

# 基于 CSO-SVM 的轴承健康状态评估研究

贾萌珊<sup>1</sup>, 齐子元<sup>1</sup>, 薛德庆<sup>1</sup>, 朱常安<sup>1,2</sup>

(1. 陆军工程大学石家庄校区, 石家庄 050003; 2. 中国人民解放军 61035 部队, 北京 100094)

**摘要:** 支持向量机是基于统计学理论的机器学习算法, 在解决高维、局部极值和结构选择问题中具有优势, 广泛应用于数据挖掘中; 但是其核宽度和惩罚因子的选择直接关系到支持向量机分类结果; 针对上述问题, 可采取优化算法对该参数进行优化, 达到提高支持向量机分类精度的目的; 鸡群优化算法是近年新提出的一种全局优化算法, 具备结构清晰, 全局搜索能力等优点, 在优化问题中得到广泛应用; 基于此, 提出一种基于鸡群优化的支持向量机模型 (CSO-SVM) 的健康状态评估方法, 并应用在轴承健康状态评估领域中; 结果表明, 基于 CSO-SVM 的轴承健康状态评估精度达到 97%, 明显优于基于传统机器学习模型的健康状态模型的评估精度, 具有更好的健康状态识别效果。

**关键词:** 支持向量机; 机器学习; 鸡群优化算法; 健康状态评估; 轴承

## Health Status Identification of Bearing Based on CSO-SVM

JIA Mengshan<sup>1</sup>, QI Ziyuan<sup>1</sup>, XUE Deqing<sup>1</sup>, ZHU Changan<sup>1,2</sup>

(1. Shijiazhuang Campus of Army Engineering University, Shijiazhuang 050003, China;  
2. Unit 61035 of PLA, Beijing 100094, China)

**Abstract:** Support vector machine is the machine learning algorithm based on statistical theory. It has the advantages of solving high-dimensional, local extremum and structure selection problems. It is widely used in machine learning and data mining. However, the selection of its kernel width and penalty factor is directly related to the classification results of support vector machine. To solve the above problems, the optimization algorithm is used to optimize the parameters and improve the classification accuracy of support vector machine to a certain extent. Chicken swarm optimization algorithm is a new global optimization algorithm proposed in recent years, It has clear structure and excellent global search ability, It is widely used in optimization problems. Based on this, a support vector machine model based on chicken swarm optimization (CSO-SVM) is proposed, which is used to evaluate the bearing health status. The results show that the accuracy of bearing health state evaluation based on the CSO-SVM is up to 97%, which is much higher than that of the health state model based on traditional machine learning model, and has better health state recognition effect.

**Keywords:** support vector machine; machine learning; chicken swarm optimization algorithm; health status assessment; bearing

## 0 引言

健康状态评估方法包括基于物理模型的方法和基于数据驱动的方法。基于物理模型的健康状态评估需要对机械系统深入了解, 这对于复杂机械系统存在着建模难度大、验证困难等问题。近几年, 随着人工智能技术的迅猛发展, 基于数据驱动的健康状态评估得到了广泛应用。

轴承在运转一段时间后会 出现一系列问题, 如磨损或疲劳剥落等, 从而影响机器的正常工作。设备大多数故障属于渐变故障, 一般代表着因为损失或磨损引起的输出参数变化导致机械产生故障的过程, 这类故障产生的概率与机械运转的时间相关, 运转时间越长, 发生故障的概率越大, 可以通过仪器事先进行测试和监控。轴承的故障多由磨损引起, 属于渐变故障, 其耗损程度与使用时间呈正相关。因此, 分析其故障机理以及研究其演化过程, 可以有效控制和预防故障的出现。对轴承进行健康状态评估能够

降低安全风险和维修成本, 保证机械系统的安全运行。越来越多的研究人员专注于轴承健康状态评估, 并取得了一系列成果。西北工业大学的史晓雪和吴亚锋等人提出了基于 AGPF 的滚动轴承性能衰退趋势预测, 准确预测出了轴承性能衰退趋势, 有利于进一步开展预测性维护<sup>[1]</sup>。

随着人工智能的发展, 许多学者将基于数据驱动的健康状态评估方法引入到了轴承的健康状态评估过程。1992 年 Boser 等人首次提出支持向量机 (support vector machine, SVM) 用于解决二分类问题。随着对 SVM 的不断改进, SVM 能够完成回归 (support vector regression, SVR) 和分类 (support vector classification, SVC) 等学习任务, 前者多用于分类, 后者多用于预测, 因此, SVM 成为了机器学习领域典型方法之一<sup>[2]</sup>。

SVM 的基本思想是将原数据样本通过一个非线性映射函数从原始空间映射到高维空间, 并在高维空间寻找一个

收稿日期: 2022-04-13; 修回日期: 2022-05-31。

作者简介: 贾萌珊 (1996-), 女, 河北人, 硕士研究生, 主要从事装备健康状态评估、故障预测方向的研究。

通讯作者: 齐子元 (1980-), 男, 山东人, 博士, 副教授, 主要从事数字信号处理、智能检测与诊断方向的研究。

引用格式: 贾萌珊, 齐子元, 薛德庆, 等. 基于 CSO-SVM 的轴承健康状态评估研究[J]. 计算机测量与控制, 2022, 30(9): 242-248, 254.

