

DOI: 10.19783/j.cnki.pspc.191093

基于 RBM 和 SVM 的风电机组叶片开裂故障预测

张鑫¹, 徐遵义¹, 何慧茹¹, 王飞²

(1. 山东建筑大学计算机科学与技术学院, 山东 济南 250100; 2. 国网瑞盈电力科技(北京)有限公司, 北京 100088)

摘要: 针对风电机组 SCADA 监测数据的非线性、高冗余等特点, 提出一种基于受限玻尔兹曼机(Restricted Boltzmann Machine, RBM)和支持向量机(Support Vector Machine, SVM)的风电机组叶片开裂故障预测方法。利用 RBM 优异的特征学习能力, 将其作为特征提取器来获得风电机组 SCADA 数据中表达能力更强的数据特征。将 RBM 的输出作为 SVM 的输入, 构建 RBM+SVM 组合预测模型。利用训练集、验证集进行预测模型构建和参数微调。为验证提出模型的有效性, 将其预测结果与 RBM+Logistic 回归、SVM 和 Logistic 回归的预测结果进行对比。实验表明, RBM+SVM 的预测准确率为 93.08%, 与三组对比模型相比具有明显的优势。研究结果可为实际风电机组叶片开裂故障预测提供重要参考。

关键词: 风电机组; 叶片开裂故障; SCADA 数据; 受限玻尔兹曼机; 支持向量机

Wind turbine blade cracking fault prediction based on RBM and SVM

ZHANG Xin¹, XU Zunyi¹, HE Huiru¹, WANG Fei²

(1. School of Computer Science and Technology, Shandong Jianzhu University, Jinan 250100, China;

2. State Grid Rayiee Electric Power Technology (Beijing) Co., Ltd., Beijing 100088, China)

Abstract: For the nonlinear, high redundancy and other characteristics of wind turbine SCADA monitoring data, this paper puts forward a wind turbine blade cracking fault prediction method based on the Restricted Boltzmann Machine (RBM) and the Support Vector Machine (SVM). The RBM's excellent feature learning ability is used as feature extractor to obtain the more expressive data features in the SCADA system data of a wind turbine. RBM's output is used as the input to the SVM to construct the combined prediction model of RBM+SVM. The prediction model is constructed and parameters are fine-tuned by using a training set and a validation set. To verify the effectiveness of the proposed model, the prediction results are compared with those of RBM+Logistic regression, SVM and Logistic regression. The experiments show that the prediction accuracy of RBM+SVM is 93.08%, which has obvious advantages over the three other compared models. The results can provide an important reference for the prediction of wind turbine blade cracking.

This work is supported by Shandong Province Key Research and Development Plan (No. 2016GGX101024) and China Huadian Corporation LTD. 2019 Annual Science and Technology Project (No. CHDKJ18-02-52).

Key words: wind turbine; blade cracking fault; SCADA data; restricted boltzmann machine; support vector machine

0 引言

风电机组工作中连续受到空气动力、惯性力等交变载荷的冲击, 使得叶片产生不规则摆动和扭曲变形, 导致叶片角度不对称、叶片开裂等故障的发生频率较高^[1]; 另一方面, 风电机组大多集中在寒冷区域或沿海潮湿地区, 并且叶片安装位置较高,

导致叶片结冰故障频发^[2]。据统计, 风电场运行 8 年中, 停机超过 7 天及以上时间的事故中约有 30% 是由叶片故障导致的^[3]。叶片故障同时会危害发电机组轴承、塔筒等部件, 叶片开裂甚至会导致倒塔, 严重影响风电机组乃至整个风电场的正常运行。如何降低因叶片故障带来的高昂维护费用、提高风电机组的运维效率成为风电行业亟需解决的关键问题之一。近年来国内外学者利用风电机组 SCADA 数据对叶片故障预测进行了相关研究。

文献[4]利用 Bagging、神经网络、PART、

基金项目: 山东省重点研发计划项目资助(2016GGX101024); 中国华电集团有限公司 2019 年度科技项目资助(CHDKJ18-02-52)

