

# 机器学习在变压器故障诊断中的应用

马洪斌, 杨 飞, 杨蕾娜, 张 冉

(国网山东省电力公司枣庄供电公司, 山东 枣庄 277000)

摘要: 当前变压器故障诊断主要采用油中气体分析法(IEC三比值法), 在实际应用中诊断的准确率偏低, 只有80%左右。通过将机器学习与变压器故障诊断相结合, 搭建了故障诊断模型, 同时采用集成学习, 综合多个基分类器, 将故障识别的准确率提高到了92.5%。

关键词: 机器学习; 支持向量机; 变压器; 故障诊断

中图分类号: TM411 文献标志码: B 文章编号: 1001-9898(2021)04-0035-04

## Application of Machine Learning in Transformer Fault Diagnosis

MA Hongbin, YANG Fei, YANG Leina, ZHANG Ran

(State Grid Shandong Electric Power Company Zaozhuang Power Supply Company, Zaozhuang 277000, China)

Abstract: At present, gas in oil analysis method (IEC three ratio method) is mainly used for transformer fault diagnosis. In practical application, the accuracy of diagnosis is low, only about 80%. By combining machine learning with transformer fault diagnosis, a fault diagnosis model is built. At the same time, integrated learning and multiple base classifiers are used to improve the accuracy of fault recognition to 92.5%.

Key words: machine learning; support vector machine; transformer; fault diagnosis

变压器作为电力系统的重要组成部分, 其运行状态决定了整个系统的安全稳定。一旦变压器发生故障, 将带来巨大的经济损失。因此, 变压器故障的有效诊断对保证变压器的安全可靠运行具有非常重要的意义。随着机器学习的兴起, 越来越多的研究人员将机器学习应用到变压器故障诊断中, 取得较好的效果。文献[1]提出了基于加权 $K$ 近邻算法的变压器故障诊断, 根据待分类样本在特征空间中 $K$ 个最近邻样本中的多数样本的类别进行分类, 因此具有直观、无需先验统计知识、无师学习等特点。文献[2]提出了基于油中溶解气体的支持向量机变压器故障诊断, 有效解决了传统学习方法的“维数灾难”和“过学习”等问题。文献[3]提出了基于知识粗糙度的多变量决策树在变压器故障诊断系统中的应用, 有效地简化了决策树, 减少诊断信息的冗余性, 诊断效率高, 结果易于理解。

### 1 机器学习方法

机器学习通过对外部环境的信息和知识进行加工改造, 使其成为有效的信息内容, 学习是机器学习的核心, 包括对信息的采集、接受监督和指导, 同时还包括对学习的推理以及修改学习系统的知识库。

机器学习具有目的性、结构性、有效性和开放性。机器学习可以了解学习的内容, 学习的行为具有一定的目的性。结构性中, 可以系统地将知识的结构和组织形式进行修改和完善。针对有效性, 机器学习系统学习的知识能够适应实践的需要, 不断改善机器学习系统。开放性表现在机器学习能够与外部环境进行信息的交互和自身的进化。

#### 1.1 支持向量机

支持向量机(Support Vector Machine, SVM)是在统计学习理论的基础上发展出的一种模式识别方法, 是一种全局最优求解算法而不是求得局

收稿日期: 2020-06-08

作者简介: 马洪斌(1991—), 男, 工程师, 主要从事变压器检修维护、故障诊断及负荷预测工作。

部极小值,该方法“以结构风险最小化原则”,根据实际需求对子集中的判别函数以及函数子集进行恰当选择。既可以保障学习的效果不会存在异常的状况,又可以保障样本的误差,且其独立测试集的测试误差仍然较小,具有很好的泛化能力。SVM 具有坚定的理论基础,在人脸识别、文本自动分类等众多领域获得广泛的应用<sup>[4]</sup>。

## 1.2 vote 集成学习

集成学习方法主要是由个体的学习器进行组合生成,通过对样本数据集的学习可以获得多个分类器,这些分类器具有多样性和准确性。根据个体的学习器之间是否是同一类型的学习算法,又可以将分类器分为“同质”和“异质”。对于个体的学习器的合并,是按照某种策略将个体学习器对新样本分类的预测结果综合考虑,从而得到一个最终的预测结果,见图 1。集成策略主要有“平均值法”和“多数投票法”等<sup>[5]</sup>。

