

基于集成学习的操作机构控制电路故障分类

张 峥, 孟海涛, 邵霆啸

盐城工学院, 江苏 盐城

收稿日期: 2022年9月11日; 录用日期: 2022年10月17日; 发布日期: 2022年10月25日

摘 要

实时监测和故障诊断是智能化高压配电系统稳定正常运行的关键, 为了实现在电力系统中准确识别故障的要求, 本文提出了一种基于集成学习的电力系统故障分类模型, 通过对特征空间中的非线性规律进行编码的多项式特征衍生, 极大地提高了特征的效用。该方法能够有效识别控制系统中各元器件是否发生故障。对提高高压断路器设备监测效率、降低安全事故率、提升配电系统智能化改造具有一定的应用价值和现实意义。

关键词

高压断路器, 操作机构, 故障分类, 集成学习

Fault Classification of Operating Mechanism Control Circuits Based on Integrated Learning

Zheng Zhang, Haitao Meng, Tingxiao Shao

Yancheng Institute of Technology, Yancheng Jiangsu

Received: Sep. 11th, 2022; accepted: Oct. 17th, 2022; published: Oct. 25th, 2022

Abstract

Real-time monitoring and fault diagnosis is the key to the stable and normal operation of the intelligent high-voltage distribution system, in order to achieve the requirements of accurate fault identification in the power system, this paper proposes a fault classification model of power system based on integrated learning, which greatly improves the utility of features by encoding the nonlinear

law in the feature space. This method can effectively identify whether each component in the control system is faulty. It has certain application value and practical significance for improving the monitoring efficiency of pressure circuit breaker equipment, reducing the safety accident rate, and improving the intelligent transformation of the power distribution system.

Keywords

High Voltage Circuit Breakers, Operating Mechanism, Fault Classification, Integrated Learning

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

电动机操作机构是高压配电系统中的重要组成部分,承担着关合、开断线路的作用[1] [2]。当电力系统故障时,电动机操作机构能够将故障部分迅速从电网中隔离出去,缩小事故范围。若未能及时发现电动机操作机构故障,将使得分合闸功能失效,将导致线路和设备损坏甚至影响居民生产[3] [4]。

近年来,许多学者通过引入模式识别[5]的方法来完成电力系统故障分类任务,何婷、乔俊强等人[6]提出了一种基于 EMD 和 SVM 的电力系统故障分类的方法,通过 EMD 将电压信号分解为 IMF,进一步通过 HHT 提取相关故障特征,最后以 SVM 建立分类模型,经对比分析,能够较为精准地识别各类故障;唐颖川、黄姣茹等人[7]提出了一种基于深度学习与注意力机制的分类模型,采用小波阈值进行去噪,并在 CNN-LSTM 模型的输出层前引入注意力机制进行特征筛选,相较于改进前有较好地提升;但是仍存在着大量问题,其一是对特征的使用不充分,其二是模型的预测效率还不够高。

针对上述情况,本文提出并实现了:

1) 针对问题 1,本文采用多项式特征衍生的办法,对特征空间中的非线性规律进行编码的多项式特征衍生,极大地提高了特征的效用。

2) 针对问题 2,本文提出了一种基于集成学习的电力系统故障分类模型,该方法将识别过程分为三层,每层扩展为一个单独的故障分类算法。第一层采用 XGBoost 算法,将其预测结果转化为第二层的输入特征,第二层采用的是 LightGBM 算法,第三层将上两层的输出线性加权融合到一起,赋予不同的权重作为超参,其中,本文采用网格搜索算法寻找权重参数最优解,与传统算法相比,该算法在准确率、精准率、召回率和 F1 值上均有提升。

2. XGBoost 算法

Xgboost 是一种高效的梯度提升决策树算法[8],它是由 m 个基模型组成的一个加法运算式:

$$\hat{y}_i = \sum_{t=1}^m f_t(x_i)$$

其中, f_m 为第 m 个基模型, \hat{y}_i 为第 i 个样布的预测值。给定一组训练数据 $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)\}$, 其中, n 为训练样本总数, x_i 为样本数据, y_i 样本数据对应的标签。对于每一个测试样本,都有目标函数为:

$$L^{(t)} = \sum_{i=1}^n \left[g_i f_t(x_i) + \frac{1}{2} h_i f_t^2(x_i) \right] + \Omega(f_t)$$

3. LightGBM 算法

LightGBM 是一种基于决策树算法的分类模型[9]，该算法没有使用大多数 GBDT 工具采用的按层生长的决策树生长决策，而采用的带有深度限制的按叶子生长算法。为了达到对类别特征的最优切分，LightGBM 采用 many-vs-many 的切分方式。假设数据集 D 的维度特征共有 n 个类别，则共有 $2^{(n-1)} - 1$ 种情况，其时间复杂度为 $O(2^n)$ ，其中， x_i 表示第 i 个样本， y_i 表示样本其对应的标签， l 表示最初目标函数， L_n 表示添加正则项后第 n 次迭代的目标函数， f^n 表示第 n 次迭代后的模型函数， T 为叶子节点数量， w_j 为对应的叶子结点输出值， r, λ 均为权重参数。

