

基于 XGBoost 与 LightGBM 集成的 电动汽车充电负荷预测模型*

吴丹¹, 雷琰¹, 李芝娟², 王宁³, 段艳³

(1. 国网上海市电力公司, 上海 200122; 2. 浦东供电公司, 上海 200122; 3. 同济大学汽车学院, 上海 201804)

摘要: 随着电动汽车规模化发展, 充电站负荷对电网造成一定影响, 为保障电网平稳运行, 提出一种基于极端梯度提升(eXtreme Gradient Boosting, XGBoost)与轻量级梯度提升机(Light Gradient Boosting Machine, LightGBM)融合的电动汽车充电负荷预测模型。该方法运用 Stacking 集成学习的策略: 首先根据时间特征与历史负荷数据采用 XGBoost 与 LightGBM 算法构建负荷预测的基学习器, 然后采用岭回归(Ridge Regression, RR)算法将基学习器的输出结果进行融合之后输出负荷预测值。为了对比多种不同的负荷预测模型, 采用上海市嘉定区的充电站订单数据进行试验, 结果表明, 该方法所构建的负荷预测模型相比单一算法模型具有更高的预测准确度, 对电网平稳运行有一定理论及实用价值。

关键词: 电动汽车; 负荷预测; Stacking 集成学习; 极端梯度提升(XGBoost); 轻量级梯度提升机(LightGBM)

中图分类号: TM910.6; U469.72

文献标识码: A

DOI: 10.16157/j.issn.0258-7998.212316

中文引用格式: 吴丹, 雷琰, 李芝娟, 等. 基于 XGBoost 与 LightGBM 集成的电动汽车充电负荷预测模型[J]. 电子技术应用, 2022, 48(9): 44-49.

英文引用格式: Wu Dan, Lei Yu, Li Zhijuan, et al. Electric vehicle charging load forecasting based on XGBoost and LightGBM integration model[J]. Application of Electronic Technique, 2022, 48(9): 44-49.

Electric vehicle charging load forecasting based on XGBoost and LightGBM integration model

Wu Dan¹, Lei Yu¹, Li Zhijuan², Wang Ning³, Duan Yan³

(1. State Grid Shanghai Municipal Electric Power Company, Shanghai 200122, China;

2. State Grid Shanghai Pudong Electric Power Supply Company, Shanghai 200122, China;

3. Institute of Automobile, Tongji University, Shanghai 201804, China)

Abstract: With the scale development of electric vehicles, the load of charging stations has a certain impact on the power grid. In order to ensure the power grid run steadily, an electric vehicle charging load forecasting model based on the integration of eXtreme Gradient Boosting(XGBoost) and Light Gradient Boosting Machine(LightGBM) is proposed. This method uses the strategy of stacking integrated learning. Firstly, the base models of load forecasting are constructed based on XGBoost and LightGBM respectively. And then Ridge Regression(RR) algorithm is used to fuse the output results of the base models, the fusion result is the load forecasting value. Based on a variety of different load forecasting models, comparative experiments are carried out with the order data of charging station located in Jiading District, Shanghai. The results show that the load forecasting model constructed by this method has higher forecasting accuracy than the model based on single algorithm, and has certain theoretical and practical value for the smooth operation of power grid.

Key words: electric vehicle; load forecasting; Stacking integrated learning; eXtreme Gradient Boosting (xGBoost); Light Gradient Boosting Machine(LightGBM)

0 引言

近年来电动汽车的保有量快速上升, 电动汽车规模化将对电网的输电网络、配电网络、充电设施等多方面

带来影响^[1-8], 因此准确的电动汽车充电负荷预测对于电网平稳运行具有重要意义。

电动汽车充电负荷预测是根据过去一段时间的用电负荷及日期类型等相关数据预测未来一段时间的用电负荷^[9], 构建准确的电动汽车充电负荷预测模型不仅

* 基金项目: 上海市科技类项目(P21-098)

