

运用在线贯序极限学习机的故障诊断方法^{*}

尹刚, 张英堂, 李志宁, 程利军

(军械工程学院车辆与电气工程系 石家庄, 050003)

摘要 针对传统的前馈神经网络学习算法泛化能力不高、训练速度慢、易出现局部最优解及无法处理随时间不断变化的信息流等问题,提出了基于在线贯序极限学习机的快速故障诊断方法。针对旋转机械故障复杂、样本少的特点,将测试过程中得到的预测数据加入训练样本,作为下一次预测的已知信息,建立在线贯序极限学习机分类模型,从而在最大程度上提高故障诊断的精度。试验结果表明,在线贯序极限学习机在故障分类准确率与支持向量机相近的条件下,参数选择简单且学习速度提高近200倍。

关键词 极限学习机; 在线神经网络; 旋转机械; 故障诊断; 支持向量机

中图分类号 TP18; TH165.3

引言

旋转机械的快速不解体检测诊断通常采用信号处理和故障模式分类相结合的方法来实现。通过对机械运行过程中产生的信息(如振动、噪声、磨粒信号)提取特征参数,采用模式识别的方法对特征向量进行分类,从而实现旋转机械的故障诊断。

前馈神经网络由于其强大的非线性辨识能力,在发动机、旋转机械等故障诊断中具有重要的应用和较好的发展。基于统计学习理论的支持向量机(support vector machine,简称SVM)以其在小样本数据上的精确拟合和分类能力,在故障诊断等领域得到了快速应用。随着人们对诊断结果要求的提高和机器学习方法的快速发展,逐渐发现基于梯度下降学习的传统前馈神经网络存在的不足,如需要人为设置大量的网络训练参数,迭代次数过多,训练速度慢,泛化能力欠缺,容易产生局部最优解等。支持向量机在实际应用过程中为了获得较为精确的分类结果,通常采用遗传算法、蚁群算法、粒子群算法等进行全局寻优,虽然通过优化提高了故障分类精度,但进化算法却耗费了大量时间调整核函数、惩罚系数、误差控制等参数。极限学习机^[1](extreme learning machine,简称ELM)作为单隐层前馈神经网络(single-hidden layer feedforward neural networks,简称SLFNs)的一种快速学习算法,在训练前只需要

设置网络隐层节点个数和激励函数,算法执行过程中不需要调整网络的输入权值向量和隐层单元偏置阈值,通过解析的方法确定连接隐层和输出层的输出权值,参数选择简单,学习速度快,具有较好的全局搜索能力。

人为设置常见故障类型是目前故障诊断的通用方法。通过采集故障产生的振动或噪声信号,并将信号经过降噪、特征选择和提取后输入到神经网络、支持向量机等分类器中进行训练,然后将训练完成的分类器用于实际运行中的机械故障的诊断^[2-3]。由于旋转机械自身结构的复杂性和其运行环境的恶劣性,导致其产生的故障往往是多种状态并存,提取到的特征参数和单一的故障类别难以一一对应;因此,将人为设置故障类型进行一次性训练得到的故障分类器用于旋转机械实际运行中的故障诊断存在缺陷。现有的故障诊断方法在实际故障的测试中无法对新的故障数据进行再学习,而在实际应用中,学习是一个连续进行的过程,运行中的机器产生的故障信息尤为重要。由于实验设备等条件的限制,人为设置故障训练的分类器和一次性学习往往会遗漏较多的故障信息,故笔者引入在线自适应学习的思想,运用在线贯序极限学习机(online sequential extreme learning machine,简称OS-ELM)^[4]训练前馈神经网络。在采用人为故障训练分类器的基础上,对旋转机械故障类型进行判别,并将旋转机械运行中的故障数据作为训练样本对分类器进行再学习,既实现了

旋转机械主要故障的识别,也提取了旋转机械运行中的故障信息,从而提高了实际运行中旋转机械的故障模式识别精度。

1 极限学习机

含有 N 个训练样本 $Z=(x_i, t_i), i \in 1, 2, \dots, N, L$ 个隐层节点,激活函数为 $g(x)$ 的单隐层前馈神经网络的数学模型为

$$y_i = \sum_{j=1}^L \beta_j g(\omega_j x_i + b_j) \quad (x_i \in R^n, \omega_j \in R^n) \quad (1)$$

