

DOI: 10.19783/j.cnki.pspc.190631

基于多层极限学习机的电能质量扰动多标签分类算法

靳果¹, 朱清智¹, 孟阳², 闫奇³

(1. 河南工业职业技术学院, 河南 南阳 473000; 2. 西安交通大学, 陕西 西安 710000;
3. 国网南阳供电公司, 河南 南阳 473000)

摘要: 电力系统中电能质量扰动分类特征选择标准不统一、泛化能力差、分类效果与分类效率有待提高。为了解决这些问题,一方面,引入多层极限学习机自编码器,优化输入权重,完成电能质量扰动信号的特征提取。另一方面,引入多标签排位分类算法,充分考虑各标签之间的相关性,完成电能质量扰动的分类。基于两种算法,设计出基于多层极限学习机的多标签分类模型,并得到多层极限学习机的最优网络结构和多标签分类的最佳分类阈值。实验结果表明,所提方法适用于电能质量单一扰动和复合扰动的分类,改善了分类效果和分类效率,具有较高的分类精度、良好的抗噪能力和泛化能力。

关键词: 电能质量; 扰动分类; 多层极限学习机; 多标签分类

Power quality disturbances multi-label classification algorithm based on a multi-layer extreme learning machine

JIN Guo¹, ZHU Qingzhi¹, MENG Yang², YAN Qi³

(1. Henan Polytechnic Institute, Nanyang 473000, China; 2. Xi'an Jiaotong University, Xi'an 710000, China;
3. State Grid Nanyang Power Supply Company, Nanyang 473000, China)

Abstract: In a power system, the characteristic selection criteria for power quality disturbance classification are not uniform. Generalization ability is weak, and the classification effect and efficiency need to be improved. In order to solve these problems, first a multi-layer extreme learning machine auto-encoder is used to optimize the input weights and extract the characteristics of electric power quality disturbances. Secondly a multi-label classification algorithm based on ranking is used to consider the correlation between labels and to classify various power quality disturbances. Combining the two algorithms, a multi-label classification model based on multi-level extreme learning machine is designed, and the optimal network structure of multi-level extreme learning machine and the optimal classification threshold of multi-label classification are obtained. The experimental results show that the proposed method can be applied to the classification of single and compound disturbances of power quality and improve the classification effect and efficiency with higher classification accuracy and excellent robustness and anti-noise ability.

This work is supported by National Natural Science Foundation of China (No. 51275084) and 2020 Science and Technology Research Program of Henan Province (No. 202102210134).

Key words: power quality; disturbances classification; multi-layer extreme learning machine; multi-label classification

0 引言

随着大量非线性、冲击性和不平衡负荷在电力系统中的投入使用,电能质量对电力系统安全稳定运行的影响更加明显,对电能质量扰动的实时有效

分类,成为进一步改善电能质量的基础。近年来,对电能质量扰动的分类得到了很好的研究^[1-5]:采用支持向量机法(SVM),可以检测复杂的电能质量问题;采用多标签K-近邻法(ML-KNN),可以清晰地检测、定位和分类不同的电能质量问题;采用多标签径向基神经网络(ML-RBF),可以提高对常见电能质量扰动的分类效率和正确率。

研究发现,现有的电能质量扰动分类方法,缺

基金项目:国家自然科学基金项目资助(51275084);河南省2020年度科技攻关计划项目资助(202102210134)

乏统一的分类标准, 易产生冗余特征, 从而导致分类准确率下降, 并且泛化能力差, 受噪声影响较大。本文以提高分类精度及分类效率为研究目标, 采用多层极限学习机自编码网络结构, 提出了一种基于排位的电能质量扰动分类模型, 并且与其他分类方法在电能质量扰动分类上进行了对比, 验证了所提方法的有效性。

1 多层极限学习机的多标签分类模型

基于自编码的多层极限学习机, 能够有效表征复杂函数、具有较高的预测精度和泛化能力; 基于排位的多标签分类算法, 考虑到各标签之间的相关性, 适合于各种电能质量扰动的分类, 具有较强的抗噪能力。本文将多层极限学习机和多标签分类算法相结合, 提出一种基于多层极限学习机的多标签分类模型。

1.1 多层极限学习机自编码器结构

极限学习机自编码器(ELM-AE)对输入权重进行优化, 提高了分类精度和泛化能力, 克服了极限学习机随机权重和隐层阈值所导致的神经元无效的问题, 并且提高了分类效率^[6-7]。

极限学习机自编码器网络结构如图 1 所示, 输入层和输出层节点数为 n , 隐层节点数为 l 。

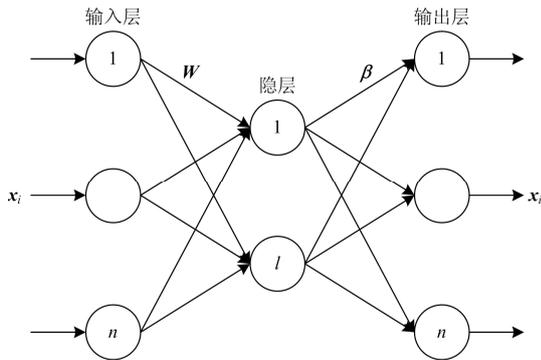


图 1 ELM-AE 网络结构

Fig. 1 Structure of ELM-AE

假设有数量为 N 的任意样本空间 \mathbf{X} , 即输入和输出数据样本集合为 $\mathbf{X}=[\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T$, 其中 $\mathbf{x}_i (1 \leq i \leq N)$ 表示第 i 组样本向量。

引入输入层与隐层之间的正交随机权值矩阵 \mathbf{W} 和隐层的正交随机偏置向量 \mathbf{b} , 即

$$\begin{cases} \mathbf{W}^T \mathbf{W} = \mathbf{I} \\ \mathbf{b}^T \mathbf{b} = 1 \end{cases} \quad (1)$$

式中, $\mathbf{W}=[w_1, w_2, \dots, w_l]$, 其中 $w_j=[w_{1j}, w_{2j}, \dots, w_{nj}]^T (1 \leq j \leq l)$ 表示所有输入层节点与隐层第 j 个节点之

间的正交随机权值向量, w_{ij} 表示输入层第 i 个节点与隐层第 j 个节点之间的权值; $\mathbf{b}=[b_1, b_2, \dots, b_l]^T$ 。

可得:

$$\mathbf{H} = \begin{bmatrix} f(\mathbf{w}_1 \cdot \mathbf{x}_1 + \mathbf{b}_1) & \dots & f(\mathbf{w}_l \cdot \mathbf{x}_1 + \mathbf{b}_l) \\ \vdots & & \vdots \\ f(\mathbf{w}_1 \cdot \mathbf{x}_N + \mathbf{b}_1) & \dots & f(\mathbf{w}_l \cdot \mathbf{x}_N + \mathbf{b}_l) \end{bmatrix} \quad (2)$$

