

# 基于大数据人工智能的充电安全评估技术

鞠强, 尹玉鹏, 项宝庆, 朱诗严, 潘博存, 郭鹏

(特来电新能源股份有限公司, 山东 青岛 266104)

**摘要:** 针对电池指标安全阈值是静态的实验室数据, 导致难以准确识别未达固定阈值而实际存在异常的车辆问题, 提出了基于同区域同车型同时段对比的分布异常检测方法、基于电池产热机理和机器学习融合进行温升异常识别的方法、基于数据驱动的高斯过程有监督异常检测方法。基于同区域同车型同时段对比的分布异常检测方法, 首先利用统计学检验方法, 对电池指标数据进行正态分布检验, 然后将非正态分布的指标数据通过数学方法转换为正态分布, 再利用统计学小概率事件原则, 并结合专家经验, 确定异常检测的阈值, 建立异常检测模型; 基于电池产热机理和机器学习融合进行温升异常识别的方法, 首先利用电池产热的物理原理计算电池充电过程中的放热, 然后利用鲁棒回归算法, 训练充电过程中电流与温升变化的拟合模型, 最后结合 boxplot 异常检测算法对同区域同车型同时段订单拟合模型的  $k$  进行异常检测, 识别相同条件下温升速率异常充电电池; 基于数据驱动的高斯过程有监督异常检测方法, 利用已有的异常电池样本, 首先通过 LightGBM 技术选择特征, 然后建立高斯过程模型并确定分类阈值, 用于异常检测。验证结果表明, 通过同区域同车型同时段电池的对比, 分布异常检测提供了指标阈值动态精确调整的解决方案; 温升异常检测方案建立了复杂指标, 并基于该指标有效地实现了异常检测; 通过高斯过程有监督模型, 充分挖掘异常电池的数据特征规律, 进一步提高了异常检测的精度。基于这三个技术, 形成了从无监督技术到有监督技术、电池产热机理与人工智能技术相结合的、全面有效的综合解决方案。

**关键词:** 充电安全; 异常检测; 正态分布; 产热机理; 高斯过程

## Charging safety assessment technology based on big data and AI

JU Qiang, YIN Yupeng, XIANG Baoqing, ZHU Shiyan, PAN Bocun, GUO Peng

(Teld New Energy Co., Ltd., Qingdao 266104, China)

**Abstract:** In view of the fact that the safety threshold of battery index is static laboratory data, which makes it difficult to accurately identify vehicles that do not reach a fixed threshold but are actually abnormal, a distributed anomaly detection method based on the comparison of the same region and the same vehicle type at the same time, a method based on battery heat generation mechanism and machine learning fusion to identify temperature rise anomalies, and a data driven Gaussian process supervised anomaly detection method are proposed. Based on the distribution anomaly detection method of comparing the same vehicle type in the same region at the same time, this paper first uses the statistical test method to test the normal distribution of the battery index data, then applies the mathematical method to convert the non normal distribution index data to the normal distribution, and then uses the statistical small probability event principle, combined with expert experience, to determine the anomaly detection threshold and establish the anomaly detection model. Based on the fusion of battery heat generation mechanism and machine learning, the method of identifying temperature rise anomaly is used. First, the physical principle of battery heat generation is used to calculate the heat release during battery charging. Then, the robust regression algorithm is used to train the fitting model of current and temperature rise changes during charging. Finally, the boxplot anomaly detection algorithm is used to detect the anomaly of  $k$  in the fitting model of orders in the same region and the same vehicle type at the same time, identify the charging battery with abnormal temperature rise rate under the same conditions. Based on the data driven Gaussian process supervised anomaly detection method, using the existing abnormal battery samples, it first selects the features through LightGBM technology, then establishes the Gaussian process model and determines the classification threshold for anomaly detection. The verification results show that, by comparing the batteries of the same region and vehicle type at the same time, the detection of abnormal distribution provides a solution for dynamic and accurate adjustment of the indicator threshold. The detection scheme of temperature rise anomaly establishes a complex index, and based on this index, the anomaly detection is effectively realized. Through the supervised model of Gaussian process, the data characteristic law of abnormal battery is fully excavated, and the accuracy of abnormal detection is further improved. Based on these three technologies, a comprehensive and effective comprehensive solution has been formed, which combines unsupervised technology with supervised technology, battery heat generation mechanism and artificial intelligence technology.

**Key words:** charging safety; abnormal detection; normal distribution; heat generation mechanism; Gaussian process

## 0 引言

随着新能源设备的飞速发展，电池的充电技术越来越受到关注，其中充电的安全性更是关注的重点。

电池安全故障种类繁多且诱发因素复杂，在电池管理系统(BMS)里针对各类故障分别开发诊断算法成为一个巨大挑战。内外研究机构利用电池电化学模型及电压、温度传感器，对故障电池的微观特征参数进行量化分析，在动力电池寿命预测与故障预警等关键技术进行了深入研究。除了机理模型方法，如华威大学等研究机构基于数据驱动的方式，对内部短路、结构故障等深层问题开展研究。现有研究尚未考虑复杂充放电工况的影响。同时，车载动力电池管理系统受制于有限的计算能力，无法应对频繁高倍率充放电工况下的技术需求。

为防止充电过程中由于电池温度过高等因素造成事故，通常厂家通过测试实验得到电池充电各变量的安全阈值并写入 BMS。但是，测试实验所得到的安全阈值通常是固定，而随着电池性能的变化，所对应的阈值也在不断变化，以一个固定的安全阈值作为异常判断标准必然导致判断结果不准确。此外，现有的安全阈值通常是量化一个静态特征，例如动力蓄电池的温度，而不考虑变量的动态发展，故无法及时识别出充电异常。因此，在新能源设备充电过程中，如何准确和及时识别出充电异常是本领域技术人员亟待解决的问题。

