



基于二次聚类的充电桩执行电价异常检测方法

薛晓慧¹, 张文², 张静², 陈雁², 周春², 陈亮², 曹增伟²

(1. 国网青海省电力公司, 青海 西宁 810008;

2. 北京中电普华信息技术有限公司, 北京 100192)

摘要: 由于电价政策复杂, 执行环节多, 监管难度大, 电价执行错误现象时有发生, 这不仅损害电力市场的公平性和效率, 也影响电力企业的经济效益和用户的用电成本。提出了一种基于二次聚类的充电桩执行电价异常检测方法, 首先进行电价执行异常分类及用电特征分析, 其次通过 K -means 聚类算法剥离出电瓶车用户, 进而在第二次聚类中采用含噪声应用的基于密度的空间聚类 (density-based spatial clustering of applications with noise, DBSCAN) 算法精确识别高价低接等更为复杂的违约情况。所提方法通过两次聚类分析, 提高电价执行的准确性和效率, 具有一定的理论意义和应用价值。

关键词: 充电桩; 执行电价异常; 聚类分析; 离群点检测; 高价低接

中图分类号: TM76

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2025004

A method for detecting abnormal electricity prices at charging stations based on two rounds of clustering

XUE Xiaohui¹, ZHANG Wen², ZHANG Jing², CHEN Yan², ZHOU Chun², CHEN Liang²,
CAO Zengwei²

1. State Grid Qinghai Electric Power Company, Xining 810008, China

2. Beijing China-Power Information Technology Co., Ltd., Beijing 100192, China

Abstract: Due to the complexity of electricity pricing policies, multiple implementation steps, and difficulty in regulation, errors in electricity pricing implementation occur from time to time. This not only damages the fairness and efficiency of the electricity market but also affects the economic benefits of power enterprises and the electricity costs of users. A method was proposed for detecting electricity price anomalies at charging stations based on secondary clustering. Firstly, the electricity price anomalies were classified and the electricity consumption characteristics were analyzed. Secondly, the K -means clustering algorithm was used to extract electric vehicle users. Then, in the second clustering, the density-based spatial clustering of applications with noise (DBSCAN) algorithm was used to accurately

收稿日期: 2024-10-16; 修回日期: 2024-12-09

通信作者: 张静, jingzh1036@163.com

基金项目: 国家自然科学基金资助项目 (No.72071070)

Foundation Item: The National Natural Science Foundation of China (No.72071070)

identify more complex default situations, such as high price and low connection. The proposed method improves the accuracy and efficiency of electricity price implementation through two rounds of cluster analysis, and has certain theoretical significance and application value.

Key words: charge pile, abnormal execution of electricity prices, cluster analysis, outlier detection, high price and low acceptance

0 引言

随着全球能源结构的转型和电动汽车产业的迅猛发展,充电桩作为电动汽车充电的重要基础设施,其建设和运营已成为电力市场的重要组成部分^[1-2]。然而,电价政策的复杂性、执行环节的多样性以及监管难度的增加,使得电价执行错误的现象在充电桩运营中屡见不鲜。这些问题不仅破坏了电力市场的公平竞争环境,降低了市场效率,还直接影响了电力企业的经济效益和广大用户的用电成本^[3-4]。因此,如何有效检测并纠正充电桩执行电价中的异常现象,成为当前电力市场管理和技术研究中亟待解决的问题。

目前,充电桩执行电价异常检测领域的专门研究相对匮乏,尽管在电力市场异常检测与用电行为分析等领域已有广泛探索,但针对充电桩电价异常的直接研究尚显不足^[5-6]。

现有的电力市场异常检测方法主要可划分为两大类:一类是基于统计学的异常识别技术,这类方法通过预设阈值或利用数据的统计分布特性来识别异常,然而其高度依赖数据的先验假设,难以灵活应对电价执行中复杂多变的实际情况^[7-9];另一类则是基于机器学习的智能检测算法,包括支持向量机、随机森林等^[10-11],这些算法能够自动学习数据特征并识别异常,但在处理充电桩产生的大规模、高维度数据时,常遇到计算负担重、模型复杂度上升及过拟合等问题^[12-13]。

