

DOI: 10.13382/j.jemi.B2103861

基于复合核 SVM 的智能电表基本误差预测方法*

王永超¹ 唐求² 马俊² 邱伟² 杨莹莹²

(1. 国网新疆电力有限公司营销服务中心 乌鲁木齐 830000; 2. 湖南大学 电气与信息工程学院 长沙 410082)

摘要:智能电表作为电网的终端设备,其退化情况与工作环境、运行时间等因素密切相关。针对复杂变量条件下智能电表退化情况难以预测的问题,提出一种基于复合核支持向量机(support vector machine, SVM)的智能电表基本误差预测方法。首先对智能电表退化数据进行分析,采用皮尔逊相关性分析找出与智能电表基本误差相关性极强的环境变量。然后,为进一步提取数据退化特征,采用模糊 C 均值聚类算法对智能电表退化数据进行聚类,确定退化特征向量。最后,基于高斯径向基核函数与多项式核函数构造一种新的复合核 SVM 模型用以预测智能电表基本误差。结合新疆地区智能电表退化数据对复合核 SVM 模型性能进行验证,实验结果表明,复合核 SVM 模型可以准确预测复杂环境下智能电表的基本误差,其预测准确率高于贝叶斯方法、神经网络方法以及经典 SVM 方法。

关键词:智能电表;复合核支持向量机;模糊 C 均值聚类;基本误差预测

中图分类号: TM933.4 **文献标识码:** A **国家标准学科分类代码:** 470.40

Prediction method of basic error of smart meter based on composite core SVM

Wang Yongchao¹ Tang Qiu² Ma Jun² Qiu Wei² Yang Yingying²

(1. State Grid Xinjiang Electric Power Co., Ltd. Marketing Service Center, Urumqi 830000, China;

2. College of Electrical and Information Engineering, Hunan University 410082, Changsha, China)

Abstract: As the terminal equipment of the power grid, the degradation of smart meters is closely related to factors such as working environment and running time. Aiming at the problem that the degradation of smart meters under complex variable conditions is difficult to predict, a smart meter basic error prediction method based on the composite core support vector machine (SVM) is proposed. First, analyze the degradation data of smart meters, and use Pearson correlation analysis to find environmental variables that are highly correlated with the basic errors of smart meters. Then, in order to further extract the data degradation features, the fuzzy C-means clustering algorithm is used to cluster the smart meter degradation data and determine the degradation feature vector. Finally, based on the Gaussian radial basis kernel function and polynomial kernel function, a new composite kernel SVM model is constructed to predict the basic error of smart meters. The performance of the composite core SVM model is verified by combining the degradation data of smart meters in Xinjiang. The experimental results show that the composite core SVM model proposed in this paper can accurately predict the basic errors of smart meters in complex environments, and its prediction accuracy is higher than that of Bayesian methods. Neural network method and classic SVM method.

Keywords: smart meter; composite core support vector machine; fuzzy C-means clustering; basic error prediction

0 引言

计量设备在不同环境下是否运行可靠、计量准确是

电力企业、用户及行业专家所关注的问题^[1]。典型环境条件下,如高温高干热环境,会加速电子式电能计量装置的集成电路、液晶显示器、晶体振荡器等电子元器件、零部件失效,造成电能计量装置不计量或电量计算错误,从

而引起电能计量纠纷甚至造成安全、生产事故^[2]。智能电表作为我国当前智能电网建设下的终端设备,其工作性能的可靠性直接关系到电网计量的准确性。智能电表退化情况可通过基本误差值反映,基本误差值升高说明退化加剧,超出一定阈值后,则说明智能电表失效。因此,对智能电表实行基本误差预测具有十分重要的意义。

随着计算机技术和传感器技术的飞速发展,基于数据驱动的设备退化预测方法得到了广泛的应用^[3-5]。近年来,包括神经网络、贝叶斯算法以及支持向量机(support vector machine, SVM)在内的数据驱动预测方法取得了较大的发展。文献[6]针对电力负荷中存在的非线性时间序列预测问题,构建了一种基于脉冲神经网络的短期负荷预测模型,并结合实际电网数据证明了预测模型的有效性。文献[7]结合不同的环境特征,提出了一种加权层次贝叶斯的电能表故障率预测方案。但针对复杂环境下智能电表的基本误差这类小样本、非线性以及未知分布的性能退化数据时,神经网络方法以及贝叶斯方法表现出了一定的局限性,神经网络方法通过减少损失进而提高预测精度,但往往需要大量的训练数据,贝叶斯方法在小样本数据预测时,可以给出预测结果的概率区间,但需要选取合适的先验分布否则易造成过拟合。