另外，目前充电安全分析多是基于直接测量指标，或是经过简单线性计算得到的二次指标。而电池充放电是复杂的非线性电化学反应过程，上述指标不能全面反映这一过程，因此，基于电化学机理建立新的指标计算模型，并且电池异常识别是非常必要的途径。

同时，车企和充电平台已经积累了较多的异常电池样本，这些样本作为宝贵的数据资产，需要基于有监督机器学习技术挖掘其中的数据规律，以发挥其应有的巨大作用，进而提升异常检测准确率。

针对这些问题，本文提出基于大数据人工智能的充电安全评估方法，从三个不同的方向分别提出对应的评估技术。基于利用正态分布统计学异常检测方法，对充电过程的原始指标和统计指标进行分析，通过同区域同车型电池的对比，确定各项指标数据的安全阈值，用于安全评估；结合电池产热机理和机器学习技术，建立充电过程中电池温升异常识别模型，用于识别充电过程中的温升异常；利用已有的安全异常电池样本，建立高斯过程有监督分类模型，学习异常电池的指标数据规律，用于电池

异常检测。最后，通过大量实验对本文提出方法的准确性和有效性进行验证分析。

## 1 正态分布统计学异常检测

### 1.1 基于正态分布的异常检测

异常检测是机器学习的一个分支。在一般的机器学习问题中，模型是对数据整体性质的描述，与整体性质表现不同的点是异常点。基于正态分布的异常检测，利用事物自身分布符合正态分布的特点，而选择的相应算法。

正态分布的“ $3\sigma$ ”原则：区间  $(\mu - 3\sigma, \mu + 3\sigma)$  是随机变量 X 实际可能的取值区间，X 落在该区间以外的概率小于 3%，在实际问题中一般认为这种事件是不会发生的。如果变量属于这千分之三，即是异常点。如图 1。

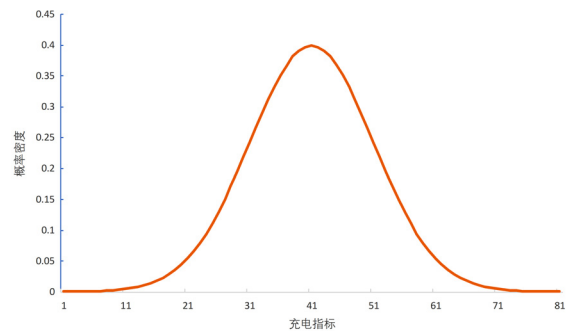


图 1 正态分布异常值检测

Fig. 1 Normal distribution outlier detection

### 1.2 正态分布异常检测步骤流程

正态分布异常检测，分为 4 个步骤：

#### 1) 数据清洗

基于正态分布的异常检测，需要基于同区域同车型同时段的数据对比，因此，样本数据需要包含区域城市、充电车型、充电时间信息。

安全分析的指标变量有：最高温度、最大温差、最大温升速率、单体压差、SOC 变化率。

#### 2) 探索分析

探索分析的主要目的是检测数据分布是否符合正态分布，对于不符合正态分布的数据需要进行转换。检测正态分布的方法有两类：统计学检验方法，包括 K-S 检验、A-D 检验、W 检验等；画图对比方法，包括 Q-Q 图、P-P 图等。基于以上方法，识别出非正态分布的数据后，需要将这类数据转换成正态分布，转化方法有对数变换、平方根变换、倒数变换、平方根反正弦变换。

(1) K-S 检测案例：以某城市某货车车型为例，该车型在某天的充电最高温度数据分布如图 2 所示。基于 K-S 检验该数据是否符合正态分布，得到

的  $pvalue = 0.12548143850323046$ , 高于阈值  $0.05$ , 则不拒绝原假设, 因此该最高温度数据符合正态分布, 可以应用正态分布异常检测方法。

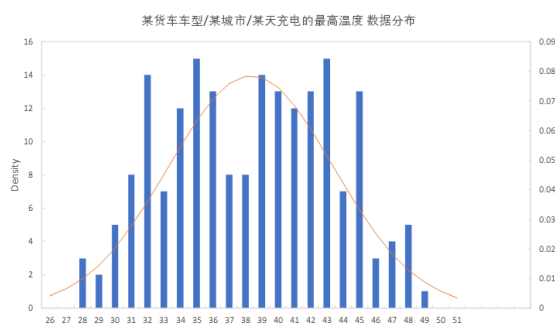


图2 某货车车型最高温度分布

Fig. 2 Maximum temperature distribution of a certain truck model

(2) Q-Q 图检测案例: 以某城市某乘用车车型为例, 该车型在某天的单体压差数据分布如图 3 所示。

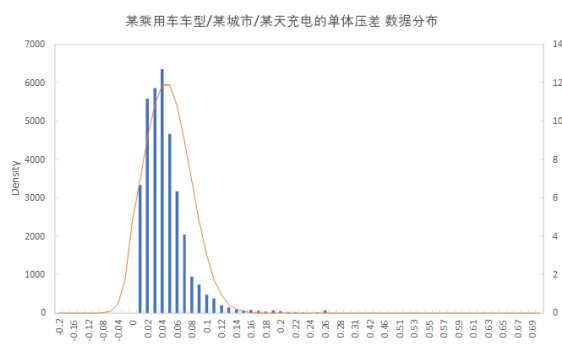


图3 某乘用车车型单体压差分布

Fig. 3 Individual pressure difference distribution of a certain passenger vehicle model

