

文章编号: 1000-4882(2006) 03-0064-06

模糊支持向量机在船舶柴油机故障诊断中的应用

朱志宇, 张 冰, 刘维亭

(江苏科技大学电子信息学院, 江苏 镇江 212003)

摘 要

探讨了支持向量机(SVM)的机理,并通过计算每一个样本的模糊隶属度,引入了模糊支持向量机(FSVM)的概念,解决了SVM中的不可分区域问题;应用FSVM对船舶电站主柴油机进行故障诊断,研究了FSVM参数的选择方法。仿真结果表明,FSVM具有较好的诊断效果和较强的抵抗噪声的能力。

关 键 词: 船舶、舰船工程;模糊支持向量机;故障诊断;模糊隶属度;柴油机

中图分类号: TP18 TP206.3; U664.121 **文献标识码:** A

1 引 言

加强船舶电站设备的故障诊断,保证其可靠运行,具有十分重要的工程意义。由于神经网络具有较强的模式识别能力,因而在故障诊断研究方面获得了一定的重视^[1,2]。但是,神经网络方法的不足在于:网络结构难以确定、局部最优、容易导致过学习或泛化能力差等。

而近几年发展起来的支持向量机(Support Vector Machine,缩写为SVM)则具有理论上可保证的最佳泛化性^[3,4]。SVM是Vapnik等人根据统计学习理论的VC维理论和结构风险最小原理而提出的一种新的机器学习方法。它根据有限的样本信息在模型的复杂性(即对特定训练样本的学习精度)和学习能力(即无错误地识别任意样本的能力)之间寻求最佳折衷,以期获得最好的推广能力,能较好地解决小样本、非线性、高维数和局部极小点等实际问题^[5],因此成为机器学习界研究的热点,并成功地应用于分类、函数逼近和故障诊断等方面^[6~9]。

故障诊断的瓶颈之一是故障样本的缺乏,鉴于SVM能在训练样本很少的情况下很好地达到分类推广的目的,作者尝试采用SVM对船舶电站主柴油机进行故障诊断,同时考虑到SVM可能出现不可分区域以及噪声对样本数据的污染,将模糊思想和SVM相结合,进一步优化SVM的分类面。

2 支持向量机

支持向量机最初用于解决模式识别问题。在模式识别中,为了发现具有推广能力的决策规则,所选择的训练数据的一些子集称为支持向量。最佳的支持向量分离等效于所有数据的分离。支持向量机是从线性可分情况下的最优分类面发展而来的。支持向量机形式上类似于一个神经网络,输出是中间节点的线性组合,每个中间节点对应于一个支持向量。

故障诊断的实质是模式识别问题。对于线性可分的两类模式识别问题,设 n 个训练样本 $(x_i, y_i), y_i \in \{-1, 1\}, i = 1, 2, \dots, n$,假设存在分离超平面 $w^T x_i + b = 0$,使得对于每个样本有

$$\begin{aligned} w \cdot x_i + b &\geq 1, y_i = 1 \\ w \cdot x_i + b &\leq -1, y_i = -1 \end{aligned} \quad (1)$$

式中 b 为常数。样本到超平面的距离,即分类间隔可定义为 $1/\|w\|$,使分类间隔最大等价于使 $\|w\|$ 最小,此时的分离超平面即为最优分类面。

考虑到有一些样本不能被超平面正确分类,引入一个松弛变量 $a_i \geq 0, i = 1, 2, \dots, n$,此时超平面约束变成:

$$y_i [w \cdot x_i + b] - 1 + a_i \geq 0 \quad (2)$$

则最优分类问题可转变为

$$\begin{aligned} \min & \left[\frac{1}{2} \|w\|^2 + C \sum_i a_i \right] \\ \text{s. t. } & y_i [w \cdot x_i + b] - 1 + a_i \geq 0 \end{aligned} \quad (3)$$

目标函数式中的第一项使分类间隔尽量大,从而提高推广能力;第二项则使误差尽量小。其中 $C > 0$ 为权值系数,它实际上控制对错分样本惩罚的程度。 C 越大,则对错分样本惩罚的程度也就越大,等价于要求实现较小的分类间隔;当 $C = \infty$ 时,表示要求无错分样本,显然这时算法的训练时间就会很长。

