

中文引用格式:范成强,夏元友,张宏伟,等. 基于 LLE-DBSCAN-SMOTE 数据处理的隧洞岩爆预测[J]. 中国安全科学学报, 2024, 34(12): 140-148.

英文引用格式:FAN Chengqiang, XIA Yuanyou, ZHANG Hongwei, et al. Tunnel rockburst prediction based on LLE-DBSCAN-SMOTE data processing [J]. China Safety Science Journal, 2024, 34(12): 140-148.

基于 LLE-DBSCAN-SMOTE 数据处理的 隧洞岩爆预测*

范成强, 夏元友** 教授, 张宏伟, 黄建

(武汉理工大学 土木工程与建筑学院, 湖北 武汉 430070)

中图分类号: X936

文献标志码: A

DOI: 10.16265/j.cnki.issn1003-3033.2024.12.1917

基金项目: 国家自然科学基金资助(42077228)。

【摘要】 为解决岩爆预测中预测指标关联以及原始数据存在离群点与数据不平衡等问题, 提出基于局部线性嵌入(LLE)-基于密度的带噪声应用空间聚类(DBSCAN)-合成少数类过采样(SMOTE)数据处理的岩爆预测方法。首先, 选取围岩最大切向应力 σ_θ 、岩石单轴抗压强度 σ_c 、岩石单轴抗拉强度 σ_t 、弹性应变能指数 W_{et} 、脆性系数 σ_c/σ_t 、应力系数 σ_θ/σ_c 和表征围岩应力梯度的应力集度值 β 构建岩爆预测指标体系; 其次, 采用LLE算法进行数据降维处理以消除指标间的交叉关联影响, 引入DBSCAN算法去除数据离群点; 然后, 引入SMOTE技术进行数据平衡化; 最后, 分别采用决策树(DT)、随机森林(RF)与梯度提升树(GBDT)算法构建3类岩爆预测模型, 对比分析数据处理前后数据训练模型的预测精度, 并通过江边水电站引水隧洞实测岩爆数据进行工程验证。结果表明: 预测指标由原始数据的7维降至4维, 以及采用分级离群值处理后的3类算法模型的预测准确率皆为同类模型中最高, 江边水电站工程岩爆预测验证了数据处理后的模型预测准确率明显高于基于原始岩爆数据建立的同类模型。

【关键词】 局部线性嵌入(LLE); 基于密度的带噪声应用空间聚类(DBSCAN); 合成少数类过采样(SMOTE); 数据处理; 岩爆预测

Tunnel rockburst prediction based on LLE-DBSCAN-SMOTE data processing

FAN Chengqiang, XIA Yuanyou, ZHANG Hongwei, HUANG Jian

(School of Civil Engineering and Architecture, Wuhan University of Technology,
Wuhan Hubei 430070, China)

Abstract: To address issues of correlation prediction indicators, outliers, and data imbalance in original data in rockburst prediction, a rockburst prediction method based on LLE-DBSCAN-SMOTE for data processing was proposed. Firstly, the maximum tangential stress of surrounding rock σ_θ , uniaxial compressive strength of rock σ_c , uniaxial tensile strength of rock σ_t , elastic strain energy index W_{et} , brittle coefficient σ_c/σ_t , stress coefficient σ_θ/σ_c , and stress concentration value β characterizing the stress gradient of surrounding rock were selected to construct a rockburst prediction indicator system. Secondly, the LLE algorithm was used for data dimensionality reduction to eliminate the cross-correlation

* 文章编号: 1003-3033(2024)12-0140-09; 收稿日期: 2024-07-14; 修稿日期: 2024-09-19

** 通信作者: 夏元友(1965—), 男, 安徽庐江人, 博士, 教授, 主要从事岩土工程安全分析等方面的研究。E-mail: xiayy1965@whut.edu.cn。

effect between indicators, and the DBSCAN algorithm was introduced to remove outliers. Then, the SMOTE technology was introduced for data balancing. Finally, three types of rockburst prediction models were proposed using Decision Tree (DT), Random Forest (RF), and Gradient Boosting Decision Tree (GBDT) algorithms. The prediction accuracy of the data training models before and after processing was compared and analyzed. Moreover, engineering verification was performed through the measurement in the diversion tunnel of Jiangbian Hydropower Station. The results show that the prediction accuracy of the three types of algorithm models which reduce the prediction index from the 7 dimensions of the original data to the 4 dimensions and adopt the graded outlier processing is the highest among the similar models. The rockburst prediction of the Jiangbian Hydropower Station demonstrates that the proposed model significantly improves prediction accuracy compared to similar models using original rockburst data.