式中: \mathbf{H} 表示隐层的输出矩阵, 其中 $h_{ij} = f(\mathbf{w}_j \cdot \mathbf{x}_i + \mathbf{b}_j)$ 表示第 i 个样本在隐层第 j 个节点上的输出映射; $f(x)$ 为激活函数。

通过重构矩阵 β , 将隐层输出矩阵重构为输入样本集合, 即 $\mathbf{X} = \mathbf{H}\beta$, β 可由式(3)求出。

$$\beta = (\frac{\mathbf{I}}{C} + \mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{X} \quad (3)$$

式中: $\beta = [\beta_1, \beta_2, \dots, \beta_n]$, 其中 $\beta_k = [\beta_{1k}, w_{2k}, \dots, w_{lk}]^T (1 \leq k \leq n)$ 表示所有隐层节点与第 k 个输出节点之间的权值; β_{jk} 表示隐层第 j 个节点与输出层第 k 个节点之间的权值; \mathbf{I} 为单位矩阵; C 为正则项系数。

将 β 的转秩矩阵作为原网络结构输入层与隐层的权值矩阵, 完成单层 ELM-AE 训练, 如图 2 所示。

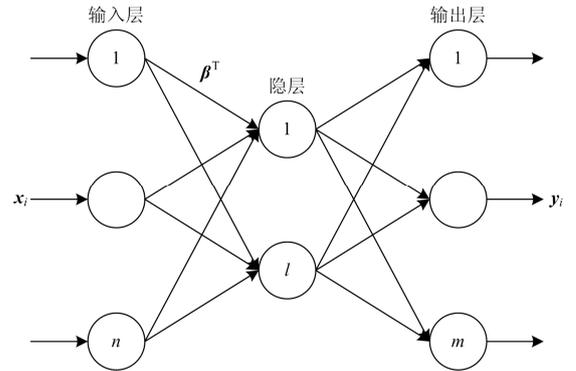


图 2 ELM-AE 重构

Fig. 2 Reconfiguration of ELM-AE

多层极限学习机自编码器(ML-ELM-AE)在 ELM-AE 的基础上进行堆叠运算^[8], 构建了含多层隐层的网络结构, 如图 3 所示, 可以提高对电能质量扰动分类的平均精度和泛化能力^[9]。

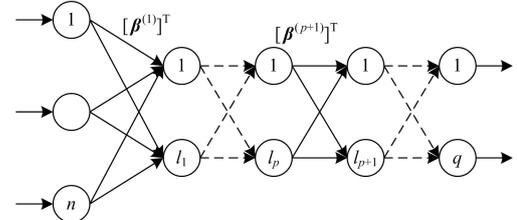


图 3 ML-ELM-AE 网络结构

Fig. 3 Structure of ML-ELM-AE

第一步, 利用输入数据样本集合, 根据式(1)~式(3), 计算第 1 层隐层的输出矩阵 $\mathbf{H}^{(1)}$ 和重构矩阵 $\boldsymbol{\beta}^{(1)}$ 。

第二步, 将 $\mathbf{H}^{(1)}$ 作为第 2 层隐层的输入, 得到第二层的输出矩阵 $\mathbf{H}^{(2)}$ 和重构矩阵 $\boldsymbol{\beta}^{(2)}$ 。

第三步, 以此类推, 以第 p 层隐层的输出矩阵 $\mathbf{H}^{(p)}$ 为输入, 得到第 $p+1$ 层隐层的输出矩阵 $\mathbf{H}^{(p+1)}$ 和权值矩阵 $[\boldsymbol{\beta}^{(p+1)}]^\top$ 。

1.2 多标签分类算法

多标签分类算法, 单个数据样本属于多个类别标签, 充分考虑了标签之间的相关性^[10-12], 本文采用排位分类法完成多标签分类问题。

假设样本空间为 \mathbf{S} , 每个样本 \mathbf{s}_i 都是 d 维的特征向量, 即 $\mathbf{s}_i = [s_{i1}, s_{i2}, \dots, s_{id}]^\top$; \mathbf{Y} 为类别标签集, 包含 q 种标签, 即 $\{y_1, y_2, \dots, y_q\}$ 。对于样本 \mathbf{s}_i , 其对应的标签集合 $\mathbf{y}_i = y(\mathbf{s}_i) \subseteq \mathbf{Y}$, \mathbf{y}_i 包含 \mathbf{Y} 中的一个或多个标签元素。多标签分类算法的目标是根据给定的训练样本 $(\mathbf{s}_i, \mathbf{y}_i)$ 学习一个分类函数, 通过学习到的分类函数来预测一个测试实例 \mathbf{s}_i 的类别标签集合 \mathbf{y}_i 。

基于排位的多标签分类算法^[13], 根据训练样本集构造一个特征值到标签所属程度值的映射 $g: \mathbf{S} \rightarrow \mathbf{Y}$ 。

对于样本 \mathbf{s}_i , g 按照样本所带标签的排位值进行排列, 即 $\mathbf{y}_1, \mathbf{y}_2 \in \mathbf{Y}$, $\mathbf{y}_1 \in \mathbf{Y}_i$, $\mathbf{y}_2 \notin \mathbf{Y}_i$ 时, $g(\mathbf{s}_i, \mathbf{y}_1) > g(\mathbf{s}_i, \mathbf{y}_2)$ 。由 g 得排位函数 g_{rank} , 当 $g(\mathbf{s}_i, \mathbf{y}_1) > g(\mathbf{s}_i, \mathbf{y}_2)$ 时, $g_{\text{rank}}(\mathbf{s}_i, \mathbf{y}_1) < g_{\text{rank}}(\mathbf{s}_i, \mathbf{y}_2)$ 。

由此可得, 样本 \mathbf{s}_i 的预测标签集合 $h(\mathbf{s}_i) = \{\mathbf{y}_j | g_{\text{rank}}(\mathbf{s}_i, \mathbf{y}_j) > t(\mathbf{s}_i), \mathbf{y}_j \in \mathbf{Y}\}$, 其中, $t(\mathbf{s}_i)$ 为阈值函数。基于排位的多标签分类, 关键在于找到映射函数 g 以及阈值函数 $t(\mathbf{s}_i)$ ^[14]。