上述优化问题可转化为一个可用 Lagrange 乘子法求解的、如下式所示的二次规划 (QP) 问题:

$$\begin{aligned} \max W(T) &= \sum_i T_i - \frac{1}{2} \sum_{i,j} T_i T_j y_i y_j (x_i \cdot x_j) \\ \text{s. t. } &\sum_i y_i T_i = 0, \quad 0 \leq T_i \leq C, i = 1, 2, \dots, n \end{aligned} \quad (4)$$

式中 $(x_i \cdot x_j)$ 表示欧式空间中的内积, T_i 为与第 i 个样本对应的 Lagrange 乘子,实质上是凸优化问题的解,与 $T_i > 0$ 对应的样本点 x_i 就是支持向量。

对于非线性问题,可以通过非线性变换 $H(x): R^d \rightarrow F, F$ 为高维内积空间,即特征空间,然后在 F 中构造最优分类面,这样便仍可用线性分类器实现分类。此时的目标函数为

$$W(T) = \sum_i T_i - \frac{1}{2} \sum_{i,j} T_i T_j y_i y_j K(x_i, x_j) \quad (5)$$

相应的判别函数为

$$y_i = \operatorname{sgn} \left[\sum_{x_j \in w} T_j K(x_i, x_j) + b \right] \quad (6)$$

式中 $K(x_i, x_j)$ 为满足 Mercer 条件的核函数,它对应某一变换空间的内积。当采用不同的核函数时,支持向量机的算法不同,目前常用的核函数主要有列三种:

阶次为 d 的多项式核函数: $K(x_i, x_j) = [(x_i \cdot x_j) + 1]^d$;

径向基核函数: $K(x_i, x_j) = \exp \left[-\frac{\|x_i - x_j\|^2}{c^2} \right]^d$;

两层神经网络核函数: $K(x_i, x_j) = \tanh[a(x_i \cdot x_j) + t]^d$; 式中 a 和 t 为常数。

3 模糊支持向量机

目前的支持向量机方法本质上是一个两类划分问题。通常在使用支持向量机时,将 n 类问题转化为 n 个两类问题来处理,因此,一般的支持向量机可能存在着不可分区域。此外,在其应用中,它要求每一个样本必须要属于某一类。然而在现实应用中,每一个样本的影响是不同的,某些样本可能要比其他的样本重要,因此更需要那些重要样本被正确分类而不在意其他样本(例如干扰)是否被错分。本文尝试将模糊思想和支持向量机相结合,通过求解每个样本对各个类的模糊隶属度,解释支持向量机的学习结果,进一步优化其分类面,因而,能够提高支持向量机的学习分类能力。同时模糊支持向量机 (FSVM) 可以忽略非重要样本或噪声对 SVM 学习的影响,从而提高了 SVM 抵抗噪声的能力。

对满足式 (4) 的每一个样本 x_i , 在得到相同的分类结果的情况下引入了样本模糊隶属度 $0 < m_i < 1$ 。

且 m_i 越大,表明其对应的样本越重要,或隶属于某一类的可能性越大。当所有样本的 m_i 都设置为 1 时,此时 FSV M 就退化为一般的 SVM。

引入了样本模糊隶属度 m_i 后,最优分类问题变为

$$\min \left[\frac{1}{2} \|w\|^2 + C \sum_i m_i \eta_i \right] \quad (7)$$

$$\text{s. t. } y_i [w \cdot x_i + b] - 1 + \eta_i \geq 0$$

$m_i \eta_i$ 用来度量对具有不同重要性的样本错分的惩罚程度。为了解决上述的优化问题,定义 Lagrange 函数 $L(w, b, \eta, T, V_i)$ 为

$$L(w, b, \eta, T, V_i) = \frac{1}{2} \|w\|^2 + \sum_i m_i \eta_i - \sum_i T_i (y_i [w \cdot x_i + b] - 1 + \eta_i) - \sum_i V_i \eta_i \quad (8)$$

式中的 $V_i \geq 0$ 为 Lagrange 乘子。Lagrange 函数对 w, b, η 分别求偏导数并令求导结果等于 0,则可得到

$$\frac{\partial L(w, b, \eta, T, V_i)}{\partial w} = 0$$

$$\frac{\partial L(w, b, \eta, T, V_i)}{\partial b} = 0 \quad (9)$$

$$\frac{\partial L(w, b, \eta, T, V_i)}{\partial \eta_i} = 0$$

从而得到 $w = \sum_i T_i y_i x_i, \sum_i T_i y_i = 0, m_i C = T_i - V_i$ 。

将式 (9) 代入式 (8), 得到式 (7) 的对偶优化问题:

$$\max W(T) = \sum_i T_i - \sum_{i,j} T_i T_j y_i y_j K(x_i, x_j) \quad (10)$$

$$\text{s. t. } \sum_i y_i T_i = 0, \quad 0 \leq T_i \leq m_i C, \quad i = 1, 2, \dots, n$$