有利于电网对充电站的充电负荷进行调度与管理,也有利于充电站制定科学的运营计划。不少国内外学者从用户端及车端出发对电动汽车的充电负荷预测展开了研究^[10-18],通过融合电动汽车出行特征、用户行为特点和道路交通状况等因素,建立电动汽车充电负荷预测模型。真实的充电过程与从车端仿真结果存在差异,所以从充电站端得到的负荷预测结果比车端更能真实反映电动汽车充电对电网造成的影响。目前从充电站端对充电负荷进行预测的相关研究较少,并多数是以深度学习算法构建负荷预测模型,具有一定局限性,例如文献[19]采用模糊聚类分析与BP神经网络相结合的方法建立电动汽车充电负荷的短期预测模型,文献[20]采用随机森林与神经网络相结合的方法建立电动汽车充电站短期负荷的预测模型,深度学习算法在输入序列较长时存在梯度消失问题,模型无法克服对异常值敏感的缺点,导致模型预测准确度变差。

针对上述问题,本文从充电站端的数据出发,通过挖掘电动汽车充电负荷随时间的变化规律,提取负荷影响因素作为模型的输入特征。为了实现较高的负荷预测准确度,本文采用数据挖掘比赛中表现优异的XGBoost与LightGBM算法分别构建负荷预测模型,再结合Stacking集成学习的策略,利用岭回归模型将XGBoost与LightGBM模型的输出结果进行融合之后再输出。实验结果表明,XGBoost与LightGBM模型实现了较高的预测准确度高,再采用Stacking集成学习方法将XGBoost与LightGBM模型的预测结果进行融合后,模型的预测准确度得到了进一步提升。

1 电动汽车充电站的负荷特性

本文分析数据源自上海市嘉定区内的充电站于2020年5月~10月份的充电订单数据。原始数据没有能直接反映负荷变化的数据,利用订单数据中的“交易电量”“充电开始时间”“充电结束时间”信息,通过下式计算充电负荷:

$$\text{Power}_{t_a, t_b} = \frac{\text{Trade_Electricity}_{t_a, t_b}}{t_b - t_a} \quad (1)$$

式中, t_a 为充电开始时间, t_b 为充电结束时间; Power_{t_a, t_b} 是 t_a 到 t_b 的时间段内的平均充电负荷,单位为kW; $\text{Trade_Electricity}_{t_a, t_b}$ 是 t_a 到 t_b 时间段内的交易电量,单位为kW·h。

图1是嘉定区所有充电站整个7月份的负荷曲线(采样间隔为1h),该月没有法定节日。从图中可以看到,充电负荷的波动幅度较大,但存在明显的以日为间隔的周期性,即充电负荷曲线在一天中的相同时段内的变化趋势相似。此外,负荷变化和日类型也密切相关,表现为负荷曲线不仅变化趋势相同,并且在同一采样时刻的数值接近,即每周第 N 天的负荷曲线具有相似性。

图2是位于嘉定区的充电站在国庆节前后的负荷曲线(采样间隔为1h)。从图中可以看到,国庆节这一周的充电负荷明显高于其他时间,所以节假日也是影响充电负荷的一个重要因素。

对于一个特定的模型,预测相似的数据往往会得到相似的结果,基于充电负荷变化呈现周期性的特点,同时考虑节假日对负荷的影响,以预测日期为 i 的 j 时刻的充电负荷 $P_{i,j}$ 为例,构建日期 $i-1\sim i-7$ 在 j 时刻的充电负荷序列 $(P_{i-1,j}\sim P_{i-7,j})$ 、日期 i 的 $j-1\sim j-7$ 时刻的充电负荷序列 $(P_{i,j-1}\sim P_{i,j-7})$,以及构建日期特征:一年中的第几周(Week_of_Year _{i})、周几(Weekday _{i})、第几月(Month _{i})、几号(Day _{i})、几点(Hour _{j}),将上述特征作为模型的输入特征。

2 模型相关算法原理

XGBoost和LightGBM都是基于梯度提升决策树原理^[21]改进而来的。XGBoost在对目标函数进行泰勒展开时,会将其展开至二阶而非一阶,此外XGBoost还在子叶权重中加入了L2正则化,即平方正则化项,上述改进使得XGBoost获得了更为优异的性能^[22]。而LightGBM使用直方图算法、带深度限制的按叶子生长策略对GBDT算法