衡量多标签分类算法效果的评价指标^[15-16]主要有汉明损失、一类错误、排位损失、覆盖率和平均精度。

汉明损失 f_{HL} 考虑错误标签和丢失标签, 衡量单个标签出现错误划分的程度。

$$f_{\text{HL}}(h) = \frac{1}{n} \sum_{i=1}^n \frac{1}{L} |h(\mathbf{s}_i) \Delta \mathbf{Z}_i| \quad (4)$$

式中: n 表示样本个数; L 表示标签总数; $h(\mathbf{s}_i)$ 表示预测标签集; \mathbf{Z}_i 表示实际标签集; Δ 表示计算去对称差分。

一类错误 f_{OE} 衡量预测标签排序第一的不是样本真实标签的概率情况:

$$f_{\text{OE}}(f) = \frac{1}{n} \sum_{i=1}^n \{f_{\text{argmax}}(\mathbf{s}_i, \mathbf{y} \in \mathbf{y}_i)\} \quad (5)$$

式中, f_{argmax} 表示使得 $f(\mathbf{s}_i, \mathbf{y})$ 最大的属于标签集 \mathbf{y}_i 的标签 \mathbf{y} 。

排位损失 f_{RL} 衡量预测标签排序与实际排序间的差异程度。

$$f_{\text{RL}}(f) = \frac{1}{n} \sum_{i=1}^n \frac{|D_i|}{|\mathbf{y}_i| \|\mathbf{y}_i\|} \quad (6)$$

式中: D_i 表示错判的标签集合; $\|\mathbf{y}_i\|$ 表示 \mathbf{y}_i 的补集。

覆盖率 f_{C} 衡量预测标签的排序序列中, 遍历真实标签集所需要的平均搜索深度:

$$f_{\text{C}}(f) = \frac{1}{n} \sum_{i=1}^n \max f_{\text{rank}}(\mathbf{x}_i, \mathbf{y}) - 1 \quad (7)$$

平均精度 f_{AP} 衡量预测标签的平均精确程度:

$$f_{\text{AP}}(f) = \frac{1}{n} \sum_{i=1}^n \frac{1}{|\mathbf{y}_i|} \sum_{\mathbf{y} \in \mathbf{y}_i} \frac{|L_i|}{f_{\text{rank}}(\mathbf{x}_i, \mathbf{y})} \quad (8)$$

式中, L_i 表示预测正确的标签样本集合。

1.3 基于多层极限学习机的多标签分类模型

本文将多层极限学习机自编码结构和基于排位的多标签分类算法相结合, 构建如图 4 所示的分类器模型基本结构^[17]。

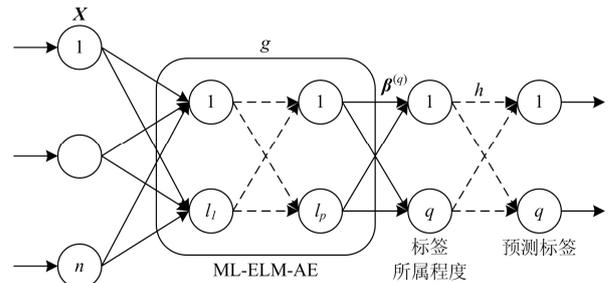


图 4 分类器模型

Fig. 4 Classifier model

根据图 4 模型, 输入样本数据经过 ML-ELM-AE 网络实现到标签的映射, 再经过分类函数的转换得到分类结果^[18]。其中, \mathbf{X} 表示输入样本数据集合; ML-ELM-AE 含有 p 层隐层, 实现样本空间到标签的映射, 即 g 的映射过程; $h(\mathbf{s}_i)$ 为分类函数; $\boldsymbol{\beta}^{(p)}$ 为第 p 层隐层的输出权值, 由式(9)得到

$$\boldsymbol{\beta} = \left(\frac{\mathbf{I}}{C} + \mathbf{H}^\top \mathbf{H} \right)^{-1} \mathbf{H}^\top \mathbf{T} \quad (9)$$

式中, \mathbf{T} 表示样本的实际类别标签。

对电能质量扰动进行分类的过程中, ML-ELM-AE 算法根据输入数据样本集合, 通过实验确定最优的网络结构, 产生各隐层的输出矩阵和权值矩阵; 基于排位的多标签分类算法通过实验确定最优的分

类阈值^[19], 并且将分类阈值归一化到[0,1]。将两种算法相结合, 设计出多层极限学习机的多标签分类算法, 适用于电能质量单一扰动的分类和复合扰动的分类, 具有良好的抗噪能力和较高分类精度^[20], 同时克服了传统 ELM 参数赋值的随机性, 具有较高的分类效率、良好的鲁棒性和泛化能力。

2 实验结果和分析

本文利用 Matlab 环境完成仿真实验, 采用平均精度、汉明损失、一类错误、排位损失、覆盖率作为分类效果的评价指标, 另外, 实验还增加了分类器的训练时间和测试时间作为衡量分类效率的评价指标。为减少误差, 每一实验数据均为算法程序在同一实验条件下运行 20 次所得的算术平均值。

2.1 电能质量扰动信号模型

常见的单一电能质量扰动信号模型^[21]如表 1 所示, 其中基波频率 f_0 为 50 Hz, T 表示工频周期, $u(t)$ 表示单位阶跃函数。实验中出现的信号类型包括标准信号、单一扰动信号和复合扰动信号共计 48 种, 其中复合扰动信号 40 种^[22]。根据复合扰动的存在形式, 第一类复合扰动由两种单一扰动构成, 第二类由三种单一扰动构成, 第三类由四种单一扰动构成。

表 1 电能质量扰动信号模型

Table 1 Models of power quality disturbance signal

类型	数学模型	参数
标准信号	$v(t) = \sin(\omega t)$	$\omega = 2\pi f_0$
电压暂降	$v(t) = \{1 + k[u(t_2) - u(t_1)]\} \sin(\omega t)$	$0.1 \leq k \leq 0.9$ $0.5T \leq t_2 - t_1 \leq 30T$
电压暂升	$v(t) = \{1 - k[u(t_2) - u(t_1)]\} \sin(\omega t)$	$0.9 \leq k \leq 1$ $0.5T \leq t_2 - t_1 \leq 30T$
电压中断	$v(t) = \{1 - k[u(t_2) - u(t_1)]\} \sin(\omega t)$	$0.1 \leq k \leq 0.9$ $0.5T \leq t_2 - t_1 \leq 30T$
谐波	$v(t) = \sin(\omega t) + \sum_{k=2}^{13} a_k \sin(k\omega t)$	$0.01 < a_k < 0.2$
暂态振荡	$v(t) = \sin(\omega t) + ae^{-\lambda(t-t_1)} \sin(\beta\omega t)[u(t-t_1)]$	$0.05 < a < 1$ $15 < \lambda < 130$ $14 \leq \beta < 30$
暂态脉冲	$v(t) = \sin(\omega t) + a[u(t_1) - u(t_2)]$	$a > 0.8$ $1 \text{ ms} < t_2 - t_1 < 3 \text{ ms}$
电压波动	$v(t) = [1 + a \sin(\beta\omega t)] \sin(\omega t)$	$0.05 < a < 0.2$ $0.1 < \beta < 0.5$