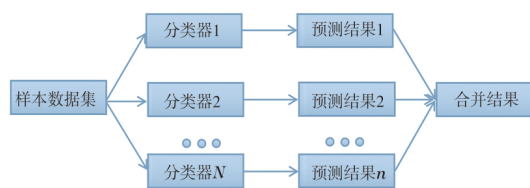


图 1 集成学习

## 2 变压器故障特性

### 2.1 变压器油及固体绝缘产生气体机理

对于大型电力变压器,目前几乎都是用油来绝缘和散热,电力变压器油与油中的固体有机绝缘材料在运行电压下因电、热、氧化和局部电弧等多种因素作用会逐渐老化、裂解,产生少量的甲烷、乙烯、乙烷、乙炔等低分子烃类,以及一氧化碳、二氧化碳和氢气等气体,并多数溶解在油中。油中溶解气体的组分和含量在一定程度上反映出电力变压器绝缘老化或故障的程度,可以作为反映电力设备异常的特征量。通过对运行中的电力变压器定期分析溶解于油中的气体组分、含量和产气速率,能够及早发现电力变压器内部存在的潜伏性故障<sup>[6]</sup>。

### 2.2 变压器典型故障分析

变压器内部故障主要为过热式和异常放电式,其中过热式主要包括低温过热( $<300\text{ }^{\circ}\text{C}$ ),中温过热( $300\sim 700\text{ }^{\circ}\text{C}$ )和高温过热( $>700\text{ }^{\circ}\text{C}$ ),异

常放电式主要包括低能量放电和高能量放电。

根据试验和变压器故障处理案例分析,油浸式变压器发生故障时产生的气体成分见表 1。

表 1 油浸式变压器不同故障类型时产生的气体组分

故障类型	主要气体组分	次要气体组分
油过热	甲烷、乙烯	氢气、乙烷
油和纸过热	甲烷、乙烯、一氧化碳 二氧化碳	氢气、乙烷
油和纸绝缘 中局部放电	氢气、甲烷 一氧化碳	乙炔、乙烷 二氧化碳
油中火花放电	氢气、乙炔	
油中电弧	氢气、乙炔	甲烷、乙烯、乙烷
油和纸中电弧	氢气、乙炔、一氧化碳 二氧化碳	甲烷、乙烯、乙烷

通过分析,将变压器的故障类型分为中低温过热、高温过热、低能量放电和高能量放电。本文将变压器的 4 种故障类型与正常情况,共 5 种状态类型作为识别结果进行模式识别。

## 3 基于机器学习的变压器故障诊断

将机器学习与变压器故障诊断相结合,提出了故障诊断流程,见图 2。

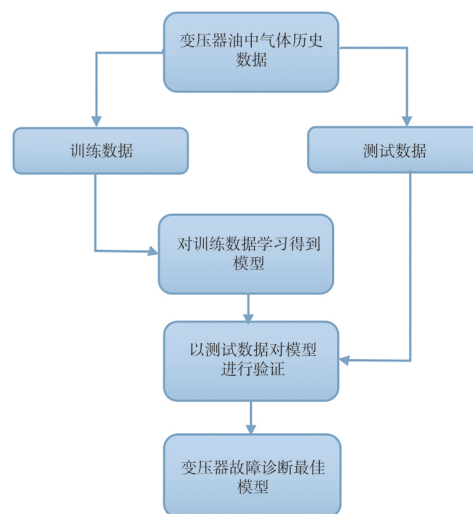


图 2 基于机器学习的变压器故障诊断流程

对采集得到的变压器油中气体数据 115 组,分成训练数据(75 组)和测试数据(40 组),通过对训练数据进行学习得到故障诊断模型,以测试数据对得到的模型进行验证,判断模型的准确率。同时采用 vote 集成学习进一步对各个分类器进行组合,提升故障识别的准确率。

将甲烷、乙烯、乙烷、乙炔、氢气作为特征向量进行变压器故障模式识别,考虑到分类问题主要是针对数值型结构,气体组分为数值型不需转换,因此将变

压器的状态正常、中低温过热、高温过热、低能量放电和高能量放电由字符型转换为数值型,分别标注为0、1、2、3、4,训练数据和测试数据见表2—3。

表2 训练数据

氢气/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	甲烷/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	乙烯/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	乙烷/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	乙炔/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	故障 类型
46.132 1	11.572 3	8.522	33.144 7	0.628 9	0
32.523 1	4.134 1	25.432 1	35.413 1	2.554 3	0
43.435 2	32.451 3	9.441 3	15.323 1	0.412 3	0
46.432 4	14.242 1	6.343 2	32.523 1	0.534 3	0
40.412 3	25.423 1	14.643 1	20.854 2	7.323 4	0
45.613 3	15.395 8	5.673 4	11.346 9	0.990 1	0
⋮	⋮	⋮	⋮	⋮	⋮
49.769 6	8.448 5	19.354 8	2.150 5	20.276 5	4
44.792 7	17.492 4	23.559 2	3.64	10.515 7	4
23.023 3	12.004 6	12	10.059 3	61.821 3	4
49.423 4	7.434 1	19.084 6	1.048 4	22.088 7	4
34.561 3	17.849 1	40.976 1	6.981 4	7.084 9	4
22.094 8	14.501 8	14.508 5	11.500 9	60.975 6	4