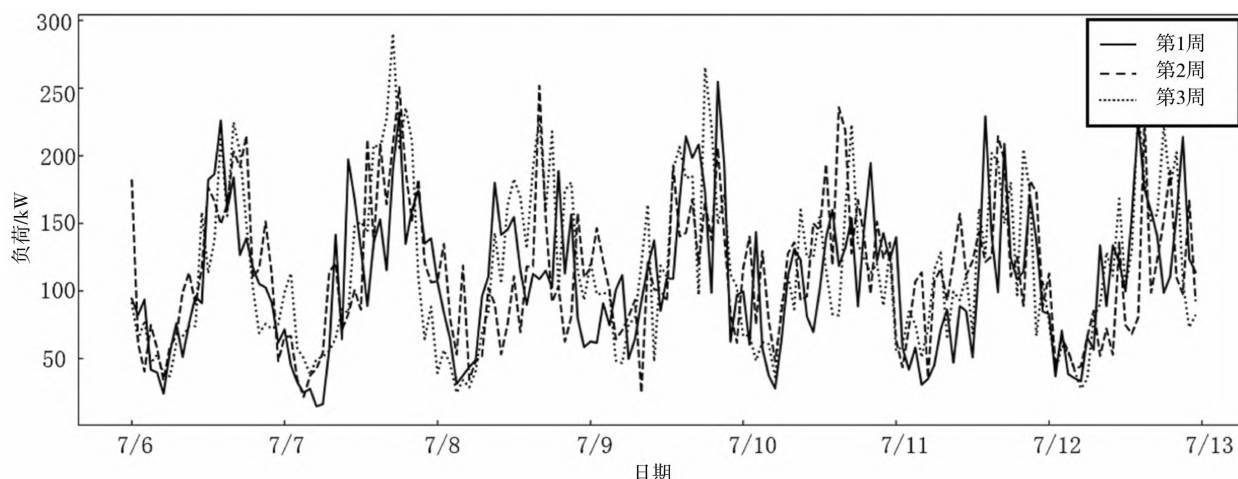


图1 嘉定区7月份的充电负荷曲线

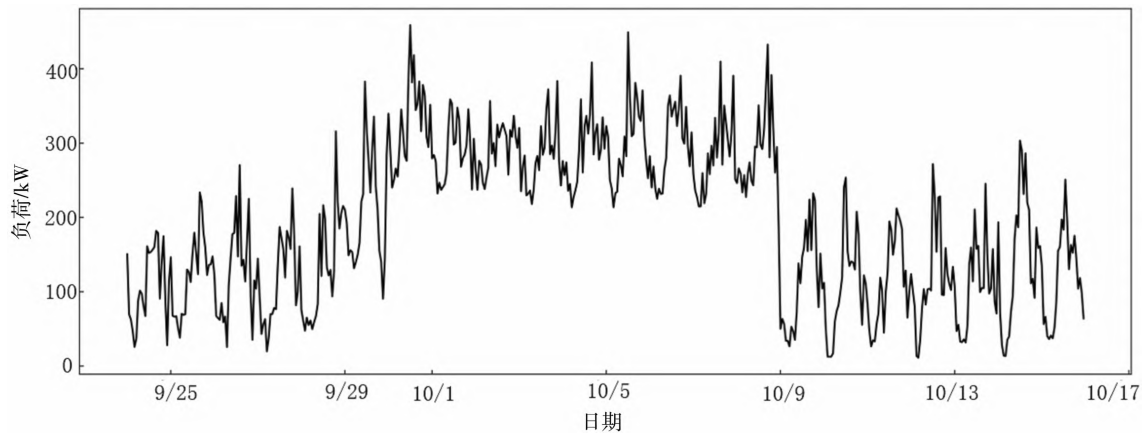


图2 嘉定区国庆节前后的充电负荷曲线

进行了改进^[23],另外通过限制最大深度来防止模型过拟合,所以 LightGBM 模型可以在不降低预测精度的同时加快预测速度。

Stacking 将多个不同类型的学习算法进行集成,从而取得优于单一学习算法的性能^[24]。模型一般采用两层式结构,第一层由 n 个基学习器构成,第二层由一个元学习器构成。Stacking 的学习方式如图 3 所示,首先采用 K 折交叉验证法训练与测试第一层模型中的 n 个基学习器;再将第一层的预测结果组合成新的数据集,作为第二层元学习器的输入数据;最后元学习器的预测结果即为最终的预测结果。元学习器通过学习基学习器的预测误差,从而达到提升预测精度的效果。