分类面。能够将样本分隔开, 与最近的样本距离最大的分类面称为最优超平面。与传统的机器学习相比, SVM 能够解决小样本问题、高维问题、局部极值问题和结构选择问题, 被广泛应用于自然语言处理、计算机视觉等多个研究领域<sup>[3]</sup>。

但利用 SVM 进行解决故障诊断和健康状态评估等问题时, 又存在着需要人工设置参数而导致自适应能力差的问题, 进而导致分类精度的下降。因此, 研究人员提出了使用优化算法对 SVM 的参数进行寻优选取, 以提高 SVM 的分类精度。

宋立业、孙琳提出了基于 EEMD-GSSA-SVM 的滚动轴承故障诊断, 结果表明, 利用全局麻雀群搜索算法能够有效提高支持向量机的分类精度, 得出了精确的滚动轴承故障诊断结果<sup>[4]</sup>。张吴飞、李帅帅等人提出了基于 IGWO-SVM 的轴承故障分类预测, 通过改进灰狼算法收敛支持向量机的参数, 并对数据进行训练优化, 与其他主流分类模型相比分类精度进一步提高, 验证了该算法的优越性<sup>[5]</sup>。时培明等人提出了基于粒子群优化的支持向量机齿轮智能故障诊断模型。该模型现了中速轴大齿轮不同故障类型的可靠识别, 验证了粒子群算法对支持向量机参数优化的可靠性<sup>[6]</sup>。

因此, 利用智能优化算法对其参数进行寻优, 能够提高支持向量机的分类精度。应用在轴承健康状态评估领域中, 可以提高对轴承健康状态评估的准确度, 使其结果更加可靠。

## 1 数据预处理

### 1.1 经验模态分解 (empirical mode decomposition, EMD)

从数据采集系统获得的数据通常为时域信号, 且由于采集系统自身或环境因素影响, 在信号中含有噪声, 需要对数据进行适当的处理后再进行特征提取。

相对于常用于信号降噪处理的小波变换, 经验模态分解不仅吸收了小波变换多分辨率的优势, 还克服了小波变换中需选取小波基与确定分解尺度的困难。EMD 是一种基于信号局部特征的信号分解方法, 该算法不需要信号的频率、幅度等先验知识即可对信号进行分解处理, 在分析非平稳非线性信号方面具有独特的优势<sup>[7]</sup>。轴承的振动信号是一种非线性、非平稳信号, 因此 EMD 分解在处理轴承振动信号时具有一定的优势。

经验模态分解的主要思想是自适应地将输入信号分解为有限的单一频率零均值基本分量, 这些基本分量称为固有模态函数 (intrinsic mode function, IMF), 且每一个 IMF 分量都是独立的。利用 EMD 分解, 可以对轴承信号按照真实存在的不同尺度或趋势分量逐级分解, 将轴承信号分解为一系列具有相同特征尺度的数据序列。与原始信号相比, 具有更强的规律性。

EMD 的分解式为:

$$x(t) = \sum_{i=1}^n b_i(t) + r_i(t) \quad (1)$$

IMF 应满足条件<sup>[8]</sup>:

1) 极值点的数目等于过零点的数目, 或者二者差值相差为 1;

2) 包络线平均值在任何点上均为零 (分别用极大值和极小值组成上包络线和下包络线)。

EMD 的核心理念是提取出具有更高原始波形的相似性, 同时也包含了更小的冲击信息的 IMF 成分, 从而消除影响判断的噪声数据。

### 1.2 峭度准则

轴承的故障信号中往往存在着瞬时冲击。峭度是一个能够反应信号瞬时冲击强弱的时域参数, 对该类冲击特征非常敏感, 与轴承的尺寸、载荷、转速无关, 因此常用于分析轴承振动成分的强弱和轴承故障诊断。

谱峭度 (spectral kurtosis, SK) 是一种能够描述轴承发生故障时的冲击特性频率变化的统计参数, 能够有效的检测出轴承的故障冲击信号<sup>[9-10]</sup>。

谱峭度的计算公式:

$$K = \frac{\langle |H(t, f)|^4 \rangle}{\langle |H(t, f)|^2 \rangle^2} - 2 \quad (2)$$

式中,  $H(t, f)$  为信号  $x(t)$  在频率  $f$  处的复包络, 由短时傅里叶变换计算。 $\langle \cdot \rangle$ ,  $|\cdot|$  分别代表数学期望和模。

其中:

$$H(t, f) = \int_{-\infty}^{+\infty} [x(\tau - t)\nu(\tau - t)]e^{-j2\pi f\tau} dt \quad (3)$$