Keywords: local linear embedding (LLE); density-based spatial clustering of applications with noise (DBSCAN); synthetic minority over-sampling technique (SMOTE); data processing; rockburst prediction

0 引言

岩爆是深地工程常见的地质灾害,严重威胁施工人员和设备的安全,岩爆预测日益成为工程热点问题之一^[1-2]。随着机器学习的快速发展,众多学者运用机器学习方法来预测岩爆^[3-5],该方法能够较好地规避人为主观因素的影响。但大多研究者重视岩爆预测理论模型的选择,而在一定程度上忽视了预测指标选取和原始数据集的优化,具体表现在选取的预测指标独立性较差,数据集中存在离群数据和数据不平衡等问题未得到有效的处理。

针对以上问题,部分学者进行了一些探索。如针对离群数据处理,谭文侃等^[6]利用局部离群因子法去除离群样本,利用改进合成少数类过采样技术(Synthetic Minority Over-sampling Technique, SMOTE)平衡样本类别,提高了模型的学习效果与预测精度;夏元友等^[7]引用集成思想,耦合 3 类算法去除离群值得到了较好的处理效果;YIN Xin 等^[8]将局部离群因子和期望最大化算法用于离群值的检测和离群值的替换。但目前学者多采用整体去除方式去除岩爆离群样本,忽视了不同等级中的离群值,如强岩爆中的离群样本会混入其他岩爆等级样本中,采用整体去除时会忽视该类型值,进而影响岩爆预测准确率。

在预测指标选取上,众多学者主要基于围岩应力与围岩力学性质提出岩爆多因素预测指标,但选取的预测指标普遍存在交叉关联等问题,如围岩力学性质指标中的岩石单轴抗压强度 σ_c 、岩石抗拉强度 σ_t 与岩石脆性系数 σ_c/σ_t 等,围岩应力指标中的最大切向应力 σ_θ 与围岩应力系数 σ_θ/σ_c 等^[8-11]。

针对指标关联问题,学者们通常采用 2 种方式处理:①尽可能选取相对独立指标,如李明亮^[12]、侯克鹏^[13]等选取 σ_c/σ_t 、 σ_θ/σ_c 和岩石弹性变形能指数 W_{et} 构建岩爆烈度等级预测指标体系;②采用指标权重或降维等处理方法提取预测指标的主要信息,如谢学斌等^[14]选取 σ_θ 、 σ_c/σ_t 、 W_{et} 和 σ_θ/σ_c 作为评价指标,运用改进的指标客观赋权法加权处理指标样本,充分考虑指标间相关性和重要度,得到较高的预测精度;陈则黄等^[15]选取 W_{et} 、 σ_θ/σ_c 、 σ_c/σ_t 、 σ_c 、 σ_t 和 σ_θ 构建岩爆预测指标体系,采用主成分分析法降维处理指标,提取指标数据主要信息,得到 3 个线性无关的主成分输入向量,提高了岩爆的预测准确率;杨小彬等^[16]采用自组织特征映射神经网络算法挖掘指标间的非线性关系以及样本间的关联,解决了指标间的交错关联导致分级模糊问题,取得了不错的效果。但由以上研究不难发现:①要选取相对独立指标仍有一定难度,如 σ_c/σ_t 与 W_{et} 仍有关联^[17],同时选取太少的指标也会丢失一些有用信息;②众多学者忽视围岩应力梯度对于岩爆的影响,夏元友等^[18]研究表明:围岩应力梯度对岩爆有重要影响;③现有文献针对数据集数据处理不系统,多针对某一方面问题进行处理。

鉴于此,笔者拟引入局部线性嵌入法(Local Linear Embedding, LLE)降维岩爆原始数据,减少岩爆数据指标间的交叉关联,引入基于密度的带噪声应用空间聚类(Density-Based Spatial Clustering of Applications with Noise, DBSCAN),按整体与分级方式分别去除离群值并对比去除效果;针对数据不平衡问题,采用 SMOTE 技术平衡训练集;同时分别采用决策树(Decision Tree, DT)、随机森林(Random Forest,