支持向量机建立在统计学习理论的 VC 维理论和结构风险最小化原则基础上,在解决小样本、非线性以及不确定问题方面具有独特优势^[8]。文献[9]采用最小二乘支持向量机实现了对机械式温度仪表测量误差的预测,模型训练速度快且预测精度较高。文献[10]基于主成分分析和最小二乘支持向量机,提出了一种轴承退化预测的新方法。文献将支持向量机的方法与随机子空间辨识的方法相结合,实现了对风电机组轴承的故障诊断并具有较高的诊断精度;文献[11]基于经典的支持向量机原理,提出了一类和多类支持向量机方法,实现了对智能电表的电能质量在线扰动检测,并结合智能电表的实际扰动数据证明了所提方法的有效性。

传统的支持向量机方法采用单一核函数解决简单数据的预测问题^[12],然而无法有效地解决如智能电表基本误差数据这类复杂的性能退化数据的预测问题,基于此,本文为实现复杂变量条件下对智能电表基本误差的预测,采取皮尔逊相关性分析法选取环境应力,充分考虑典型环境条件下环境因素对智能电表基本误差的影响。选用模糊 C 均值聚类方法处理智能电表退化数据,确定用于训练模型的退化特征向量。提出一种融合高斯核与多项式核的新的复合核 SVM 模型用以预测智能电表基本误差,提高了对智能电表基本误差的预测精度。结合复合核函数,提出一种基于复合核 SVM 的智能电表基本误差预测框架。

1 数据预处理

1.1 皮尔逊相关性分析

智能电表使用时长是影响智能电表退化情况的根本原因,其他环境因素如温度、湿度等也会影响智能电表的工作状态^[13]。在建立智能电表基本误差预测模型前,需充分考虑环境因素对智能电表基本误差的影响。

设智能电表退化数据集 $P = \{(X_1, y_1), (X_2, y_2), \dots, (X_n, y_n)\}$, 其中 X 为环境变量, y 为智能电表基本误差。采用皮尔逊相关系数确定多个环境因素中与智能电表基本误差密切相关的变量:

$$COR(X, y) = \frac{\sum_{i=1}^n (X_i - \bar{X})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (X_i - \bar{X})^2 \sum_{i=1}^n (y_i - \bar{y})^2}} \quad (1)$$

式中: $COR(X, y)$ 为环境变量与智能电表基本误差之间的皮尔逊相关系数, $COR(X, y)$ 的绝对值介于 0.8~1.0 表示两者极强相关; 介于 0.6~0.8 表示强相关, 介于 0.4~0.6 表示中等程度相关; 介于 0.2~0.4 表示弱相关; 0.0~0.2 表示极弱相关或无相关^[14]。

经皮尔逊相关性分析后,智能电表退化数据集处理为 $P' = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 其中 x 为与智能电表基本误差极强相关的环境变量。

1.2 基于智能电表退化数据的模糊 C 均值聚类

对智能电表进行基本误差预测时,采用较多的样本训练模型可以提高模型预测精度,但样本数量增多会提升样本空间复杂度,导致模型学习效率降低。为在保证模型预测精度的同时提高学习效率,采用模糊 C 均值聚类算法(fuzzy C means, FCM)对训练样本进行降维。FCM 聚类是一种常用的模糊聚类算法,它可将一组样本对象划分几类,使得不同类之间的相似性最小化,且同一类内的对象具有一定相似性。

FCM 聚类算法将智能电表退化数据集 P' 划分为 k 个模糊组,并求得每组聚类中心 k_i 。FCM 聚类的目标函数如下:

$$J(U, k_1, k_2, \dots, k_k) = \sum_{i=1}^k J_i = \sum_{i=1}^k \sum_{j=1}^n u_{ij}^m d_{ij}^2 \quad (2)$$

式中: U 为隶属度矩阵, U 中元素 u_{ij} 的值介于 0~1, 意义为 x_j 与类 i 之间的相关程度; m 是决定群模糊性的参数; d_{ij} 为第 i 个聚类中心与第 j 个数据点间的欧氏距离。采用 Lagrange 乘法法可求得使式(2)达到最小的解为^[15]:

$$k_i = \sum_{j=1}^n u_{ij}^m x_j / \sum_{j=1}^n u_{ij}^m \quad (3)$$

$$u_{ij} = \sum_{q=1}^k \left(\frac{d_{ij}}{d_{iq}} \right)^{2/(1-m)} \quad (4)$$

基于式(3)和(4)两个条件,FCM 聚类算法的流程表示如下:

- 1) 设定迭代次数最大值 T_{\max} 和迭代精度 ξ ;
- 2) 初始化隶属度矩阵 U ;
- 3) 利用式(3)计算聚类中心 $k_i, i=1, 2, \dots, k$;
- 4) 利用式(2)计算目标函数值,当相对上次目标函数值的改变量小于 ξ 或迭代次数大于 T_{\max} 时,算法停止;否则采用式(4)计算新的隶属度矩阵 U ,并返回步骤 3)。

