

基于机器学习的智能电网数据分析研究

文 ◆ 国网湖北省电力有限公司直流公司 郑潇涵 高梦露

引言

在新型电力系统建设过程中，实现电网运行的高可靠性、高效率和高智能化已成为主要目标。然而，智能电网以数据驱动为核心，其运行严重依赖于海量时序数据的有效采集、处理与建模分析。为了提升对电能质量扰动、电网负荷波动以及设备状态变化的识别与响应能力，构建以模式识别和预测建模为核心的数据分析体系成为研究重点。

1 机器学习在智能电网中应用方向

目前，机器学习算法在智能电网数据分析领域被广泛应用。电能质量扰动分类任务中，常采用的监督学习算法包括支持向量机（Support Vector Machine, SVM）、随机森林（Random Forest, RF）与反向传播神经网络（Back Propagation Neural Network, BPNN）等。典型的数据处理流程包括数据采集、特征提取、数据标注、数据划分、模型训练以及模型评估^[1]。

（1）数据采集。收集包含各种电能质量扰动的实时数据，如电压波形、电流波形、频谱分析等。

（2）特征提取。从原始数据

中提取有意义的特征，这些特征包括频率、振幅、谐波含量等，而特征提取的目的是将复杂的波形数据转化为计算机可以理解的数字特征数据标记。

（3）数据标注。对采集的数据进行标注，将不同类型的电能质量干扰分类。例如，对电压暂降、电压暂升、电流谐波等进行分类标注。

（4）数据划分。将标注好的数据分为训练集和测试集，训练集用于训练分类器模型，测试集用于评估模型的性能。

（5）模型训练。使用机器学习算法，根据训练集的特征和标签，训练一个分类器模型，模型将学习不同类型电能质量干扰的模式和规律。

（6）模型评估。使用测试集评估模型的分类性能，如准确率、召回率、F1 分数等，一旦分类器模型训练完毕，就可以用于实时数据的分类。同时，将实时采集的电能质量数据输入模型，模型会根据学习到的模式和规律将其分类为不同的干扰类型。

针对实际运行环境中存在的数据规模庞大、维度冗余、分布复杂等问题，传统分析方法在实时性和稳定性方面面临诸多挑战。为此，可按数据种类与应用场景分级建模，采用特征筛选、模型压缩、集成策略等方



【作者简介】郑潇涵（1995—），男，山东寿光人，本科，工程师，研究方向：高压直流输电、智能电网。

法对分析系统进行结构化优化,从而提升模型泛化能力与处理效率^[2]。

2 智能电网数据分析方法

2.1 基于小波分析的数据处理方法

在智能电网复杂运行环境下,原始信号往往同时具有突变性、周期性与随机性,传统信号处理方法难以兼顾时域与频域特性。离散小波变换(Discrete Wavelet Transform, DWT)作为一种多尺度分析方法,能够将电力信号在不同时间尺度上进行分解与重构,实现对扰动特征的有效捕捉。

DWT通过构造尺度函数与小波函数,将原始信号逐层分解为低频与高频子信号。低频部分描述信号的趋势变化,高频部分保留信号中的尖峰、跳变与短时干扰,是电能质量扰动检测的核心信息来源。

在实际电网应用中,电压暂降、谐波畸变等扰动表现为信号在短时频率或幅值上的异常变化,这类特征往往集中在DWT的高频部分。因此,可将各层小波系数作为输入特征用于后续建模。模型准确性、稳定性与分解层数密切相关,需综合考虑采样频率、信号带宽与扰动持续时间来设置最优层级。

DWT方法能够提升分类模型在复杂数据下的计算效率与预测准确性,特别在数据维度较高、扰动特征模糊的任务中更具优势。其中,“电压暂变记录”与“谐波分析”两类任务对时频精度要求较高,DWT与RF或SVM的组合能够有效提升系统响应的灵敏度与鲁棒性。故小波变换作为一种可扩展、低冗余的特征提取手段,可显著提升电网扰动分析任务中的建模效率与分类性能,具有较强的工程适应性与泛化能力。

2.2 基于 RF 的数据处理方法

RF是一种集成学习方法,通过构建多棵决策树并采用多数投票机制完成分类,具备高鲁棒性和强泛化能力。针对电网数据中存在的非线性、多噪声特征,RF能够有效降低过拟合风险,并在处理中等规模数据集时具有良好的计算效率。

RF模型利用自助采样法构建子训练集,并在每棵决策树的训练过程中随机选择特征子集,从而提升模型多样性与稳定性。在参数设置上,模型性能主要受决策树的数量和最大深度影响。通过实验发现,适度增加决策树的数量可提升模型稳定性,但也会带来训练时间增加的问题,因此需平衡精度与效率。

应用结果表明,RF在处理电网数据时具备良好的分类精度、可扩展性和抗干扰能力,特别适用于扰动类型复杂、信号特征不明显的场景,是构建智能电网数据分析系统的重要技术手段。

2.3 基于 SVM 的数据分析方法

SVM是一种基于统计学习理论的分类方法,具有良好的泛化能力,特别适用于样本规模有限、特征维度高且分布复杂的问题。在电网扰动识别、状态评估等任务中,SVM凭借其构建最优分类边界的能力,广泛应用于非线性问题的建模。