此外,尽管深度学习技术在处理复杂数据方面展现出卓越能力,但在充电桩电价执行异常检

测的具体应用中仍属少见,且面临着数据标注成本高、模型解释性不足等挑战^[14]。

鉴于上述技术存在的局限性,本文提出一种更高效且准确的充电桩执行电价异常检测方法,主要围绕以下方面展开研究。首先,本文深入剖析充电桩电价执行异常的根源与表现形式,并综合考虑电价政策、用户行为模式及时间因素等多维度特征;其次,针对现有技术的不足,设计并实施一种优化的聚类分析方法,通过采用精细的聚类策略与特征选择策略,提升电价异常检测的精度与效率;最后,通过实验验证所提方法的有效性,并评估其在实际应用中的可行性与推广潜力。本文的研究不仅丰富了电价异常检测领域的研究成果,也为电动汽车产业的可持续发展贡献了新的技术思路。

1 电价执行异常分类及用电特征分析

某地区电力公司基层人员在定期对电动汽车充电桩用户进行现场稽查时,发现主要存在两种执行电价错误的情况。一是电力公司工作人员在办理业务中电价知识欠缺,导致大量小区物业车棚中的电瓶车充电桩或者个人电瓶车充电桩执行错误电价,应将其更改为享受居民类电价。如果是营利性电车充电桩,则应执行商业电价。二是客户违约用电的原因。有的用户私自在充电桩上接入其他家用电器,有的用户现场接入自家超市作为超市商业用电,有的用户现场接入家中作为居民家庭用电,有的用户私自拉线,将电表另外外接两处,一条接回家中使用,一条接入仓库使用。这些高价低接的违约行为应及时整改。因此,应针对电瓶



车和违约用电两种异常行为分别进行用电特征分析,以区分充电桩正常用电。

1.1 电瓶车用电特征分析

不同用电模式的日用电量时间序列如图1所示。从图1中可以看出,电瓶车和电动汽车用电均表现为间断用电特点,即充电时用电量非零,充电桩空闲时用电量为零或值很小。空闲状态不能反映用户充电特性,因此,去除用电序列中零用电量和低用电量,对余下电量求平均值,该电量值反映了每次充电状态的用电量水平,记为充电用电量,计算充电谷电量与充电用电量比值,记为充电谷电量占比。

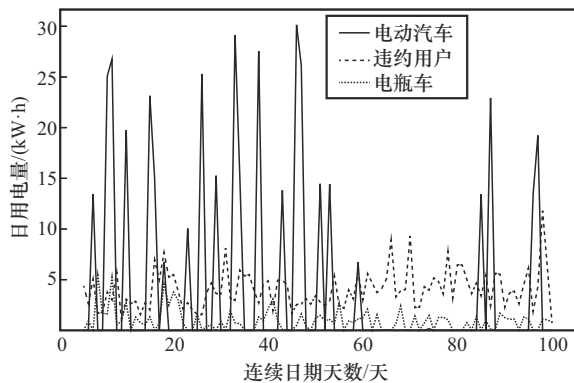
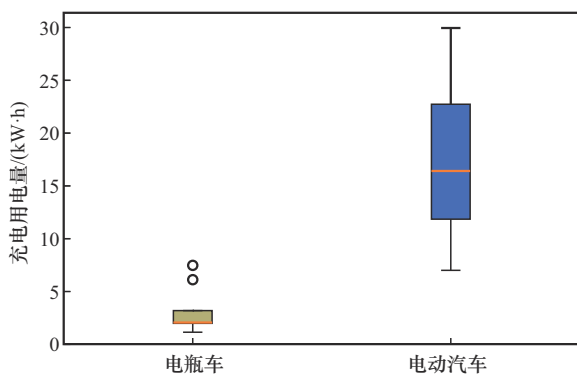


图1 不同用电模式的日用电量时间序列

电瓶车与电动汽车用电特征对比如图2所示。从图2可以看出,电动汽车用电量显著高于电瓶车,且重叠交叉区域较少。同时,电动汽车充电谷电量占比高。因此,电瓶车与电动汽车在充电用电