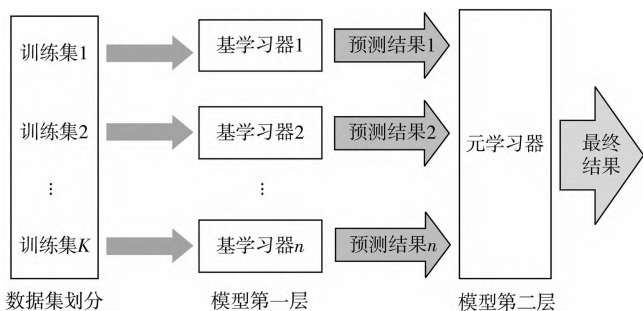


图3 Stacking 学习方式

3 模型框架设计

不同学习算法具有各自的误差缺陷。XGBoost 算法在 GBDT 算法的基础上做了优化,但是效率降低,LightGBM 算法相比 XGBoost 算法在效率和准确率方面有了提升,但与 XGBoost 算法一样,都可能会长出比较深的决策树,容易过拟合,并对噪声比较敏感。Stacking 框架下的多模型融合方法可以弱化单一基学习器的误差影响^[25],通过对 XGBoost 模型和 LightGBM 模型进行融合,以达到提升预测准确度的效果。

Stacking 集成模型设计为 2 层结构:第一层由 XGBoost 算法和 LightGBM 算法构成融合系统的基学习器层,第

二层将第一层基学习器的输出作为输入,对于预测目标相同的模型,它们的输出结果可能存在多重共线性,所以第二层元学习器采用了岭回归算法。Stacking 集成模型的框架如图 4 所示,模型融合过程可以描述为以下 2 个步骤:

(1)使用 5 折交叉验证的策略来训练 XGBoost 和 LightGBM 基学习器,即训练集中的 4/5 用于训练,剩余的 1/5 用于验证,利用这两个基学习器生成训练集与测试集的 2 组预测值;

(2)第一层基学习器输出的训练集的 2 组预测值用作第二层元学习器的训练集,而第一层基学习器输出的测试集的 2 组预测值用作第二层元学习器的测试集,作为预测集的最终预测结果。

将融合 XGBoost 算法和 LightGBM 算法的 Stacking 集成模型用于充电桩负荷预测流程,如图 5 所示。

4 算例分析与验证

4.1 数据处理与评价指标

实验数据是位于上海市嘉定区的充电站的负荷信息,时间为 2020 年 5 月~2020 年 10 月。考虑到原始数据的负荷序列变化幅度较大,如图 1 和图 2 所示,负荷数据的噪声太多,为便于模型拟合数据,按照如下公式对负荷数据进行平滑处理:

$$\ln_Power = \ln(Power + 1) \quad (2)$$

式中,Power 表示原始负荷值,ln_Power 表示经过对数平滑处理后的负荷值。

选取平均绝对百分误差(Mean Absolute Percentage Error, MAPE)、归一化均方根误差(Normalized Root Mean Squared Error, NRMSE)作为负荷预测模型的评价指标,计算公式如下:

$$MAPE = \frac{1}{N} \sum_{n=1}^N \left| \frac{\hat{y}_m - y_m}{y_m} \right| \times 100\% \quad (3)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{n=1}^N (\hat{y}_m - y_m)^2} \quad (4)$$

