

大数据在电力负荷预测中的应用分析

魏大伟

新疆兵安电力建设有限责任公司 新疆 图木舒克 844000

摘要: 本文聚焦大数据在电力负荷预测中的应用。阐述了大数据与电力负荷预测相关理论,分析大数据环境下电力负荷数据特点与处理方式。介绍基于大数据的多种预测方法,包括传统、机器学习、深度学习及组合预测方法。最后从数据、模型、技术支撑层面提出优化策略,旨在提升电力负荷预测精准度与可靠性,为电力系统规划、运行与调度提供有力支持。

关键词: 大数据; 电力负荷预测; 数据挖掘; 机器学习; 预测模型

引言: 电力负荷预测是电力系统稳定运行的关键,精准预测对保障电网安全、提高能源利用效率意义重大。随着大数据技术兴起,其为电力负荷预测带来新契机。大数据具有海量、多样等特征,电力领域数据来源广泛,蕴含丰富负荷变化信息。基于大数据的预测理论融合多学科知识,数据预处理技术可优化模型性能。在此背景下,深入研究大数据在电力负荷预测中的应用具有重要的现实意义。

1 大数据与电力负荷预测相关理论概述

电力负荷预测是电力系统规划、运行与调度的重要依据,精准预测对保障电网安全稳定、提高能源利用效率意义重大。大数据具有海量、多样、高速和价值密度低等特征。在电力领域,大数据来源广泛,涵盖智能电表采集的实时用电数据、气象数据、经济数据、用户行为数据等。这些多源异构数据蕴含着丰富的负荷变化信息,为提高预测精度提供了坚实基础。基于大数据的电力负荷预测理论融合了多种学科知识。统计学方法如回归分析、时间序列分析等,能从历史数据中挖掘负荷变化的统计规律,建立预测模型^[1]。机器学习算法,像支持向量机、神经网络等,凭借强大的非线性拟合能力,可处理复杂的数据关系,提升预测准确性。深度学习作为机器学习的延伸,通过构建深层神经网络模型,能自动学习数据中的深层次特征,在处理大规模、高维度电力数据时表现卓越。另外,大数据环境下的数据预处理技术,如数据清洗、特征提取与选择等,能有效去除噪声、提取关键特征,进一步优化预测模型性能,从而实现电力负荷更精准、可靠的预测。

2 大数据环境下电力负荷数据的特点与处理

2.1 电力负荷数据的来源与类型

大数据环境下,电力负荷数据的来源呈现多维度、广覆盖的特征,核心来源可分为三大类。一是电力系统内

部数据源,包括智能电表、SCADA系统、配电自动化终端等采集的实时负荷数据、电压电流数据,以及电力调度系统的历史运行数据,这类数据具有时序性强、精度高的特点,是负荷预测的核心基础数据^[2]。二是用户侧数据源,涵盖工业企业生产计划、居民用户用电行为、商业建筑运营数据等,通过物联网、移动终端等渠道采集,反映用户用能需求的个性化特征。三是外部关联数据源,包括气象数据、地理信息数据、社会经济数据(GDP、产业结构、节假日安排等)。从数据类型看,包含结构化的时序负荷数据、半结构化的用户档案数据与非结构化的文本、图像类关联数据,多源异构的数据特征是大数据环境下电力负荷数据的显著标识。

2.2 电力负荷数据的特点

大数据环境下的电力负荷数据具备区别于传统数据的显著特征,核心表现为四方面。其一,数据体量呈指数级增长,智能电表等终端的高频采集使得单日负荷数据量可达百万级甚至亿级,海量数据对存储与处理能力提出极高要求。其二,数据类型复杂多样,除传统的结构化时序负荷数据外,还包含大量非结构化的用户行为数据、气象图像数据等,多源异构数据的融合难度大。其三,数据时序性与动态性突出,电力负荷随时间呈现日、周、月、季及年度的周期性变化,同时受突发因素(如极端天气、重大事件)影响呈现剧烈波动,数据动态变化特征显著。其四,数据价值密度分布不均,海量数据中蕴含用户用能规律、系统运行特征的核心价值信息,但与大量冗余、噪声数据并存,需通过高效处理挖掘有效价值。这些特点决定了电力负荷数据处理需依托大数据技术实现精准筛选与有效利用。