经聚类后训练样本数量减少,样本空间复杂度降低,可以在保证精度的同时减少训练时间、提高模型学习效率^[16]。设聚类后训练样本数据量由 n 压缩为 s ,则聚类后智能电表退化数据集为 $P'' = \{(x_1, k_1), (x_2, k_2), \dots, (x_s, k_s)\}$,取 P'' 作为训练数据集训练 SVM 模型,其中环境变量 x 为模型输入数据, k 为模型学习目标。

2 基于复合核 SVM 的智能电表基本误差预测模型构建

2.1 支持向量机基本原理

经皮尔逊相关性分析与模糊 C 均值聚类后,得到处理后的数据集 P'' ,需求一种拟合方法对智能电表测量误差进行评估与预测。SVM 模型可通过核函数将非线性可分问题从原始的特征空间映射至更高维的希尔伯特空间(Hilbert space),从而转化为线性可分问题,适用于小样本分析求解。结合模型训练数据集 $P'' = \{(x_1, k_1), (x_2, k_2), \dots, (x_s, k_s)\}$,此时在特征空间中划分超平面所对应的模型可表示为:

$$f(\mathbf{x}) = \boldsymbol{\omega}^T \boldsymbol{\varphi}(\mathbf{x}) + b \quad (5)$$

式中: $\boldsymbol{\omega}$ 为超平面的法向量; b 为超平面截距; $\boldsymbol{\varphi}: X' \rightarrow F$ 为原空间映射到高维空间的映射函数。问题转化为:

$$\begin{cases} \max_{\alpha} \sum_{i=1}^s \alpha_i - \frac{1}{2} \sum_{i=1}^s \sum_{j=1}^s \alpha_i \alpha_j k_i k_j \kappa(\mathbf{x}_i, \mathbf{x}_j) \\ \text{s. t. } \sum_{i=1}^s \alpha_i k_i = 0, \alpha_i \geq 0, i = 1, 2, \dots, s \end{cases} \quad (6)$$

式中: α 为拉格朗日乘子。对式(6)求解后可得到分类函数如式(7)所示。

$$f(\mathbf{x}) = \sum_{i=1}^s \alpha_i k_i \kappa(\mathbf{x}, \mathbf{x}_i) + b \quad (7)$$

式中:核函数 $\kappa(\mathbf{x}, \mathbf{x}_i) = \boldsymbol{\phi}(\mathbf{x}_i)^T \boldsymbol{\phi}(\mathbf{x})$ 为映射函数 $\boldsymbol{\phi}$ 的内积。

将 SVM 推广至回归问题可以得到支持向量回归(support vector regression, SVR)。SVR 具有稀疏性,给定回归模型的间隔边界,经 FCM 聚类后的样本数据落入间隔边界内时,样本损失可不计。SVR 问题可形式化为:

$$\min_{\boldsymbol{\omega}, b} \frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \sum_{i=1}^s l_{\varepsilon}(f(\mathbf{x}_i), k_i) \quad (8)$$

式中: C 为正则化参数; l_{ε} 是 ε 为不敏感损失函数。

$$l_{\varepsilon}(z) = \begin{cases} 0, & |z| \leq \varepsilon \\ |z| - \varepsilon, & \text{其他} \end{cases} \quad (9)$$

引入松弛变量 ξ_i 和 $\hat{\xi}_i$, 可将式(8)重写为:

$$\begin{cases} \min_{\boldsymbol{\omega}, b, \xi_i, \hat{\xi}_i} \frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \sum_{i=1}^s (\xi_i, \hat{\xi}_i) \\ \text{s. t. } f(\mathbf{x}_i) - k_i \leq \varepsilon + \xi_i, \\ k_i - f(\mathbf{x}_i) \leq \varepsilon + \hat{\xi}_i, \\ \xi_i \geq 0, \hat{\xi}_i \geq 0, i = 1, 2, \dots, s \end{cases} \quad (10)$$

采用拉格朗日乘子法求解,可得到单一核函数的 SVR 决策函数为:

$$f(\mathbf{x}) = \sum_{i=1}^s (\hat{\alpha}_i - \alpha_i) k_i \kappa(\mathbf{x}, \mathbf{x}_i) + b \quad (11)$$

2.2 基于复合核 SVM 的智能电表误差预测

核函数 $\kappa(\mathbf{x}, \mathbf{x}_i)$ 的选取直接影响着智能电表基本误差的预测效果。目前大多数支持向量机预测模型中采用的是人为选择的单个核函数,是基于单个特征空间的。为了充分描述智能电表退化数据的全局特性和局部特点,本文将多项式核函数与高斯核函数线性组合后构造一种新的复合核函数:

$$\kappa_{mk}(\mathbf{x}_i, \mathbf{x}_j) = \lambda \kappa_{poly} + (1 - \lambda) \kappa_{rbf}(\mathbf{x}_i, \mathbf{x}_j) \quad (12)$$