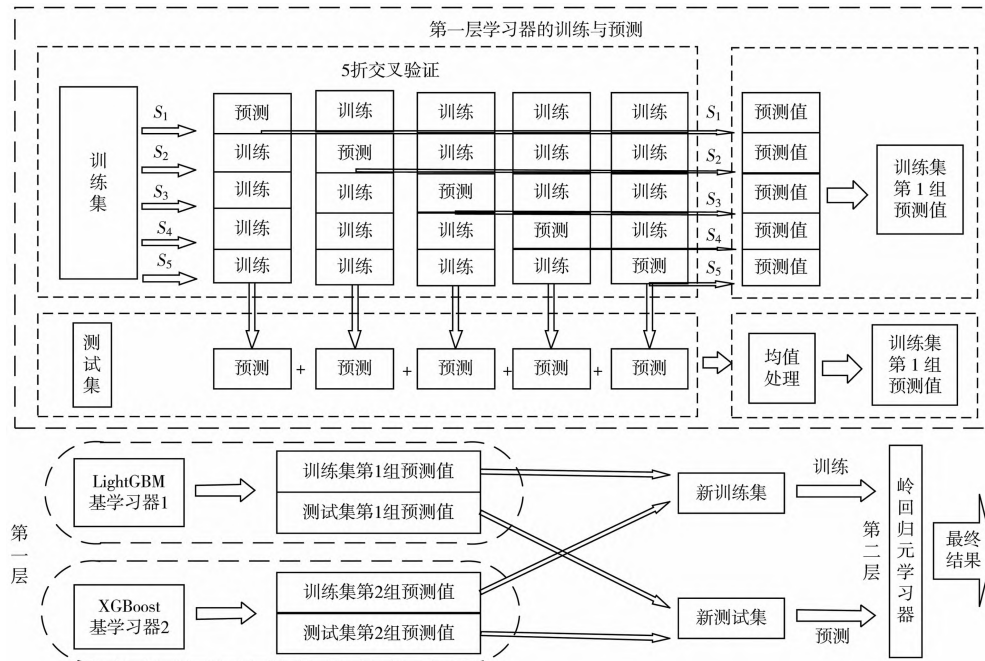


图4 Stacking集成模型框架

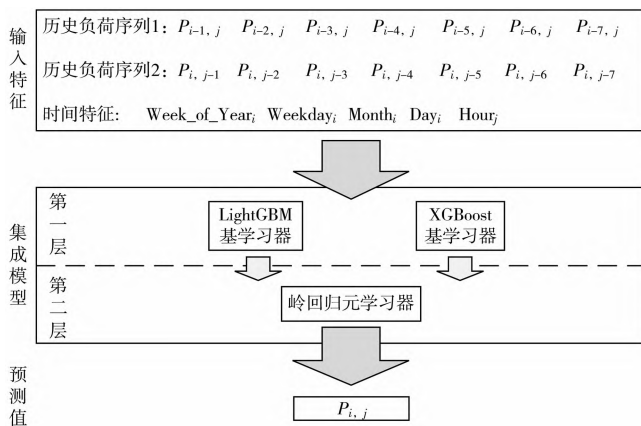


图5 基于集成模型的充电负荷预测

式中, N 为测试样本数量, \tilde{y} 为测试样本真实值序列, \tilde{y}_m 和 y_m 表示第 m 个测试样本的真实值和预测值。

4.2 模型的预测效果对比分析

为验证基于集成 XGBoost 和 LightGBM 的负荷预测模型的测试效果的有效性, 将实验数据根据按照 2:1 划分为训练集和测试集, 并对比基于随机森林(Random Forest, RF)、K 近邻(K-Nearest Neighbor, KNN)、岭回归(Ridge Regression, RR)、神经网络算法的负荷预测值, 模型参数采用随机搜索的方法进行调整, 各个模型的测试结果如表 1 所示。

以上基于单一算法的 6 种负荷预测模型中, XGBoost 和 LightGBM 模型的均方根误差(RMSE)都低于 0.337, 绝对百分误差(MAPE)也下降到了 5.84%, 相比其他机器学习回归算法以及神经网络模型的预测准确度高, 体现了 XGBoost 和 LightGBM 作为单一算法模型的优越性。