对实验中的信号以 3 200 Hz 为采样频率进行采样, 共采样 30 个周波, 总采样点数为 1 921。所有信号分别根据数学模型随机生成 200 个样本, 取 100 个作为训练数据, 另 100 个作为测试数据, 训练数

据集和测试数据集各包含 4 800 个样本。

2.2 分类器模型参数设计

本文通过实验确定 ML-ELM-AE 的最优网络结构, 采用 ML-ELM-AE 得到多标签映射函数 $g(s_i, y_j)$, 阈值函数 $t(s_j)$ 通过多次测试实验进行确定^[23]。

2.2.1 ML-ELM-AE 网络结构的确定

本文将 ML-ELM-AE 隐层层数限制在 9 层以内, 各层节点数设置为 50~2 000, 以 50 为节点数变化步长。将汉明损失作为分类效果评价依据, 同时参考训练时间和测试时间, 选择出分类效果好并且效率高的多层极限学习机网络结构参数。

为了得到不同总层数时, 各层最优节点数, 设计如下的实验步骤:

① 设置 ML-ELM-AE 的隐层层数为 1。

② 设置隐层节点数按照 $50+L \times 50$ (L 为从 0 到 39 递增的整数) 的规律变化, 得到每一个 L 值对应的汉明损失、训练时间和测试时间, 选择最优的 L 值, 即该层的最优节点数为 $50+L \times 50$ 。

③ 设置 ML-ELM-AE 的层数为 2, 其中第 1 层的节点数已经由步骤①和②固定, 重复步骤②, 得到第 2 层的最优 L 值, 得到该层的最优节点数。

④ 依次设置 ML-ELM-AE 的层数为 3~9, 得到各层的最优节点数。

表 2 为 ML-ELM-AE 在总层数为 1~9 时, 汉明损失最优值对应的各层节点数, 可以发现: 从 4 层开始, 新增隐层的最优节点数都为 750。

表 2 ML-ELM-AE 在不同总层数下的最优节点数

Table 2 Optimal node numbers of ML-ELM-AE with

different total layers

总层数	第 k 层								
	$k=1$	$k=2$	$k=3$	$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$
1	350	—	—	—	—	—	—	—	—
2	350	1 150	—	—	—	—	—	—	—
3	350	1 150	350	—	—	—	—	—	—
4	350	1 150	350	750	—	—	—	—	—
5	350	1 150	350	750	750	—	—	—	—
6	350	1 150	350	750	750	750	—	—	—
7	350	1 150	350	750	750	750	750	—	—
8	350	1 150	350	750	750	750	750	750	—
9	350	1 150	350	750	750	750	750	750	750

图 5 为 ML-ELM-AE 隐层层数与汉明损失的关系, 可以发现: 汉明损失呈现出先减小, 再趋于平稳的变化规律, 当隐层层数超过 4 层时, 汉明损失较低。

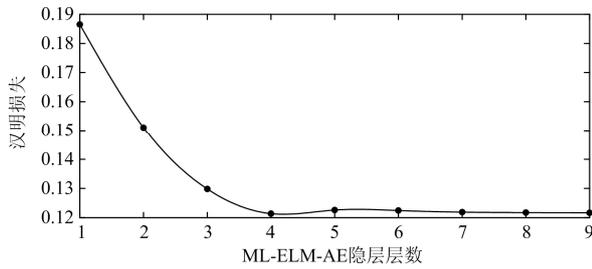


图 5 ML-ELM-AE 层数与汉明损失的关系
Fig. 5 Relation between ML-ELM-AE layer number and Hamming Loss

图 6 和图 7 分别为 ML-ELM-AE 隐层层数与训练时间和测试时间的关系，可以发现：随着隐层层数的增加，训练时间和测试时间整体均呈现出增长趋势。

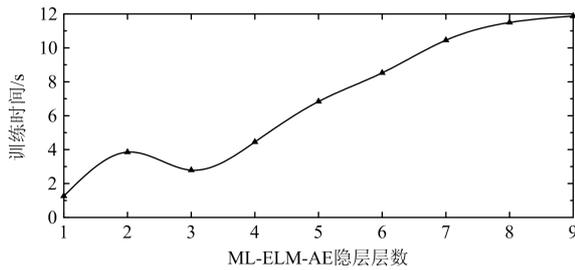


图 6 ML-ELM-AE 层数与训练时间的关系
Fig. 6 Relation between ML-ELM-AE layer number and training time

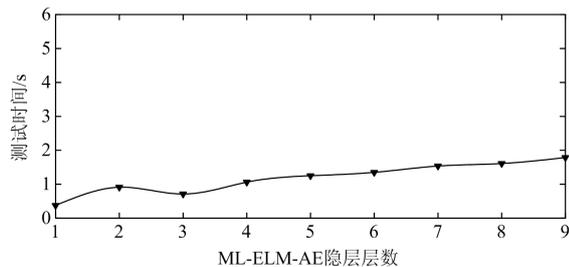


图 7 ML-ELM-AE 层数与测试时间的关系
Fig. 7 Relation between ML-ELM-AE layer number and testing time

综合上述实验结果，可以得到：ML-ELM-AE 隐层层数及节点个数与分类结果之间并非单一的数量关系，存在最优的隐层层数和节点个数，在保证分类效果的同时，节省训练和测试时间，提高分类效率。本文中，ML-ELM-AE 隐层层数为 4 时，电能质量扰动分类效果最优，并且效率较高，因此，本文将多层极限学习机设定为“输入层—隐层(350—1150—350—750)—输出层”的结构。

2.2.2 分类阈值的确定

实验采用最优的 4 层 ML-ELM-AE 网络结构，多标签分类器的分类阈值设置为 0.05~0.95 范围内，以 0.05 为步进间隔。以汉明损失、排位损失、一类错误、覆盖率、平均精度作为分类效果的评价依据。