基于该样本数据绘制 Q-Q 图, 如图 4 所示。如果 Q-Q 图上的数据点近似在一条直线附近, 就表明该数据近似于正态分布。由于图 4 中的数据点分布明显偏离直线, 因此该样本数据不符合正态分布, 需要转换成正态分布后, 才能利用正态分布异常检测方法。

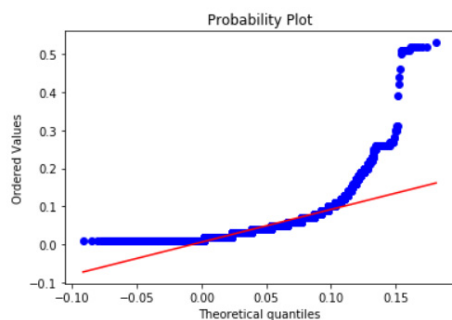


图4 原始样本绘制 Q-Q 图

Fig. 4 Q-Q diagram for the original sample

依据该数据的长尾分布规律, 对数据进行对数变换操作。基于对数变换后的数据再次绘制 Q-Q 图, 如图 5 所示。其数据点分布在直线附近, 该数据近似正态分布。

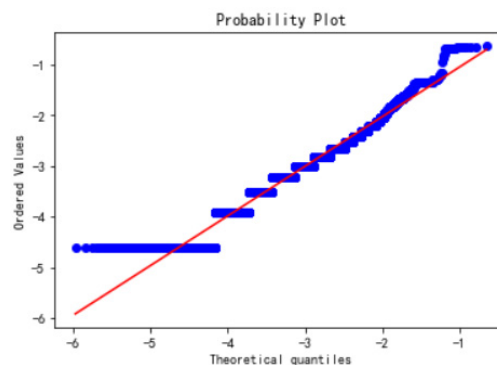


图5 样本对数变换后绘制 Q-Q 图

Fig. 5 Q-Q diagram after sample logarithmic transformation

对数变换后的样本数据分布如图 6 所示。基于该数据, 可以应用正态分布异常检测方法。

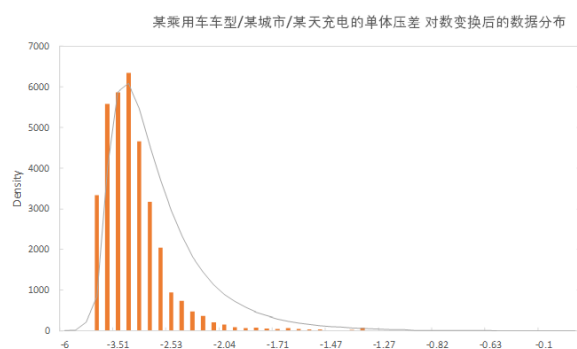


图6 样本对数变换后数据分布

Fig. 6 Data distribution after sample logarithmic transformation

### (3) 建立模型

结合专家经验, 选择  $\mu + 3\sigma$  作为充电安全电池指标的异常阈值, 当指标数据超过  $\mu + 3\sigma$  时, 即判定为电池异常。其中  $\mu$  为同区域同车型的指标数据均值,  $\sigma$  为同区域同车型的指标数据标准差。

### (4) 模型验证

将上述阈值方案应用到充电安全相关的电池指标上, 发现电池异常时, 异常电池跟踪确认。

最大温升速率异常检测案例: 车辆在 2019 年 8 月 6 号之后, 最大温升速率大部分超过了相同条件下的  $\mu + 3\sigma$  值。该车辆在 2019 年 10 月 2 日, 经主机厂确认为电池问题, 并进行维修, 维修之后, 车辆的最大温升速率数据恢复至  $\mu$  值附近, 远低于  $\mu + 3\sigma$  阈值。如图 7 所示。

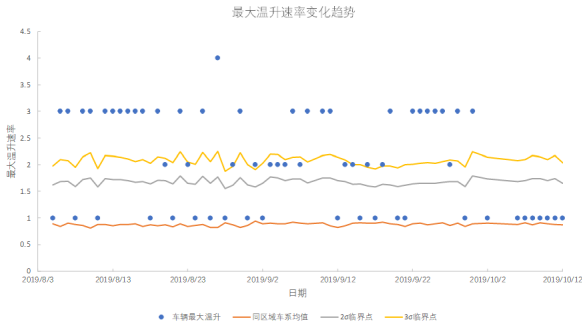


图7 最大温升速率异常检测案例

Fig. 7 Abnormal detection case of maximum temperature rise rate

最大温异常检测案例：车辆在2019年8月6号之后，最大温差呈增长趋势，先是超过了相同条件(区域、车型、时间)下的 $\mu + 2\sigma$ 值，在2019年8月25号之后又超过 $\mu + 2\sigma$ 值。该车辆于2019年10月2日，经主机厂确认为电池问题，并进行维修，维修之后，车辆的最大温差恢复到 $\mu + 2\sigma$ 以内。如图8所示。维修后，仍有部分数据点接近 $\mu + 2\sigma$ ，表明以 $\mu + 3\sigma$ 作为阈值，是更稳定可行的方案。

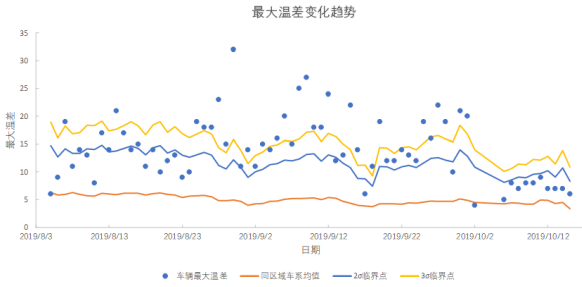


图8 最大温差异常检测案例

Fig. 8 Anomaly detection case of maximum temperature difference

根据上述验证分析，可以看出，基于正态分布的异常检测方法，通过同区域同车型的对比，能够快速有效地检出指标异常的电池。

## 2 产热机理和机器学习相结合的温升异常识别

### 2.1 产热机理

在实际生产中，充电过程中的电池温度变化主要受电池类型、电芯连接件内阻以及电池散热系统性能等内部因素影响，同时还受电池充电电流、用户充电习惯以及环境温度等外部因素影响。影响因素多而且复杂，因此，在电池温升异常检测过程中，需要研究不同外部因素和内部因素情况下温升的变