对应于 $0 < T_i < m_i C$ 的支持向量(样本)位于超平面 $1/\|w\|$ 的距离处;而 $T_i > m_i C$ 所对应的支持向量就是被错分的样本。

在具体的应用中,首先要依据训练数据的具体特征,确定样本的模糊隶属度的下界 m_{inf} 。然后再选择一个适当的隶属度函数,建立起样本 x_i 与隶属度 m_i 之间的联系,最后,根据 FSV M 理论,求出最优分类面和分类函数。

4 船舶主柴油机故障诊断

船舶柴油机是一个复杂的机电系统,采用整机诊断模型非常复杂和困难。如果仅采用一个支持向量机会使网络规模庞大,学习训练困难。考虑到系统故障具有层次性的特点,于是将整机分成增压系统、燃烧系统、润滑系统、燃油系统、冷却系统和传动系统等子系统^[10],如图 1 所示。因此,本文采用层次诊断模型方法:首先由整机性能参数的恶化状态判断故障可能发生在哪个子系统,称为第一层次诊断;然后对该子系统进一步诊断出故障的原因、位置和程度,称为第二层次诊断。

限于篇幅,本文仅介绍对冷却子系统进行故障诊断的结果。

4.1 船舶主柴油机冷却系统故障机理

冷却系统是船舶主柴油机的一个子系统。该系统属于船舶机舱高温淡水冷却回路。它用海水作为冷介质,通过热交换器去冷却高温淡水;经冷却后的高温淡水再去冷却主柴油机。该子系统的故障征兆可由以下六个可测的工艺参数表征: T_1 为淡水温度 ($^{\circ}\text{C}$); T_2 为海水温度 ($^{\circ}\text{C}$); P_1 为淡水压力 (MPa);

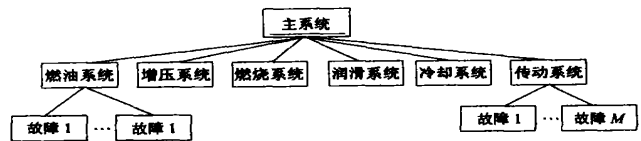


图 1 主柴油机故障诊断系统的层次分类模型

P_2 为海水压力(MPa); Q_1 为淡水流量(t/h); Q_2 为海水流量(t/h)以上六个参数作为SVM输入变量。考虑该系统的正常情况和常见的15种故障,并表示如下:

F_0 :系统正常; F_1 :汽缸超载; F_2 :汽缸后燃; F_3 :淡水阀关闭或损坏; F_4 :淡水泵压力低; F_5 :淡水管系泄漏; F_6 :淡水泵压力高; F_7 :淡水旁通阀开度小; F_8 :淡水管系堵塞; F_9 :冷却器管系堵塞; F_{10} :海域气候; F_{11} :海水泵压力低; F_{12} :海水滤器堵塞; F_{13} :海水泵压力高; F_{14} :海水旁通阀开度小; F_{15} :海水阀堵塞

每个输入数据用3种状态来表示,即Normal High Low,分别记为 N H L 。为了防止测量值的波动对SVM的学习过程的影响,对样本数据做如下归一化处理: H 对应1, N 对应0, L 对应-1对每一种故障采集3个样本数据为了反映机组运行负荷与故障之间的对应关系,参考文献[11],考虑额定负荷(100% MCR)、部分负荷(90% MCR 75% MCR)和半负荷(50% MCR)四种工况;再考虑船舶航行环境的大气温度变化,选取有代表性的大气温度为288K 300K 312K,最后共得到576个样本,将前384个数据样本用于训练,后192个数据样本作为检验样本

4.2 模糊支持向量机及其参数的确定

关于如何设置SVM的各个参数,理论上还没有统一的论述,往往要针对待解决的实际问题,具体选择。在本文的故障诊断中,采用RBF核函数,也称为高斯函数:

$$k(x_i,x_j)=\exp\left\{-\frac{\|x_i-x_j\|^2}{\epsilon^2}\right\},\quad i=1,2,\cdots,n$$