2.3 电力负荷数据的预处理方法

电力负荷数据的预处理是提升负荷预测精度的关键前置环节,需针对数据的噪声、缺失、异常等问题开展

系统性处理,核心方法涵盖四大类。一是数据清洗,通过异常值检测算法(如 3σ 准则、箱线图法)识别并剔除极端异常数据,采用插值法、均值法等填补缺失数据,同时去除重复采集的冗余数据,保障数据的完整性与准确性。二是数据集成,针对多源异构数据,通过数据对齐、格式统一、属性融合等操作,将不同来源、不同格式的负荷数据与关联数据整合为统一的数据集,解决数据孤岛问题。三是数据变换,包括数据标准化(如Z-score标准化、最小-最大归一化)、离散化与特征编码,将不同量纲、不同类型的数据转换为适配预测模型的格式,同时提取负荷周期特征、气象关联特征等关键特征变量。四是数据降维,针对高维特征数据,采用主成分分析(PCA)、t-分布邻域嵌入(t-SNE)等算法降低维度,剔除冗余特征,提升模型训练效率与预测精度。预处理后的高质量数据为后续预测模型的构建提供坚实基础。

3 基于大数据的电力负荷预测方法

3.1 传统电力负荷预测方法回顾

传统电力负荷预测方法以统计分析与数学建模为核心,形成成熟技术体系,主要有三类。时间序列分析法,以ARMA、ARIMA模型为代表,分析负荷数据时序相关性,构建线性关联模型,适用于短期、平稳场景,原理简单、计算量小,但难捕捉非线性、非平稳特征。回归分析法,如线性回归、多元线性回归,构建负荷与影响因素的回归方程,量化影响程度,适用于中期预测,但依赖线性关系,适配性有限。专家系统法,依托专家经验知识构建规则库进行预测,灵活性强,但依赖经验积累,难以自动化与规模化应用。在大数据时代,传统方法面临处理能力不足、精准度受限等问题,需与大数据技术结合实现升级。

3.2 基于机器学习的电力负荷预测方法

基于机器学习的电力负荷预测是大数据与负荷预测的核心结合点,让模型从海量数据自主学习特征规律以精准预测,主流方法有四类。支持向量机(SVM)用核函数将低维非线性数据映射到高维空间构建超平面,适用于小样本、高维数据预测,泛化能力强,但参数调优难。决策树与集成学习算法,决策树易实现但易过拟合,集成学习(如随机森林、梯度提升树)组合弱学习器提升精度、降低过拟合风险^[1]。朴素贝叶斯算法基于贝叶斯定理与特征独立假设计算负荷类别概率,适用于短期、多分类预测,计算效率高。K近邻(KNN)算法通过计算相似度选取近邻样本预测,原理简单,但对数据分布敏感。机器学习方法弥补了传统方法不足,成为主流技术方向。

3.3 基于深度学习的电力负荷预测方法

基于深度学习的电力负荷预测是大数据驱动的技术升级,通过构建多层神经网络模型自动提取深层特征与非线性关系,提升预测精度,核心方法有三类。循环神经网络(RNN)与长短期记忆网络(LSTM),RNN用循环结构捕捉时序依赖关系,但易梯度消失;LSTM通过门控机制解决长序列依赖问题,适配长时序特征,常用于短期、超短期预测。卷积神经网络(CNN)用卷积层、池化层提取局部特征,适用于处理多维关联数据,挖掘负荷与外部因素的空间关联特征。Transformer模型基于自注意力机制,并行处理长序列数据,精准捕捉全局特征与长距离依赖,在中长期预测中广泛应用,适配新型电力系统复杂场景。深度学习是大数据环境下核心预测技术,但模型训练需海量数据与强大计算资源。

3.4 组合预测方法

组合预测方法是通过整合多种单一预测模型的优势,弥补单一模型的局限性,实现预测精度与鲁棒性提升的技术路径,是大数据环境下电力负荷预测的重要发展方向,核心构建逻辑与应用场景如下。从组合逻辑看,主要分为模型层组合与特征层组合两类:模型层组合是将不同类型的预测模型(如LSTM与ARIMA、随机森林与SVM)通过加权平均、stacking、Blending等方式融合,结合各模型的优势;特征层组合是先通过不同方法提取负荷数据的不同特征,再输入不同模型进行预测,最后整合结果。从应用场景看,针对负荷的多尺度特征,组合方法可实现短期、中长期预测的适配:短期预测中,结合LSTM的时序特征提取能力与SVM的小样本拟合优势,提升精准度;中长期预测中,融合回归分析的趋势拟合能力与Transformer的长序列特征捕捉能力,适配复杂经济与社会环境下的负荷预测。另外,组合预测方法还可结合大数据的多源数据优势,引入外部关联特征优化模型融合逻辑,进一步提升预测的稳定性与可靠性,成为应对复杂负荷预测场景的关键技术。