RF)和梯度提升树(Gradient Boosting Decision Tree, GBDT) 3种算法,构建岩爆预测模型,并验证数据处理效果,以期提升岩爆预测模型的准确率。

1 岩爆预测样本库的建立

1.1 预测指标选取

众多学者主要围绕围岩应力与围岩力学性质提出岩爆多因素预测指标,如 σ_θ 、 σ_θ/σ_c 、 σ_c 、 σ_t 、 σ_c/σ_t 、 W_{et} 、冲击能指数 W_{ef} 、围岩完整性系数 K_V 等指标^[19-22]。考虑收集样本数据的指标数据缺失情况及围岩应力梯度的影响,选取 σ_c 、 σ_t 、 W_{et} 、 σ_c/σ_t 、 σ_θ 、 σ_θ/σ_c 和围岩应力集度 β 构建岩爆预测指标体系。围岩切向应力分布如图1所示。

围岩应力集度 β ^[23]反映围岩二次应力在围岩扰动区的集中程度(图1),是围岩应力的平均变化率,具体计算如下式:

$$\beta = \frac{\int_0^{3D_0} f'(x) dx}{3D_0} = \frac{\sigma_\theta - \sigma_{\theta 0}}{3D_0} \quad (1)$$

式中: $f'(x)$ 为围岩应力分布曲线函数表达式 $f(x)$

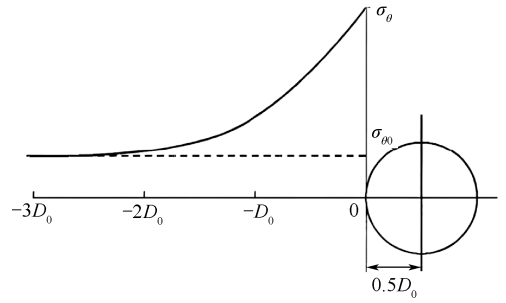


图1 围岩切向应力分布

Fig.1 Tangential stress distribution of surrounding rock

的一阶导数; D_0 为开挖洞室直径,对于非圆形洞室,采用当量方法来计算其等效洞径^[24]; $\sigma_{\theta 0}$ 与 σ_θ 为隧道开挖前后横截面切向初始应力。

1.2 样本数据库建立

通过文献调研并整理课题组前期成果,收集到309组无重复和缺失值岩爆案例样本^[24]。其中,无岩爆样本39条(岩爆烈度1),弱岩爆样本94条(岩爆烈度2),中等岩爆样本112条(岩爆烈度3),强岩爆样本64条(岩爆烈度4)。部分岩爆案例数据见表1。

表1 岩爆案例数据集

Table 1 Rockburst case dataset

| 序号 | σ_θ/MPa | σ_c/MPa | σ_t/MPa | σ_θ/σ_c | σ_c/σ_t | W_{et} | $\beta/(\text{MPa} \cdot \text{m}^{-1})$ | 岩爆烈度 | 工程名称 |
|-----|----------------------------|-----------------------|-----------------------|--------------------------|---------------------|----------|--|------|------------|
| 1 | 18.80 | 178.00 | 5.70 | 0.11 | 31.23 | 7.40 | 1.04 | 1 | 龙羊峡水电站地下洞室 |
| 2 | 11.00 | 115.00 | 5.00 | 0.10 | 23.00 | 5.70 | 0.26 | 1 | 李家峡水电站地下洞室 |
| 3 | 34.00 | 150.00 | 5.40 | 0.23 | 27.78 | 7.80 | 0.70 | 1 | 鲁布革水电站地下隧洞 |
| 4 | 13.90 | 124.00 | 4.20 | 0.11 | 29.52 | 2.00 | 0.19 | 1 | 括苍山隧道 |
| ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ | ⋮ |
| 307 | 65.20 | 114.38 | 5.79 | 0.57 | 19.74 | 0.62 | 1.27 | 4 | 西康铁路秦岭隧道 |
| 308 | 85.50 | 144.91 | 7.89 | 0.59 | 18.35 | 0.81 | 1.31 | 4 | 新建川藏铁路某隧道 |
| 309 | 78.10 | 144.62 | 7.26 | 0.54 | 19.92 | 0.71 | 2.17 | 4 | 秦岭终南山公路隧道 |

2 岩爆数据处理

岩爆数据处理包含如下步骤:首先,降维处理岩爆数据集,消除指标间的交叉关联影响;其次,对降维处理后的数据进行数据离群值的处理,消除少数离群值对模型学习效果的影响;最后,平衡化处理数据,提高模型收敛与学习效果。