式中: λ 为权值, $\lambda \in [0, 1]$ 。通过调节 λ 的值,可以改善模型的学习和泛化能力,从而适应不同的输入^[8]。

其中, κ_{poly} 与 κ_{rbf} 为支持向量机常用的核函数,多项式核函数 κ_{poly} 具有全局特性,可以刻画退化数据的趋势特征,其表达式如下

$$\kappa_{poly} = (\mathbf{x}_i^T \mathbf{x}_j)^d \quad (13)$$

高斯径向基核函数 κ_{rbf} 具有局部特性,有助于提取退化数据的局部特征,其表达式为:

$$\kappa_{rbf}(\mathbf{x}_i, \mathbf{x}_j) = \exp\left(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2}\right) \quad (14)$$

式中: d 和 σ 是人工设置的参数。 d 为多项式核函数的最高此项次数, σ 为高斯径向基核函数的宽度参数。

结合提出的复合核函数,本文进一步提出一种复合核 SVM 模型,该模型的优化目标可表示为:

$$\begin{cases} \min_{\boldsymbol{\omega}, b, \xi_i, \hat{\xi}_i} \frac{1}{2} \|\boldsymbol{\omega}\|^2 + C \sum_{i=1}^s (\xi_i, \hat{\xi}_i) \\ \text{s. t. } f_{mk}(\mathbf{x}_i) - k_i \leq \varepsilon + \xi_i, \\ k_i - f_{mk}(\mathbf{x}_i) \leq \varepsilon + \hat{\xi}_i, \\ \xi_i \geq 0, \hat{\xi}_i \geq 0, i = 1, 2, \dots, s \end{cases} \quad (15)$$

复合核对应的决策函数为:

$$f_{mk}(\mathbf{x}) = \sum_{i=1}^s (\hat{\alpha}_i - \alpha_i) k_i \kappa_{mk}(\mathbf{x}, \mathbf{x}_i) + b \quad (16)$$

相比于单核 SVM,复合核 SVM 能够充分发挥各个基本核的不同特征映射能力,从而提升拟合能力。

在构造的复合核 SVM 中,惩罚参数 c 的取值直接影响模型的拟合程度,核函数参数 g 的取值影响着训练与预测的速度^[17],权重系数 λ 则关系到复合核函数的泛化性能。为提高复合核 SVM 模型的训练速度与预测精度,需对参数 c 、 g 、 λ 进行参数寻优。

为简化参数寻优过程,本文先构建高斯径向基核函数 SVM 模型并进行训练,找出 (c, g) 参数组合最优值。在保持 (c, g) 参数值设置不变的基础上将核函数线性组合构建新的复合核 SVM,最后对核函数权重 λ 进行寻优。为保证搜索速度和准确率,采用网格搜索法作为寻找最佳参数的算法。网格搜索算法(grid search, GS)具有搜索准确率高的优点,在搜索参数较少时耗时相对较短,且每个 (c, g) 对相互独立,可并行性高^[18]。

3 基于复合核 SVM 模型的智能电表基本误差预测框架

基于复合核 SVM 模型的智能电表基本误差预测整体过程如图 1 所示,其具体步骤如下。

1) 数据预处理:对智能电表退化数据进行皮尔逊相关性分析,找出与智能电表基本误差极强相关的环境变量,并对退化数据进行 FCM 聚类,最终确定模型训练数据。模型输入数据为智能电表使用时间、温度、湿度,学习目标为智能电表基本误差。

2) 复合核 SVM 模型构建:将多项式核函数与高斯径向基核函数组合为新的复合核函数并分配权重,训练复合核 SVM 模型。采取网格搜索算法确定核函数最佳参数 c 、 g 及核函数权重 λ 。

3) 基本误差预测:采用训练完成的复合核模型预测测试样本的基本误差,并与真实数据进行对比,以均方误差(mean squared error, MSE)和决定系数(R-squared, R^2)作为模型评价指标。

4 智能电表基本误差预测实例分析

4.1 智能电表退化数据处理

为验证所建立复合核 SVM 模型对智能电表基本误差预测的效果,本文选取 A、B 两公司生产的智能电表样本进行分析,依据其在新疆甲、乙两地区 2017 年 12 月~2019 年 10 月的智能电表退化数据进行实验。数据来源于新疆实验基地智能电表数据采集器。A、B 公司各选取 12 组智能电表样本,每组样本包含 23 个数据点(每月 1 个数据点),其中甲地区 8 组数据作训练样本,乙地区 4