KNN(k-Nearest Neighbor) 和遗传规划 (Genetic Programming, GP)等数据挖掘模型对叶片角度不对称的故障进行了预测, 结果表明 GP 的预测效果较好并且易于理解。文献[5]提出基于三轴加速度传感器网络模型的叶片故障预测方法, 利用叶片振动监测数据, 对振幅超限的低频或超低频信号进行时域和频域分析, 成功对风电机组叶片故障实现了预测。文献[6]提出一种基于 Relief 特征选择和 XGBoost 的风电机组叶片结冰预测方法。文献[7]通过分析 SCADA 数据中的风电机组工况参数、环境参数等对叶片结冰的影响, 提出了基于栈式自编码(Stacked Auto Encoder, SAE)的叶片结冰预测方法, 实现了对风电机组叶片结冰故障的早期预警。

现有研究多集中在对风电机组叶片结冰、叶片角度不对称等故障进行预测, 对叶片开裂故障进行预测的相对较少。针对风电机组 SCADA 监测数据的高冗余、非线性、数据量大等特点, 近年来主要进行数据预处理和特征提取等方面的研究^[8-11]。与传统的主成分分析(Principle Component Analysis, PCA)、Pearson 相关性分析等特征提取方法相比, RBM 可以自适应提取高维数据的特征, 能克服多层网络结构梯度消失和容易陷入局部极值等缺陷^[12]。

本文通过分析风电机组 SCADA 数据中的工况参数、环境参数等属性与叶片开裂之间的关系, 提出基于 RBM 和 SVM 的风电机组叶片开裂故障预测模型。利用 RBM 对风电机组 SCADA 数据进行自适应特征学习; 将其输出作为 SVM 的输入, 构建 RBM+SVM 的故障预测模型; 利用无标签数据对 RBM+SVM 模型进行训练, 再利用有标签数据对模型进行参数微调; 基于训练好的模型对风电机组叶片开裂故障进行早期预警。

1 受限玻尔兹曼机和支持向量机简介

1.1 受限玻尔兹曼机

受限玻尔兹曼机是玻尔兹曼机的一种变体。近年来, 其在电力系统短期负荷预测^[13-14]、Web 服务质量预测^[15]等领域有较多的应用。RBM 是一种无监督、由一层可视层(visible layer)和一层隐含层(hidden layer)组成、对称连接且无自反馈的随机神经网络模型, 其网络结构如图 1 所示。

由图 1 可以看出, RBM 是一种层内无连接、层间全连接的无向图。其中, $v = (v_1, v_2, v_3, \dots, v_m)$ 和 $h = (h_1, h_2, h_3, \dots, h_n)$ 表示可视层和隐含层内的结点值, m 表示可视层结点数, n 表示隐含层结点数, ω 表示连接可视层和隐含层之间的权值, $b = (b_1, b_2, b_3, \dots, b_m)$ 表示可视层的偏置, $c = (c_1, c_2, c_3, \dots, c_n)$

表示隐含层的偏置。

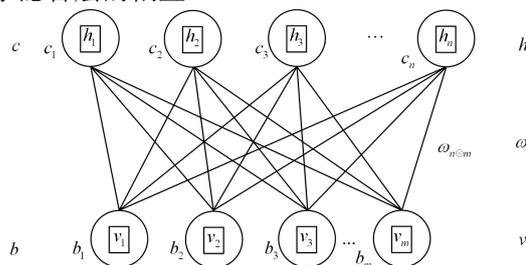


图 1 受限玻尔兹曼机网络结构图

Fig. 1 Restricted Boltzmann machine network structure

RBM 是一种基于能量的模型, 能量最小时网络达到理想状态, 网络的训练过程就是最小化能量函数的过程。RBM 能量函数的定义为

$$E(v, h | \theta) = -\sum_{i=1}^m \sum_{j=1}^n v_i \omega_{ij} h_j - \sum_{i=1}^m b_i v_i - \sum_{j=1}^n c_j h_j \quad (1)$$

其中: $\theta = (b_i, \omega_{ij}, c_j)$; ω_{ij} 为可视层结点 v_i 与隐含层结点 h_j 之间的连接权重; b_i 为可视层结点的偏置, c_j 为隐含层结点的偏置。v 和 h 的联合概率分布为

$$P(v, h | \theta) = \frac{e^{-E(v, h | \theta)}}{Z(\theta)} \quad (2)$$

其中, $Z(\theta)$ 为分配函数, 如式(3)所示。

$$Z(\theta) = \sum_{v, h} e^{-E(v, h | \theta)} \quad (3)$$

根据 RBM 层内无连接的结构特点, 第 i 个可视层结点的激活概率为

$$P(v_i = 1 | h, \theta) = \sigma(b_i + \sum_{j=1}^n \omega_{ij} h_j) \quad (4)$$