(11)

式中 ϵ 是感知变量,其数值可以自由选择,它决定了该高斯函数围绕中心点的宽度; ϵ 的选择对SVM的学习至关重要,其数值可以控制支持向量机的个数,当支持向量机个数太多时,可以适当减小 ϵ 的数值,反之亦然。由于目前还没有统一的方法来确定 ϵ 的数值,往往要根据特征数据分析的结果取经验数值。为简单起见,本文采用以下经验公式确定 ϵ 的数值:

$$\epsilon^2=E(\|x_i-x_j\|^2)$$

式中 E 表示数学期望。

参数 C 的选择: $C>0$ 为权值系数,它控制对错分样本的惩罚程度。 C 越大,等价于越小的分类间隔,但是训练时间将延长。目前还没有指导如何选择 C 的理论,只能折衷考虑诊断精度和学习时间之间的矛盾,通过试验找到对特定训练集最好的 C 的数值,本文取 $C=120$

模糊隶属度的选择:样本的模糊隶属度的下界 m_{inf} 的确定方法可以参见文献[12]。将模糊隶属度视作时间的单调线性函数,即

$$m_i=f(t)=at+b$$

(12)

式中 $t_1\leq t_2\leq\cdots\leq t_n$ 是数据采集的时刻,时间越靠后,数据越重要,即第一个样本的隶属度为 $f(t_1)=m_{inf}>0$,最后一个样本的隶属度 $f(t_n)=1$ 。将上述两个边界条件代入式(12)后,得到 a 和 b 的数值,最后得到:

$$m_i=f(t)=\frac{1-m_{inf}}{t_n-t_1}t+\frac{t_nm_{inf}-t_1}{t_n-t_1}$$

(13)

4.3 仿真结果

采用获得的样本数据对FSVM进行训练和测试,表1给出了部分故障诊断的结果

表1 FSVM 故障诊断结果

序号	模拟故障	特 征 征 兆						诊断结果
		T_1	P_1	Q_1	T_2	P_2	Q_2	
1	F_1F_2	H	N	N	H	L	L	F_1F_2
2	F_1F_2	H	0.9N	1.2N	0.7H	0.8L	L	F_1F_2
3	F_1F_2	0.6H	0.9N	N	0.8H	L	1.3L	F_1F_2
4	F_6	H	1.3L	N	0.8H	N	N	F_6

续 表 1

序号	模拟故障	特 征 征 兆						诊断结果
		T_1	P_1	Q_1	T_2	P_2	Q_2	
5	$F_3F_4F_5$	H	1. 2L	N	0. 7N	0. 7N	0. 6N	$F_3F_4F_5$
6	F_6F_7	1. 1L	0. 8H	0. 8N	0. 8N	0. 8N	0. 8N	F_6F_7
7	F_8F_9	0. 8H	1. 2N	0. 8L	0. 8N	0. 8N	0. 8N	F_8F_9
8	$F_{11}F_{12}$	H	0. 8V	0. 8N	0. 8N	0. 8L	0. 8N	$F_{11}F_{12}$
9	$F_{14}F_{15}$	1. 2H	0. 8V	1. 1N	1. 1N	1. 1N	1. 1L	$F_{14}F_{15}$
10	F_{10}	0. 8H	0. 9V	0. 8N	0. 8H	0. 8N	0. 8N	F_{10}

表 1的结果表明 , FSV M 能够正确对故障样本进行分类 ,同时对偏离训练样本的输入数据依然具有较强的处理能力。例如 ,当学习输入样本为 $X = \{T_1, P_1, Q_1, T_2, P_2, Q_2\} = \{H, N, N, H, L, L\}$ 时 ,故障诊断的结果为复合故障 F_1F_2 ,在测试样本中叠加噪声 ,使该样本偏离上述标准输入数值 ,如表 1中第 2行和第 3行所示 ,重新测试后 ,依然能够得到正确的诊断结果 ,说明 FSV M 具有较强的抵抗噪声的能力。

为了说明 FSV M 的故障诊断能力 ,分别采用 FSV M、SVM 和 RBF网络 ,对同一批训练样本和测试样本 (叠加噪声)进行学习 ,将出现误诊的样本数目作为最终结果 ,如表 2所示。表 2中的结果表明 , FSV M 对复合故障样本有较高的诊断准确度 ,而 RBF神经网络对复合故障样本诊断的准确度不高 ,这也从另一个侧面说明 FSV M 具有很强的分类能力 ;同时 FSV M 的分类能力也比 SVM 强。