式中,  $\gamma(t)$  为时间长度极小的窗函数。

根据 EMD 方法判断, 滚动轴承正常工作的峭度值约为 3, 因此, 当 IMF 分量峭度值  $K$  大于 3 时, 该 IMF 分量中含有较多的故障信息。当  $K$  的值不断增大时, 说明故障程度在不断加深。因此, 在优选 IMF 分类时, 选择  $K$  值较大的分类, 即为需要保留的有用 IMF 分量。利用峭度准则, 筛选出有效的 IMF 分量, 并对有效 IMF 分量进行重构, 达到剔除原始信号中噪声信号的目的<sup>[11]</sup>。

## 2 CSO-SVM 健康状态评估模型

### 2.1 鸡群优化算法 (chicken swarm optimization algorithm, CSO)

鸡群优化算法 (CSO) 是 Meng 等人于 2014 年根据鸡群的等级制度和觅食行为提出的一种优化算法<sup>[12-13]</sup>。鸡群优化算法集成了粒子群算法、遗传算法、蝙蝠算法等的优化特性。对鸡群生活规律抽象化得到了一种新的群智能全局优化算法。鸡群优化算法的主体思路是模拟鸡群等级制度和鸡群行为。该算法模拟了鸡群行为和鸡群的等级制度, 根据不同鸡所遵循的差异化移动规律和鸡群中存在的等级制度、竞争关系, 母鸡孵化后代以及小鸡成长为公鸡或母鸡等真实鸡群行为规律, 实现了群智优化算法。

鸡群优化算法能够实现多子群协同搜索，并且具有优秀的自适应能力，广泛应用于计算科学、管理科学、工程科学等领域中。鸡群优化算法为全局优化问题提供了新的思路和解决途径<sup>[14]</sup>。

在鸡群的等级制度中，具有较强觅食行为的公鸡占优势，其周围围绕着小鸡和母鸡。这种等级制度会在觅食和搜索食物的过程中不断更新。在鸡群算法中，按照鸡的适应度值，将其分为若干子种群和类型，包括公鸡、母鸡、小鸡。其中，公鸡具有最高适应度值、并作为每一个子种群的领头。紧接着的几只鸡，作为母鸡的适应度值。而适应度值最低的鸡被随机分配到一个子种群中，相应地，确定了母鸡与小鸡之间的母子亲密关系。其等级制度一旦确立将数代保持不变，等级制度每隔  $G$  ( $G \in [2, 20]$ ) 代更新一次。每一组内，母鸡跟随公鸡觅食，同时也随机偷取其他组内食物，小鸡跟随母鸡觅食。每一只鸡的位置对应优化问题的一个解，而公鸡、母鸡、小鸡的位置迭代方式也有所不同，其流程图如图 1 所示。

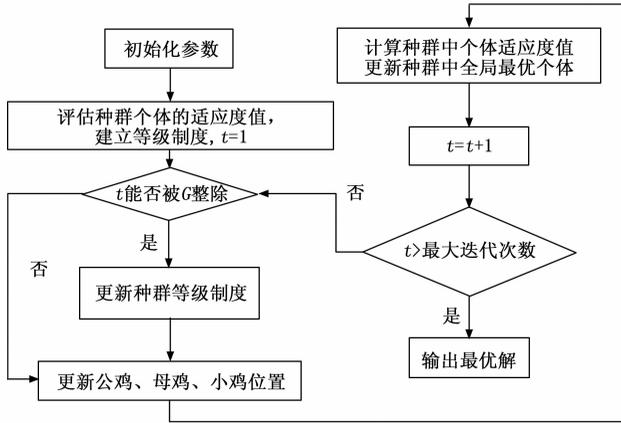


图 1 鸡群优化算法流程图

公鸡位置更新公式：

公鸡的觅食领域随着其适应度值增大而增大。适应度值较大的公鸡比适应度值较小的公鸡在食物竞争中更具有竞争性，能够在更广泛的领域内寻找食物。

$$x_{i,j}(t+1) = x_{i,j}(t) * (1 + \text{randn}(0, \sigma^2)) \quad (4)$$

$$\sigma^2 = \begin{cases} 1 & f_i < f_k \\ \exp\left(\frac{f_k - f_i}{|f_i| + \epsilon}\right) & f_i \geq f_k \end{cases} \quad (5)$$

$k \in [1, N_R], k \neq i$

其中： $\text{randn}(0, \sigma^2)$  是高斯分布函数，其均值为 0、标准差为  $\sigma^2$ ， $f_i$  为个体适应度值， $f_k$  为个体  $k$  的适应度分配值，为个体  $k$  从公鸡的群体中随机选择出来的公鸡。 $\epsilon$  为无穷小数。

母鸡更新公式：

母鸡可以在同一组中公鸡处获取食物，还可以窃取其他鸡所寻觅的食物，因此其在小组中具有独特的优势，因

此适应度值较高的母鸡比适应度值较低的母鸡具有优势。

$$x_{i,j}(t+1) = x_{i,j}(t) + S_1 * \text{rand}(x_{r1,j}(t) - x_{i,j}(t)) + S_2 * \text{rand} * (x_{r2,j}(t) - x_{i,j}(t)) \quad (6)$$

$$S_1 = \exp\left(\frac{f_i - f_{r1}}{|f_i| + \epsilon}\right) \quad (7)$$

$$S_2 = \exp(f_{r2} - f_i) \quad (8)$$

其中： $r1$  为第  $i$  只母鸡所配公鸡的位置， $r2$  为任一公鸡或母鸡个体编号，且  $r1 \neq r2$ ， $\text{rand}$  为  $[0, 1]$  内的随机数。