表 1 单一预测模型的测试效果对比

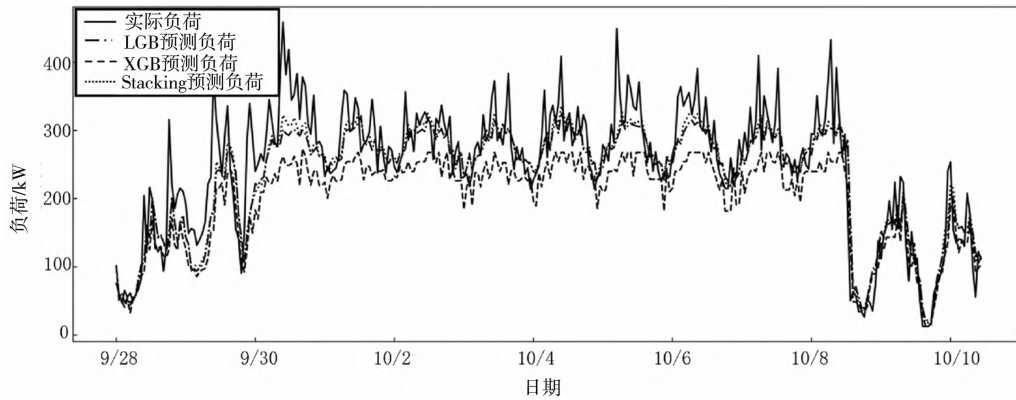
模型	评价指标	
	RMSE	MAPE/%
XGBoost	0.336 3	5.841 0
LightGBM	0.335 0	5.824 3
RF	0.339 5	5.982 0
KNN	0.354 4	6.117 8
RR	0.351 4	6.156 1
神经网络	0.341 2	6.031 0

由表 1 可知, 前 6 种模型中效果最优的是 LightGBM 模型, 采用 Stacking 集成策略将 XGBoost 和 LightGBM 模型融合之后的模型与基学习器的预测效果对比如表 2 所示, Stacking 模型的 RMSE 值相比 LightGBM 下降了约 3.28%, 其 MAPE 值相比 LightGBM 降低 0.0442%, 证明 Stacking 模型可有效提高预测准确度。

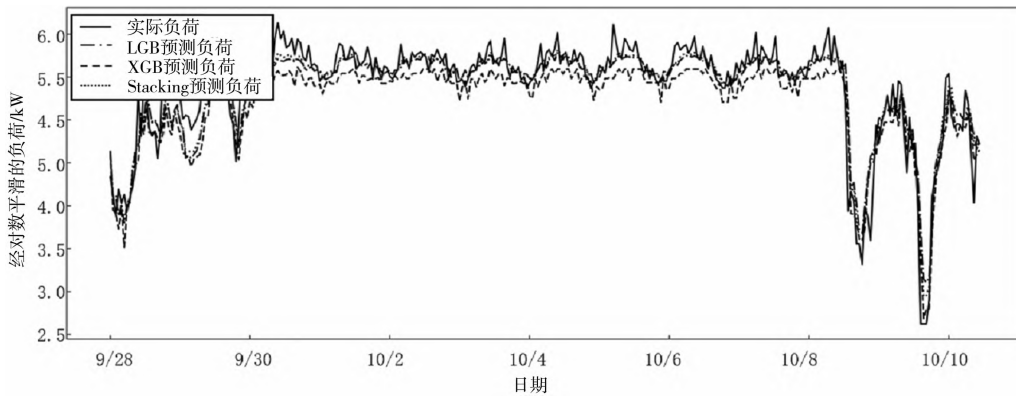
表 2 Stacking 模型与基学习器的测试效果对比

模型	评价指标	
	RMSE	MAPE/%
XGBoost	0.336 3	5.841 0
LightGBM	0.335 0	5.824 3
Stacking 集成模型	0.324 0	5.780 1

Stacking 模型与基学习器的预测曲线如图 6(a)和(b)所示, 即便是国庆期间的负荷预测, 抛开噪点的影响下, XGBoost、LightGBM 模型以及 Stacking 集成模型都较好地拟合了数据, 凸显了模型的泛化性能, 尤其是 Stacking 模型很好地学习了负荷的变化规律, 预测值比任何一个