该算法的目标函数公式如下所示：

$$L_n = \sum_{i=1}^n l(y_i, y_i^{n-1} + f^n(x_i)) + rT + \frac{1}{2} \lambda \sum_{j=1}^T w_j^2$$

4. 基于集成学习的电力系统故障分类模型

4.1. 总体框架

基于集成学习的思想进行模型融合，增加了容错率，避免了单一模型进行电力故障分类的错误情况，通过线性加权融合的方式，综合考虑了每一层分类模型的预测结果，进一步纠正了错误率。如图 1 为基于 Boosting 的电力系统故障分类模型结构图。

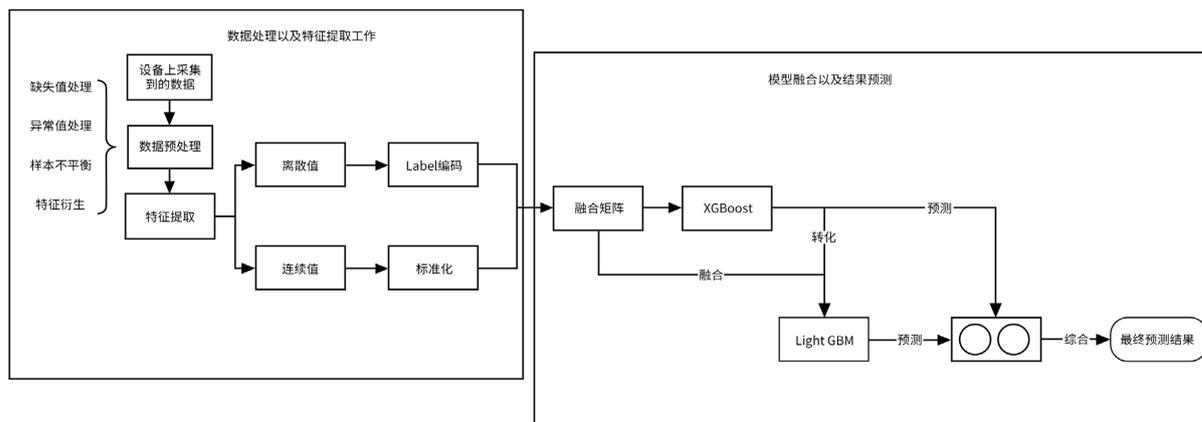


Figure 1. Structure diagram

图 1. 结构图

图 1 主要展示了一种基于集成学习的电力系统故障分类模型结构图，总体来说该框架总共分为三部分：

1) 数据预处理：对实验设备采集的工业化数据进行数据处理如缺失值处理、异常值处理、样本不平衡问题以及特征衍生等相关处理。

2) 特征工程：主要是对处理后的数据进行特征提取，为了提升模型的收敛速度以及模型精度，本文在这里为了选择对离散值如 M、N1、N2_0-5、N2_5-10 等属性进行 Label 编码，对连续型数值如 1-4_0-5、1-4_5-10、1-1_0_5 等属性进行标准化。

3) 模型融合：采用集成学习的思想，将多种分类器如 XGBoost、LightGBM 通过线性加权融合的方式集成在一起构造强分类器，以此达到对故障分类的效果。

4.2. 数据预处理及特征工程

本文采用的数据集为图 2 中各信号采集点所采集的电压、电流、电量等信息数据，采集时间间隔为

5 s。电路原理图中分闸按钮 SB1、合闸按钮 SB2、和停止按钮 SB3 保持常开；行程开关 SP1、SP2、SP3 保持常闭。S1、S2 分别为分闸控制回路信号输入点；M1 至 M10 为信号采集点。