其中: β_j 为连接第 j 个隐层节点和输出节点的权值向量; $\omega_j = [\omega_{j1}, \omega_{j2}, \dots, \omega_{jn}]$, 为连接输入节点和第 j 个隐层节点的权值向量; b_j 为第 j 个隐层节点的阈值; $g(\omega_j x_i + b_j)$ 为第 j 个隐层节点在激活函数为 $g(x)$ 、输入向量为 x_i 时的输出。

模型如图 1 所示。

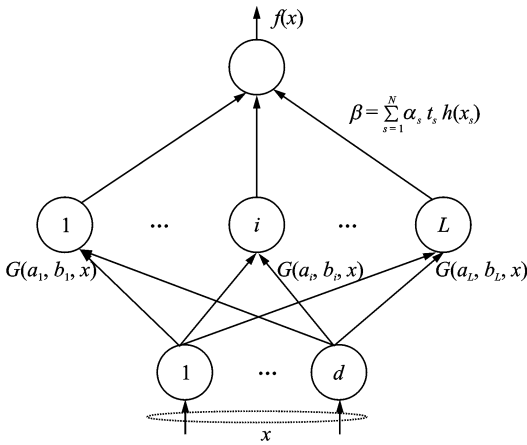


图 1 SLFNs 网络结构

定义式(1)建立的 SLFNs 模型的输出能够以零误差无限接近上述 N 个训练样本,即

$$\sum_{i=1}^N ||y_i - t_i||^2 = 0 \quad (2)$$

因此存在 ω_j, b_j, β_j , 使得

$$\sum_{j=1}^L \beta_j G(\omega_j x_i + b_j) = t_i \quad (i = 1, 2, \dots, N) \quad (3)$$

式(3)可以简化为

$$H\beta = T \quad (4)$$

其中

$$H(\omega_1, \omega_2, \dots, \omega_L; b_1, b_2, \dots, b_L; x_1, x_2, \dots, x_N) = \begin{bmatrix} G(\omega_1, b_1, x_1) \dots G(\omega_L, b_L, x_1) \\ \vdots \quad \dots \quad \vdots \\ G(\omega_1, b_1, x_N) \dots G(\omega_L, b_L, x_N) \end{bmatrix}_{N \times L}; \beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix};$$

$$T = \begin{bmatrix} T_1^T \\ \vdots \\ T_N^T \end{bmatrix}_{N \times m}; H \text{ 为 ELM 的隐层输出矩阵}; H(i, j)$$

为第 i 个训练数据在第 j 个隐层节点的输出。

网络调整的目标为找到一组最优参数 $\hat{\omega}_j, \hat{b}_j, \hat{\beta}_j$, 使得

$$\| \hat{H}\hat{\beta} - T \| = \min \| H\beta - T \| \quad (5)$$

Huang 等^[5]证明,若激活函数 $g(x)$ 无限可微,则网络参数并不需要全部进行调整,其中输入与隐层的连接权值向量 ω_j 以及隐层的偏置值 b_j 可以随机选择。输出权值向量 β_j 可通过求解式(5)的最小二乘解 $\hat{\beta} = H^* T$ 来获得。其中: H^* 表示隐含层输出矩阵 H 的 Moore-Penrose 广义逆。

2 在线贯序极限学习机

传统的神经网络只能处理静止不动的模式。在旋转机械故障诊断中,学习是一个连续进行的过程,故障数据是不断变化的信息流。为了通过不断变化的输入和输出的信息流对旋转机械运行状态实时监控,笔者在 SLFNs 的基础上,将时间的概念引入网络的训练中。采用 OS-ELM 方法训练网络,构建基于在线神经网络的快速诊断方法,其算法流程^[4]如下。

1) 初始化阶段。取 $K=0$, 其中 k 为送到网络的数据段的个数。给定激活函数 $g(x)$, 隐层节点数目 L 和初始训练数据 $Z_0 = \{(x_i, t_i)\}_{i=1}^{N_0} \in N_0 \times L$, 随机产生输入权值向量 ω_j 和阈值 b_j , 初始化网络, 求得初始隐层输出矩阵 H_0 和输出权值向量 $\beta^{(0)} = P_0 H_0^T T_0$ 。其中

$$P_0 = (H_0^T H_0)^{-1}$$

$$T_0 = [t_1, t_2, \dots, t_{N_0}]^T$$

2) 在线学习阶段。给定第 $k+1$ 组数据段, 计算隐层输出矩阵为 H_{k+1} , 输出权值向量为

$$\beta^{(k+1)} = \beta^{(k)} + P_{k+1} H_{k+1}^T (T_{k+1} - H_{k+1} \beta^{(k)})$$