化。首先结合焦耳定律，假设：

$$\begin{aligned} T(k) - T(0) \\ \propto Q_{\text{产热}}(k) - Q_{\text{散热}}(k) \\ \propto \sum_{i=0}^k I^2 R \Delta t - Q_{\text{散热}}(k) \end{aligned} \quad (1)$$

式中， $T(0)$ 表示电池组充电初始最高温度， $T(k)$ 表示电池组第 $k$ 个采样时间的最高温度， $Q_{\text{散热}}(k)$ 表示电池从0时刻到 $k$ 时刻累计总散热， $Q_{\text{产热}}(k) - Q_{\text{散热}}(k)$ 表示电池从0时刻到 $k$ 时刻电池净产热。 $Q_{\text{散热}}(k)$ 为时间 $k$ 的函数，但是精确的解析式无法获取。电池温度升高正比于电池净产热，而电池总产热的最重要一部分为电池电阻上的热损耗，其可以表示为 $\sum_{i=0}^k I^2 R \Delta t$ ， $R$ 为电池组内阻。 $Q_{\text{散热}}(k)$ 很难建模，主要原因是电池散热功率主要受电池采用的散热介质、环境温度以及电池组是否开启散热系统等影响。因此，可以通过经验将订单按照环境温度、电池是否开启散热系统进行分流，在此情况下可以认为电池热耗散功率为恒定值。假定电池内阻在充电过程中不会发生变化，可以通过式(2)对电池热动态行为进行拟合。

$$\Delta T(k) = k \sum_{i=0}^k I^2 \Delta t + B \quad (2)$$

式中， $\Delta T(k) = T(k) - T(0)$ ， $k$ 为电池温升速率，单位为 $[^{\circ}\text{C} \cdot \text{A}^{-2} \cdot \text{s}^{-1}]$ ， $B$ 为表征电池散热性能的常值。利用订单实测数据拟合式(2)，可分析出每个订单对应的温升速率与散热性能指标。 $K$ 值大小代表电池在吸收单位热量所升高的温度，数值越大代表温升越快，电池异常的可能性越高，当电池多次出现温升异常后，则需要重点关注电池的安全状况。

### 2.2 机器学习方法

#### 1) 鲁棒回归

充电数据在收集和传输过程中，存在异常数据和丢失数据的情况，传统的回归算法会将离群值的误差累加到目标函数中，从而严重影响模型整体的拟合效果，所以本文利用鲁棒回归算法，降低离群值对回归模型的影响，提高模型泛化能力。

鲁棒回归的优化目标函数为：

$$\min_{w, \sigma} \sum_{i=1}^n \left( \sigma + H_{\varepsilon} \left( \frac{X_i w - y_i}{\sigma} \right) \sigma \right) + \alpha \|w\|_2^2 \quad (3)$$

其中：

$$H_{\varepsilon}(z) = \begin{cases} z^2 & \text{if } |z| < \varepsilon \\ 2\varepsilon \|z\| - \varepsilon^2 & \text{otherwise} \end{cases} \quad (4)$$

#### 2) boxplot 异常检测

异常值一般指数据远离大多数观察值，有多种异常值检测方式，如基于变量空间距离度量，把距



离太远的观察值标记为异常值。还有基于数据分布的检测方法, **boxplot** 异常检测是基于分布的异常检测方法。该检测方法先假定数据的预期分布, 把背离这种分布的数值标记为异常值。首先计算数据集的四分位数据  $Q_1$ 、 $Q_3$ , 利用公式(5)计算四分位距:

$$IQR = Q_3 - Q_1 \quad (5)$$

结合式(5)的计算结果计算两边的阈值:

$$\begin{aligned} \text{最大阈值: } & Q_3 + 1.5 * IQR \\ \text{最小阈值: } & Q_1 - 1.5 * IQR \end{aligned} \quad (6)$$

对比每笔订单的  $k$  值, 超过最大阈值的订单判定为温升异常。

### 2.3 温升异常识别的步骤流程

温升异常检测流程分为 3 个步骤, 如图 9 所示。

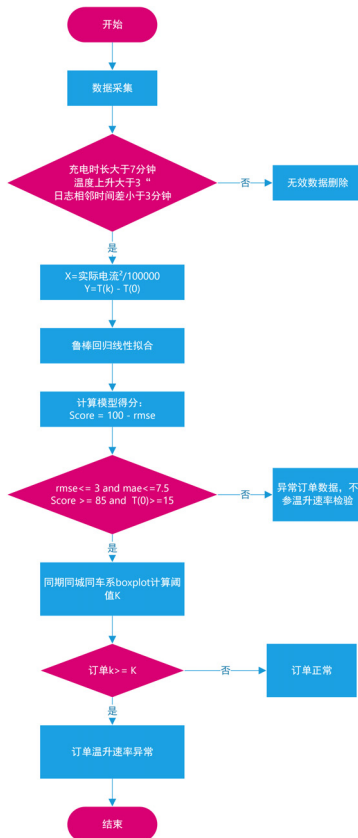


图 9 温升异常检测全流程

Fig. 9 Entire process of temperature rise anomaly detection

#### 1) 数据预处理

以一次完整的充电订单为样本, 首先判断订单充电订单时长是否超过 7 min, 判断温升是否大于 3 °C, 相邻记录时间差是否超过 3 分钟, 若充电时长少于 7 min, 或者温升小于 3 °C, 处理后参与建模的数据量过小, 造成模型不稳定, 拟合效果差; 若相邻记录时间差超过 3 min, 有数据丢失, 也将降低模型的准确度。