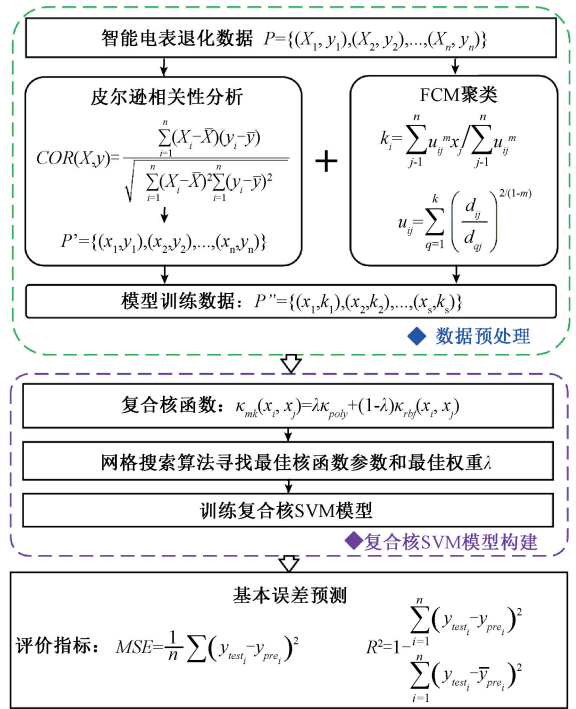


图 1 复合核 SVM 模型整体框架图

Fig. 1 Overall framework of multi-core SVM model

组数据作测试样本,验证复合核 SVM 模型的泛化能力。A、B 公司训练样表的退化曲线如图 2、3 所示。

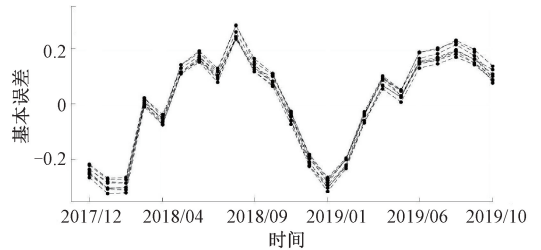


图 2 A 公司智能电表退化曲线

Fig. 2 Degradation curve of A company smart meter

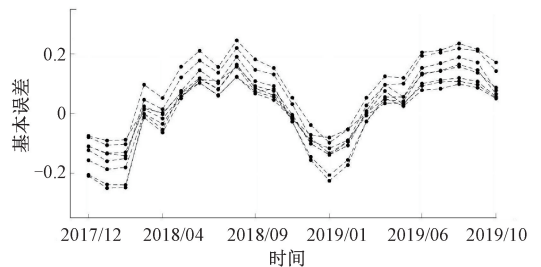


图 3 B 公司智能电表退化曲线

Fig. 3 Degradation curve of B company smart meter

从图 2 和 3 可以看出,智能电表的基本误差值在 1~8 月份呈上升趋势,达到峰值后开始逐渐下降,整体呈现周期性变化。A 公司智能电表基本误差差异性较小且周

期性趋势明显, B 公司智能电表基本误差的差异性相对较大。A 公司智能电表个体差异性相对于 B 公司较小, 但基本误差总体上大于 B 公司。两公司智能电表的基本误差与环境因素之间的皮尔逊相关系数如表 1 所示。

表 1 智能电表基本误差与环境因素间皮尔逊相关系数

Table 1 Pearson correlation coefficient between basic error of smart meter and environmental factors

变量 X	变量 Y	皮尔逊相关系数	
		公司 A	公司 B
温度	电表基本误差	0.986 92	0.959 23
湿度	电表基本误差	-0.807 09	-0.823 73
风速	电表基本误差	0.486 29	0.422 66
紫外线	电表基本误差	0.282 50	0.213 80

从表 1 可知, 智能电表基本误差与温度和湿度极强相关, 与风速和紫外线的相关性相对较弱。

为进一步提取数据退化特征, 对两公司训练样表的退化数据进行 FCM 聚类, 将 8 组数据聚类为 3 组。聚类后数据量减少 5/8, 从而将加权计算量减少 5/8。取智能电表使用时长、温度、湿度作为自变量, 有效聚类中心作为学习目标, 训练复合核支持向量机模型。

4.2 复合核 SVM 模型参数选择与结果分析

为选取合适的复合核 SVM 模型参数, 本文设置惩罚因子 c 的取值范围为 $[2^{-5}, 2^5]$, 高斯径向基核函数参数 g 的取值范围为 $[2^{-5}, 2^5]$, 多项式核函数的幂次数为 4, 给定参数 c 和 g 的搜索空间和步长, 设定 3 折交叉验证。采用网格搜索和交叉验证找出 c 和 g 的最优值。再采用网格搜索算法对核函数权重进行寻优, 即寻找式 (12) 中 λ 的最优值。

对 A 公司 FCM 聚类后的数据进行参数寻优, 参数 c 、 g 寻优的结果如图 4 所示。图 4 中, z 轴数值表示拟合均方误差值, 使得均方误差值最小的 c 、 g 值即为所求最佳参数值, 由图 4 可知, 搜索得到的最佳参数为 $c = 32, g = 0.031 25$, 此时取得最小均方误差值 0.005 699 9。核函数 λ 的最佳寻优结果为 0.91。