2.1 数据降维

引入LLE法降维处理原始岩爆数据。LLE是将整体指标降维成综合指标的一类降维算法,即利用原始数据对新特征不同的贡献度来构成新的特征。

2.1.1 算法原理

算法步骤为:①寻找每个样本点的 k 个近邻点;②由每个样本点的近邻点计算出该样本点的权重系数;③将高维权重系数对应的线性关系与降维后的低维保持一致。

运用 k 近邻算法确定近邻点个数,则对每个样本点可以用其近邻样本点的线性组合来表示,即:

$$x_i = \sum_{j=1}^k w_{ij} x_j \quad (2)$$

式中: x_i 为第 i 个样本点; w_{ij} 为重构系数; x_j 为 x_i 的 k 个近邻点 ($1 \leq j \leq k$)。

求解如下优化问题可得到重构系数 $J(w)$:

$$J(w) = \sum_{i=1}^m x_i - \sum_{j \in Q(i)} w_{ij} x_j^2 \quad (3)$$

$$\sum_{j \in Q(i)} w_{ij} = 1 \quad (4)$$

式中: $Q(i)$ 为 i 个样本点的集合; m 为样本点数。

得到重构系数后,将数据集 X 映射低维空间 Y 中,保持同样的重构关系,即:

$$J(Y) = \sum_{i=1}^m y_i - \sum_{j=1}^m w_{ij} y_j^2 \quad (5)$$

通过求解上述问题就可得到映射后的数据集 Y 。

2.1.2 算法参数

算法需调整的参数主要包括近邻点个数 k 与需要降维的维数 n 。 k 是指选择 k 个近邻点来表示数据集 X 中每个样本 x_i , n 是指通过算法所降至的维数。参数选择不同, LLE 降维的结果也会不同。

1) n 的选择。引入重构误差概念, 重构误差是指重建的数据点与原始数据点之间的欧式距离之和。重构误差越小, 说明重构后的数据与原始数据差距越小, 降维效果越好。重构误差可以表示为:

$$\gamma = \min \sum \sqrt{(x - \hat{x})^2} \quad (6)$$

式中: x 为原始数据; \hat{x} 为降维后的数据; γ 为重构误差最小值。

基于表 1 原始岩爆数据, 取 $k=20$, 基于 LLE 算法可得到不同维度 n 的重构误差值, 见表 2。由表 2 可知: $n=4$ 时重构误差最小, 说明由原始数据 7 维降到 4 维能更好地保留原始数据中的信息。

表 2 重构误差

Table 2 Reconstruction error

| n | 最小重构误差值 |
|-----|------------------------|
| 2 | 1.53×10^{-16} |
| 3 | 8.13×10^{-17} |
| 4 | 4.86×10^{-17} |
| 5 | 5.15×10^{-17} |
| 6 | 9.09×10^{-16} |

2) k 的选择。如果近邻数 k 大于输出数据的维度, 上述的权重系数不是满秩, 因此, 采用变种的 LLE 找到最近 k 近邻的同时考虑近邻的分布权重, 不管降到几维前后的权重系数都是保持不变的。基于表 1 数据, 取 $n=4$, 以降至 4 维中的 2 个指标 (lle1, lle2) 为例, 通过 python 中的 matplotlib 模块显示 k 值不同的降维效果, 如图 2 所示, 其中, 灰色颜色越深代表岩爆烈度越高。

由图 2 可知: 随着 k 值的增大, 降维效果越好,

但 k 值越大, 计算量会越大; 当 k 达到 20 时, 各等级数据点分布已经较均匀, 因此, 结合计算量考虑, 取 $k=20$ 较为合理。

2.1.3 Person 关联性分析

降维的主要目的是为了消除指标间的交叉关联, 为评估降维的效果, 采用 Person 关联性分析降维前后各指标, 得到相关系数热力图如图 3 所示。从图 3 可以看出, 降维前各指标之间存在明显的关联, 降维后的指标之间关联程度极低。

2.2 离群样本处理

引入 DBSCAN 方法并分别采用整体去除和分级去除的处理方式处理降维后岩爆数据集。其中, 整体去除是指不区分岩爆烈度等级去除离群数据, 分级去除是指按岩爆烈度等级分级去除离群数据, 以期找出混杂在岩爆数据内部的离群点。