小鸡位置更新公式：

小鸡跟随母鸡寻找食物。

$$x_{i,j}(t+1) = x_{i,j}(t) + FL * (x_{m,j}(t) - x_{i,j}(t)) \quad (9)$$

其中： $x_{i,j}(t)$  为第  $t$  次迭代时第  $i$  只小鸡追随的母鸡的位置， $FL$  ( $FL \in [0, 2]$ ) 为跟随系数。

## 2.2 支持向量机 (support vector machine, SVM)

支持向量机是一种基于统计学习理论的有监督机器学习，其主要思想是对结构风险最小化的近似实现<sup>[15]</sup>。支持向量机因其构造简便容易运行，在处理小样品数据分析、处理不确定性问题中具有一定的优势。

当支持向量机用于分类时，其目的是构建样本之间最优的分类界限，即边界函数，当样本为二维时，边界函数为一个线性函数或一组线性函数，在高维空间中，边界函数为一组超平面。

设一组带标签的样本集为：

$$Data = \{(x_1, y_1), (x_2, y_2), (x_3, y_3), \dots, (x_m, y_m)\}$$

式中， $x_i, y_i$  为第  $i$  个样本输入值和输出值， $x_i \in R^d, y_i \in R$ 。SVM 的回归函数为：

$$g(x) = w^T \psi(x) + v \quad (10)$$

式中， $w$  为权值， $(\cdot)$  为映射函数，其目的是将输入映射到高维空间， $v$  为偏移量。

拉格朗日函数<sup>[16-17]</sup>：

$$L(\omega, b, \alpha) = \frac{1}{2} \|\omega\|^2 - \sum_{i=1}^n \alpha_i [y_i(\omega \cdot x_i + b) - 1] \quad (11)$$

式中， $\alpha_i$  为拉格朗日乘子；

最终回归函数：

$$g(x) = \sum_{i=1}^m (\alpha_i^* - \alpha_i) K(x_i, x) + v \quad (12)$$

式中， $\alpha_i$  和  $\alpha_i^*$  为拉格朗日乘子；

其中： $K(\cdot)$  为径向基核函数：

$$K(x_i, x) = \exp\left[-\frac{\|x_i - x\|^2}{2\sigma^2}\right] \quad (13)$$

式中， $\sigma$  为核宽度。

## 2.3 鸡群优化算法改进的支持向量机 (CSO-SVM)

由支持向量机的结构可知，核宽度  $\sigma$  和惩罚因子  $t$  直接关系到支持向量机的分类效果。其中，惩罚  $t$  的作用是平衡支持向量的复杂度与误差率。 $t$  取值越大，支持向量越多，模型越复杂；反之亦然， $t$  取值越小，支持向量越少，模型越简单。核宽度  $\sigma$  反映单个样本对超平面的影响。 $\sigma$  越小，

单个样本对超平面的影响越小; 反之,  $\sigma$  越大, 单个样本对超平面的影响越大。

因此, 不合适的核宽度  $\sigma$  和惩罚系数  $t$  取值会使降低 SVM 的回归效果。采用优化算法对参数  $\sigma$  和  $t$  寻优能提高回归效果, 进而提高 SVM 的分类精度。

基于 CSO-SVM 的轴承健康状态评估流程如图 2 所示。

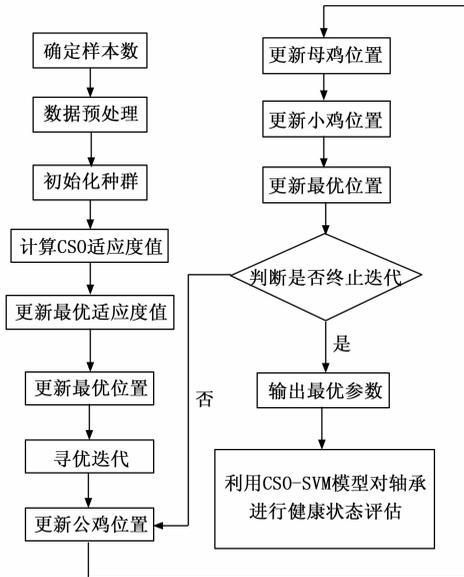


图 2 基于 CSO-SVM 轴承健康状态评估流程

基于上述 EMD、CSO 和 SVM 模型, 构建 CSO-SVM 健康状态识别模型, 步骤如下。

Step 1: 对原始信号进行经验模态分解, 根据公式 (2) 计算各 IMF 分量的峭度值, 保留峭度值大于 3 的 IMF 分量, 并重构信号, 达到振动信号降噪的目的。

Step 2: 降噪后的重构信号进行特征提取, 利用熵权法进行特征降维, 筛选出最能表现出轴承退化趋势的特征。然后对特征进行归一化处理, 得到轴承的退化趋势曲线。