模型训练数据准备, 根据式(7)处理筛选后的数据, 生成模型自变量、因变量。

$$\begin{aligned} X(k) &= \sum_{i=0}^k I_k^2 / 100000 \\ Y(k) &= T(k) - T(0) \end{aligned} \quad (7)$$

#### 2) 回归模型训练

结合鲁棒回归算法, 对  $X$ 、 $Y$  进行回归, 设置参数  $\max\_iter = 100$ ,  $\epsilon = 1.35$ ,  $\alpha = 0.0001$ 。通过模型训练, 获取拟合模型参数  $k$ 、 $b$ , 计算拟合模型的最大绝对误差、均方根误差, 根据公式:

$$score = 100 - 16 * rmse \quad (8)$$

计算拟合模型得分。

#### 3) 异常检测计算

获取同区域同车型同时段模型的  $k$  值, 根据公式(6)计算阈值  $K$ , 对比订单  $k$  与阈值  $K$ , 若  $k$  大于  $K$ , 则判定订单温升异常。

### 2.4 温度异常识别结果

#### 1) 案例 1

某车辆的订单  $k$  值变化趋势如图 10 所示, 基于正态分布统计学异常检测评价趋势, 如图 11 所示。

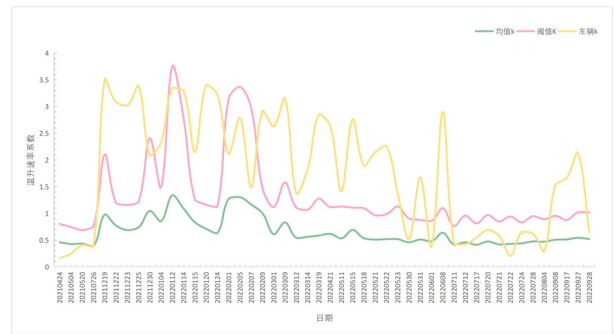


图 10 某车辆的订单  $k$  值变化趋势

Fig. 10 Trend of  $k$ -value for a certain vehicle

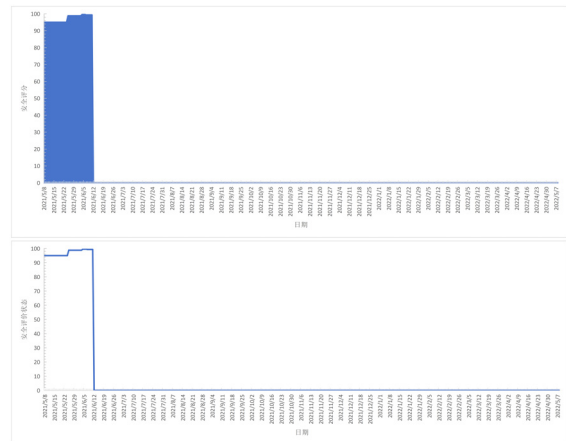


图 11 正态分布统计学异常检测评价趋势

Fig. 11 Normal distribution anomaly detection and evaluation trend

2021年12月19日电池开始出现温升异常,同时基于正态分布的异常检测评估模型评估电池为亚健康状态,后期温升异常频繁出现,车辆状态出现了高危状态。综合该车辆的两种异常检测结果,可以判定基于正态分布的异常检测与温升异常检测结果基本一致,说明温升异常检测算法可以检测出电池异常。

2) 案例 2

某车辆的订单  $k$  值变化趋势如图 12 所示,基于正态分布统计学异常检测评价趋势如图 13 所示。



图 12 某车辆的订单  $k$  值变化趋势

Fig.12 Trend of  $K$ -value for a certain vehicle

该电池可判定为安全隐患,需重点观察。因此,温升异常检测模型也可以检测出未达到高危状态的电池,所以温升异常检测结果可以作为基于正态分布统计学异常检测方法的补充。

3) 案例 3

某车辆近两年订单的温升模型  $k$  值曲线如图 14 所示,可以观察到,随着车辆充电次数的增加,温升模型  $k$  值逐渐靠近阈值  $K$ ,说明随车辆充电次数的增加,温升速率越来越高,电池安全状态需要重点关注。



图 14 某车辆订单  $k$  值曲线

Fig. 14 The  $k$ -value curve of a certain vehicle

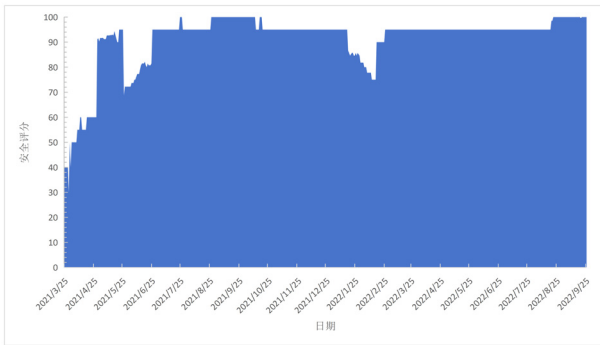


图 13 正态分布统计学异常检测评价趋势

Fig. 13 Normal distribution anomaly detection and evaluation trend

2022年08月23日以后,订单  $k$  值频繁超过阈值  $K$ ,即频繁温升异常。基于正态分布统计学异常检测评价分数在亚健康状态。综合两个方法的结论,

某车辆历年每月平均  $k$  值变化曲线如图 15 所示;图中历年每月  $k$  值趋势可以看出,  $k$  值随季节存在明显的周期性变化,因此异常检测的阈值不能是固定值,也需要随季节周期性变化;同时电池每月的  $k$  值也存在逐年递增趋势(2020年2月因疫情原因,车辆没有充电订单),说明随电池运行年限的增加,  $k$  值也在逐年升高,运行越久的电池安全状态越需要关注。