2.2.1 DBSCAN 算法原理

DBSCAN 算法是为识别样本空间内部低密度的异常样本^[25]。DBSCAN 算法流程如下:

1) 选择一个数据点作为起点, 再寻找与这个数据点指定范围内的所有数据点。

2) 假如指定范围内的数据点的数量超过指定阈值, 则将这些数据点标为核心点。

3) 对于所有标记的核心点, 将其距离指定范围内的数据点归为同一族内, 如果 2 个核心点之间存在重叠的数据点, 则将它们归为同一族中。

4) 对于所有的非核心点, 将它们标记为噪点。

DBSCAN 算法的主要参数是半径 (eps) 和最少核心点数 (points), 经过多轮参数优化, 最终选定 eps 为 0.6 和 points 为 8。

2.2.2 处理效果

为直观比较整体去除与分级去除 2 种处理方式对离群样本的去除效果, 基于 LLE 降维后的岩爆数据, 选取 lle1 与 lle2 指标的二维散点图作为离群点去除效果的对比示例, 如图 4、图 5 所示。

由图 4、图 5 可知: 整体去除方式处理后, 识别出 21 条离群样本; 通过分级去除方式处理后, 去除 37 条离群样本, 其中, 无岩爆去除 12 例, 弱岩爆去除 9 例, 中等岩爆去除 12 例, 强岩爆去除 4 例, 相比整体去除方式多去除 16 例。