Step 3: 利用 CSO 优化 SVM 的核宽度  $\sigma$  和惩罚因子  $t$ 。分别对归一化后的数据进行训练和测试, 得到最终识别结果。

Step 4: 计算识别精度, 并将上述结果与传统的分类模型进行对比。

### 3 实验验证

本文的数据来源为 IEEE 可靠性协会和 FEMTO-ST 研究所组织的 IEEE PHM 2012 数据挑战赛轴承全寿命数据<sup>[18]</sup>。数据采集过程采用两个振动传感器, 其安装位置呈 90°, 第一个安装在纵向轴上, 第二个安装在水平轴上。传感器的类型为微型加速度计, 沿径向放置在轴承的外圈上。采样频率为 25.6 kHz。采样间隔为 10 秒, 一次采样时间持续 0.1 s, 即每次采集 2 560 个点, 共 2 802 个训练样本。本文选用 Bearing1\_1 数据的纵轴方向数据, 对轴承进行健康状态评估。

### 3.1 轴承特征提取

轴承是旋转元件, 且大多数振动数据包含周期性运动, 但随着故障的进行, 轴承退化的过程会存在大量的噪声。相较于其他的信号处理方法, EMD 在处理非平稳非线性信号表现出了一定的优越性, 因此, 采用 EMD 对振动信号进行降噪处理。首先对信号进行 IMF 分解, 挑选出峭度大于 3 的 IMF 分量。

根据公式 (2), 得出各 IMF 分量的峭度值。正常状态下轴承 IMF 分量的峭度值和故障状态下轴承 IMF 分量的峭度值分别如表 1、表 2 所示。

表 1 正常轴承信号各 IMF 分量峭度值

IMF	K	IMF	K
1	2.878 6	5	2.590 3
2	3.748 1	6	1.787 8
3	3.241 4	7	1.765 6
4	3.020 2	8	1.866 6

表 2 故障轴承信号各 IMF 分量峭度值

IMF	K	IMF	K
1	15.787 3	5	4.828 9
2	19.813 8	6	2.471 8
3	9.258 3	7	2.657 7
4	5.2941		

由表 1~2 可见, 失效后的信号中包含故障信息的 IMF 分量逐渐增多。以正常的振动信号为例, 提取出 IMF2、IMF3、IMF4 分量进行信号重构。正常状态下原始信号和重构信号如图 3 所示。

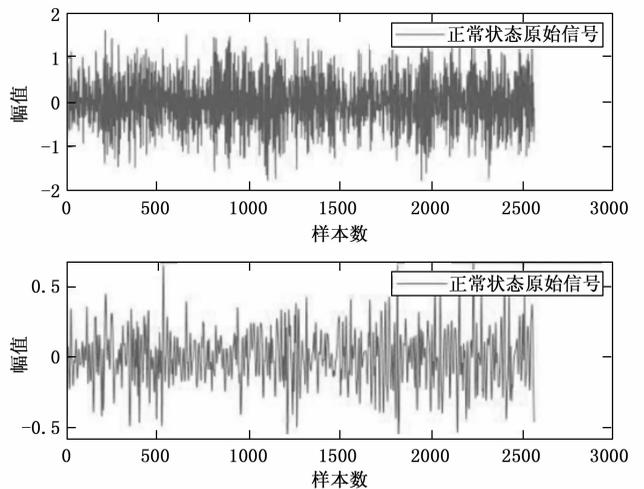


图 3 正常状态下原始信号和重构信号

对振动信号时间域进行处理所得到的特征成为时域特征<sup>[19]</sup>。当轴承出现故障时, 时域信号的概率和幅值会随之发生变化。但是, 时域信号的统计特征计算简单, 适用于在线监测, 缺乏稳定性和敏感性。当信号呈现非周期和故

障不典型的问题时，难以从时域波形上看出设备的故障状态信息。因此，在对轴承振动信号进行特征提取时，还应考虑频域特征，对信号进行频谱分析，通过频谱分析，可以观察到信号在频域上的分布情况。当轴承发生故障时，其振动信号的频率结构发生改变，体现在不同频率成分的能量以及频谱的主能量谱峰位置的变化<sup>[20-21]</sup>。通过对轴承振动信号的时域信息和频域信息的描述，可以全面反映轴承的运行状态。

通过傅里叶变换，时域信号转化为频域信号，再根据频域信号反映出的频域特征进行统计分析。傅里叶变换公式为：

$$X(f) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi ft} dt \quad (14)$$

经过 EMD 和谱峭度法对轴承振动信号进行降噪处理，然后对轴承信号进行时域特征提取和频域特征提取。时域特征和频域特征能够反映振动信号的概貌，从全局的角度对振动信号进行描述。因此，本文综合利用时域特征和频域特征的参数，共提取了 27 维的时域、频域特征。

高维数据包含着大量的无关信息和冗余信息，对机器学习算法的性能产生负面影响。因此，面对高维数据时，特征降维能够有效地消除无关和冗余特征，提高挖掘任务的效率，增强机器学习的准确性，增强学习结果的易理解性。