其中, σ 为 Sigmoid 函数。第 j 个隐含层结点的激活概率为

$$P(h_j = 1 | v, \theta) = \sigma(c_j + \sum_{i=1}^m \omega_{ij} v_i) \quad (5)$$

最终通过最大化 v 和 h 的联合概率分布函数 $P(v, h | \theta)$ 来求取参数 θ 。但当样本数量很大时要对 v 和 h 的所有可能取值都进行遍历, 计算复杂度较高。较好的解决方法是通过 Gibbs 采样来拟合 RBM 网络的输入, 其优点在于当使用训练样本进行 RBM 初始化的时候, 仅需要较少的抽样次数就可以得到训练数据足够好的近似结果。

1.2 支持向量机

支持向量机是 Vapnik 针对数据分类和回归问题提出的一种基于统计的机器学习方法。其在解决小样本、非线性以及高维模式识别等问题中表现出了许多特有的优势, 近年来在风电机组故障诊断^[16-18]、电力系统稳态评估^[19-20]、短期负荷预测^[21-22]、发电功率预测^[23-24]等领域有较多的应用。

对于二分类问题，假定 n 个样本的训练集 $D = \{(x_i, y_i) | i = 1, 2, \dots, n\}$ ， $x_i \in R^m$ 为输入向量， $y_i \in \{+1, -1\}$ 为类别标号， m 为输入向量的维数， n 为样本数目。当数据线性可分时，输入数据集可以被某个超平面正确地分开，并且离超平面最近的向量与超平面之间的距离是最大的，该超平面就称为最优超平面，记为

$$H: w \cdot x + b = 0 \quad (6)$$

式(6)中： $w = (w_1, w_2, w_3, \dots, w_n)$ 为输入向量的权值； b 为线性判别函数的偏移值。若样本能够按照各自类别被正确地区分开，即对于 $y_i = +1$ ，有 $w \cdot x_i + b \geq +1$ ；对于 $y_i = -1$ ，有 $w \cdot x_i + b \leq -1$ 。

在样本空间中，点到超平面 H 距离定义为

$$\frac{|w \cdot x_i + b|}{\sqrt{w_1^2 + w_2^2 + w_3^2 + \dots + w_n^2}} \geq \frac{1}{\|w\|} \quad (7)$$

当点到超平面之间的距离最大时，说明不同类别数据集之间的间隔最远，即模型的分类效果最好。

对于线性不可分问题，可以通过选取合适的核函数将样本数据 x 映射到某个高维空间 p ，然后在高维空间 p 中使用线性分类器。SVM 常用的核函数有线性内积函数、多项式内积函数、双层神经网络内积函数和径向基内积函数等。

2 风电机组叶片开裂故障预测模型构建

风电机组叶片开裂故障预测整体流程如图2所示。

由图2可以看出，风电机组叶片开裂故障预测的实现过程如下：

- 1) 对原始风电机组 SCADA 数据进行 0-1 归一化处理，然后将处理之后的数据按照 6:2:2 的比例随机分成训练集、验证集和测试集；
- 2) 构建 RBM+SVM 的故障预测模型，用训练数据进行模型训练；
- 3) 利用验证数据对构建的故障预测模型的学习率、隐含层结点数、 C 、 γ 等参数进行微调；
- 4) 利用构建好的 RBM+SVM 模型进行故障预测，并与其他模型的预测结果进行对比。

3 算例分析

3.1 数据集介绍

本文所用数据为内蒙古某风电场单台 1.5 MW 风机一年的共约 5 万条 SCADA 监测数据。为了尽可能地描述叶片的运行状态，根据行业知识和人工经验初步筛选出原始 SCADA 数据中的 26 个属性变量作为研究属性^[25]，包括叶片工况参数、环境参数和相关组件工况参数等。数据属性描述如表 1 所示。