2.3 数据平衡化

数据不平衡会严重影响预测模型的准确率, 因此, 在建立岩爆预测模型之前, 要将数据平衡化。引入 SMOTE 算法对训练集进行数据平衡化。SMOTE

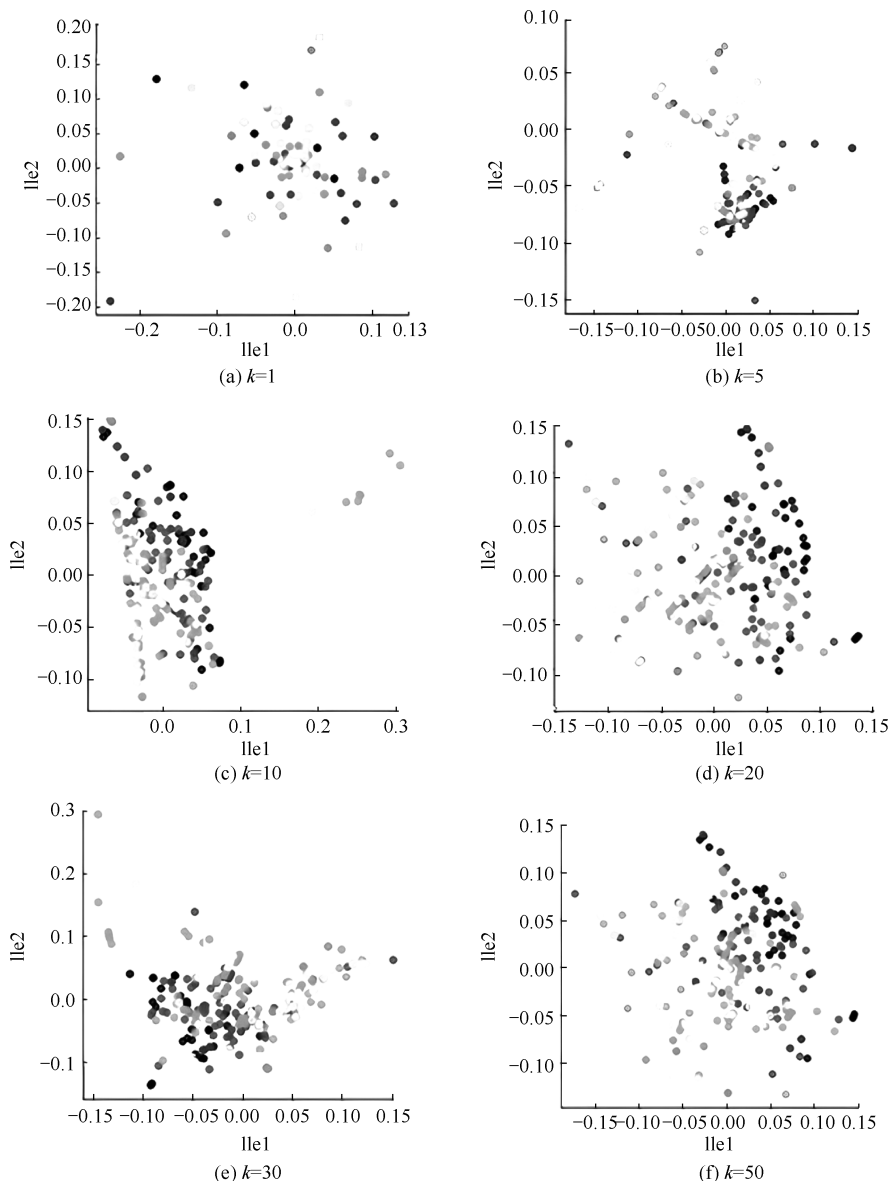


图2 不同 k 值下降维效果的散点图

Fig. 2 Scatter plot of dimensionality reduction effects with different k values

的核心是在少数类别的样本之间通过插值来产生额外的样本,进而补充少数类别样本^[14]。将开发环境 Pycharm 中训练集设置成整体数据的 80%,测试集设置成整体数据的 20%,按比例随机抽取出训练集的数据 216 条,测试集的数据 56 条,且保持测试集中每一等级岩爆数据均为 14 条。为防止模型信息泄露,数据平衡化只针对训练集。由于训练集中中等岩爆样本最多,为 86 条,因此,通过过采样处理,其他等级样本均增加至 86 条。

3 数据处理效果的模型验证

采用非线性学习性能优秀的 DT、RF 与 GBDT 这 3 类算法构建岩爆预测模型。通过对比采用原始

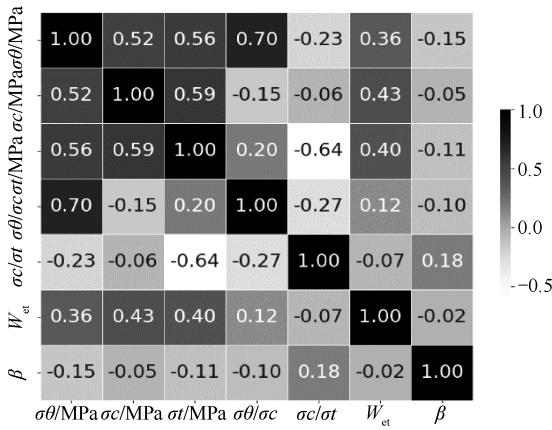
数据集以及经过数据处理后的数据集所训练出的岩爆预测模型的预测效果,说明数据处理的必要性与效果。

3.1 降维的模型验证

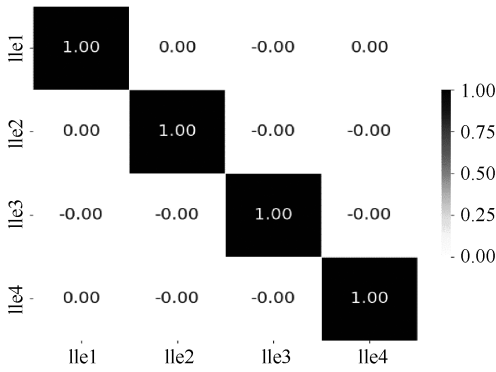
采用不同降维数 n 的 3 类岩爆预测模型的准确率如图 6 所示。由图 6 可知:3 类模型都是在 $n = 4$ 时预测准确率最高,分别为 78.5%,80.3%,82.1%,效果最好,结论与 2.1.2 节的重构误差值分析结论一致; $n = 4$ 的 3 类模型预测准确率较原始数据 $n = 7$ 的同类模型的准确率分别提高 8.9%、5.3%、7.1%。