图 15 某车辆历年每月平均  $k$  值变化曲线

Fig. 15 Monthly average  $k$ -value curve of a certain vehicle over the years

根据上述验证分析可以看出,基于温升异常的检测方法,对存在安全隐患电池的检测效果明显,可以作为基于正态分布统计学异常检测结果的补充,提高电池异常检出率。

### 3 高斯过程有监督电池异常检测模型

#### 3.1 高斯过程有监督异常检测方法

传统的电池异常检测方法, 多是基于单个电池指标如电压、温度、压差、温差等, 通过无监督方法进行异常识别, 比如通过同区域同车型同时段之间的对比, 当超过某个阈值时, 即判定为异常。

车企和充电平台已经积累了一定数量的安全异常电池样本, 这些样本作为宝贵的数据资产, 无监督技术难以利用起来。需要借助有监督机器学习技术, 挖掘其中的数据规律, 才能发挥其应有的价值, 带来异常检测准确率的提升。高斯过程(GP)分类模型, 即是其中之一。

机器学习中的很多技术都是基于参数化的函数, 并对这些参数进行建模。GP 却是直接对函数建模的非参数模型, 这种方法的好处是: 既能对任意函数进行拟合, 又能对不确定性进行拟合。对不确定性的量化极其有价值, 这也符合贝叶斯优化的思想。正是 GP 具有严格的统计学理论基础, 使其对于解决高维非线性的小样本复杂问题, 具有良好的适应能力。

#### 3.2 高斯过程有监督异常检测的步骤流程

高斯过程有监督异常检测分为 4 个步骤。

##### 1) 数据清洗

以一次充电过程的电池指标统计数据作为样本。采用 Z-score 标准化方法对样本数据进行标准化处理, 达到消除量纲的目的。经过处理后的数据符合标准正态分布, 即均值为 0, 标准差为 1。标准化方法如式(9)所示。

$$z = \frac{x - \mu}{\sigma} \quad (9)$$

式中:  $x$  是指标原始值,  $\mu$  是指标数据的均值,  $\sigma$  是指标数据的标准差,  $z$  是标准化的结果。

样本数据的标签定义: 针对电池安全异常车辆, 其标签定义为 1, 作为正样本; 电池正常的车辆, 标签定义为 0, 作为负样本。

经过上述清洗操作得到训练集和测试集, 训练集共 1243 辆车, 其中正样本 186 辆, 负样本 1057 辆; 测试集 1812 辆, 没有确定标签。在本实验中, 将有监督模型应用于测试集识别出其中的异常车辆后, 对部分车辆进行回访确认。

##### 2) 特征选择

采用 LightGBM 算法进行特征选择, 确定用于有监督分类模型的特征。LightGBM(Light Gradient Boosting Machine)是基于决策树算法的分布式梯度提升框架, 它提供了一个快速高效、低内存占用、

高准确度、支持并行和大规模数据处理的数据科学工具。

利用 LightGBM 算法, 选择与电池安全温度异常相关性最高的指标作为特征: 最大温升速率、最小温升速率、最高温度最大值、最高温度最小值、最高温度平均值、最高温度标准差、最低温度最大值、最低温度最小值、最低温度平均值、最低温度标准差、最大温差、起始温度、环境温度、电流平方积分值、充电时长。

LightGBM 参数设置为: 学习率=0.1, L1 正则化参数=0.1, L2 正则化参数=0.2, 叶节点最少样本数量=100, 树的最大深度=10, 目标函数=binary。

特征重要性排名如图 16 所示。

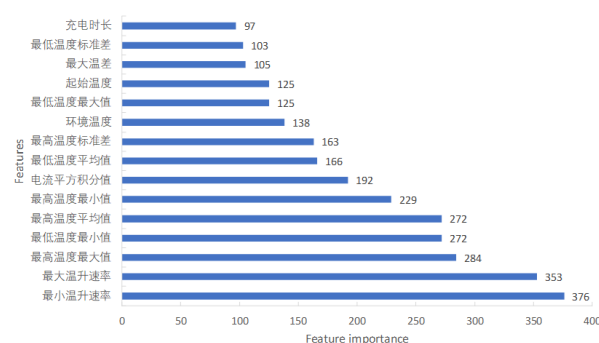


图 16 特征重要性排序

Fig. 16 Feature importance ranking

##### 3) 建立模型

##### (1) 高斯过程参数设置

核函数 kernel = RBF, 内核长度尺度参数  $l = 1$ 。其中, RBF 内核如式(10)所示。

$$k(x_i, y_j) = \exp\left(-\frac{d(x_i, y_j)^2}{2l^2}\right) \quad (10)$$

式中:  $d$  是欧几里得距离;  $l$  是内核的长度尺度。

##### (2) 评价指标

考虑大多数车辆电池为正常电池, 只有极少数车辆存在电池异常, 因此, 异常检测模型不能只使用准确率作为评价指标, 为对模型进行有效评估, 本文基于表 1 所示的混淆矩阵进行评价。

表 1 混淆矩阵

Table 1 Confusion matrix

电池类别	检测为异常	检测为正常
实际异常	TP	FN
实际正常	FP	TN

如表 1 所示, 利用混淆矩阵获得多个评价指标, 本文选择准确率 ACC、召回率 TPR 进行评价, 其定义分别为  $ACC = TP / (TP + FP)$ 、 $TPR = TP / (TP + FN)$ 。其中, ACC 表示检出真实异常电池占检出异常电池