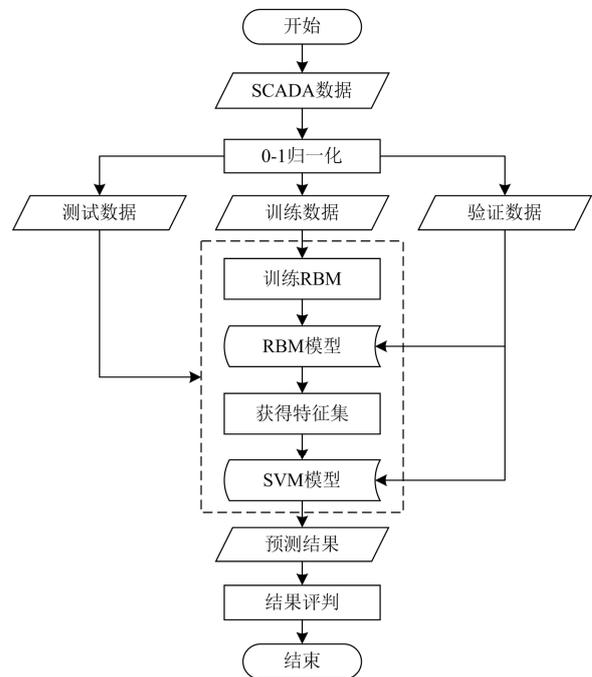


图2 风电机组叶片开裂故障预测流程

Fig. 2 Wind turbine blade cracking fault prediction process

表1 风电机组 SCADA 数据属性描述

Table 1 Wind turbine SCADA data attribute description

属性名	含义
wheel_speed	轮毂转速/(r/min)
pitch1_angle	叶片1角度/(°)
pitch2_angle	叶片2角度/(°)
pitch3_angle	叶片3角度/(°)
pitch1_moto_DC	叶片1变桨机电流/A
pitch2_moto_DC	叶片2变桨机电流/A
pitch3_moto_DC	叶片3变桨机电流/A
vibration_x	x方向振动值
vibration_y	y方向振动值
wind_speed	机舱气象站风速/(m/s)
bearing_tmp1	主轴承温度1/°C
bearing_tmp2	主轴承温度2/°C
wheel_tmp	轮毂温度/°C
int_tmp	机舱温度/°C
pitch1_bat_tmp	叶片1电池箱温度/°C
pitch2_bat_tmp	叶片2电池箱温度/°C
pitch3_bat_tmp	叶片3电池箱温度/°C
pitch1_moto_tmp	叶片1变桨机电温度/°C
pitch2_moto_tmp	叶片2变桨机电温度/°C
pitch3_moto_tmp	叶片3变桨机电温度/°C
pitch1_tran_tmp	叶片1变频器箱温度/°C
pitch2_tran_tmp	叶片2变频器箱温度/°C
pitch3_tran_tmp	叶片3变频器箱温度/°C
pitch1_cap_vol	叶片1超级电容电压/V
pitch2_cap_vol	叶片2超级电容电压/V
pitch3_cap_vol	叶片3超级电容电压/V

风电机组 SCADA 数据中正常与开裂故障数据数量如表 2 所示。

表 2 风电机组正常与开裂故障数据数量

Table 2 Number of normal and cracking fault data of wind turbine

叶片状态	样本例数	标签
正常	24 686	0
开裂	23 653	1

3.2 数据 0-1 归一化处理

由表 1 可知, 风电机组 SCADA 数据的属性量纲大多不同。为了统一数据量纲, 并提高预测模型的收敛速度和预测精度, 本文采用 0-1 归一化方法对监测数据进行标准化处理。

0-1 归一化原理为

$$d^* = \frac{d - d_{\min}}{d_{\max} - d_{\min}} \quad (8)$$

其中: d_{\max} 为样本数据 d 的最大值; d_{\min} 为样本数据 d 的最小值。

3.3 RBM 模型构建

RBM 模型的构建主要是迭代次数(Epochs)、学习率(Learning Rate, LR)和隐含层结点数(Number of Hidden Nodes, N_H)确定的过程。