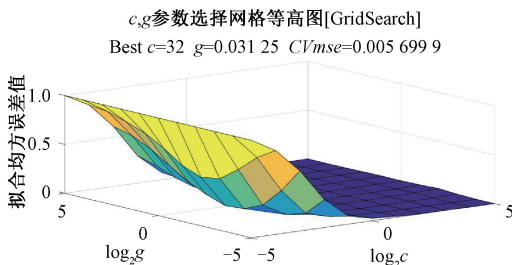


图 4 参数选择网格等高图(A 公司)

Fig. 4 Parameter selection grid contour map(Company A)

结合训练完成的复合核 SVM 模型预测 A 公司在乙地区测试样表的退化情况, 并将预测结果与真实数据进行比较, 结果如图 5 所示。从图 5 可以看出, 复合核模型的预测结果在数值与趋势方面与测试样表基本误差的真实数据高度符合。

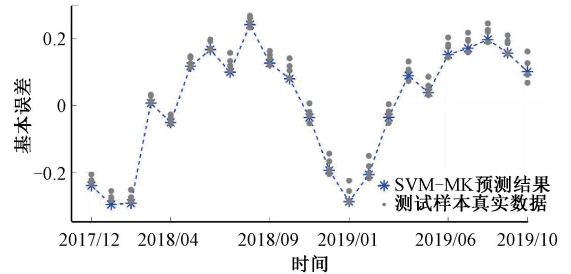


图 5 A 公司复合核模型预测结果与真实数据对比

Fig. 5 Comparison between the prediction results of company A's multi-core model and real data

再对 B 公司测试样表进行实验, 参数 c 、 g 的寻优结果如图 6 所示。由图 6 可知, 搜索得到的最佳参数为 $c = 4, g = 0.031 25$, 此时取得最小均方误差值 0.036 987。核函数 λ 的最佳值为 0.72。复合核模型对 B 公司测试样表退化情况的预测结果如图 7 所示。由图 7 可以看出, B 公司测试样表基本误差的差异性相对 A 公司较大, 复合核模型的预测结果较好符合了测试样本真实数据。

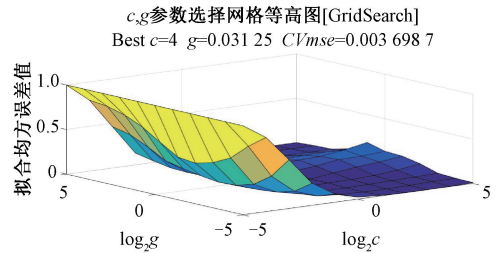


图 6 参数选择网格等高图(B 公司)

Fig. 6 Parameter selection grid contour map (Company B)

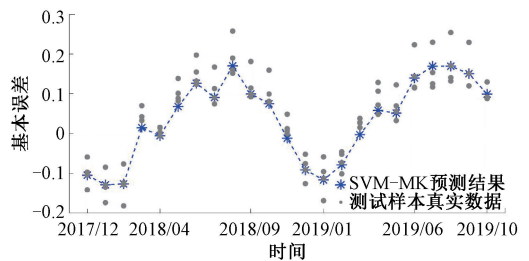


图 7 B 公司复合核模型预测结果与真实数据对比

Fig. 7 Comparison between the prediction results of company B's multi-core model and real data

4.3 模型评价指标

为验证所建立复合核模型对智能电表基本误差预测的有效性和准确性,选取均方误差 MSE 和决定系数 R^2 作为模型评价指标。 MSE 可用来衡量预测值同真值之间的偏差, MSE 的值越小,说明预测模型描述实验数据具有更好的精确度,其定义如式(17)所示。决定系数 R^2 的取值范围一般为 $[0,1]$,越接近于1表示模型拟合效果越优秀,其定义如式(18)所示。

$$MSE(y_{test}, y_{pre}) = \frac{1}{n} \sum_{i=1}^n (y_{test_i} - y_{pre_i})^2 \quad (17)$$

$$R^2(y_{test}, y_{pre}) = 1 - \frac{\sum_{i=1}^n (y_{test_i} - y_{pre_i})^2}{\sum_{i=1}^n (y_{test_i} - \bar{y}_{test})^2} \quad (18)$$

式中: y_{test} 为真实数据, y_{pre} 为预测数据, \bar{y}_{test} 为真实数据的平均值。将复合核模型预测结果与单核模型预测结果进行比较,表2给出了复合核模型和单核模型预测性能评价指标的对比值。

表2 单核模型和复合核模型预测性能评价指标对比

Table 2 The pair ratio of prediction performance evaluation indexes between the single-core model and the multi-core model

预测模型	MSE		R ²	
	公司 A	公司 B	公司 A	公司 B
SVM-Gauss	0.000 50	0.000 75	0.982 94	0.931 20
SVM-Poly	0.000 42	0.000 57	0.985 58	0.947 99
复合核 SVM	0.000 41	0.000 55	0.986 04	0.949 36