SVM法将不同类别的数据划分开来。为应对现实中数据不可线性可分的问题,引入核函数将输入数据映射至高维特征空间,实现非线性可

分问题的线性求解。

针对配电网中面临的数据维度高、特征密度低等问题^[3],SVM融合小波特征后建模效果更佳。在电压暂降分类任务中,SVM模型在准确率、召回率与F1分数等多个指标上均优于传统方法。

从工程角度看,SVM适用于对精度要求较高但计算资源受限的应用场景,尤其在扰动特征清晰、样本较少的任务中表现优越。在融合特征提取方法与参数优化技术后,SVM可成为电网数据分析中高精度模型构建的重要手段。

2.4 基于 BPNN 的数据分析方法

BPNN是一种典型的多层前馈神经网络结构,适用于建模强、非线性、高耦合关系的复杂系统。其基本思想是信号正向传播、误差反向传播,通过不断调整权重优化模型输出,使之更贴合目标输出。

在智能电网中,系统状态与负荷波动常表现出复杂的非线性关系^[4],BPNN因其结构灵活、非线性拟合能力强,广泛应用于负荷预测、设备状态识别以及异常检测等场景。

BPNN法通常需要构建三层感知器网络,输入层与输出层神经元数量分别依据特征维度与扰动类别设置,隐藏层通过经验法结合试验确定。激活函数采用双曲正切函数,在提高非线性表达能力的同时保持梯度稳定。模型训练采用误差反向传播算法,优化目标为最小化均方误差。

学术界用以下4个统计参数来评估模型的性能,分别是召回率(Recall)、精确率(Precision)、准确率(Accuracy)和F1分数,召回率定义如式(1)所示,精确

率定义如式（2）所示。

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (2)$$

召回率表示模型，可在所有真实正类中识别出正类的比例。式（1）~（2）中，*TP*（True Positive）表示将正类正确预测的数量；*TN*（True Negative）表示将负类正确预测的数量；*FP*（False Positive）表示将负类错误预测为正类的数量；*FN*（False Negative）表示将正类错误预测为负类的数量。

准确度定义为正确分类的样本与总样本的比率，如式（3）所示。

$$Accuracy = \frac{TN + TP}{TP + TN + FN + FP} \times 100\% \quad (3)$$

另一个重要参数是 F1 分数，它是准确率与召回率的调和平均值，适用于正负类样本不平衡时的综合评估。F1 分数定义如式（4）所示。

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \times 100\% \quad (4)$$

研究结果显示，RF 在训练效率和准确率上占优，SVM 在小

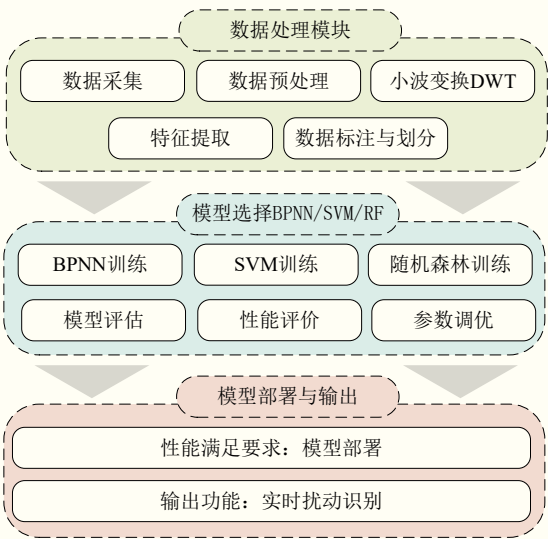


图 1 算法分析框架图

样本精度控制方面表现稳定，而 BPNN 在召回率与非线性建模方面表现良好，适合应用于扰动形态复杂、时间序列特征显著的电网数据。

3 智能电网数据分析应用架构

随着电网规模的持续扩大和运行复杂性的不断提升，数据驱动的分析技术在智能电网中的应用需求日益增强。算法分析框架图如图 1 所示。

通过构建统一的数据处理流程，结合小波变换的特征提取能力与 3 种主流分类模型（RF、SVM、BPNN）的建模优势，实现了对典型电网扰动的高精度分类。其中，随机森林在处理高维数据和实现快速训练方面表现突出，支持向量机在样本较少、特征边界清晰的任务中具备良好的判别能力，BPNN 则适用于复杂非线性状态建模，适合动态性强的场景，而配电网中常见的数据维度大问题，利用 DWT 方法能够较好解决。

结语

构建以特征工程为基础、多算法融合的电网数据建模框架，是支撑智能电网数据治理与状态感知能力提升的重要路径。未来研究可进一步拓展至多源异构数据融合、自适应模型调节与在线识别等方向，以提升系统的实时性与动态响应能力。

引用

[1] 姚梓豪,栗远明,马自强,等.基于机器学习的多目标缓存侧信道攻击检测模型[J].计算机应用,2024,44(6):1862-1871.

[2] 谷毅,富子豪,王登政,等.基于机器学习的配用电场景信号覆盖优化技术[J].电波科学学报,2024,39(3):518-525.

[3] 曹瑞,刘燕斌,裔扬.基于物理信息机器学习的复杂系统长时间演化分析[J].控制理论与应用,2024,41(11):2041-2052.

[4] 石磊,李天,高宇飞,等.基于机器学习的数据库系统参数优化方法综述[J].郑州大学学报(工学版),2024,45(1):1-11+28.