4 大数据在电力负荷预测应用中的优化策略

4.1 数据层面优化策略

数据层面的优化是提升大数据驱动电力负荷预测效果的基础,核心围绕数据质量、数据整合与数据安全三大维度展开。优化数据采集与治理体系,部署高精度、低延迟的智能采集终端,提升数据采集频率与精度;建立全流程数据治理机制,通过数据清洗、去重、补全、标准化等操作,保障数据的完整性、准确性与一致性,降低噪声数据对预测模型的干扰。深化多源数据整合与特征工程,打破电力系统内部、用户侧与外部关联数据的壁

垒,构建统一的电力负荷数据中台;针对不同预测场景,提取差异化的核心特征,如短期预测的时序特征、中长期预测的经济与社会特征,通过特征选择与融合算法剔除冗余特征,提升特征有效性。强化数据安全性与隐私保护,采用数据加密、访问控制、匿名化处理等技术,保障用户用电隐私与电力系统数据安全,同时遵循数据合规使用原则,实现数据价值挖掘与安全保障的平衡。高质量的数据体系为预测模型的训练与优化提供核心支撑。

4.2 模型层面优化策略

模型层面的优化是提升电力负荷预测精准度与适应性的核心环节,需结合大数据技术特性与负荷数据特征,从模型选择、参数优化与模型融合三方面开展。其一,精准适配模型选择,针对不同预测时间尺度(超短期、短期、中长期)与数据特征(线性/非线性、平稳/非平稳),选取最优基础模型:超短期预测优先采用LSTM、Transformer等时序特征捕捉能力强的模型;短期预测结合机器学习算法(如随机森林、XGBoost)与传统时间序列模型,平衡精度与效率;中长期预测融合回归分析与深度学习模型,适配复杂趋势变化。其二,开展模型参数智能优化,利用网格搜索、遗传算法、贝叶斯优化等算法,对模型的关键参数进行自动调优,避免人工调优的主观性与局限性,提升模型的拟合效果与泛化能力。其三,深化多模型融合优化,构建集成学习框架,通过加权融合、动态融合等方式,组合多个互补模型的预测结果,例如将LSTM的时序预测结果与随机森林的特征拟合结果融合,同时引入误差修正模型,进一步降低预测误差。另外,持续迭代模型结构,结合新的负荷特征与大数据技术,优化网络层数、激活函数等模型参数,提升模型对复杂场景的适配性。

4.3 技术支撑层面优化策略

技术支撑层面的优化是保障大数据在电力负荷预测中高效应用的关键,涵盖算力支撑、技术架构与人才体系三大核心维度。升级算力支撑体系,依托云计算、边缘计算与超算中心构建混合算力架构:边缘计算部署于

电力终端与配电节点,实现实时负荷数据的本地预处理与短期预测,降低数据传输延迟;云计算与超算中心支撑海量历史数据的存储与中长期预测模型的训练,保障大规模计算需求。同时采用分布式计算框架(如Spark、Flink)提升数据处理与模型训练的效率,适配大数据的并行计算特性^[4]。优化技术架构体系,搭建“数据-算法-应用”一体化技术架构,构建统一的数据中台、算法中台与应用中台:数据中台实现多源数据的整合与治理;算法中台提供标准化的预测模型库与优化工具包,降低模型开发门槛;应用中台针对电力调度、规划设计等不同场景,提供定制化的预测服务与可视化展示。完善人才支撑体系,培养跨领域复合型人才,涵盖电气工程、数据科学、计算机技术等多领域知识;开展产学研合作,联合高校、科研机构与电力企业,攻克大数据负荷预测的关键技术难题,同时加强技术培训与推广,提升行业技术人员的应用能力。完善的技术支撑体系为大数据技术在负荷预测领域的落地应用提供坚实保障。

结束语

大数据为电力负荷预测带来变革,通过多源数据融合与先进算法应用,预测精度与可靠性显著提升。然而,数据质量、模型适应性及技术支撑等方面仍面临挑战。未来,需持续优化数据采集治理、模型选择与融合以及算力技术架构等,加强跨领域人才培养与产学研合作,推动大数据在电力负荷预测中深度应用,助力电力系统智能化、高效化发展。

参考文献

- [1]邹翔,王星.大数据分析在电力系统负荷预测与规划优化中的应用[J].消费电子,2025(15):101-103.
- [2]张明杰,邓志东,刘鲲鹏.基于历史大数据分析的电力负荷区间组合预测系统[J].电子设计工程,2024,32(3):166-169.
- [3]高嘉廉.基于大数据分析的电力系统负荷预测与规划优化[J].光源与照明,2025(1):78-80.
- [4]王泽众,李涛,王可嘉.基于大数据技术的电力负荷预测与优化控制分析[J].集成电路应用,2023,40(10):130-131.