基于数据驱动的健康状态评估方法中，大多采用将时域特征、频域特征笼统的进行模型训练，未注意到每个特征量来表征轴承健康状态。而特征指标的权重系数能够较为客观地反映轴承时域、频域特征的重要程度，因此本文利用熵权法来优选特征，达到特征降维的目的<sup>[22]</sup>。

熵权法赋权步骤。

1) 归一化处理：

$$x_{ij} = \frac{x_{ij} - \max(x_{ij})}{\max(x_j) - \max(x_j)} \quad (15)$$

$x_{ij}$  表示第  $i$  个采样点的第  $j$  个特征。

2) 各指标的信息熵：

$$E_j = -K \sum_{i=1}^n p_{ij} \ln p_{ij} \quad (16)$$

$$p_{ij} = \frac{x_{ij}}{\sum_{i=1}^n x_{ij}} \quad (17)$$

$E_j$  为第  $j$  个指标的信息熵， $p_{ij}$  为第  $i$  个采样点的第  $j$  个特征比重。其中  $K$  为常数， $K = \frac{1}{\ln(m)}$ 。

3) 确定各指标权重：

$$\omega_j = \frac{1 - E_j}{m - \sum E_j} \quad (j = 1, 2, \dots, m) \quad (18)$$

利用熵权法计算时域和频域信号的权重，得到权重最大的三个特征，分别为偏斜度、方差、标准差。时域和频域特征权重如图 4 所示。

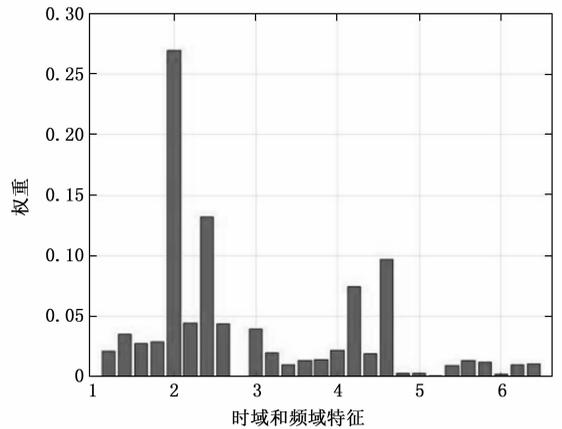


图 4 时域和频域特征权重

偏斜度：

$$\alpha = \frac{1}{N} \sum_{n=1}^N x^3(n) \quad (19)$$

方差：

$$s^2 = \frac{1}{N-1} \sum_{n=1}^N (x(n) - \bar{X})^2 \quad (20)$$

标准差：

$$S = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}} \quad (21)$$

根据所提取的特征，构建如图 5 所示的滚动轴承的退化趋势曲线。

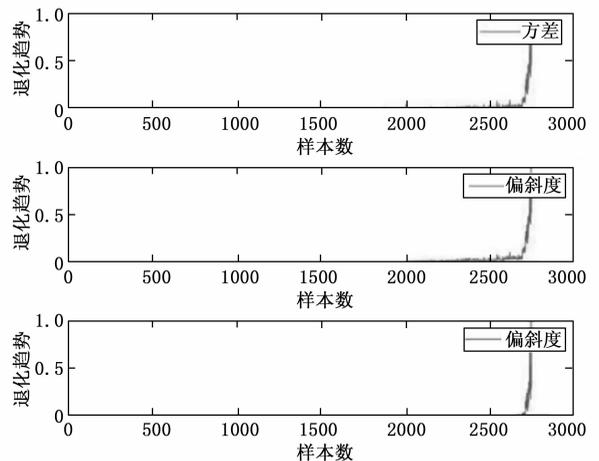


图 5 轴承退化趋势曲线

健康状态评估本质是识别轴承退化状态和不同退化阶段的模式。根据轴承退化曲线图可知，轴承的退化趋势从 0~1 代表着从健康状态逐步退化到故障状态<sup>[21]</sup>。可以确认轴承在 0~1 400 组样本时，轴承退化趋势相对稳定，认为轴承此时出具健康状态；在 1 401~2 769 组样本点时，退化趋势曲线的幅值小幅度高于正常状态下的幅值，并且以较低的斜率增加，此时轴承处于退化状态。在 2 770~2 802 组样本中，退化趋势曲线的幅值明显高于正常状态，并且

以较大的斜率快速增加, 此时认为轴承处于故障状态。根据上述分析, 以退化趋势曲线的斜率为判断准则, 将轴承退化状态分为三个阶段: 健康、亚健康 and 故障。

### 3.2 基于 CSO-SVM 的轴承健康状态评估

首先对实验数据划分训练集和数据集, 利用训练样本对 CSO-SVM 模型进行训练。利用训练好的模型对测试样本进行分类, 并对模型进行分类效果进行评估。进而得到基于 CSO-SVM 模型的轴承健康状态评估结果。

基于 CSO-SVM 的轴承健康状态评估结果如图 6 所示, 对照组分别为传统的机器学习模型, 基于 SVM 的轴承健康结果如图 7 示, 基于极限学习机 (ELM) 的轴承健康状态评估结果如图 8 所示。