A 公司数据预测结果显示,复合核模型预测准确率可达到 98.60%,其均方误差值为 0.000 41,相比于高斯径向基核模型减少了 18.00%,相比于多项式核模型减少了 2.38%;B 公司数据预测结果显示,复合核模型预测准确率可达到 94.93%,其均方误差值为 0.000 55,相比于高斯径向基核模型减少了 26.67%,相比于多项式核模型减少了 3.51%。实验结果表明复合核模型预测性能更优,预测准确率高于单核模型。

在对智能电表的退化研究和故障预测方法中,常用的有贝叶斯、神经网络等方法^[19-20]。为验证本文所提出的方法的有效性,采用对应一致的实验数据,将复合核模型与贝叶斯模型、神经网络模型的预测结果进行比较,实验对比结果如表3所示。

从表3可以看出,在A公司与B公司实验数据条件下,复合核模型的均方误差值分别为 0.000 41 和 0.000 55,小于贝叶斯模型和神经网络模型;决定系数 R^2 值为 0.986 04 和 0.949 36,相比贝叶斯模型分别提高了 1.17% 和 3.08%,相比神经网络模型分别提高了 10.21%

和 15.66%。实验结果验证,在同等实验条件下,本文提出的复合核模型的预测性能优于贝叶斯模型和神经网络模型。

表3 复合核模型与贝叶斯模型和神经网络模型预测性能评价指标对比

Table 3 Comparison of prediction performance evaluation indexes between multi-core model, Bayesian model and neural network model

预测模型	MSE		R ²	
	公司 A	公司 B	公司 A	公司 B
贝叶斯	0.000 74	0.000 86	0.974 59	0.920 96
神经网络	0.003 06	0.001 95	0.894 73	0.820 78
复合核 SVM	0.000 41	0.000 55	0.986 04	0.949 36

5 结论

为解决复杂变量条件下智能电表退化情况难以预测的问题,本文提出了一种复合核 SVM 模型用于预测智能电表基本误差。模型分析了智能电表基本误差与环境因素之间的相关性,充分融合温度、湿度等环境信息。采用 FCM 聚类算法提取数据退化特征,减少了 5/8 的加权计算量。选取网格搜索算法优化核函数参数,并将核函数线性组合建立复合核核函数。同等实验条件下的模型预测结果表明,本文构建的复合核 SVM 模型能够有效预测智能电表的基本误差,预测准确率高于单核 SVM 模型、贝叶斯模型及神经网络模型,验证了本文提出的基于复合核 SVM 的智能电表基本误差预测方法的有效性。

参考文献

- [1] 袁金灿,马进,王思彤,等. 智能电能表可靠性预计技术[J]. 电力自动化设备,2013,33(7):161-166.
YUAN J C, MA J, WANG S T, et al. Reliability prediction technology of smart electric energy meters [J]. Electric Power Automation Equipment, 2013, 33 (7): 161-166.
- [2] GAO J, HUANG D, LI X. Accelerative effect analysis of quantitative environmental stress screening [C]. 11th International Conference on Reliability, Maintainability and Safety (ICRMS), 2016, doi: 10.1109/ICRMS.2016.8050169.
- [3] CAI B, HUANG L, XIE M. Bayesian networks in fault diagnosis [J]. IEEE Transactions on Industrial Informatics, 2017, 13(5): 2227-2240.
- [4] JIANG Q C, YAN X F, HUANG B A. Performance-driven distributed PCA process monitoring based on fault-relevant variable selection and bayesian inference [J]. IEEE Transactions on Industrial Electronics, 2016, 63(1): 377-386.
- [5] SHENG H M, XIAO J, CHENG Y H, et al. Short-term