表3 测试数据

氢气/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	甲烷/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	乙烯/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	乙烷/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	乙炔/ ( $\mu\text{L} \cdot \text{L}^{-1}$ )	故障 类型
46.13	11.57	8.52	33.14	0.63	0
38.19	31.03	22.91	7.88	0	1
0.96	15.79	70.25	12.3	0.71	2
56.98	10.48	13.57	1.28	17.68	4
24.32	16.36	30.18	1.67	27.47	3
33.66	2.9703	27.72	33.17	2.48	0
0.98	43.83	28.39	26.8	0	1
44.11	24.55	24.14	6.66	0.54	4
⋮	⋮	⋮	⋮	⋮	⋮
45.3	17.98	4.51	3.4	9.09	3
60.04	15.99	4.13	5.13	11.45	3
84.91	7.4	2.01	4.96	0	3
21.94	18.98	25.88	2.11	40.61	4

分别采用支持向量机、逻辑回归、最近邻分类、贝叶斯分类、决策树和随机森林对训练数据进行学习,得到故障诊断模型。利用测试数据对各个故障诊断模型进行模型验证,得到各个模型的识别结果见图3—8,蓝色圆圈表示的是实际的故障类型,红色圆圈表示的是预测的故障类型。

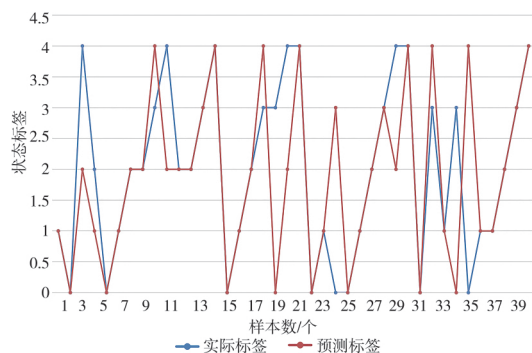


图3 决策树分类结果

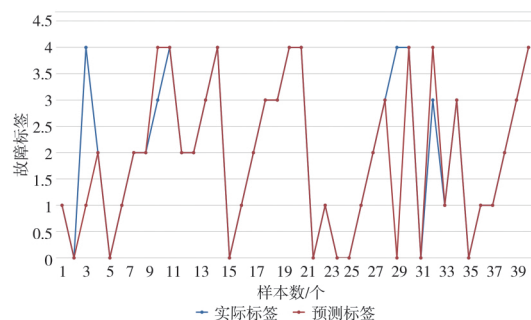


图4 支持向量机分类结果

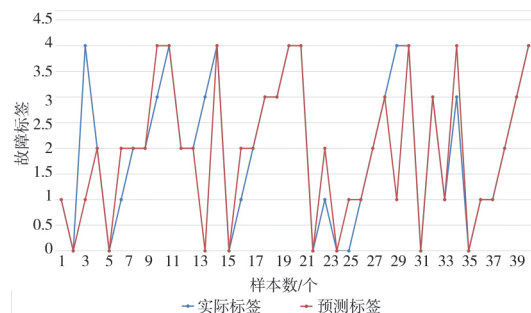


图5 最近邻分类结果

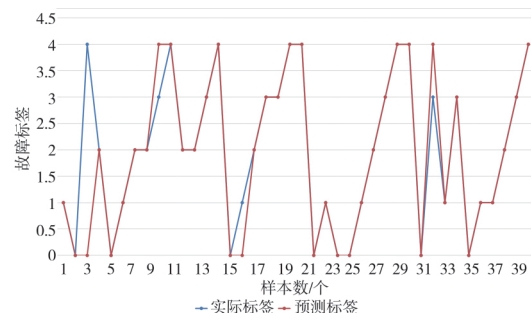


图6 随机森林分类结果

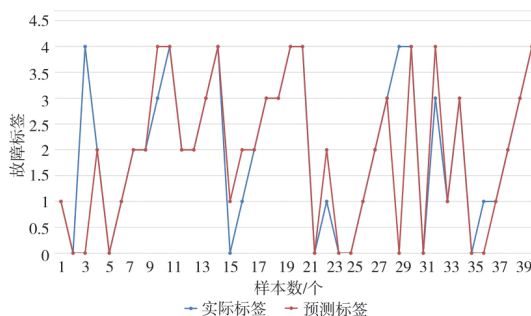


图7 逻辑回归分类结果

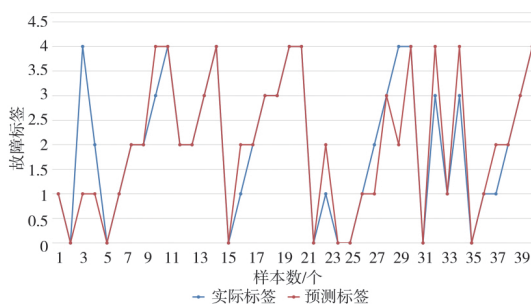


图8 贝叶斯分类结果

由 40 组测试数据对以上 6 个分类器的识别验证结果,可以得到各分类器的识别准确率,同时得到各分类器对每一种故障的识别准确率如表 4—5 所示。

表 4 不同分类器的模式识别结果

分类器	识别正确个数/个	识别准确率/%
决策树	28	70
支持向量机	36	90
最近邻分类	31	77.5
随机森林	36	90
逻辑回归	32	80
贝叶斯分类	30	75