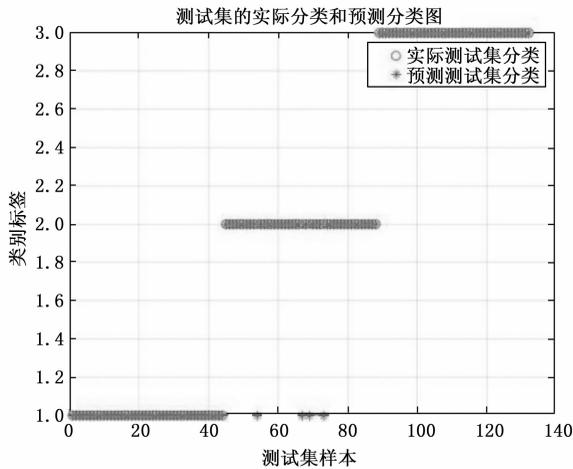


图 6 CSO-SVM 康状态评估结果

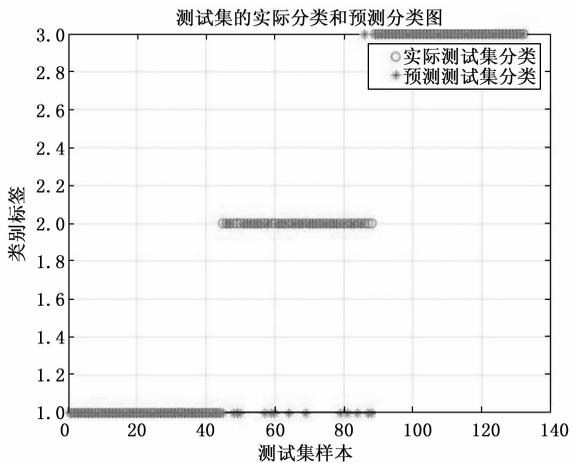


图 7 SVM 健康状态评估结果

其中, 1、2、3 分别代表健康、亚健康和故障三个阶段。利用 CSO-SVM 的健康状态评估准确率为 97%, 而利用 SVM 和 ELM 的健康状态评估的准确率为分别为 88.6% 和 87.5%。结果表明, 基于鸡群优化算法的支持向量机 (CSO-SVM) 的评估精度远高于传统的机器学习模型。

在机器学习中, 常用混淆矩阵来分析、总结分类模型

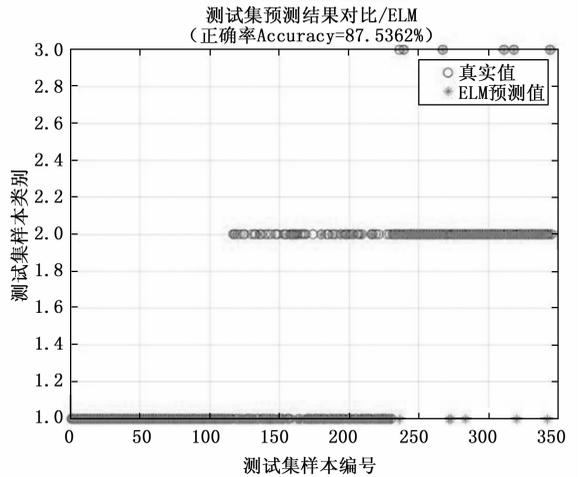


图 8 ELM 健康状态评估结果

的预测结果。其作用机理是, 对集中记录的数据按照真实类别和分类模型预测的结果以矩阵的形式汇总。其中, 混淆矩阵的行向量表示真实值, 列向量表示预测值。CSO-SVM、SVM 和 ELM 的混淆矩阵分别如图 9、图 10、图 11 所示。

	1	2	3	
1	44 33.3%	4 3.0%	0 0.0%	91.7% 8.3%
2	0% 0.0%	40% 30.3%	0% 0.0%	100% 0.0%
3	0% 0.0%	0% 0.0%	44% 33.3%	100% 0.0%
	100% 0.0%	90.9% 9.1%	100% 0.0%	97.0% 3.0%
	1	2	3	
	Target Class			

图 9 CSO-SVM 混淆矩阵图

## 4 结束语

本文的数据来源为 IEEE 可靠性协会和 FEMTO-ST 研究所组织的 IEEE PHM 2012 数据挑战赛轴承全寿命数据。构建健康状态评估模型时, 以支持向量机为基本模型, 利用鸡群优化算法对支持向量机的超参数进行优化, 进而达到提高支持向量机的分类能力的目的。最后应用轴承全寿命数据进行健康状态评估并得出以下结论:

- 1) 利用经验模态分解 (EMD) 对轴承的原始振动信号进行分解、重构能够达到降噪目的。分解后的 IMF 分量中, 峭度值大于 3 的 IMF 分量进行重构, 这种方法对处理非平