1) 迭代次数(Epochs)的确定

随机设定多个学习率 LR 和 RBM 隐含层结点数 N_H 的组合, 通过大量实验分析确定 RBM 训练的最佳迭代次数, 其中代表性的实验结果如图 3 所示。图中横轴代表迭代次数, 纵轴代表损失值, 实线代表验证集的损失值随迭代次数的变化情况, 虚线代表训练集的损失值随迭代次数的变化情况。实验中的损失值均采用均方误差(Mean Square Error, MSE)来表示。MSE 的计算方法为

$$MSE = \frac{\sum_{i=1}^n (h_i - h'_i)^2}{n} \quad (9)$$

其中: n 为元组数; h_i 为实际值; h'_i 为预测值。

由图 3 可以看出, 随着迭代次数的不断增加, 训练损失不断降低, 这意味着训练精度不断提升, 即 RBM 模型的能量函数值不断减小、模型越来越稳定; 但验证损失在大约第 19 次迭代时达到最小值, 之后不断增加。为了有效防止 RBM 训练模型的过拟合, 选择在第 19 次迭代之后停止模型训练。其他 LR 和 N_H 组合的训练结果基本与上述变化趋势相符合, 因此 RBM 模型的最佳迭代次数设定为 Epochs=19。

2) 学习率(LR)和隐含层结点数(N_H)的确定

在 Epochs=19 的前提下, 通过改变 LR 和 N_H 的取值, 根据训练集和验证集的损失值来确定学习率和隐含层结点的最佳组合。

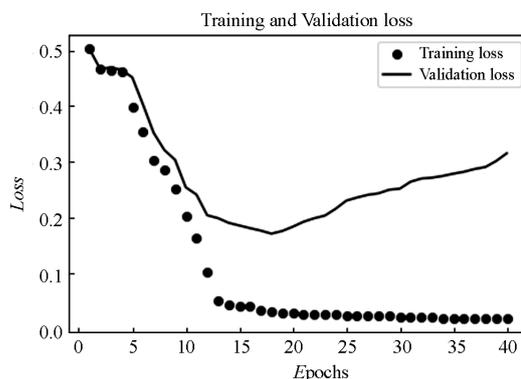


图 3 不同迭代次数下训练集和验证集的损失变化
Fig. 3 Loss variation of training sets and validation sets under different epochs

(1) 固定 RBM 隐含层结点数 N_H 为 10, 改变学习率 LR 。

由表 3 可知, 在迭代次数为 19、隐含层结点数 10 的情况下, 当学习率为 0.001 和 0.000 1 时, 训练集和验证集的损失较小。并且两种学习率情况下的 MSE 相差不大, 但是与 0.001 相比, $LR=0.0001$ 时的算法运行时间复杂度较高。综上, 在 Epochs=19、 $N_H=10$ 的前提下, $LR=0.001$ 时模型的训练效果较好。

表 3 $N_H=10$, 不同 LR 条件下训练集和验证集 MSE 变化

Table 3 MSE variation of training set and validation set under $N_H=10$ and different LR

LR	训练集 MSE	验证集 MSE
0.1	0.266 2	0.236 3
0.07	0.198 9	0.198 3
0.01	0.167 0	0.142 0
0.007	0.082 5	0.087 5
0.004	0.093 9	0.090 3
0.001	0.021 9	0.019 8
0.0001	0.020 6	0.018 7

(2) 固定学习率 LR 为 0.001, 改变 RBM 隐含层结点数 N_H 。

由表 4 可知, 在迭代次数为 19、学习率为 0.001 的情况下, 当 RBM 隐含层结点数 10 时, 训练集和验证集的损失最小。

综上, RBM 模型最终参数选择为: Epochs=19、 $LR=0.001$ 、 $N_H=10$ 。

3.4 RBM+SVM 模型构建

基于 3.3 节的 RBM 特征学习模型, 将其输出作为 SVM 的输入来构造 RBM+SVM 的叶片开裂故障预测模型。实验中所用 SVM 模型基于 Python 的 scikit-learn 库进行实现。SVM 模型主要有 C 、kernel、

表 4 $LR=0.001$, 不同 N_H 条件下训练集和验证集 MSE 变化
Table 4 MSE variation of training set and validation set under $LR=0.001$ and different N_H

N_H	训练集 MSE	验证集 MSE
40	0.093 9	0.090 3
30	0.075 4	0.083 5
20	0.063 2	0.059 5
10	0.021 9	0.019 8
5	0.160 1	0.027 6

γ 等参数。其中, C 为错误项的惩罚参数或称为松弛系数, $kernel$ 为核函数, γ 为核函数的核系数。 C 越大, 说明模型对于误差的容错性越小, 模型容易过拟合; 反之, C 越小, 模型容易欠拟合; C 过大或过小都会导致模型的泛化能力变差。 γ 越大, 支持向量越少; γ 越小, 支持向量越多; 支持向量的个数会影响模型训练与预测的速度。实验中选择径向基函数(Radial Basis Function, RBF)作为核函数, 通过改变 C 和 γ 的取值, 根据预测精确度来确定松弛系数和核系数的最佳组合。