表 5 不同分类器对不同故障的识别结果 %

分类器	正常 (0)	中低温 过热(1)	高温 过热(2)	低能量 放电(3)	高能量 放电(4)
决策树	75	100	87.5	37.5	50
支持向量机	100	100	100	75	75
最近邻分类	87.5	62.5	100	62.5	75
随机森林	100	87.5	100	75	87.5
逻辑回归	87.5	62.5	100	75	75
贝叶斯分类	100	62.5	75	62.5	75

将决策树、支持向量机和随机森林作为基分类器采用 vote 集成学习形成新的集合分类器,采用多数投票法的方式输出故障识别结果见图 9 和表 6—7。



图 9 vote 集成学习识别结果

表 6 vote 集成学习故障识别结果

分类器	识别正确个数	识别准确率/%
vote 集成学习	37	92.5

表 7 vote 集成学习对各故障类型识别结果 %

分类器	正常 (0)	中低温 过热(1)	高温过热 (2)	低能量 放电(3)	高能量 放电(4)
vote 集成学习	100	100	100	75	87.5

vote 集成学习相对于单一的分类器,将同质分类器和异质分类器进行合并,其故障识别准确率得到提升,故障识别准确率为 92.5%。

## 4 应用效果

通过实际现场中油色谱在线监测装置采集的 5

种特征气体进行变压器故障诊断,诊断结果见表 8。

表 8 220 kV 主变压器故障诊断结果

时间	氢气/ ( $\mu\text{L}\cdot\text{L}^{-1}$ )	甲烷/ ( $\mu\text{L}\cdot\text{L}^{-1}$ )	乙烯/ ( $\mu\text{L}\cdot\text{L}^{-1}$ )	乙烷/ ( $\mu\text{L}\cdot\text{L}^{-1}$ )	乙炔/ ( $\mu\text{L}\cdot\text{L}^{-1}$ )	预测 故障	实际 故障
3.23	14.3	1.78	0.38	0.43	0	温过 中低	温过 中低
3.15	124.25	17.13	1.99	6.51	0.32	热 低能 量放 电	热 低能 量放 电

第 1 组变压器的预测结果为中低温过热,通过实际的现场红外测温检测,发现主变压器本体下方连接处存在过热现场。第 2 组变压器的预测结果为低能量放电,通过对现场的变压器进行巡视检查,发现裸引线对套管导电管之间有放电的痕迹。

## 5 结论

基于机器学习搭建变压器故障诊断模型对变压器的故障类型进行识别,故障准确率由原来气体分析 IEC 三比值法的 80% 提高至 90%,通过 vote 集成学习进一步组合单一分类器形成新的故障分类器,故障识别的准确率进一步提高为 92.5%。采用机器学习方法对变压器进行故障诊断相比于传统的油中气体分析可以实现实时、随地的故障诊断,不需要进行现场油样采集和油化试验,大大提高了工作效率,减少了劳动时间。在实际变压器故障诊断中,具有实用性和可靠性,同时具有广泛的推广应用价值和实际意义。但目前只是通过简单的机器学习对故障进行诊断,存在一定的局限性,下一步将继续优化模型,增加训练数据量,增强模型的学习能力和适应能力,进一步提高故障识别的准确率。

参考文献:

- [1] 刘君,游家训,梁薇,等.基于加权 K 近邻算法的变压器故障诊断[J].电力系统及其自动化,2010,32(5):59-62.
- [2] 祖文超.基于油中溶解气体的支持向量机变压器故障诊断[D].北京:华北电力大学,2013.
- [3] 黎静华,栗然.基于知识粗糙度的多变量决策树在变压器故障诊断系统中的应用[J].电力自动化设备,2005,25(10):40-43.
- [4] 肖燕彩,张清.基于模糊支持向量机的变压器故障诊断[J].北京交通大学学报,2012,36(1):117-121.
- [5] 唐波.基于 Spark 和集成学习的分类方法研究[D].桂林:桂林电子科技大学,2017.
- [6] 李岳峰,赵志刚.变压器油中溶解气体的检测与故障诊断技术分析[J].科学与财富,2017(32):73.

本文责任编辑:靳书海