其中

$$P_{k+1} = P_k - P_k H_{k+1}^T (I + H_{k+1} P_k H_{k+1}^T)^{-1} H_{k+1} P_k$$

令 $k=k+1$, 返回到在线学习阶段, 不断更新参数 H 和 β , 直到数据学习完毕。

在线贯序极限学习机包含两个阶段: 初始化阶段, 基于 ELM 方法对小部分故障数据进行学习, 然后丢弃训练数据, 并将训练得到的参数 H 和 β 保存在网络中; 在线学习阶段, 通过变化的故障数据实时更新参数 H 和 β , 从而强化网络的分类和泛化能力, 提高故障诊断的准确度。训练得到的 OS-ELM 参数

存储在网络节点中,可以进行跨平台的移植。将网络的权值参数 β 加载到新的PC机、DSP,ARM等嵌入式环境下,对于新的故障测试数据只需要重新生成测试数据对应的隐层输出矩阵 H ,就可以完成故障模式再学习和分类功能。

3 多故障分类器的建立

引入了时间概念的OS-ELM,可以快速处理大量不断变化的信息流,增加了故障分类的精度和实时性。为了检验OS-ELM在诊断中的效率和泛化性能,笔者构建了4种分类器进行对比试验,OS-ELM与ELM采用相同的多分类方法,因此不再单独构建。

3.1 SVM分类器

旋转机械的故障诊断为多分类问题,由于SVM在分类上只考虑了二值分类的问题,故笔者采用目前较为常用的“一对一”方法进行多故障分类^[6]。对于 n 个故障模式构造 $\lceil n(n-1) \rceil / 2$ 个SVM分类器进行两两分类,根据每个分类器的识别结果对相应模式进行投票,最后票数最多的故障模式即为测试故障数据的所属模式。

3.2 ELM分类器

ELM通过设定与分类类别数相同的输出结点数实现多分类。给定训练样本 $Z=(x_i, t_i), t_i \in R^m$ 表示样本属于 m 类,对于1个训练样本 x ,其对应的输出值 t 有 m 个属性: $(t_1, t_2, \dots, t_m)^T$,若 x 属于故障模式 j ,则 $t_j=1$,其余为 -1 。通过求解目标函数的最小二乘解训练网络,最终得到的输出值 y 同样含有 m 个属性。若 $y_j = \max_{k=1,2,\dots,m} y_k(x)$,则样本 x 为故障模式 j 。

3.3 ELM-OAA分类器

ELM用于多分类需要建立具有多个输出结点的网络结构,若用于二分类则仅需要建立具有两个输出结点的网络结构。由式(4)可知,矩阵 T 的维数从 m 维降到二维,其计算时间复杂度急剧下降,故笔者采用“一对多”的方法建立ELM-OAA分类器,实现ELM多分类。若分类类别数为 m ,则构造 m 个ELM二分类器,每一个ELM二分类器将每一模式与剩下的所有模式区分开来。每一个二分类器的输入数据相同,目标输出不同。对于第 q 个ELM二分类器,类别 q 的目标输出设为1,其余设为 -1 。任意

选择训练样本 x ,ELM-OAA从 m 个二分类器中产生 m 个目标输出值,为确定样本 x 的最终类别,采用最小损失的编码方法^[7]构造判别函数。样本 x 标记为类别 i 的损失为

$$d_L(M(i), y(x)) = \sum_{q=1}^m L(M(i, q) y_q(x)) \quad (6)$$

其中: M 为对角线为1,其余元素为 -1 的 $m \times m$ 矩阵。

样本 x 基于上述损失函数的最终输出类别为:若 $y_k = \min_{i=1,2,\dots,m} d_L(M(i), y(x))$,则样本 x 属于类别 k 。

4 实验与分析

为验证OS-ELM方法的有效性,笔者采用美国凯斯西储大学轴承数据中心的实验数据^[8]进行实验分析。对滚动轴承正常状态、内圈故障、外圈故障、滚动体故障4种状态模式下的振动信号进行模式分类。每种状态提取100组信号,每个模式对应的信号降噪后,经过db6小波包3层分解,得到8个子频带,计算每个子频带的能量,进行归一化后作为该模式最终的特征向量。