3.2 去除离群样本的模型验证

同样采用上述 3 类模型,基于 $n = 4$ 的降维数据



(a) 降维前



(b) 降维后

图 3 指标相关系数热力图

Fig. 3 Indicator correlation coefficient heat map

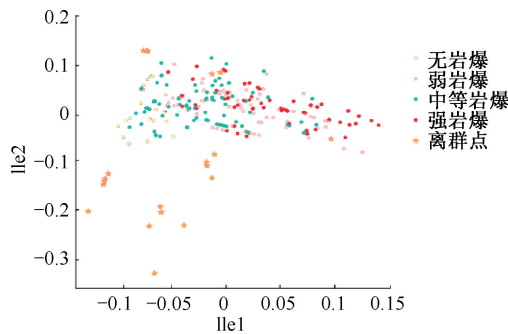
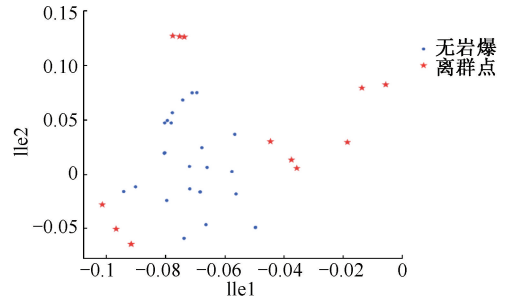


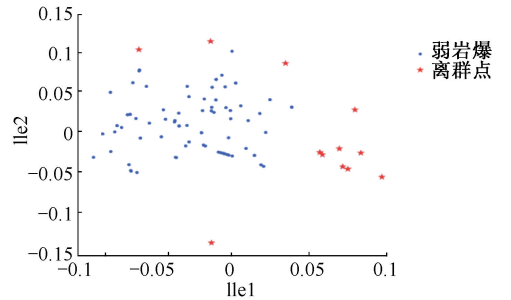
图 4 整体方式去除离群值

Fig. 4 Outlier removal using a holistic manner

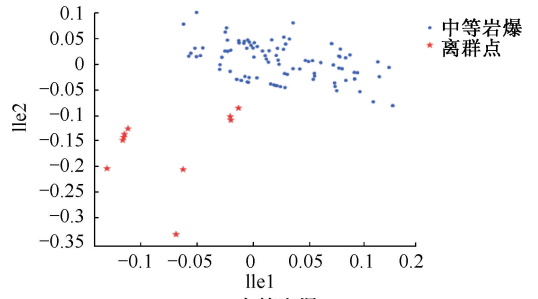
集,采用整体与分级去除方式处理离群数据,对比 3 类模型的预测准确率,如图 7 所示。由图 7 可知:不去除离群值的 3 类模型的预测准确率均为同类模型中的最低;整体去除方式的 3 种模型的预测准确率得到了提升;分级去除离群值的 3 类模型预测准确率则为最高,分别为 82.1%、85.7%、87.5%。对比同类仅做 LLE 降维($n=4$)处理的模型,其预测准确率分别提高 3.6%、5.4%、5.4%。



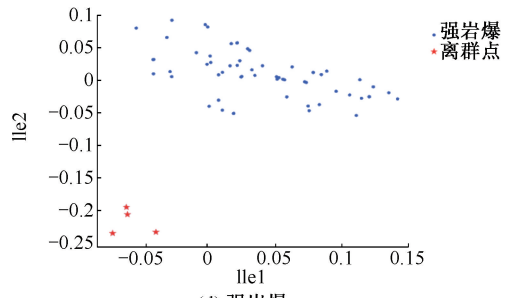
(a) 无岩爆



(b) 弱岩爆



(c) 中等岩爆



(d) 强岩爆

图 5 分级方式去除离群值

Fig. 5 Outlier removal using a graded approach

4 工程验证

工程验证数据来自江边水电站引水隧洞岩爆数据,该隧洞开挖长度 8 600 m,岩爆主要发生的位置在拱顶和拱肩,岩性有黑云母花岗岩、黑云母石英片岩,最大的爆坑深度为 4 m,洞径为 7 m×8.4 m 平底马蹄形段。现场部分岩爆数据集见表 3。