图 8 为分类阈值与汉明损失的关系，可以发现：随着分类阈值的增大，汉明损失迅速降低并趋于稳定，然后略有升高，分类阈值在 0.55 附近变化时，汉明损失值较小。

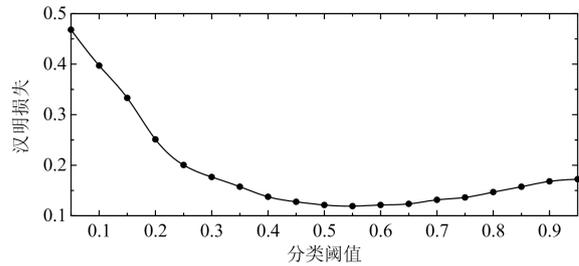


图 8 分类阈值与汉明损失的关系
Fig. 8 Relation between classification threshold value and Hamming Loss

图 9 和图 10 分别为分类阈值与一类错误、排位损失的关系，可以发现：分类阈值的变化对这两项评价指标的影响较小，一类错误值在 0.086 附近浮动，排位错误值在 0.063 附近浮动，且变化幅度很小。

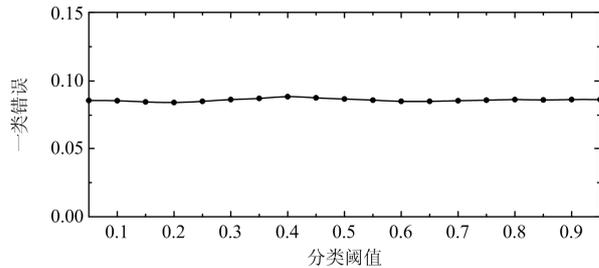


图 9 分类阈值与一类错误的关系
Fig. 9 Relation between classification value and one-type error

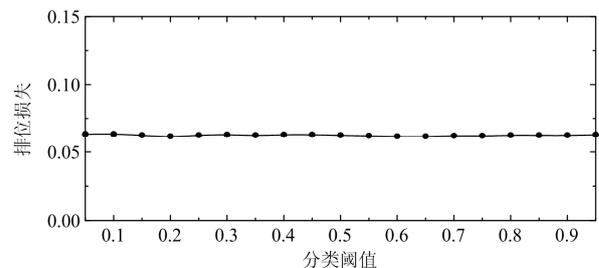


图 10 分类阈值与排位损失的关系
Fig. 10 Relation between classification threshold value and ranking loss

图 11 为分类阈值与覆盖率的关系, 可以发现: 随分类阈值的变化, 覆盖率在 2.006~2.017 之间浮动变化, 当分类阈值在 0.6 附近时, 覆盖率较小。

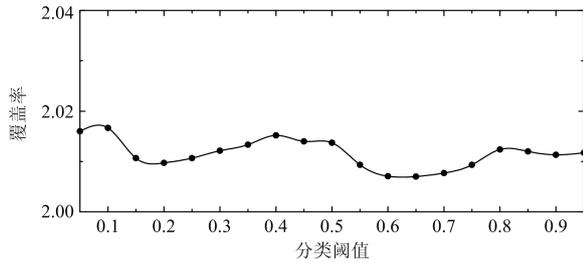


图 11 分类阈值与覆盖率的关系

Fig. 11 Relation between classification threshold value and coverage

图 12 为分类阈值与平均精度的关系, 可以发现: 平均精度值变化范围为 0.9075~0.9085, 当分类阈值在 0.65 附近变化时, 平均精度值较高。

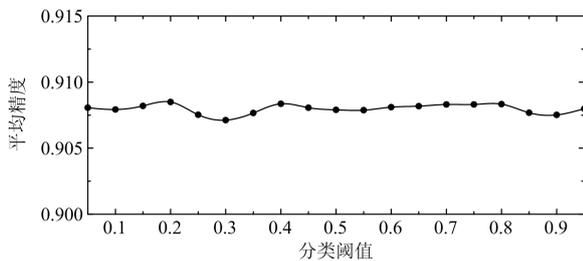


图 12 分类阈值与平均精度的关系

Fig. 12 Relation between classification threshold value and average precision

综合上述实验结果, 可以得到: 分类阈值对汉明损失的影响较为明显, 覆盖率和平均精度在一定程度上受到分类阈值变化的影响, 一类错误和排位错误受分类阈值的影响不明显, 本文将多标签分类器的分类阈值设定为 0.6。

2.3 分类实验结果和对比

2.3.1 分类过程

利用多层极限学习机的多标签分类模型, 对电能质量扰动信号所包含的扰动类别进行预测^[24-25]。

第一步, 将电能质量扰动信号进行离散小波变换, 得到各层的分解系数, 并把分解系数分为训练数据集和测试数据集。

第二步, 随机产生正交化的 ML-ELM-AE 网络各层权值。

第三步, 将训练数据集作为输入, 对多层极限学习机网络进行训练, 调整网络各层权值。

第四步, 根据 ML-ELM-AE 的映射关系, 得到多标签分类算法的分类函数。

最后, 将测试数据集作为输入, 利用训练好的多层极限学习机的多标签分类模型进行扰动类别预测。

2.3.2 实验结果和对比

实验研究对象为单一扰动信号和复合扰动信号, 对每种电能质量扰动信号分别叠加信噪比(SNR)为 50 dB、40 dB、30 dB 及 20 dB 的噪声干扰。实验数据为同一噪声强度下, 各类扰动信号的实验结果平均值。为了进一步验证本文所提电能质量扰动分类方法的性能, 分别采用 SVM、ML-KNN 和 ML-RBF 完成扰动信号分类, 并将结果与本文所提方法进行对比。

表 3 为不同噪声水平下各类干扰信号分类结果的汉明损失, 图 13 为不同分类方法的汉明损失对比结果, 可以发现: 噪声强度由 50 dB 减弱到 20 dB 的过程中, 本文方法的汉明损失值虽然呈现增大的变化趋势, 但是均低于另外三种方法。

表 3 不同噪声下分类结果的汉明损失

扰动信号	干扰噪声/dB			
	50	40	30	20
单一扰动	0.118	0.117	0.138	0.161
第一类复合扰动	0.114	0.121	0.137	0.169
第二类复合扰动	0.125	0.127	0.145	0.195
第三类复合扰动	0.128	0.132	0.148	0.215
平均值	0.121	0.124	0.142	0.185