将样本数据按照7:3的比例随机分成训练样本集和测试样本集,在2.2GHz的CPU和1G RAM环境下进行实验。鉴于上述提到的SVM参数选择困难问题,笔者采用文献^[6]提到的径向基核函数作为其核函数,核参数 γ 和惩罚参数 c 采用交叉验证的方法搜索选取,得到最优参数为 $c=256, \gamma=1$ 。ELM只需要选择隐层节点数和激励函数,参数的选择相对简单,为了验证隐层节点数对网络学习时间和分类准确率的影响,选择初始节点数为10,以2为周期增加隐层节点数,直至达到与训练样本数相同,激励函数选择为Hardlim函数,网络性能随隐层节点的变化如图2,3所示。实验结果表明,ELM分类器的时间消耗随隐层节点数的增加而快速增加,分类准确率也随隐层节点的增加而增大,且当准确率增加到一定程度时基本趋于平稳。对含有噪声的故障数据降噪不彻底和ELM算法仅考虑经验风险最小化,训练时受到部分离群点的影响导致了测试分类准确率不能进一步提高。为验证激励函数对网络性能的影响,笔者选择sine, sigmoidal, hardlim3种函数进行比较。如图4所示,Hardlim函数在隐层节点相对较少的情况下,分类误差与其他激励函数相比较较大。

OS-ELM分类器是在ELM分类器的基础上,增加时间的概念,对连续出现的故障数据进行分类

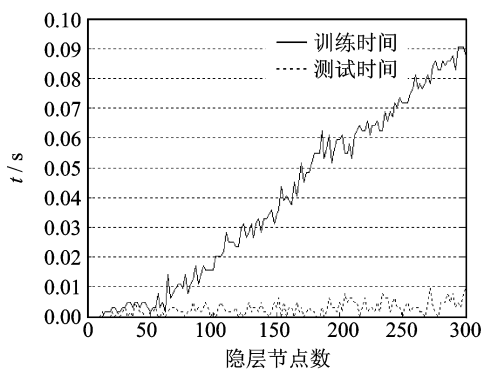


图2 时间消耗随隐层节点的变化

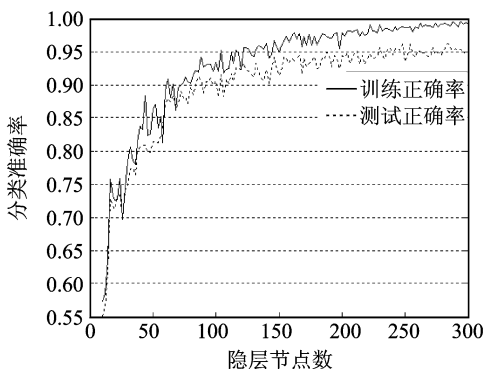


图3 分类准确率随隐层节点的变化

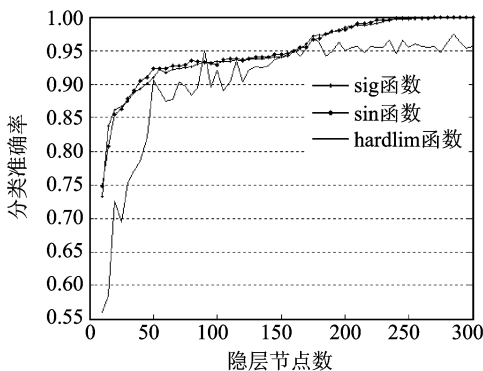


图4 不同激励函数对分类准确率的影响

ELM-OAA 分类器是重复运用ELM的二分类功能进行多故障模式的分类,因此其参数的选择在一定程度上与ELM相同。初始训练样本数设为50,以50为数据段样本数,模拟连续变化的故障信息流训练OS-ELM,选择sig函数作为后两种分类器的激励函数。结合以上对4种分类器参数的分析,选择针对每种分类器的最优参数,得到分类器的性能对比如表1,2所示。

由表1可知,OS-ELM分类准确率和ELM, SVM基本相同,说明该算法可用于旋转机械的故障

表1 4种分类器分类准确率对比

分类器	训练准确率	测试准确率	节点数
SVM	1.000 0	0.967 0	30
ELM	0.991 0	0.949 5	280
ELM-OAA	0.993 5	0.908 9	20
OS-ELM	0.995 0	0.965 0	30

表2 4种分类器时间消耗对比

分类器	训练时间/s	测试时间/s	节点数
SVM	16.296 9	0.100 0	30
ELM	0.087 5	0.009 4	280
ELM-OAA	0.001 6	0.000 8	20
OS-ELM	0.015 6	0.001 6	30

诊断。通过对比3种分类器的训练时间可知,OS-ELM的分类效率远远大于另外两种分类器。由表2可知,ELM消耗的时间是OS-ELM的5倍左右。这是因为OS-ELM采用连续训练的方式,需要的隐层节点数较少,从而算法的时间复杂度明显减小,提高了分类效率。ELM-OAA在效率上明显优于其他分类器,但由于“一对多”多分类方法存在如图5所示的分类盲区,因此其分类准确率还有待提高。