实验中的精确度(Precision)表示方法为

$$Precision = \frac{TP + TN}{P + N} \times 100\% \quad (10)$$

其中: TP 为真正例(被分类器正确分类的正元组); TN 为真负例(被分类器正确分类的负元组); P 为测试集中正元组的数量; N 为测试集中负元组的数量; 感兴趣的主要类的元组为正元组, 其它为负元组。

在 RBM 的迭代次数为 19、学习率为 0.001、隐含层结点数为 10, SVM 的核函数为 RBF 的情况下, 不同松弛系数和核系数条件下的模型预测精确度变化情况如表 5、表 6 所示。

表 5 $C=1$, 不同 γ 值条件下, 精度变化

Table 5 Precision variation under $C=1$ and different γ

γ	精度/%
1	82.34
10	90.25
20	91.75
40	91.72
60	91.74
80	90.04

由表 5 可知, 在 $C=1$ 的条件下, 当 γ 为 20、40、60 时预测精确度较高, 均为 91.74% 左右。基于前文对 γ 取值含义的介绍, 在预测精确度相同的情况下, γ 越大、支持向量越少, 相应的模型运算速度越快。故在 $C=1$ 的条件下, γ 值设定为 60。

表 6 $\gamma=60$, 不同 C 值条件下, 精度变化

Table 6 Precision variation under $\gamma=60$ and different C

C	精度/%
0.1	82.09
1	91.54
2	92.38
3	93.52
5	93.54
10	93.56
20	93.57

由表 6 可知, 在 $\gamma=60$ 的条件下, 当 C 为 3、5、10、20 时预测精确度较高, 均为 93.56% 左右。基于前文对 C 取值含义的介绍, 在预测精确度相同的情况下, C 值不宜过大或过小, 过大容易过拟合、过小容易欠拟合。故在 $\gamma=60$ 的条件下, C 值设定为 10。

综上, RBM+SVM 组合模型的参数设定为:

Epochs=19、 $LR=0.001$ 、 $N_H=10$ 、 $C=10$ 、 $\gamma=60$ 。

3.5 实验结果对比

为验证本文提出模型的有效性, 分别设计了 RBM+Logistic 回归、Logistic 回归和 SVM 三组对比实验。四种模型的预测结果如表 7 所示。

表 7 四种模型预测效果比较

Table 7 Comparison of prediction effects among four models

预测模型	测试集精度/%	运行时间/s
RBM+SVM	93.08	20
RBM+Logistic 回归	74.32	10
SVM	56.25	73
Logistic 回归	52.17	40

由表 7 四种模型的预测效果可知:

1) 组合预测模型的精确度要高于非组合预测模型;

2) 与 RBM+Logistic 回归预测模型相比, RBM+SVM 的预测精确度更高;

3) 由于 SVM 模型本身的复杂度要高于 Logistic 回归, 所以 RBM+SVM 的时间复杂度要高于 RBM+Logistic 回归;

4) 虽然 RBM+SVM 的时间复杂度较高, 但仍远小于风电机组 10 min 的采样间隔。

4 结论

利用风电机组 SCADA 数据研究风电机组健康状态演化趋势并进行故障预测, 根据预测结果合理安排备件储备、调整检修维护计划, 对于保障风电机组安全稳定运行、避免突发事件导致的二次事故以及降低风电机组运维成本具有重要的工程意义。

以RBM作为特征提取器对原始SCADA数据的26维属性进行特征学习,实现数据特征的自适应提取和降维;基于RBM构建RBM+SVM的组合预测模型;同时设计三组对比实验。由实验可得,非组合模型的预测精确度要远低于组合模型;RBM+SVM的预测精确度最高,其时间复杂度也相对较高,但其时间复杂度远低于风电机组SCADA数据采样时间间隔,可满足生产实际的实时性要求。结果表明,本文提出的RBM+SVM预测模型具有一定的可行性,对风电场风电机组叶片开裂故障预测具有较好的指导意义,能够应用于工程实际。