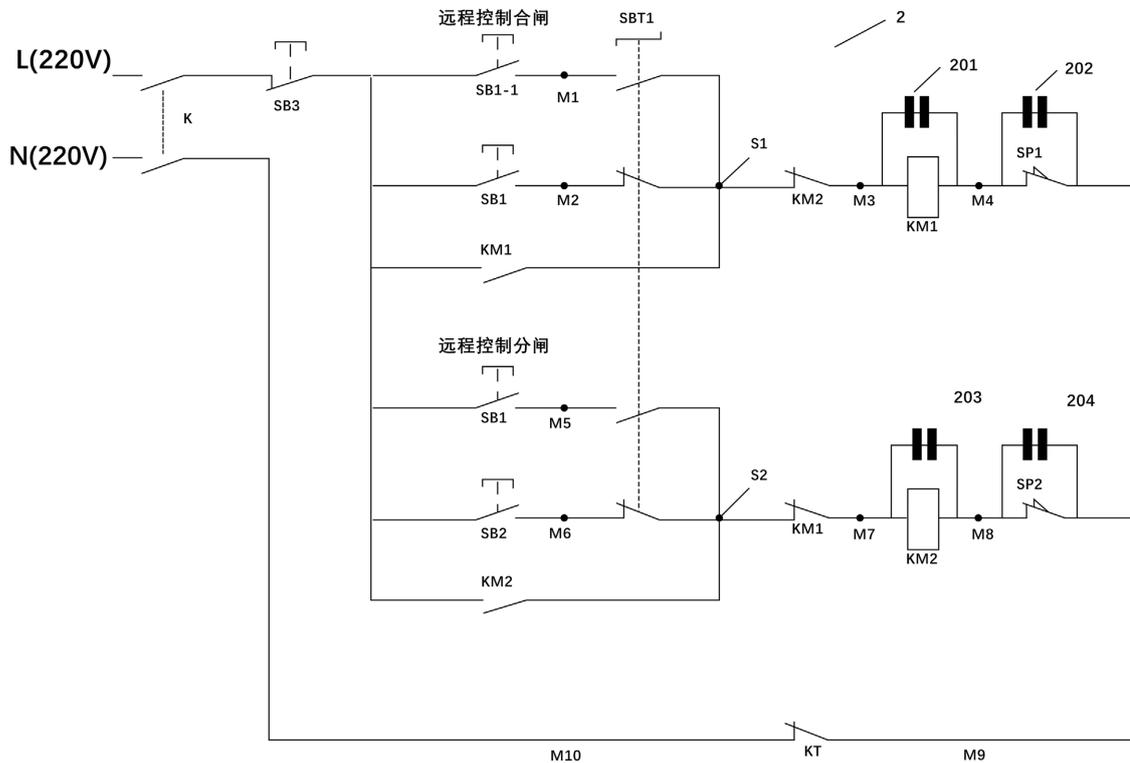


Figure 2. Circuit schematic diagram
图 2. 电路原理图

首先，本文针对数据异常、数据缺失等情况进行数据预处理，由于数据的不间断性，本文对其进行剔除以及均值修正，避免噪声的影响。

针对于特征工程部分，本文通过分析非线性规律进行分析，主要采用特征组合的方式对类别特征进行特征衍生，通过对类别特征进行独热编码后组合独热特征矢量，将独热特征矢量的特征组合视为逻辑链接，生成具备多元特征的矢量，以此来挖掘表达特征和因变量之间的非线性关系，并提高变量的效用。

4.3. 模型融合以及算法评价指标

本文在模型融合方案上选择基于加权融合的思想，将多个分类器的预测进行相结合以此来降低方差，进而提高模型的鲁棒性，而加权融合考虑了预测类别的可能性，计算每个分类器预测的每个类别的平均概率，对每个分类器的预测概率进行权重分配，进行计算得到最终预测结果。其具体计算过程如下：

$$\text{VottingClass}_{\text{pre}} = \alpha * \text{LightGBM}_{\text{pre}} + \beta * \text{XGBoost}_{\text{pre}}$$

其中， $\text{VottingClass}_{\text{pre}}$ 表示最终投票结果的概率， $\text{LightGBM}_{\text{pre}}$ 表示分类器 LightGBM 的预测概率， $\text{XGBoost}_{\text{pre}}$ 表示 XGBoost 的预测概率， α 、 β 分别为对应模型的权重参数。

通过引入一组广泛使用的评价指标来对模型预测结果进行估计：准确率、精确率、召回率、F1 值。

- ① 准确率(Accuracy): 分类正确的所有结果占总预测与真实总和的比重。
- ② 精确率(Precision): 模型预测到结论的特征与预测总之的比重。

③ 召回率(Recall): 真实值中预测正确的特征占预测总数的比例。

④ F1 值(F-score): 在尽可能地提高精确率(Precision)和召回率(Recall)的同时, 也希望两者之间的差异尽可能小。

上述评价指标的计算公式如下所示:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$\text{F-score} = \frac{2 * \text{Precision} * \text{Recall}}{\text{Precision} + \text{Recall}}$$

