decision, 2015, 30(6): 1078-1084.)
- [23] Dietterich T G. Machine learning research; four current directions[J]. AI Magazine, 1997, 18(4): 97-136.
- [24] Merz C J. Using correspondence analysis to combine classifiers[J]. Machine Learning, 1999, 36(1): 33-58.
- [25] Gao S S, Zhong Y M, Li W. Random weighting method for multi-sensor data fusion[J]. IEEE Sensors J, 2011, 11(9): 1955-1961.
- [26] 王爱平, 万国伟, 程志全, 等. 支持在线学习的增量式极端随机森林分类器[J]. 软件学报, 2011, 22(9): 2059-2074.
(Wang A P, Wan G W, Cheng Z Q, et al. Incremental learning extremely random forest classifier for online learning[J]. J of Software, 2011, 22(9): 2059-2074.)
- [27] Li G, Qin S G, Zhou D H. Geometric properties of partial least squares for process monitoring[J]. Automatica, 2010, 46(1): 204-210.
- [28] Ye L, Cao Y, Yuan X, et al. Retrofit self-optimizing control of tennessee eastman process[C]. IEEE Int Conf on Industrial Technology. Taipei: IEEE, 2016: 866-871.
- [29] 石向荣. 面向过程监控的非线性特征提取方法研究[D]. 杭州: 浙江大学控制科学与工程学院, 2014.
(Shi X R. Study on nonlinear feature extraction for process monitoring[D]. Hangzhou: College of Control Science and Engineering, Zhejiang University, 2014.)

(责任编辑: 郑晓蕾)