本文利用RBM进行特征提取,避免人为提取特征的复杂性和不确定性等问题,充分保证了数据质量。但只利用了RBM一种特征学习模型进行特征提取,可以考虑深度玻尔兹曼机(Deep Boltzmann Machine, DBM)进行特征学习。相比于RBM, DBM具有更多的隐含层,其特征提取能力更强。利用DBM进行特征学习,然后结合相应的预测模型进行故障预测是未来的研究方向之一。

参考文献

- [1] HERBERT G M J, INIYAN S, SREEVALSAN E, et al. A review of wind energy technologies[J]. *Renewable & Sustainable Energy Reviews*, 2007, 11(6): 1117-1145.
- [2] KUSIAK A, LI W Y, SONG Z. Dynamic control of wind turbines[J]. *Renewable Energy*, 2010, 35(2): 456-463.
- [3] WYMORE M L, DAM J E V, CEYLAN H, et al. A survey of health monitoring systems for wind turbines[J]. *Renewable & Sustainable Energy Reviews*, 2015, 52: 976-990.
- [4] KUSIAK A, VERMA A. A data-driven approach for monitoring blade pitch faults in wind turbines[J]. *IEEE Transactions on Sustainable Energy*, 2011, 2(1): 87-96.
- [5] 张保钦, 雷保珍, 赵林惠, 等. 风机叶片故障预测的振动方法研究[J]. *电子测量与仪器学报*, 2014, 28(3): 285-291.
ZHANG Baoqin, LEI Baozhen, ZHAO Linhui, et al. Research on vibration method of fan blade fault forecasting[J]. *Journal of Electronic Measurement and Instrumentation*, 2014, 28(3): 285-291.
- [6] 曹渝昆, 朱萌, 王晓飞. 基于特征选择和 XGBoost 的风机叶片结冰预测[J]. *电气自动化*, 2019, 41(3): 31-33.
CAO Yukun, ZHU Meng, WANG Xiaofei. Wind turbine blade icing forecast based on feature selection and XGBoost[J]. *Electrical Automation*, 2019, 41(3): 31-33.
- [7] 刘娟, 黄细霞, 刘晓晓. 基于栈式自编码网络的风机叶片结冰预测[J]. *计算机应用*, 2019, 39(5): 1547-1550.
LIU Juan, HUANG Xixia, LIU Xiaoli. Icing prediction of wind turbine blade based on stacked auto-encoder network[J]. *Journal of Computer Applications*, 2019, 39(5): 1547-1550.
- [8] ZHAO H S, ZHANG X T. Early fault prediction of wind turbine gearbox based on temperature measurement[C] // 2012 IEEE International Conference on Power System Technology (POWERCON), October 30-November 2, 2012, Auckland, New Zealand: 1-5.
- [9] KUSIAK A, VERMA A. A data-mining approach to monitoring wind turbines[J]. *IEEE Transactions on Sustainable Energy*, 2012, 3(1): 150-157.
- [10] SHIN J H, LEE Y S, KIM J O. Fault prediction of wind turbine by using the SVM method[C] // 2014 International Conference on Information Science, Electronics and Electrical Engineering, April 26-28, 2014, Sapporo, Japan: 1923-1926.
- [11] CHENG F Z, QU L Y, QIAO W. Fault prognosis and remaining useful life prediction of wind turbine gearboxes using current signal analysis[J]. *IEEE Transactions on Sustainable Energy*, 2018, 9(1): 157-167.
- [12] LI F, GAO X G, WAN K F. Training restricted Boltzmann machine using gradient fixing based algorithm[J]. *Chinese Journal of Electronics*, 2018, 27(4): 694-703.
- [13] 李若晨, 朱帆, 朱永利, 等. 结合受限玻尔兹曼机的递归神经网络电力系统短期负荷预测[J]. *电力系统保护与控制*, 2018, 46(17): 83-88.
LI Ruochen, ZHU Fan, ZHU Yongli, et al. Short-term power load forecasting using recurrent neural network with restricted Boltzmann machine[J]. *Power System Protection and Control*, 2018, 46(17): 83-88.
- [14] OUYANG T, HE Y, LI H, et al. Modeling and forecasting short-term power load with copula model and deep belief network[J]. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 2019, 3(2): 127-136.
- [15] 王兴菲, 万健, 陈璐, 等. 基于受限玻尔兹曼机的 Web 服务质量预测方法[J]. *中南大学学报: 自然科学版*, 2018, 49(11): 2745-2752.
WANG Xingfei, WAN Jian, CHEN Lu, et al. Web service QoS prediction based on restricted Boltzmann machines[J]. *Journal of Central South University: Science and Technology*, 2018, 49(11): 2745-2752.
- [16] 李东东, 周文磊, 郑晓霞, 等. 基于多重分形谱和支持向量机的风电机组行星齿轮箱故障诊断与研究[J]. *电力系统保护与控制*, 2017, 45(11): 43-48.
LI Dongdong, ZHOU Wenlei, ZHENG Xiaoxia, et al. Diagnosis and research of wind turbine planetary gearbox faults based on multifractal spectrum support vector