(a) 电瓶车和电动汽车充电用电量对比

量、充电谷电量占比两个用电特征上表现出显著区别。

1.2 电价违约用户用电特征分析

从图1可以看出,与电动汽车间断用电特点不同,商业用电或者居民用电均表现为连续用电,用户的高价低接违约行为一般用于商业用电或者电器等居民用电,因此表现为连续用电模式。

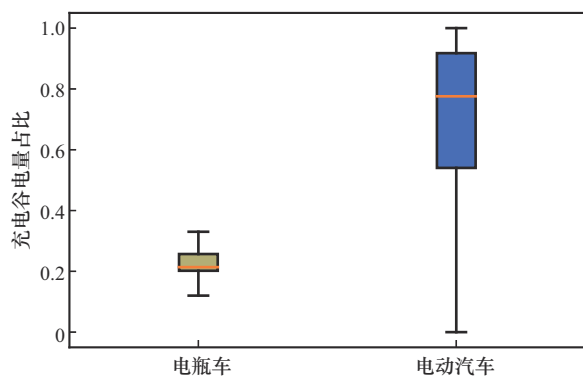
违约用户和电动汽车用电特征对比如图3所示。从图3中的日用电量、变异系数、谷电量占比的箱线图分布可以看出,电动汽车用电量显著高于违约用户,电动汽车的变异系数大,说明用电随机波动较大,违约用户变异系数较小,用电稳定。同时,电动汽车谷电量占比高。因此,违约用户与电动汽车在日用电量、变异系数、谷电量占比3个用电特征上表现出显著区别。

2 基于K-means的电瓶车用户识别

本文计算用户的充电用电量、充电谷电量占比,对这两个特征向量进行K-means聚类分析,识别出电瓶车用户群体,将其从总用户群体中剥离出来。

2.1 K-means 算法

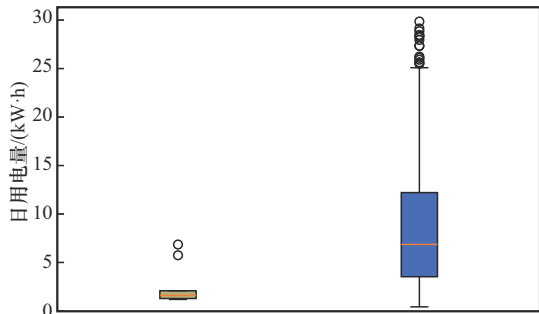
K-means 算法是一种被广泛使用的聚类算法,其核心思想是将数据集中的样本点划分为 K 个簇,使得每个簇内的样本点尽可能相似,而不同簇之间的样本点尽可能不同。这个算法通过迭



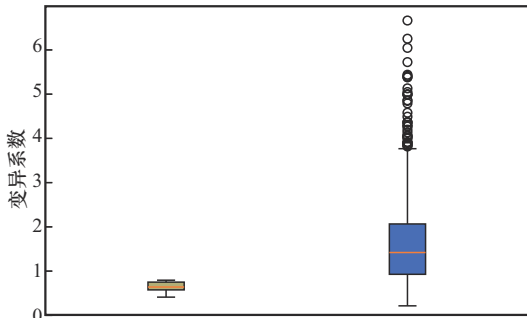
(b) 电瓶车和电动汽车充电谷电量占比对比

图2 电瓶车与电动汽车用电特征对比

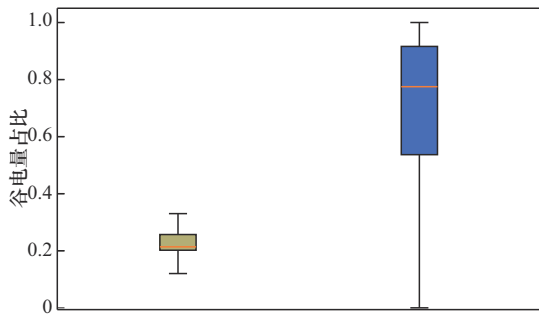
代的方式寻找簇的中心（即簇内所有点的均值），并据此将数据点分配到最近的簇中^[15]。