的比例，TPR 表示检出真实异常电池占实际异常电池的比例。本文以 ACC 和 TPR 这两个指标作为衡量学习器优劣的性能指标，当这两个指标达到平衡时，模型的分类效果最好。

(3) 模型评估

采用了交叉验证的评估方法，进行 10 轮 5 折交叉验证。交叉验证评估结果如下。

训练集：ACC = 95.79%，TPR = 93.18%；验证集：ACC = 2.32%，TPR = 48.00%。

经过以上评估，验证了模型的可行性。接下来，训练集全部参与训练，并基于训练集评估正负样本的异常概率分类阈值。当异常概率阈值为 0.57 时，精准率和召回率达到平衡，分类效果最优。如图 17 所示。

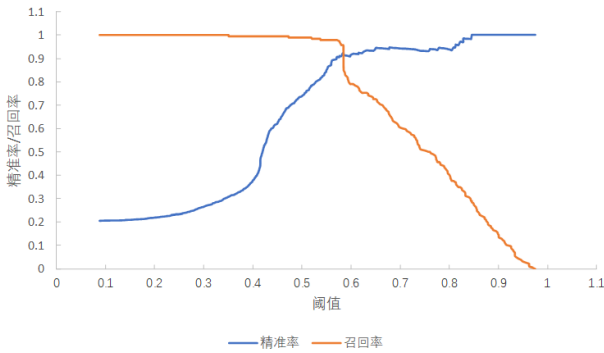


图 17 不同阈值下的精准率与召回率

Fig. 17 Accuracy and recall under different thresholds

在训练集上，当异常概率阈值为 0.57 时，分类效果最优，如图 18 所示，其中，橙色点为实际异常的车辆电池，蓝色点为正常的车辆电池。



图 18 阈值为 0.57 时训练集上的分类结果

Fig. 18 Classification results on the training set at threshold 0.57

(4) 模型验证

将上述模型应用到测试集，根据异常概率阈值 0.57，从 1812 辆车中检测出 43 辆电池存在安全异常的车辆。针对其中异常概率最大的前 10 辆车，逐车深入分析并对车主进行回访。其中 8 辆车的车主反馈电池故障，另 2 辆车主不予反馈或不接电话，

但这 2 辆车电池的某些指标在当时或随后达到危险阈值，后不再有充电记录。明细结果如表 3 所示。

表 2 异常电池车辆分析结果

Table 2 Abnormal battery vehicle analysis results

电池异常 Top10	充电日期	电池指标分析及回访结果
1	2020-10-13	最高温度异常；回访：客户接到平台通知后去检测，是一组电池烧了。
2	2020-10-19	温差趋势异常；回访：客户接到平台通知后去检测，发现电池流液。
3	2020-10-22	充电过程中SOC不变化，温差异常、不均匀性异常，10月份后不再充电；回访：不予反馈
4	2020-11-05	温差和最高温度持续增长趋势，2021.1后不再充电；回访：不接电话
5	2021-10-12	压差异常；回访：客户表示知道电池有问题，并计划去维修
6	2021-10-13	温差异常；回访：客户反馈应该是有问题，最近充电一直温度升高，但是充电速度不行，已预约检查
7	2021-10-23	其他安全故障；回访：(客户回复)电池电阻有问题，4S店更换后恢复正常
8	2021-11-10	压差趋势异常；回访：客户表示确实存在异常，并计划去检修
9	2021-11-13	压差趋势异常；回访：客户感觉车充不上电，跑不了多远，计划去4S店检修。
10	2022-01-07	温升异常；回访：客户反馈电池有问题，具体情况不便告知。

根据上述验证分析，可以看出，以已有的异常电池作样本，通过高斯过程有监督异常检测模型，可以有效检出存在安全异常的电池。

4 结论

本文针对电池指标安全阈值是静态的实验室数据问题，从 3 个不同的技术方向，分别提出了相应的解决方案，并基于大量实际业务数据进行了实验验证。基于正态分布统计学异常检测方法及其验证结果，表明在同区域同车型同时段的数据对比基础上，采用 3σ 原则可以提供了精确有效的指标安全阈值，实现了在工况下的阈值动态调整；基于电池产热机理和机器学习融合的温升异常识别方法及其验证结果，表明该方法可以挖掘新的产热指标，从而有效提高电池异常检测的精度；基于数据驱动的高斯过程有监督异常检测方法及其验证结果，表明该方法能充分利用已有的异常样本，学习电池充电过程的复杂特征规律，作为无监督学习的有效补充，提高电池异常检测的检出率。

基于这 3 个技术，形成了从无监督技术到有监督技术、电化学机理与人工智能技术相结合的、全面有效的综合解决方案。

参考文献

[1] 电动汽车非车载传导式充电机与电池管理系统之间的通信协议: GB/T 27930—2015[S]. 北京: 中国标准出版社, 2016.  
Communication protocol between off board conductive charger and battery management system of electric vehicle: GB/T 27930—2015[S] Beijing: China Standards Press, 2016.