- machine (SVM)[J]. Power System Protection and Control, 2017, 45(11): 43-48.
- [17] 李自强, 魏磊, 韩大鹏, 等. 基于相空间重构的高压断路器机械故障诊断研究[J]. 电力系统保护与控制, 2018, 46(9): 129-135.
LI Ziqiang, WEI Lei, HAN Dapeng, et al. Research on mechanical fault diagnosis of high voltage vacuum circuit breaker based on phase space reconstruction[J]. Power System Protection and Control, 2018, 46(9): 129-135.
- [18] SANTOS P, VILLA L F, et al. An SVM-based solution for fault detection in wind turbines[J]. Sensors, 2015, 15(3): 5627-5648.
- [19] HU W, LU Z, WU S, et al. Real-time transient stability assessment in power system based on improved SVM[J]. Journal of Modern Power Systems and Clean Energy, 2019, 7(1): 26-37.
- [20] YOU D, WANG K, YE L, et al. Transient stability assessment of power system using support vector machine with generator combinatorial trajectories inputs[J]. International Journal of Electrical Power & Energy Systems, 2013, 44(1): 318-325.
- [21] 熊军华, 牛珂, 张春歌, 等. 基于小波变异果蝇优化支持向量机短期负荷预测方法研究[J]. 电力系统保护与控制, 2017, 45(13): 71-77.
XIONG Junhua, NIU Ke, ZHANG Chungu, et al. LSSVM in short-term load forecasting based on wavelet transform and mutant fruit fly optimization algorithm[J]. Power System Protection and Control, 2017, 45(13): 71-77.
- [22] 李霄, 王昕, 郑益慧, 等. 基于改进最小二乘支持向量机和预测误差校正的短期风电负荷预测[J]. 电力系统保护与控制, 2015, 43(11): 63-69.
LI Xiao, WANG Xin, ZHENG Yihui, et al. Short-term wind load forecasting based on improved LSSVM and error forecasting correction[J]. Power System Protection and Control, 2015, 43(11): 63-69.
- [23] 张雨金, 杨凌帆, 葛双冶, 等. 基于 Kmeans-SVM 的短期光伏发电功率预测[J]. 电力系统保护与控制, 2018, 46(21): 118-124.
ZHANG Yujin, YANG Lingfan, GE Shuangye, et al. Short-term photovoltaic power forecasting based on Kmeans algorithm and support vector machine[J]. Power System Protection and Control, 2018, 46(21): 118-124.
- [24] YE L, LIU P. Combined model based on EMD-SVM for short-term wind power prediction[J]. Proceedings of the CSEE, 2011, 31(31): 102-108.
- [25] 黄忠山, 田凌, 向东, 等. 基于 PCA 和 SPC-动态神经网络的风电机组齿轮箱油温趋势预测[J]. 清华大学学报: 自然科学版, 2018, 58(6): 539-546.
HUANG Zhongshan, TIAN Ling, XIANG Dong, et al. Prediction of oil temperature variations in a wind turbine gearbox based on PCA and SPC-dynamic neural network hybrid[J]. Journal of Tsinghua University: Science and Technology, 2018, 58(6): 539-546.

收稿日期: 2019-09-06; 修回日期: 2019-09-23

作者简介:

张鑫(1994—), 男, 硕士, 研究方向为风电机组故障预测、数据挖掘; E-mail: 18263434037@163.com

徐遵义(1969—), 男, 通信作者, 博士, 副教授, 研究方向为风电机组故障预测、火电脱硫优化、数据挖掘; E-mail: zunyixu@sjzu.edu.cn

何慧茹(1995—), 女, 硕士, 研究方向为深度学习、数据挖掘。E-mail: 15966657208@163.com

(编辑 葛艳娜)