	1	2	3	
1	44 33.3%	14 10.6%	0 0.0%	75.9% 24.1%
2	0 0.0%	29 22.2%	0 0.0%	100% 0.0%
3	0 0.0%	1 0.8%	44 33.3%	97.8% 2.2%
	100% 0.0%	65.9% 34.1%	100% 0.0%	88.6% 11.4%
	1	2	3	

图 10 SVM 混淆矩阵图

	1	2	3	
1	178 51.6%	42 12.2%	1 0.3%	80.5% 19.5%
2	0 0.0%	119 34.5%	0 0.0%	100% 0.0%
3	0 0.0%	0 0.0%	5 1.4%	100% 0.0%
	100% 0.0%	73.9% 21.1%	83.3% 16.7%	87.5% 12.5%
	1	2	3	

图 11 ELM 混淆矩阵图

稳、非线性的信号的健康状态识别的精度较高,能实现对振动信号的降噪。

2) 鸡群优化算法 (CSO) 具有良好的参数优化能力。利用鸡群优化算法对支持向量机 (SVM) 模型中的核宽度和惩罚因子的取值进行寻优。结果表明,与传统的 SVM 模型相比,该优化算法提高了 SVM 的分类精度,使得模型的优化效果更加显著。

3) 实验结果表明,基于 CSO-SVM 的识别模型能够准确、快速地识别轴承的健康状况,为识别其健康状况提供了一种新思路。

参考文献:

[1] 史晓雪, 吴亚锋. 基于 AGPF 的滚动轴承性能衰退趋势预测 [J]. 计算机测量与控制, 2017, 25 (10): 228-231.  
 [2] 杨保海, 陈 栋, 郑冬喜, 等. 基于小波包分解和 EMD-SVM

的轴承故障诊断方法 [J]. 计算机测量与控制, 2015, 23 (4): 1118-1120.  
 [3] 岳晓峰, 邵海贺. 基于 DE-LSSVM 的滚动轴承故障诊断 [J]. 计算机测量与控制, 2015, 23, 207 (12): 3933-3935.  
 [4] 宋立业, 孙 琳. EEMD-GSSA-SVM 滚动轴承故障诊断方法研究 [J]. 传感器与微系统, 2022, 41 (4): 56-59.  
 [5] 张吴飞, 李帅帅, 李嘉成. 基于 IGWO-SVM 的轴承故障分类预测 [J]. 轻工机械, 2022, 40 (2): 86-91.  
 [6] 时培明, 梁 凯, 赵 娜, 等. 基于深度学习特征提取和粒子群支持向量机状态识别的齿轮智能故障诊断 [J]. 中国机械工程, 2017, 28 (9): 1056-1061.  
 [7] 周建民, 黎 慧, 张 龙, 等. 基于 EMD 和逻辑回归的轴承性能退化评估 [J]. 机械设计与研究, 2016, 32 (5): 72-75, 79.  
 [8] 张传凯. 基于 EMD 降噪和 LSTM 网络的地铁风机轴承寿命预测 [J]. 风机技术, 2020, 62 (3): 77-82.  
 [9] 曹玲玲, 李 晶, 彭 镇, 等. 基于 EEMD 和快速谱峭度的滚动轴承故障诊断研究 [J]. 机电工程, 2021, 38 (10): 1311-1316.  
 [10] 聂永辉, 徐明文, 张译丹. 基于多小波和峭度准则的风力发电机滚动轴承故障检测 [J]. 东北电力大学学报, 2019, 39 (6): 15-23.  
 [11] ANTONI J. The spectral kurtosis: a useful tool for characterising non-stationary signals [J]. Mechanical Systems and Signal Processing, 2004, 20 (2): 282-307.  
 [12] MENG X B, LIU Y, GAO X Z, et al. A New Bio-inspired Algorithm: Chicken Swarm Optimization [C] // International Conference in Swarm Intelligence. Germany Springer International Publishing, 2014: 86-94.  
 [13] SANCHARI DEB, GAO X Z, KARITAMMI ETC. Recent Studies on Chicken Swarm Optimization algorithm: a review (2014-2018) [C] // 2020, 53: 1737-1765.  
 [14] 黄 娟. 基于多目标优化与鸡群算法的跨境电商物流配送优化研究 [J]. 经济研究导刊, 2021 (3): 34-37.  
 [15] 李 素, 袁志高, 王 聪, 等. 群智能算法优化支持向量机参数综述 [J]. 智能系统学报, 2018, 13 (1): 70-84.  
 [16] HUSSAIN S F. A novel robust kernel for classifying high-dimensional data using support vector machines [J]. Expert Systems with Applications, 2019, 131: 116-131.  
 [17] PARHIZKAR E, SAEEDZADEH H, ALUNADI F, et al. Partial least squares-least squares-support vector machine modeling of ATR-IR as a spectrophotometric method for detection and determination of iron in pharmaceutical formulations [J]. Iranian Journal of Pharmaceutical Research, 2019, 18 (1): 72-79.  
 [18] NECTOUX P, GOURIEAU R, MEDJAHER K, et al. Prognostics: an experimental platform for bearings accelerated degradation tests [C] // IEEE International Conference on Prognostics and Health Management. Denver: IEEE, 2012.

(下转第 254 页)