(a) 违约用户和电动汽车用电量对比



(b) 违约用户和电动汽车变异系数对比



(c) 违约用户和电动汽车谷电量占比对比

图3 违约用户与电动汽车用电特征对比

K-means 算法的主要步骤包括初始化、分配簇、更新质心和迭代优化，具体步骤如下。

步骤 1 初始化：随机选择 K 个数据点作为初始的簇质心。为了改进初始质心的选择，可以使用 K-means++ 算法等策略。

步骤 2 分配簇：对于数据集中的每个点，计算它到所有簇质心的距离。将每个点分配到距

离最近的簇质心所对应的簇中：

$$c_i = \arg \min_j \left(\sqrt{\sum_{m=1}^M (x_{im} - \mu_{jm})^2} \right) \quad (1)$$

其中， c_i 表示点 x_i 所属的簇的索引， j 表示遍历所有簇的索引， M 表示数据点的维度， x_{im} 表示点 x_i 在第 m 维上的值， μ_{jm} 表示簇 j 的质心在第 m 维上的值。

步骤 3 更新质心：对于每个簇，重新计算其质心。新的质心是簇内所有点的均值：

$$\mu_j = \frac{1}{|C_j|} \sum_{x_i \in C_j} x_i \quad (2)$$

其中， μ_j 表示簇 j 的新质心， C_j 表示簇 j 内所有点的集合， $|C_j|$ 表示簇 j 内点的数量。

步骤 4 迭代优化：重复步骤 2 和步骤 3，直到满足某个停止条件。停止条件可以是质心的变化小于某个阈值，或者达到预设的迭代次数。

K-means 算法的优化目标是最小化所有数据点到其所属簇质心的距离平方和，也称为聚类内的总平方误差。优化目标函数可以表示为：

$$J = \sum_{j=1}^K \sum_{x_i \in C_j} |x_i - \mu_j|^2 \quad (3)$$

其中， K 表示簇的数量， C_j 表示簇 j 内所有点的集合， μ_j 表示簇 j 的质心， x_i 表示簇 j 内的点。

2.2 聚类结果分析

图 4 展示了某地区执行充电桩电价用户的 K-means 聚类结果，可以看出，类别 1 的用户用电量小、谷电量占比低，符合电瓶车用电特征，因此将此用户识别为电瓶车用户。

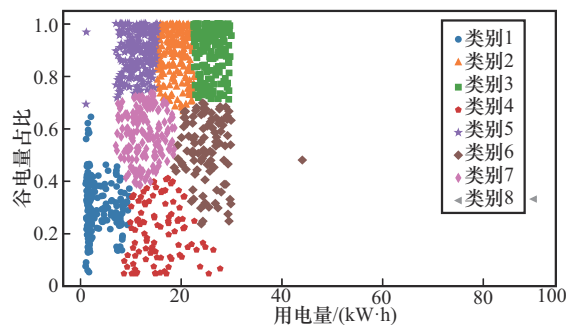


图4 某地区执行充电桩电价用户的 K-means 聚类结果



3 基于DBSCAN算法的违约用户识别

本文通过 K -means 算法进行的电瓶车用户识别, 将电瓶车用户剥离出去后, 针对非电瓶车用户进行基于含噪声应用的基于密度的空间聚类 (density-based spatial clustering of applications with noise, DBSCAN) 算法的违约用户识别。

3.1 DBSCAN 算法

基于 DBSCAN 算法的离群点检测方法可用于从数据集中识别出与大多数数据点显著不同的点^[15]。DBSCAN 算法是一种基于密度的聚类算法, 它不需要事先指定聚类簇的数量, 就能够发现任意形状的簇, 并且能够有效识别出噪声点 (即离群点), 其核心原理是通过识别数据中的高密度区域来形成聚类, 并将低密度区域或孤立点视为噪声。

DBSCAN 算法通过两个主要参数来定义密度: 邻域半径 Eps 和最小样本数 $MinPts$ 。对于数据集中的每一个点, 算法会检查以该点为圆心、 Eps 为半径的圆内是否包含至少 $MinPts$ 个点。如果是, 则该点被认为是核心点, 并有可能扩展成一个簇; 如果不是, 则该点被标记为噪声点 (离群点)。