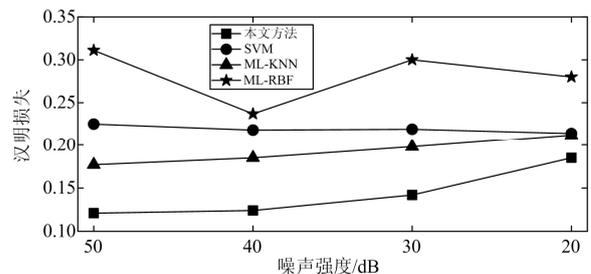


图 13 汉明损失的对比结果

Fig. 13 Comparison results of Hamming Loss

表 4 为不同噪声水平下各类干扰信号分类结果的一类错误, 图 14 为不同分类方法的一类错误对比

表 4 不同噪声下分类结果的一类错误

扰动信号	干扰噪声信噪比/dB			
	50	40	30	20
单一扰动	0.076	0.069	0.085	0.126
第一类复合扰动	0.077	0.077	0.094	0.135
第二类复合扰动	0.087	0.080	0.102	0.160
第三类复合扰动	0.092	0.087	0.108	0.178
平均值	0.083	0.078	0.097	0.15

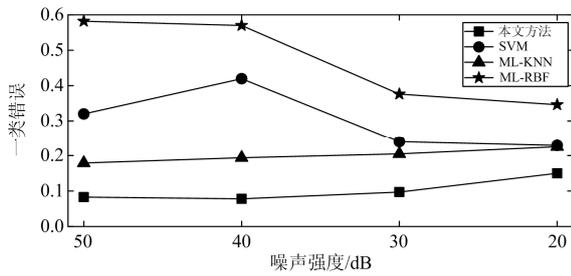


图 14 一类错误的对比结果

Fig. 14 Comparison results of one-type error

结果, 可以发现: 在不同强度的噪声干扰下, 本文方法的一类错误值均为最低, 并且不同环境下的实验结果相差不大。

表 5 和表 6 为不同噪声水平下各类干扰信号分类结果的排位损失和覆盖率, 图 15 和图 16 为不同分类方法的排位损失和覆盖率对比结果, 可以发现: 本文方法的排位损失值和覆盖率均低于另外三种方法。

表 5 不同噪声下分类结果的排位损失

Table 5 Ranking loss under different SNR values

扰动信号	干扰噪声信噪比/dB			
	50	40	30	20
单一扰动	0.058	0.062	0.070	0.120
第一类复合扰动	0.059	0.063	0.069	0.104
第二类复合扰动	0.062	0.066	0.074	0.134
第三类复合扰动	0.064	0.069	0.075	0.150
平均值	0.061	0.065	0.072	0.127

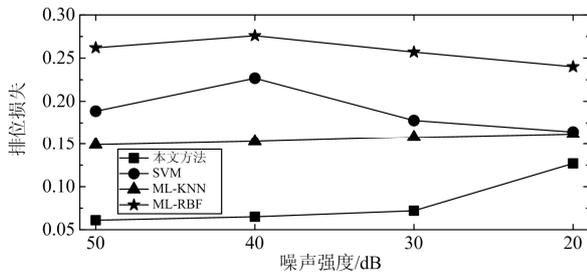


图 15 排位损失的对比结果

Fig. 15 Comparison results of ranking loss

表 6 不同噪声下分类结果的覆盖率

Table 6 Coverage under different SNR values

扰动信号	干扰噪声信噪比/dB			
	50	40	30	20
单一扰动	1.899	1.823	1.851	2.278
第一类复合扰动	1.946	1.852	2.059	2.390
第二类复合扰动	2.045	2.100	2.150	2.762
第三类复合扰动	2.113	2.201	2.326	3.039
平均值	2.001	1.994	2.097	2.617

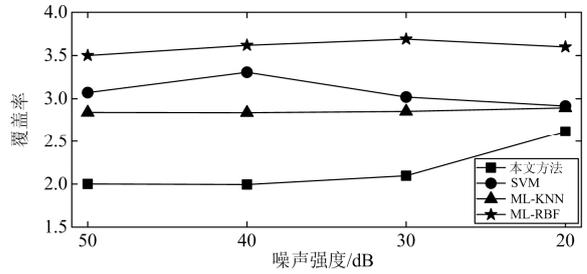


图 16 覆盖率的对比结果

Fig. 16 Comparison results of coverage

表 7 为不同噪声水平下各类干扰信号分类结果的平均精度, 图 17 为不同分类方法的平均精度对比结果, 可以发现: 随着干扰强度的减弱, 本文方法的平均精度值虽然呈下降趋势, 但是高于另外三种方法。

表 7 不同噪声下分类结果的平均精度

Table 7 Average precision under different SNR values

扰动信号	干扰噪声信噪比/dB			
	50	40	30	20
单一扰动	0.959	0.963	0.933	0.925
第一类复合扰动	0.936	0.932	0.915	0.882
第二类复合扰动	0.858	0.887	0.863	0.778
第三类复合扰动	0.884	0.866	0.868	0.766
平均值	0.909	0.912	0.895	0.838

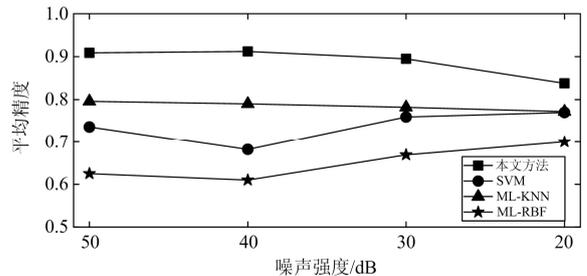


图 17 平均精度的对比结果

Fig. 17 Comparison results of average precision

表 8 和表 9 为不同噪声水平下各类干扰信号分类结果的训练时间和测试时间, 图 18 和图 19 为不同分类方法的训练时间和测试时间对比结果, 可以发现: 本文方法可以在 10 s 内完成训练和测试, 分类效率优势明显, 表明基于多层极限学习机的多标签分类模型训练好后, 可以实现对电能质量扰动的快速分类, 适用于大量干扰信号的分类识别。

综合上述实验结果可得, 本文所提的分类器模型与其他分类方法相比, 汉明损失、一类错误、排位损失、覆盖率均较低, 平均精度较高, 具有良好的分类效果, 并且分类的训练时间和测试时间较低, 表现出较高的分类效率。

表 8 不同噪声下分类结果的训练时间

Table 8 Training time under different SNR values

扰动信号	干扰噪声信噪比/dB			
	50	40	30	20
单一扰动/s	4.816	4.484	4.838	4.606
第一类复合扰动/s	4.673	4.594	4.694	4.720
第二类复合扰动/s	5.098	4.828	5.121	4.960
第三类复合扰动/s	5.226	4.989	5.250	5.125
平均值/s	4.953	4.724	4.976	4.853