(a)未平滑数据结果



(b)平滑数据结果

图 6 Stacking 模型与基学习器的预测结果对比

基学习器更接近真实值,验证了 Stacking 模型的灵活性及泛化性。

5 结论

提高充电负荷预测准确度对保障电网平稳运行具有重要意义,所以本文基于 XGBoost 和 LightGBM 算法,构建了 Stacking 融合预测模型。结合上海市嘉定区的电动汽车充电站的订单数据,采用历史负荷序列和时间特征作为模型输入,来预测未来短期内充电站的负荷值,实验结果如下:

(1)单一模型算法中,LightGBM 和 XGBoost 模型预测的均方根误差为 0.337,绝对百分误差为 5.84%,相比 RF、KNN、RR 以及神经网络算法,负荷预测准确度更高。

(2)Stacking 集成模型考虑了不同基学习器预测效果的差异性,与单一模型相比,获得了更好的泛化能力,准确度高于 LightGBM 和 XGBoost 基学习器。

上述结果表明,本文所提集成 XGBoost 和 LightGBM 算法的短期负荷预测模型能够较精确预测未来时刻的充电负荷,有效地证明了本文所提负荷预测方法的有效性。该负荷预测模型利于电网对充电站进行合理调度和管理,对于保障电网平稳运行具有一定的理论及实践价值。由于本文并未对 Stacking 集成模型的设计结构做深入研究,后期可通过优化 Stacking 模型的层级结构设置、调

整基学习器的参数等方式提高预测的准确度。

参考文献

- [1] 高赐威,张亮.电动汽车充电对电网影响的综述[J].电网技术,2011,35(2):127-131.
- [2] 李慧玲,白晓民.电动汽车充电对配电网的影响及对策[J].电力系统自动化,2011,35(17):38-43.
- [3] 胡泽春,宋永华,徐智威,等.电动汽车接入电网的影响和利用[J].中国电机工程学报,2012,32(4):1-10.
- [4] 王锡凡,邵成成,王秀丽,等.电动汽车充电负荷与调度控制策略综述[J].中国电机工程学报,2013,33(1):1-10.
- [5] FERNANDEZ L P, ROMAN T G S, OSSENT R, et al. Assessment of the impact of plug-in electric vehicles on distribution networks[J]. IEEE Transactions on Power System, 2011, 26(1): 206-213.
- [6] SOARES F J, LOPES J A P, ALMEIDA P M R, et al. A monte carol method to evaluate electric vehicles impacts in distribution networks[C]//IEEE Conference on Innovation Technologies for Efficient and Reliable Electricity Supply(CITRES). Waltham, 2010: 365-372.
- [7] KAZEROONI M, KAR N C. Impact analysis of EV battery charging on the power system distribution transformers[C]// Electric Vehicle Conference(IEVC). Greenville, IEEE International, 2012: 1-6.