- solar power forecasting based on weighted Gaussian process regression [J]. IEEE Transactions on Industrial Electronics, 2018, 65(1): 300-308.
- [6] 洪朝飞, 王江. 基于脉冲神经网络的短期负荷预测模型[J]. 电力系统及其自动化学报, 2020, 32(10): 139-144. .
HONG CH F, WANG J. Short-term load forecasting model based on spike neural network [J]. Journal of Electric Power System and Automation, 2020, 32(10): 139-144.
- [7] QIU W, TANG Q, TENG Z S, et al. Failure rate prediction of electrical meters based on weighted hierarchical Bayesian [J]. Measurement, 2019, 142: 21-29.
- [8] 康守强, 许林虎, 王玉静, 等. 基于 FOA-MKSVM 的滚动轴承故障分类方法[J]. 仪器仪表学报, 2015, 36(5): 1186-1192.
KANG SH Q, XU L H, WANG Y J, et al. Rolling bearing fault classification method based on FOA-MKSVM [J]. Chinese Journal of Scientific Instrument, 2015, 36(5): 1186-1192.
- [9] 叶永伟, 陆俊杰, 钱志勤, 等. 基于 LS-SVM 的机械式温度仪表误差预测研究[J]. 仪器仪表学报, 2016, 37(1): 57-66.
YE Y W, LU J J, QIAN ZH Q, et al. Research on error prediction of mechanical temperature meter based on LS-SVM [J]. Chinese Journal of Scientific Instrument, 2016, 37(1): 57-66.
- [10] DONG S J, LUO T H. Bearing degradation process prediction based on the PCA and optimized LS-SVM model [J]. Measurement, 2013, 46(9): 3143-3152.
- [11] PARVEZ I, AGHILI M, SARWAT A I, et al. Online power quality disturbance detection by support vector machine in smart meter [J]. Journal of Modern Power Systems and Clean Energy, 2019, 7(5): 1328-1339.
- [12] 史庆军, 郭晓振, 刘德胜. 基于特征量融合和支持向量机的轴承故障诊断[J]. 电子测量与仪器学报, 2019, 33(10): 104-111.
SHI Q J, GUO X ZH, LIU D SH. Bearing fault diagnosis based on feature fusion and support vector machine [J]. Journal of Electronic Measurement and Instrumentation, 2019, 33(10): 104-111.
- [13] 张景元, 何玉珠, 崔唯佳. 基于多应力退化模型的智能电表可靠寿命预估[J]. 北京航空航天大学学报, 2017, 43(8): 1662-1669.
ZHANG J Y, HE Y ZH, CUI W J. Reliable life estimation of smart meters based on multi-stress degradation model [J]. Journal of Beijing University of Aeronautics and Astronautics, 2017, 43(8): 1662-1669.
- [14] 闫富松, 李朝晖. 一种基于皮尔逊相关系数识别变压器励磁涌流的方法[J]. 高压电器, 2016, 52(8): 52-56, 63.
YAN F S, LI ZH H. A method for identifying magnetizing inrush current of transformers based on Pearson correlation coefficient [J]. High Voltage Apparatus, 2016, 52(8): 52-56, 63.
- [15] 康宁宁, 李川, 曾虎, 等. 采用 FCM 聚类与改进 SVR 模型的窃电行为检测[J]. 电子测量与仪器学报, 2017, 31(12): 2023-2029.
KANG N N, LI CH, ZENG H, et al. Electricity theft detection using FCM clustering and improved SVR model [J]. Journal of Electronic Measurement and Instrumentation, 2017, 31(12): 2023-2029.
- [16] CHENG G, TONG X. Fuzzy clustering multiple kernel support vector machine [C]. International Conference on Wavelet Analysis and Pattern Recognition (ICWAPR), 2018: 7-12.
- [17] 周绍磊, 廖剑, 史贤俊. RBF-SVM 的核参数选择方法及其在故障诊断中的应用[J]. 电子测量与仪器学报, 2014, 28(3): 240-246.
ZHOU SH L, LIAO J, SHI X J. RBF-SVM kernel parameter selection method and its application in fault diagnosis [J]. Journal of Electronic Measurement and Instrumentation, 2014, 28(3): 240-246.
- [18] 刘峥, 黄真银, 徐成良, 等. 基于网格搜索优化的主成分分析-支持向量机算法的冷水机组能耗预测[J]. 制冷技术, 2019, 39(6): 15-20.
LIU ZH, HUANG ZH Y, XU CH L, et al. Principal component analysis based on grid search optimization-support vector machine algorithm for energy consumption prediction of chillers [J]. Refrigeration Technology, 2019, 39(6): 15-20.
- [19] 杜蜀薇, 成达, 邱伟, 等. 基于分层贝叶斯的电能表短期故障评估与预测[J]. 电子测量与仪器学报, 2018, 32(12): 178-184.
DU SH W, CHENG D, QIU W, et al. Evaluation and prediction of short-term failure of electric energy meter based on hierarchical Bayes [J]. Journal of Electronic Measurement and Instrumentation, 2018, 32(12): 178-184.
- [20] 金阳忻. 基于模糊神经网络的电能表误差超差风险预测模型[J]. 电网技术, 2019, 43(6): 1962-1968.
JIN Y X. Electricity meter error outlier risk prediction model based on fuzzy neural network [J]. Grid Technology, 2019, 43(6): 1962-1968.

作者简介



王永超, 2015 年于北京科技大学获得硕士学位, 目前在国网新疆电力公司营销服务中心从事计量器具质量检测和新型营销业务研究。

E-mail: 329647745@qq.com

Wang Yongchao received M. Sc. from Beijing University of Science and Technology in 2015. Now he is working in Marketing Service Center of Xinjiang Electric Power Company of State Grid. His main research interests include quality testing of measuring instruments and new marketing

business.



杨莹莹(通信作者), 2019 年于湖南大学获得学士学位, 现为湖南大学硕士研究生, 主要研究方向为智能电网设备的可靠性分析。

E-mail: yangyingying@hnu.edu.cn

Yang Yingying (Corresponding author) received B. Sc. from Hunan University in 2019. Now she is a M. Sc. candidate at Hunan University. Her main research interest includes reliability analysis of smart grid equipment.