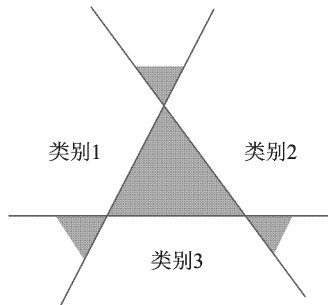


图5 “一对多”分类盲区

故障信息流采用OS-ELM算法学习完成后,学习得到的网络参数存储在 H 和 β 中。若新的故障信息流产生,或故障数据存储在其他处理终端,只需将已训练得到的 H 和 β 移植过来,而不需要对旧的故障数据进行再学习。在故障信息学习过程中既不需要存储故障数据,又可利用原有故障数据学习得到的先验知识,提高了分类器学习的效率和分类准确率。

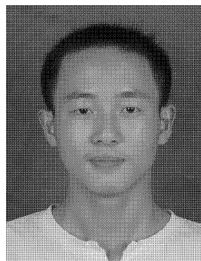
5 结束语

笔者针对旋转机械故障样本少、故障模式种类多、故障数据为不断变化的信息流等问题,建立了基于在线贯序极限学习机的快速故障分类模型。在

励函数确定的情况下,OS-ELM只需确定前馈神经网络的隐层节点数,参数选择过程快速简单。在算法执行过程中不需要调整输入权值向量、隐层阈值等变量,算法简单易学。当网络学习完成后,可将其训练得到的权值参数 H 和 β 移植到新的环境下,对新的故障数据流进行学习和判别,从而有效提高了分类准确率和泛化能力。将其与SVM算法相比较,在分类准确率基本相同的情况下,参数选择简单,学习消耗的时间和空间远远小于SVM;因此,将OS-ELM算法用于机械故障的快速诊断是有效可行的,且总体性能优于SVM。

参 考 文 献

- [1] Huang Guangbin, Zhu Qinyu, Siew C K. Extreme learning machine: theory and applications[J]. *Neurocomputing*, 2006, 70(1-3): 489-501.
- [2] 姚良,李艾华,王涛. 运用改进二叉树SVM算法的柴油机振动诊断[J]. *振动、测试与诊断*, 2010, 30(6): 689-693.
Yao Liang, Li Aihua, Wang Tao. Diesel vibration diagnosis using improved binary tree support vector machine algorithm[J]. *Journal of Vibration, Measurement & Diagnosis*, 2010, 30(6): 689-693. (in Chinese)
- [3] 李辉,郑海起,唐力伟. 基于倒谱分析的轴承故障诊断研究[J]. *振动、测试与诊断*, 2010, 30(4): 353-356.
Li Hui, Zheng Haiqi, Tang Liwei. Application of bi-cepstrum technique to bearing fault detection[J]. *Journal of Vibration, Measurement & Diagnosis*, 2010, 30(4): 353-356. (in Chinese)
- [4] Liang Nanying, Huang Guangbin, Saratchandran P, et al. A fast and accurate online sequential learning algorithm for feedforward networks[J]. *IEEE Transactions on Neural Networks*, 2006, 17(6): 1411-1423.
- [5] Huang Guangbin, Zhu Qinyu, Siew C K. Extreme learning machine: a new learning scheme of feedforward neural networks[C]// 2004 IEEE International Joint Conference on Neural Networks. Budapest, Hungary: [s. n.], 2004.
- [6] 范红波,张英堂,孙焯. 小波包和SVM在发动机故障诊断中的应用[J]. *车用发动机*, 2006(4): 49-53.
Fan Hongbo, Zhang Yingtang, Sun Ye. Application of wavelet packet and support vector machines in the engine fault diagnosis[J]. *Vehicle Engine*, 2006(4): 49-53. (in Chinese)
- [7] Allwein E L, Schapire R E, Singer Y. Reducing multiclass to binary: a unifying approach for margin classifiers [J]. *The Journal of Machine Learning Research*, 2001(1): 113-141.
- [8] Kenneth A L. Bearing data center [EB/OL]. (2010-12-20). <http://www.eecs.cwru.edu/Laboratory/bearing>.



第一作者简介:尹刚,男,1988年8月生,硕士研究生。主要研究方向为车辆试验分析与智能故障诊断。曾发表《基于在线半监督学习的故障诊断方法研究》(《振动工程学报》2012年第25卷第6期)等论文。

E-mail: gang.gang88@163.com