(1) 邻域半径 Eps : 以某一点为中心, Eps 为半径的圆内所有的数据点称为该点的邻域。邻域半径的选取通常依赖数据的具体分布, 可以通过绘制 k -距离曲线 (其中 k 建议取 $2 \times \text{维度} - 1$) 来找到拐点, 拐点处的距离即为 Eps 的值。

(2) 最小样本数 $MinPts$: 在邻域半径为 Eps 内的最小数据点数。 $MinPts$ 的取值为 $k+1$, 其中 k 为 k -距离图中选取的 k 值。

算法执行步骤如下。

步骤 1 从数据集中选取一个未访问的点 p 。

步骤 2 检查点 p 的邻域内是否有至少 $MinPts$ 个点。如果是, 则创建一个新的簇, 并将 p 的邻域半径内的所有点添加到该簇中。如果不是, 则将 p 标记为噪声点 (离群点)。

步骤 3 对于簇中的每一个点, 递归地检查其邻域半径内的点, 如果满足条件, 则将其添加到簇中。

步骤 4 重复上述过程, 直到所有点都被访问。

步骤 5 聚类完成后, 所有被标记为噪声点的数据点即离群点。

3.2 参数选择

需要注意的是, DBSCAN 算法的效果很大程度上取决于邻域半径 ϵ 和 $MinPts$ 的选择。通常根据 k -距离曲线选择出合适的邻域半径 ϵ , 计算每个样本与所有样本的距离, 选择第 k 个最近邻的距离, 将每个样本的 k 近邻距离从大到小排序, 得到 k -距离曲线, 在 k -距离曲线中明显拐点位置所对应的纵坐标为较好的半径参数值。根据经验, 一般选择 $MinPts = \text{dim} + 1$, 其中 dim 为数据集的维数。

在针对非电瓶车用户进行基于 DBSCAN 算法的违约用户识别时, 构建特征向量 (日电量均值、变异系数、谷电量占比), 设 $\text{dim} = 3$, $MinPts = 4$ 。同时, 根据 k -距离曲线 (DBSCAN 算法的 k -距离曲线如图 5 所示), 拐点位置的纵坐标在 0.4 左右, 因此选择邻域半径 ϵ 为 0.4。

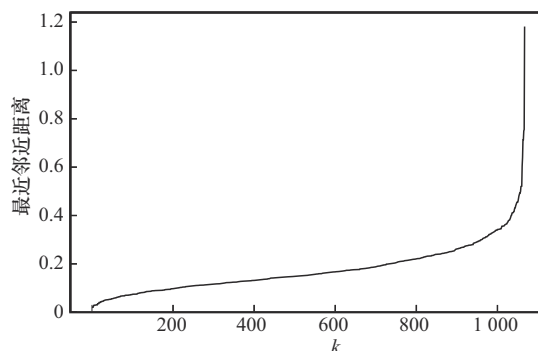


图5 DBSCAN算法的 k -距离曲线

DBSCAN 算法识别结果如图 6 所示。由电价违约用户用电特征分析可知, 违约用户变异系数较小, 因此选择图 6 中变异系数小于 1 的噪声点为目标违约用户, 同时类别 9 偏离大部分群体, 符合违约用户特征, 因此也将其选定为目标违约用户。

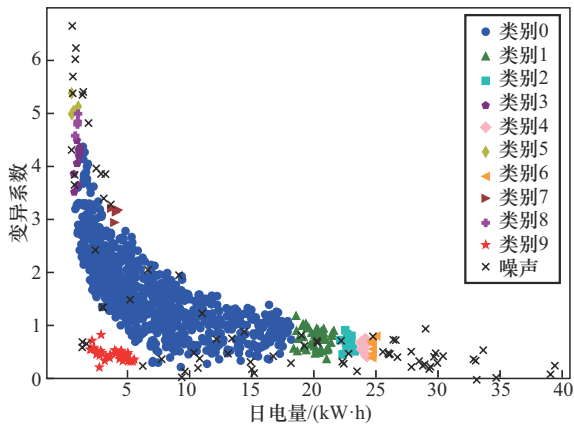


图6 DBSCAN算法识别结果

4 实例分析

4.1 数据来源

本文数据取自某地区供电公司营销系统及稽查系统，系统中该地区有1237个用户执行充电桩电价，抽取这些用户2021年12月至2022年3月期间每日的日用电量、峰用电量、平用电量、谷用电量时间序列数据，进行实验分析。