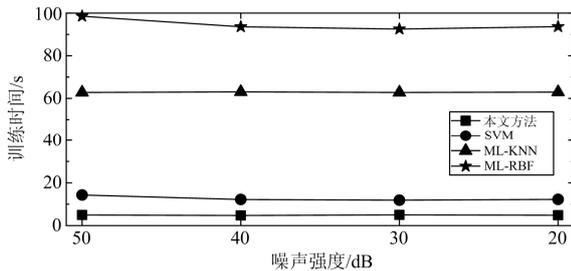


图 18 训练时间的对比结果

Fig. 18 Comparison results of training time

表 9 不同噪声下分类结果的测试时间

Table 9 Testing time under different SNR values

扰动信号	干扰噪声信噪比/dB			
	50	40	30	20
单一扰动/s	1.077	1.119	1.085	1.122
第一类复合扰动/s	1.104	1.086	1.112	1.089
第二类复合扰动/s	1.160	1.185	1.168	1.188
第三类复合扰动/s	1.199	1.214	1.207	1.218
平均值/s	1.135	1.151	1.143	1.154

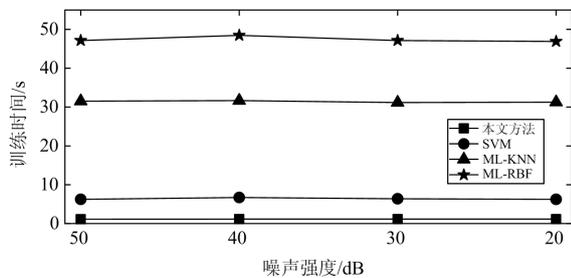


图 19 测试时间的对比结果

Fig. 19 Comparison results of testing time

第一, ML-ELM-AE 包括多个隐层, 具有较强的非线性映射能力和泛化能力, 可以有效表征复杂干扰信号, 提高分类的平均精度。

第二, ML-ELM-AE 可以提取出大量特征信息, 避免了算法随机赋值的偶然性, 在不同强度的噪声干扰下, 各项性能指标变化幅度较小, 具有良好的

泛化能力和鲁棒性。

第三, 基于排位的多标签分类算法, 充分考虑到各标签之间的相关性, 适合于单一扰动和复合扰动的分类, 具有良好的抗噪能力和分类准确率。

第四, ML-ELM-AE 算法无需迭代, 减少了训练时间和测试时间, 在处理大量电能质量干扰时优势尤为明显。

3 应用案例

南阳地区电网电力负荷承担南阳全市 13 个县市市区供电和豫鄂陕三省网间部分电力交换的任务, 拥有 1 000 kV 变电站 1 座、500 kV 变电站 4 座、220 kV 变电站 28 座、110 kV 变电站 103 座、35 kV 变电站 182 座, 供电能力 550 万 kW; 35~220 kV 线路 621 条, 线路长 8 443.83 km。

本文以 2018 年南阳地区某县电能扰动为例, 研究分类算法在实际电网中的应用效果。根据文中所示的分类模型与参数范围, 对每一类电能质量扰动信号随机选取 100 个训练样本和 100 个测试样本。

表 10 为电能质量扰动分类结果, 图 20 为采用不同分类方法的分类结果对比情况, 图 21 为不同分类方法的测试时间对比情况, 可以发现: 本文分类方法在实际电网中的分类准确度和测试时间与仿真预测结果基本一致, 分类效果和分类效率明显优于其他方法。

表 10 电能质量扰动分类结果

Table 10 Classification results of power quality disturbances

扰动信号	分类结果			
	单一扰动	第一类复合扰动	第二类复合扰动	第三类复合扰动
单一扰动	95	5	0	0
第一类复合扰动	5	92	2	1
第二类复合扰动	0	11	85	4
第三类复合扰动	0	6	9	85

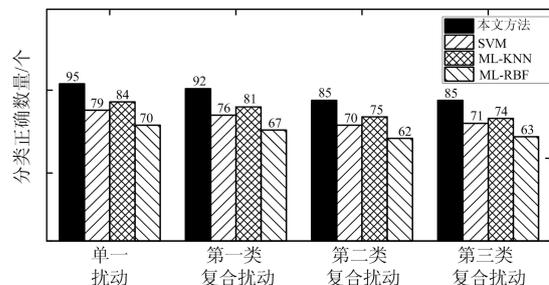


图 20 分类结果对比情况

Fig. 20 Comparison of classification results

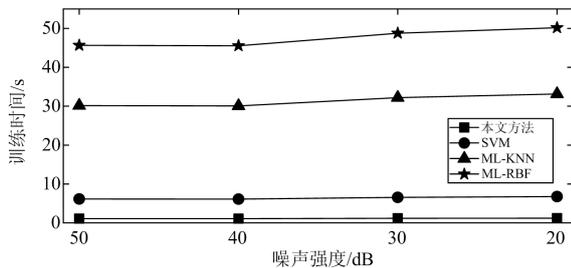


图 21 测试时间对比情况

Fig. 21 Comparison results of testing time

4 结论

本文将基于自编码的多层极限学习机和基于排位的多标签分类算法相结合,提出了一种新的电能质量扰动分类方法,并且对分类器的结构模型和分类过程做了说明。实验表明,本文方法在分类结果的汉明损失、平均精度、覆盖率等方面表现良好,并且具有较好的抗噪性和鲁棒性。本文采用的ML-ELM-AE结构,使训练时间和测试时间明显减少,当干扰信号数量较多时优势明显。