4.4. 实验结果分析

本文将上述模型训练过程的损失值(Loss)与准确率(Accuracy)进行可视化, 如图 3、图 4 所示。

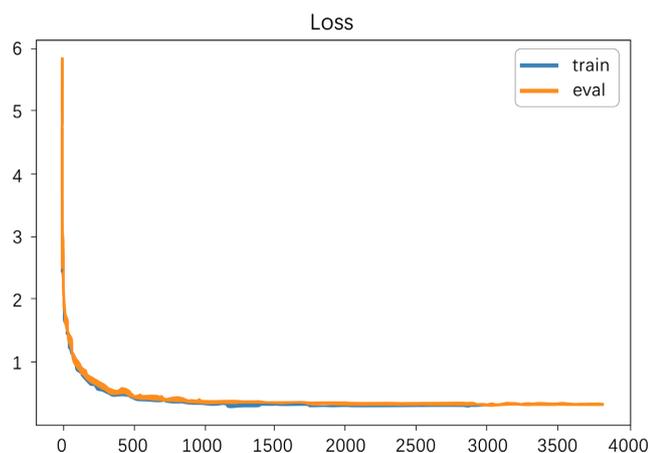


Figure 3. Visualization of loss values

图 3. 损失值可视化图

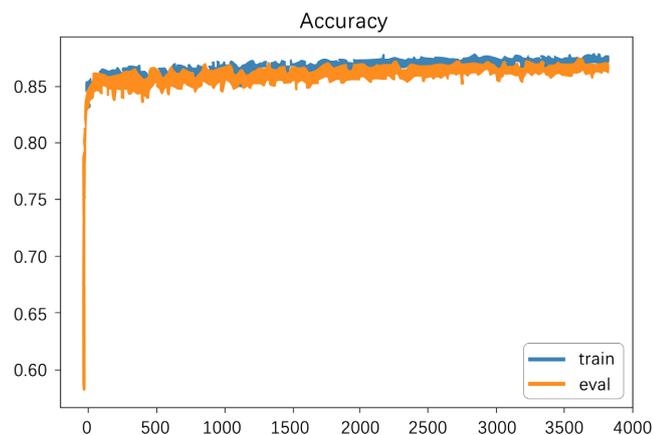


Figure 4. Visualization of accuracy

图 4. 准确率可视化图

基于集成学习的电力系统故障分类模型最终的实验对比如下表 1 所示, 单一的 LightGBM 模型与 XGBoost 模型在准确率、召回率、F1 值等评价指标上均低于本文模型。

Table 1. Model final experimental comparison

表 1. 模型最终实验对比

算法	Accuracy	Precision	Recall	F-score
LightGBM	0.841	0.864	0.826	0.845
XGBoost	0.711	0.836	0.667	0.743
本文模型	0.890	0.904	0.881	0.892

5. 结论与展望

在本文中, 我们设计了一种基于集成学习的电力系统故障分类模型, 该模型首先采用多项式特征衍生的方式对特征进行特征组合, 进而提高特征的使用效率和价值, 将处理后的特征向量融合矩阵输入到融合模型进行分类训练, 最终的预测结果通过加权融合的思想进行输出类别的最高概率, 并通过一组常用的算法评价指标对其进行估计, 经实验对比分析, 本文模型较于单一模型有较好地提升。

参考文献

- [1] 周博文, 陈麒宇, 杨东升. 巴西大停电的思考[J]. 发电与空调, 2018, 39(2): 97-105.
- [2] 黄雅罗, 黄树红. 发电设备状态检修[M]. 北京: 中国电力出版社, 2000.
- [3] Carvalho, A., Cormenzana, M.L., Furuta, H., Grieshaber, W., Hyrczak, A., Kopejtkova, D., Krone, J.G., *et al.* (2012) CIGRÉ Technical Brochure No. 512: Final Report of the 2004-2007 International Enquiry on Reliability of High Voltage Equipment, Part 4—Instrument Transformers. CIGRÉ, Paris.
- [4] 赵莉华, 金浩文, 黄小龙, 等. 基于电机电流的负荷开关操作机构状态诊断研究[J]. 高压电器, 2020, 56(6): 302-308.
- [5] 徐晶冉, 徐雯, 王栋. 基于模式识别技术的变压器故障诊断[J]. 中州煤炭, 2022, 44(2): 252-256.
- [6] 何婷, 乔俊强, 包建勤, 张亚东. 基于 EMD 和 SVM 的电力系统故障分类识别[J]. 仪表技术, 2022(4): 64-69. <https://doi.org/10.19432/j.cnki.issn1006-2394.2022.04.001>
- [7] 唐颖川, 黄姣茹, 钱富才. 基于深度学习与注意力机制的化工故障分类[J]. 电子测量技术, 2022, 45(4): 168-174. <https://doi.org/10.19651/j.cnki.emt.2108237>
- [8] 郝俊虎, 胡毅. 基于 XGBoost 和自回归模型的轴承故障诊断和预警方法研究[J]. 组合机床与自动化加工技术, 2020(2): 140-142+157.
- [9] 赵书涛, 王亚潇, 孙会伟, 等. 基于自适应权重证据理论的断路器故障诊断方法研究[J]. 中国电机工程学报, 2017, 37(23): 7040-7046.