4.2 实验结果

通过K-means算法识别出电瓶车用户188户，经验证，其中159户确实为电瓶车用户，算法识别准确率达到85%。在第二次聚类中，采用DBSCAN算法识别出电价违约用户90户，经验证其中55户确实有异常违约情况，算法识别准确率达61%，大大提高了稽查效率和准确率。

5 结束语

本文提出了一种基于二次聚类的充电桩执行电价异常检测方法。该方法旨在通过两次聚类分析，结合不同算法的优势，实现对电价执行异常的精准识别。所提方法通过电价执行异常分类及用电特征分析，明确电价执行过程中可能存在的错误类型及其特征表现，为后续聚类分析提供基础。随后，该方法利用K-means聚类算法，根据电瓶车用户的用电特征，将其从总用户群体中初步剥离出来。这一步骤有效降

低了后续分析的复杂度，提高了检测效率。

在第二次聚类中，本文采用DBSCAN算法，针对非电瓶车用户群体进行更为精细的聚类分析。DBSCAN算法基于密度的聚类思想，能够自动发现任意形状的簇，并有效识别出噪声点，即那些不符合大多数用户用电行为特征的异常用户。通过这一步骤，本文能够精确识别出高价低接等更为复杂的违约情况，进一步提高电价执行异常的检测精度。本文提出的基于二次聚类的充电桩执行电价异常检测方法，不仅具有理论上的创新性和前瞻性，还具备较高的应用价值。

参考文献：

- [1] 赵云龙, 孔庚, 李卓然, 等. 全球能源转型及我国能源革命战略系统分析[J]. 中国工程科学, 2021, 23(1): 15-23.
ZHAO Y L, KONG G, LI Z R, et al. Strategic analysis of global energy transition and China's energy revolution[J]. Strategic Study of CAE, 2021, 23(1): 15-23.
- [2] 黄学良, 刘永东, 沈斐, 等. 电动汽车与电网互动: 综述与展望[J]. 电力系统自动化, 2024, 48(7): 3-23.
HUANG X L, LIU Y D, SHEN F, et al. Vehicle to grid: review and prospect[J]. Automation of Electric Power Systems, 2024, 48(7): 3-23.
- [3] 杨镜司, 秦文萍, 史文龙, 等. 基于电动汽车参与调峰定价策略的区域电网两阶段优化调度[J]. 电工技术学报, 2022, 37(1): 58-71.
YANG J S, QIN W P, SHI W L, et al. Two-stage optimal dispatching of regional power grid based on electric vehicles' participation in peak-shaving pricing strategy[J]. Transactions of China Electrotechnical Society, 2022, 37(1): 58-71.
- [4] 李锦辉, 吴毓峰, 余涛, 等. 数据孤岛下基于联邦学习的用户电价响应刻画及其应用[J]. 电力系统保护与控制, 2024, 52(6): 164-176.
LI J H, WU Y F, YU T, et al. Characterization of user price response behavior and its application based on federated learning considering a data island[J]. Power System Protection and Control, 2024, 52(6): 164-176.
- [5] 朱峰, 单超, 吴宁, 等. 电力市场环境下的电力用户电价特征提取和异常识别方法[J]. 上海交通大学学报, 2024: 1-23.
ZHU F, SHAN C, WU N, et al. Electricity price feature extrac-