表 2 FSV M、SVM 和 RBF网 络的误 诊断 样本数 目 比较

故 障 类 型		F_1	F_6	F_{10}	F_1F_2	F_6F_7	F_8F_9	$F_{11}F_{12}$	$F_{14}F_{15}$	$F_3F_4F_5$	$F_{12}F_{13}F_{14}$
误诊断 样本数	RBF	0	0	1	2	3	3	5	5	10	12
	SVM	0	0	0	1	1	2	2	2	3	4
	FSVM	0	0	0	0	0	0	0	0	0	1

5 结 语

本文探讨了支持向量机的机理 ,尝试将模糊思想和支持向量机相结合 ,通过求解每个样本对各个类的模糊隶属度 ,解释各个样本的重要性程度 ,从而能够提高支持向量机的学习分类能力 ,解决了 SVM 中的不可分区域问题。应用 FSV M 对船舶电站主柴油机进行故障诊断 ,并讨论了模糊隶属度和 FSV M 参数的选择方法。仿真结果表明 , FSV M 具有较好的诊断效果和较强的抵抗噪声的能力。

参 考 文 献

1 马立玲. 一类基于神经网络非线性观测器的鲁棒故障检测和诊断 [J]. 控制与决策, 2003, 18(3): 309-313.
2 郝而娜,王 伟,吴光宇,等. 粗糙集-神经网络故障诊断方法研究 [J]. 东北大学学报(自然科学版), 2003, 24(3): 252-255.
3 Cambell C. An introduction to Kernel methods[A]. Radial Basis Function Networks: Design and Application[C]. Berlin: Springer V erlag, 2001. 155-192.
4 Scholkopf B. Statistical Learning and Kernel methods[R]. MSR-TR-2000-23.
5 Vapik V N. 统计学习理论的本质 [M]. 张学工,译. 北京: 清华大学出版社, 2000.
6 郝 英,孙健国,白 杰. 航空燃气涡轮发动机气路故障诊断现状与展望 [J]. 航空动力学报, 2003, 18(6): 753-760.

7 翟永杰,韩 埃,王东风,王国鹏. 基于损失函数的 SVM 算法及其在轻微故障诊断中的应用 [J]. 中国电机工程学报, 2003, 23(9): 198-203.

8 郝 英,孙建国,白 杰. 基于支持向量机及油中溶解气体分析的大型电力变压器故障诊断模型研究 [J]. 中国电机工程学报, 2003, 23(7): 88-92.

9 胡寿松,王 源. 基于支持向量机的非线性系统故障诊断 [J]. 控制与决策, 2001, 15(5): 617-620.

10 周东华,叶银忠. 现代故障诊断与容错控制 [M]. 北京: 清华大学出版社, 2000 144-156.

11 黄加亮,蔡振雄. 船舶柴油机运行工况诊断仿真研究 [J]. 船舶工程, 2002, (6): 31-34.

12 Lin C F, Wang S D. Fuzzy support vector machines[J]. IEEE Trans. on Neural Networks, 2002, 13(2): 464-471.

Application of Fuzzy Support Vector Machine in Fault Diagnosis of Ship Diesel Engine

ZHU Zhi-yu, ZHANG Bing, LIU Wei-ting

(Department of Electronics and Information,
Jiangsu University of Science and Technology, Zhenjiang Jiangsu 212003, China)

Abstract

The mechanism of support vector machine (SVM) is discussed in this paper. Through calculating fuzzy membership of every sample, the conception of fuzzy support vector machine (FSVM) is introduced, which can resolve the problem of unclassifiable regions in SVM. Then FSVM is applied to diagnose the fault in ship diesel engine, and the methods to choose parameters of FSVM is studied. Simulation results show that FSVM has better diagnosis effect and stronger ability to resist noise in samples.

Key words ship engineering; fuzzy support vector machine; fault diagnosis; fuzzy membership; diesel engine

作者简介

朱志宇 男, 1971年生,博士,副教授. 主要从事智能控制、信号处理方面的研究工作。

张 冰 女, 1967年生,博士,副教授. 主要从事智能信号处理、船舶运动控制方面的研究工作。

刘维亭 男, 1966年生,博士,教授. 主要从事鲁棒控制、电力传动控制方面的研究工作。