参考文献

- [1] 黄南天, 徐殿国, 刘晓胜. 基于 S 变换与 SVM 的电能质量复合扰动识别[J]. 电工技术学报, 2011, 26(10): 23-30.
HUANG Nantian, XU Dianguo, LIU Xiaosheng. Identification of power quality complex disturbances based on S-transform and SVM[J]. Transactions of China Electrotechnical Society, 2011, 26(10): 23-30.
- [2] ZHANG Minling, ZHOU Zhihua. ML-KNN: a lazy learning approach to multi-label learning[J]. Pattern Recognition, 2007, 40(7): 2038-2048.
- [3] ZHANG Minling. ML-RBF: RBF neural networks for multi-label learning[J]. Neural Processing Letters, 2009, 29(2): 61-74.
- [4] CHANDRASEKAR P, KAMARAJ V. Detection and classification of power quality disturbance waveform using MRA based modified wavelet transform and neural networks[J]. Journal of Electrical Engineering, 2010, 61(4): 235-240.
- [5] 李泰, 韩强, 黄银龙, 等. 基于 kNN 算法的红外测温图谱的温度数字识别研究[J]. 电力信息与通信技术, 2019, 17(6): 14-19.
LI Tai, HAN Qiang, HUANG Yinlong, et al. Research on digital recognition of infrared temperature map based on kNN algorithm[J]. Electric Power Information and Communication Technology, 2019, 17(6): 14-19.
- [6] 康松林, 刘乐, 刘楚楚, 等. 多层极限学习机在入侵检测中的应用[J]. 计算机应用, 2015, 35(9): 2513-2518.
KANG Songlin, LIU Le, LIU Chuchu, et al. Intrusion detection based on multiple layer extreme learning machine[J]. Journal of Computer Applications, 2015, 35(9): 2513-2518.
- [7] 文云峰, 赵荣臻, 肖友强, 等. 基于多层极限学习机的电力系统频率安全评估方法[J]. 电力系统自动化, 2019, 43(1): 133-140.
WEN Yunfeng, ZHAO Rongzhen, XIAO Youqiang, et al. Frequency safety assessment of power system based on multi-layer extreme learning machine[J]. Automation of Electric Power Systems, 2019, 43(1): 133-140.
- [8] 李德利, 张曦煌. 基于并行学习的多层极限学习机[J]. 计算机应用研究, 2018, 35(2): 459-461.
LI Deli, ZHANG Xihuang. Multi-layer extreme learning machine based on parallel learning[J]. Application Research of Computers, 2018, 35(2): 459-461.
- [9] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications[J]. Neurocomputing, 2006, 70(13): 489-501.
- [10] ZHAO Wenjing, SHANG Liqun, SUN Jinfan. Power quality disturbance classification based on time-frequency domain multifeature and decision tree[J]. Protection and Control of Modern Power Systems, 2019, 4(4): 337-342. DOI: 10.1186/s41601-019-0319-z.
- [11] ELISSEEFF A, WESTON J. A kernel method for multi-labelled classification[C] // The 14th International Conference on Neural Information Processing Systems: Natural and Synthetic, December 3-8, 2001, Vancouver, Canada: 681-687.
- [12] 瞿合祚, 刘恒, 李晓明, 等. 一种电能质量多扰动分类中特征组合优化方法[J]. 电力自动化设备, 2017, 37(3): 146-152.
QU Hezuo, LIU Heng, LI Xiaoming, et al. Feature combination optimization for multi-disturbance classification[J]. Electric Power Automation Equipment, 2017, 37(3): 146-152.
- [13] 张巧革, 刘志刚, 朱玲, 等. 基于多标签 Rank-WSVM 的复合电能质量扰动分类[J]. 中国电机工程学报, 2013, 33(28): 114-120.
ZHANG Qiaoge, LIU Zhigang, ZHU Ling, et al. Recognition of multiple power quality disturbances using multi-label wavelet support vector machine[J]. Proceedings of the CSEE, 2013, 33(28): 114-120.
- [14] 周维维, 管春, 卢伟国. 多标签分类法在电能质量复合扰动分类中的应用[J]. 中国电机工程学报, 2011, 31(4): 45-50.
ZHOU Luwei, GUAN Chun, LU Weiguo. Application of multi-label classification method to categorization of multiple power quality disturbances[J]. Proceedings of

- the CSEE, 2011, 31(4): 45-50.
- [15] YANG Q L, SHAO J M, SCHOLZ M, et al. Multi-label classification models for sustainable flood retention basins[J]. *Environmental Modelling & Software*, 2012, 32: 27-36.
- [16] ZHANG M L, JOSEM P, VICTOR R. Feature selection for multi-label naive Bayes classification[J]. *Information Sciences*, 2009, 179: 3218-3229.
- [17] INJETI S K, THUNUGUNTLA VK. Optimal integration of DGs into radial distribution network in the presence of plug-in electric vehicles to minimize daily active power losses and to improve the voltage profile of the system using bioinspired optimization algorithms[J]. *Protection and Control of Modern Power Systems*, 2020, 5(5): 21-35. DOI: 10.1186/s41601-019-0/49-x.
- [18] 许立武, 李开成, 罗奕, 等. 基于不完全 S 变换与梯度提升树的电能质量复合扰动识别[J]. *电力系统保护与控制*, 2019, 47(6): 24-31.
- XU Liwu, LI Kaicheng, LUO Yi, et al. Classification of complex power quality disturbances based on incomplete S-transform and gradient boosting decision tree[J]. *Power System Protection and Control*, 2019, 47(6): 24-31.
- [19] 单东, 许新征. 基于径向基神经网络和正则化极限学习机的多标签学习模型[J]. *模式识别与人工智能*, 2017, 30(9): 833-840.
- SHAN Dong, XU Xinzhen. Multi-label learning model based on multi-label radial basis function neural network and regularized extreme learning machine[J]. *Pattern Recognition & Artificial Intelligence*, 2017, 30(9): 833-840.
- [20] VINCENT P, LAROCHELLE H, BENGIO Y, et al. Extracting and composing robust features with denoising autoencoder[C] // *The 25th International Conference on Machine Learning*, June 5-9, 2008, Helsinki, Finland: 1096-1103.
- [21] 陈华丰, 张葛祥. 基于决策树和支持向量机的电能质量扰动识别[J]. *电网技术*, 2013, 37(5): 1272-1278.
- CHEN Huafeng, ZHANG Gexiang. Power quality disturbance identification using decision tree and support vector machine[J]. *Power System Technology*, 2013, 37(5): 1272-1278.
- [22] UYAR M, YILDIRIM S, GENCOGLU M T. An effective wavelet-based feature extraction method for classification of power quality disturbance signals[J]. *Electric Power Systems Research*, 2008, 78(10): 1747-1755.
- [23] YOUSSEF A M, ABDEL GALIL T K, ELSAADANY E F, et al. Disturbance classification utilizing dynamic time warping classifier[J]. *IEEE Transactions on Power Delivery*, 2004, 19(1): 272-278.
- [24] 谢善益, 肖斐, 艾芊, 等. 基于并行隐马尔科夫模型的电能质量扰动事件分类[J]. *电力系统保护与控制*, 2019, 47(2): 80-86.
- XIE Shanyi, XIAO Fei, AI Qian, et al. Parallel hidden Markov model based classification of power quality disturbance events[J]. *Power System Protection and Control*, 2019, 47(2): 80-86.
- [25] 王玥, 肖斐, 艾芊, 等. 基于深度置信网络的电能质量扰动事件分类[J]. *供用电*, 2019, 36(1): 40-45.
- WANG Yue, XIAO Fei, AI Qian, et al. Deep belief networks based classification of power quality disturbance events[J]. *Distribution & Utilization*, 2019, 36(1): 40-45.

收稿日期: 2019-06-02; 修回日期: 2019-08-22

作者简介:

靳 果 (1985—), 男, 通信作者, 硕士研究生, 讲师, 研究方向为电气自动化, 电子信息. E-mail: 27452154@qq.com

(编辑 张爱琴)