- tion and abnormal identification method for power users in the electricity market environment[J]. Journal of Shanghai Jiao Tong University, 2024: 1-23.
- [6] 谢敬东, 卢浩哲, 陆池鑫, 等. 基于分阶段离群点检测的电力市场异常辨识[J]. 科学技术与工程, 2021, 21(9): 3633-3641.
XIE J D, LU H Z, LU C X, et al. Identification of abnormal behavior in power market based on phased outlier detection[J]. Science Technology and Engineering, 2021, 21(9): 3633-3641.
- [7] 王威, 王兰君. 基于用电行为特征大数据的异常用户识别模型研究与应用[J]. 电力大数据, 2021, 24(12): 19-26.
WANG W, WANG L J. Research and application of abnormal user identification model based on big data of electricity behavior characteristics[J]. Power Systems and Big Data, 2021, 24(12): 19-26.
- [8] 刘宣, 唐悦, 卢继哲, 等. 基于概率预测的用电采集终端电量异常在线实时识别方法[J]. 电力系统保护与控制, 2021, 49(19): 99-106.
LIU X, TANG Y, LU J Z, et al. Online real time anomaly recognition method for power consumption of electric energy data acquisition terminal based on probability prediction[J]. Power System Protection and Control, 2021, 49(19): 99-106.
- [9] 张小斐, 耿俊成, 孙玉宝. 图正则非线性岭回归模型的异常用电行为识别[J]. 计算机工程, 2018, 44(6): 8-12.
ZHANG X F, GENG J C, SUN Y B. Abnormal electricity behavior recognition of graph regularization nonlinear ridge regression model[J]. Computer Engineering, 2018, 44(6): 8-12.
- [10] 钱旭盛, 朱萌, 翟千惠, 等. 基于改进孤立森林算法的异常用电行为识别方法[J]. 沈阳工业大学学报, 2023, 45(6): 601-606.
QIAN X S, ZHU M, ZHAI Q H, et al. Abnormal electrical behavior recognition method based on improved isolated forest algorithm[J]. Journal of Shenyang University of Technology, 2023, 45(6): 601-606.
- [11] 万伟, 刘红旗, 孙洪昌, 等. 用电异常行为预警方法[J]. 哈尔滨理工大学学报, 2022, 27(4): 53-62.
WAN W, LIU H Q, SUN H C, et al. Early warning method of abnormal electricity consumption behavior based on data driven[J]. Journal of Harbin University of Science and Technology, 2022, 27(4): 53-62.
- [12] 彭显刚, 林利祥, 刘艺, 等. 数据挖掘技术在电价执行稽查中的应用研究[J]. 电气应用, 2016, 35(11): 62-67.
PENG X G, LIN L X, LIU Y, et al. Research on the application of data mining technology in electricity price execution inspection[J]. Electrotechnical Application, 2016, 35(11): 62-67.
- [13] 路洁, 郑小贤, 杨辉, 等. 基于时序数据挖掘的电价执行异常分析模型设计[J]. 自动化技术与应用, 2022, 41(5): 171-174.
LU J, ZHENG X X, YANG H, et al. Design of electricity price execution anomaly analysis model based on time series data mining[J]. Techniques of Automation and Applications, 2022, 41(5): 171-174.
- [14] 万磊, 陈成, 黄文杰, 等. 基于BRB和LSTM网络的电力大数据用电异常检测方法[J]. 电力建设, 2021, 42(8): 38-45.
WAN L, CHEN C, HUANG W J, et al. Power abnormality detection method based on power big data applying BRB and LSTM network[J]. Electric Power Construction, 2021, 42(8): 38-45.
- [15] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.
ZHOU Z H. Machine learning [M]. Beijing: Tsinghua University Press, 2016.

[作者简介]



薛晓慧 (1971-), 女, 国网青海省电力公司高级工程师, 主要从事电网信息化应用、用电信息采集等工作。



张文 (1973-), 男, 北京中电普华信息技术有限公司高级工程师, 主要从事电力市场营销、电力信息化等工作。

张静 (1987-), 女, 北京中电普华信息技术有限公司高级工程师, 主要从事电力营销数据分析、机器学习等工作。

陈雁 (1981-), 女, 北京中电普华信息技术有限公司高级工程师, 主要从事电网大数据、人工智能应用等工作。

周春 (1983-), 男, 北京中电普华信息技术有限公司高级工程师, 主要从事电力工程信息化、人工智能应用等工作。

陈亮 (1979-), 男, 北京中电普华信息技术有限公司高级项目经理, 主要从事电力工程信息化、项目管理等工作。

曹增伟 (1984-), 男, 北京中电普华信息技术有限公司工程师, 主要从事电力市场营销、人工智能应用等工作。