- [8] AL-MUSAWI L, TRAN R. The impact of EV/PHEV chargers on residential loads—A case study[C]//Transportation Electrification Conference and Expo. Detroit, 2013: 1–4.
- [9] 王增平, 赵兵, 贾欣, 等. 基于差分分解和误差补偿的短期电力负荷预测方法[J]. 电网技术, 2021, 45(7): 2560–2568.
- [10] 朱慧婷, 杨雪, 陈友媛. 电动汽车充电负荷预测方法综述[J]. 电力信息与通信技术, 2016, 14(5): 44–47.
- [11] 陈静鹏, 朴龙健, 艾芊. 基于改进贪心算法的大规模电动汽车充电行为优化[J]. 电力自动化设备, 2016, 36(10): 38–44.
- [12] 胡宇航, 皮一晨, 崔静安, 等. 电动汽车充电站负荷建模研究[J]. 电力系统保护与控制, 2017, 45(8): 107–112.
- [13] 陈静鹏, 艾芊, 肖斐. 基于用户出行需求的电动汽车充电站规划[J]. 电力自动化设备, 2016, 36(6): 34–39.
- [14] GHASNEZHAD N, FILIZADEH S. Location-based forecasting of vehicular charging load on the distribution system[C]//2014 IEEE PES General Meeting/Conference & Exposition. National Harbor, MD, USA: IEEE, 2014: 1.
- [15] 王睿, 高欢, 李军良, 等. 基于聚类分析的电动汽车充电负荷预测方法[J]. 电力系统保护与控制, 2020, 48(16): 37–44.
- [16] AZHAR U H, CARLO C, SAADANYE. Probabilistic modeling of electric vehicle charging pattern in a residential distribution network[J]. Electric Power Systems Research, 2018, 10: 126–133.
- [17] ARIAS M B, BAE S. Electric vehicle charging demand forecasting model based on big data technologies[J]. Applied Energy, 2016, 183: 327–339.
- [18] 陈丽丹, 聂涌泉, 钟庆. 基于出行链的电动汽车充电负荷预测模型[J]. 电工技术学报, 2015, 30(4): 216–225.
- [19] 张维戈, 颀飞翔, 黄梅, 等. 快换式公交充电站短期负荷预测方法的研究[J]. 电力系统保护与控制, 2013(4): 61–66.
- [20] 常德政, 任杰, 赵建伟, 等. 基于 RBF-NN 的电动汽车充电站短期负荷预测研究[J]. 青岛大学学报(工程技术版), 2014, 29(4): 44–48.
- [21] FRIEDMAN J H. Greedy function approximation: a gradient boosting Machine[J]. Annals of Statistics, 2001, 29(5): 1189–1232.
- [22] CHEN T, GUESTRIN C. XGBoost: a scalable tree boosting system[C]//The 22nd ACM SIGKDD International Conference, 2016.
- [23] 颜诗旋, 朱平, 刘钊. 基于改进 LightGBM 模型的汽车故障预测方法研究[J]. 汽车工程, 2020, 42(6): 815–819, 825.
- [24] 魏书荣, 张鑫, 符杨, 等. 基于 GRA-LSTM-Stacking 模型的海上双馈风力发电机早期故障预警与诊断[J]. 中国电机工程学报, 2021, 41(7): 2373–2383.
- [25] 刘波, 秦川, 鞠平, 等. 基于 XGBoost 与 Stacking 模型融合的短期母线负荷预测[J]. 电力自动化设备, 2020, 40(3): 147–153.

(收稿日期: 2021-11-10)

作者简介:

吴丹(1981-), 男, 硕士, 主要研究方向: 电力系统自动化、电力营销。

雷珽(1986-), 男, 硕士, 主要研究方向: 电力系统自动化。

李芝娟(1991-), 女, 硕士研究生, 主要研究方向: 电力系统自动化。



扫码下载电子文档

(上接第 43 页)

- [32] CARROLL R, GEER R. Fault localization in a 3D test assembly using voltage contrast imaging[J]. Microelectronics Journal, 2019, 87: 73–80.
- [33] FU W, CHIEN C F, TANG L. Bayesian network for integrated circuit testing probe card fault diagnosis and troubleshooting to empower Industry 3.5 smart production and an empirical study[J]. Journal of Intelligent Manufacturing, 2022, 33: 785–798.
- [34] FAN W, HE J, HAN Z, et al. Reconfigurable fault-tolerance mapping of ternary N-cubes onto chips[J]. Concurrency and Computation: Practice and Experience, 2020, 32(11): e5659.
- [35] CHANG Y C, GONG C S A, CHIU C T. Fault-tolerant mesh-based NOC with router-level redundancy[J]. Journal

of Signal Processing Systems, 2020, 92(4): 345–355.

(收稿日期: 2022-03-07)

作者简介:

周永忠(1981-), 男, 硕士, 工程师, 主要研究方向: 处理器功能安全系统方案设计与验证、处理器芯片的故障注入、故障测试技术以及自我修复方法等。

洪晟(1981-), 通信作者, 男, 副教授, 主要研究方向: 信息网络安全、集成电路安全性、复杂系统安全性, 软件安全, E-mail: shenghong@buaa.edu.cn。

姜义初(1979-), 男, 硕士, 验证工程师, 主要研究方向: CPU 核的功能安全验证, 主要是 CPU 安全验证方案的制订、验证环境的搭建以及测试点的分解等。



扫码下载电子文档