[2] 工业和信息化部, 国家能源局. 电动汽车安全指南 (2019 版)[EB/OL]. (2019-10-28)[2020-06-16].  
Ministry of Industry and Information Technology, National Energy Administration. Safety guide for electric vehicles



- (2019 edition)[EB/OL]. (2019-10-28)[2020-06-16]
- [3] 郑岳久. 车用锂离子动力电池组的一致性研究[D]. 北京: 清华大学, 2014.  
ZHENG Yuejiu. Study on the consistency of lithium-ion power battery pack for vehicles[D]. Beijing: Tsinghua University, 2014.
- [4] 冯旭宁. 车用锂离子动力电池热失控诱发与扩展机理、建模与防控[D]. 北京: 清华大学, 2016.  
FENG Xuning. Induction and expansion mechanism, modeling and prevention and control of thermal runaway of vehicle lithium-ion power battery[D] Beijing: Tsinghua University, 2016.
- [5] 余承其, 张照生, 刘鹏, 等. 大数据分析技术在新能源汽车行业的应用综述—基于新能源汽车运行大数据[J]. 机械工程学报, 2019, 55(14): 3-16.  
SHE Chengqi, ZHANG Zhaosheng, LIU Peng, et al. Overview of the application of big data analysis technology in the new energy vehicle industry—based on the big data of new energy vehicle operation [J]. Journal of Mechanical Engineering, 2019, 55(14): 3-16.
- [6] 汪伟. 新能源电动客车在大数据领域的探索[J]. 汽车工程学报, 2020, 10(增刊 1): 20-23.  
WANG Wei. Exploration of new energy electric buses in the field of big data[J]. Journal of Automotive Engineering, 2020, 10(S1): 20-23.
- [7] 张坤, 魏东坡, 尹文荣, 等. 基于大数据的纯电动汽车系统故障规律研究[J]. 济宁学院学报, 2019(5): 12-16.  
ZHANG Kun, WEI Dongpo, YIN Wenrong, et al. Research on fault law of pure electric vehicle system based on big data[J]. Journal of Jining University, 2019(5): 12-16.
- [8] 刘木林, 卜凡涛, 林辉, 等. 电动汽车动力电池热失控过程分析及预警机制设计[J]. 汽车实用技术, 2020(5): 15-17.  
LIU Mulin, BU Fangtao, LIN Hui, et al. Analysis of thermal runaway process of electric vehicle power battery and design of early warning mechanism[J]. Automotive Practical Technology, 2020(5): 15-17.
- [9] 王文斌. 汽车企业的大数据分析模型研究与实现[J]. 信息与电脑, 2019(14): 149-151.  
WANG Wenbin. Research and implementation of big data analysis model for automobile enterprises[J]. Information and Computer, 2019(14): 149-151.
- [10] 吴凯, 张耀, 曾毓群, 等. 锂离子电池安全性能研究[J]. 化学进展, 2011, 23(2): 401-409.  
WU Kai, ZHANG Yao, ZENG Yuqun, et al. Study on safety performance of lithium ion battery[J]. Chemical Progress, 2011, 23(2): 401-409.
- [11] 贺春, 陈卓, 冯瑾涛, 等. 电动汽车充电安全分析与解决方案[J]. 公用电, 2017(1): 12-18.  
HE Chun, CHEN Zhou, FENG Jintao, et al. Electric vehicle charging safety analysis and solution[J]. Utilities, 2017(1): 12-18.
- [12] 刘鹏, 李阳, 李达, 等. 新能源汽车安全预警模型研究[J]. 汽车工程学报, 2020, 10(增刊 1): 40-45.  
LIU Peng, LI Yang, LI Da, et al. Research on new energy vehicle safety early warning model[J]. Journal of Automotive Engineering, 2020, 10(S1): 40-45.
- [13] 姜久春, 高洋, 张彩萍, 等. 电动汽车锂离子动力电池健康状态在线诊断方法[J]. 机械工程学报, 2019, 55(20): 60-72.  
JIANG Jiuchun, GAO Yang, ZHANG Caiping, et al. On line diagnostic method for health status of lithium-ion power battery of electric vehicle[J]. Journal of Mechanical Engineering, 2019, 55(20): 60-72.
- [14] 高飞, 刘皓, 吴从荣, 等. 锂离子电池热安全防控技术的研究进展[J]. 新能源进展, 2020, 8(1): 15-21.  
GAO Fei, LIU Hao, WU Congrong, et al. Research progress in thermal safety prevention and control technology of lithium ion batteries[J]. New Energy Progress, 2020, 8(1): 15-21.
- [15] BARNETT B. Technologies for detection and intervention of internal short circuits in Li-ion batteries[C] // 5th Annual Battery Safety, 2014, Washington D.C., USA.
- [16] WALDMANN T, WILKA M, KASPER M, et al. Temperature dependent ageing mechanisms in lithium-ion batteries—a post-mortem study[J]. Journal of Power Sources, 2014, 262: 129-135.
- [17] FENG X, PAN Y, HE X, et al. Detecting the internal short circuit in large-format lithium-ion battery using model-based fault-diagnosis algorithm[J]. Journal of Energy Storage, 2018, 18: 26-39.
- [18] KEATES, ANDREW W. Short circuit detection for batteries: United States, US7795843B2[P]. 2010.
- [19] ZHANG Xiongwen. Thermal analysis of a cylindrical lithium-ion battery[J]. Electrochimica Acta, 2011, 56(3): 1246-1255.
- [20] LING L, CHAO S, XIA Y, et al. A research on cell inconsistency prediction of power battery using Gaussian process regression[C] // The Second International Conference on Image, Video Processing and Artificial Intelligence, 2019.

收稿日期: 2023-08-10; 修回日期: 2024-01-20

作者简介:

鞠强(1975-), 男, 本科, 高级工程师, 特来电新能源股份有限公司首席科学家, 长期从事海量数据处理、大数据分析挖掘、电池安全异常检测等领域; E-mail: juqiang@teld.cn

尹玉鹏(1985-), 男, 硕士, 特来电新能源股份有限公司安全总监, 研究大规模电动汽车充电安全技术; E-mail: yiyp@teld.cn

项宝庆(1982-), 男, 本科, 高级工程师, 特来电新能源股份有限公司架构师, 研究方向为电动汽车安全检测与人工智能应用; E-mail: xiangbq@teld.cn