表 4 是采用表 1 原始数据集和经文中提出的处理方法处理后的数据集训练模型得到预测结果,不

表 3 江边水电站引水隧洞指标参数

Table 3 Index parameters of diversion tunnel of Jiangbian Hydropower Station

| 里程 | σ_θ/MPa | σ_c/MPa | σ_t/MPa | σ_θ/σ_c | σ_c/σ_t | W_{et} | $\beta/(\text{MPa} \cdot \text{m}^{-1})$ | 岩爆烈度 |
|---------|----------------------------|-----------------------|-----------------------|--------------------------|---------------------|----------|--|------|
| 0 + 250 | 33.15 | 106.94 | 5.84 | 0.31 | 18.31 | 2.15 | 0.24 | 2 |
| 0 + 550 | 91.43 | 157.63 | 11.96 | 0.58 | 13.17 | 6.27 | 1.32 | 3 |
| 0 + 900 | 51.50 | 132.05 | 6.33 | 0.39 | 20.86 | 4.63 | 0.96 | 2 |
| 2 + 230 | 23.39 | 106.32 | 2.92 | 0.22 | 36.41 | 1.75 | 0.70 | 1 |
| 3 + 320 | 12.96 | 117.81 | 3.21 | 0.11 | 36.70 | 2.49 | 0.09 | 1 |
| 3 + 710 | 89.52 | 146.75 | 7.54 | 0.61 | 19.46 | 4.70 | 3.11 | 3 |
| 7 + 330 | 121.09 | 159.33 | 11.29 | 0.76 | 14.11 | 11.62 | 1.94 | 4 |
| 7 + 625 | 121.09 | 135.67 | 9.02 | 0.45 | 15.04 | 11.2 | 0.11 | 4 |

表 4 基于原始数据集和处理后数据集的模型预测结果

Table 4 Model prediction results based on original and processed datasets

| 模型类别 | | 0 + 250 | 0 + 550 | 0 + 900 | 2 + 230 | 3 + 320 | 3 + 710 | 7 + 330 | 7 + 625 | 准确率/% |
|----------|------|---------|---------|---------|---------|---------|---------|---------|---------|-------|
| 基于原始数据集 | DT | 2 | 3 | 3 | 4 | 1 | 3 | 4 | 4 | 75.0 |
| | RF | 2 | 3 | 3 | 4 | 1 | 3 | 4 | 4 | 75.0 |
| | GBDT | 2 | 3 | 3 | 4 | 1 | 3 | 4 | 4 | 75.0 |
| 基于处理后数据集 | DT | 2 | 3 | 2 | 1 | 1 | 3 | 4 | 4 | 100.0 |
| | RF | 2 | 3 | 2 | 1 | 1 | 3 | 4 | 2 | 87.5 |
| | GBDT | 2 | 3 | 2 | 1 | 1 | 3 | 4 | 4 | 100.0 |

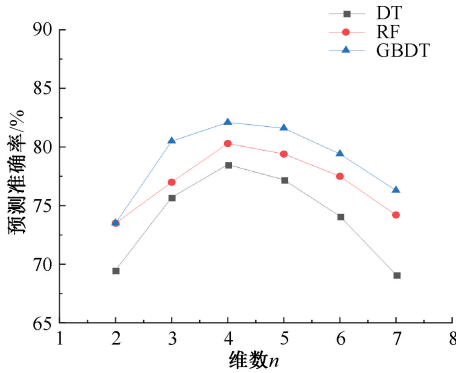


图 6 不同降维数的模型预测效果

Fig. 6 Model prediction performance of different dimensionality reduction

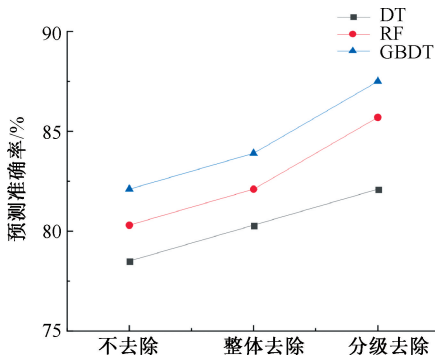


图 7 不同离群值去除方式的模型预测效果

Fig. 7 Model prediction performance of different outlier removal methods

难发现,数据处理后模型获得的预测结果与现场岩爆数据集的情况更符合,显著提高了模型的预测精度。

5 结 论

1) 基于 LLE-DBSCAN-SMOTE 的数据处理方法,能够有效减弱岩爆案例特征交叉关联、数据离群值与数据不平衡对岩爆预测效果的影响。

2) 通过 LLE 方法降维岩爆数据,有效提高 DT、RF 与 GBDT 这 3 类岩爆预测模型的性能,其预测准确率相较于原始数据模型分别提高 8.9%、5.3%、7.1%。

3) 通过 DBSCAN 法分级去除降维后的岩爆数据集的离群样本,提高了 3 类岩爆预测模型的性能,DT、RF 与 GBDT 模型预测准确率相较于仅做降维处理模型分别提高 3.6%、5.4%、5.4%。

4) 采用处理后数据训练获得的 DT、RF 与 GBDT 模型预测江边水电站引水隧洞岩爆实测样本的准确率分别达到 100%、87.5%、100%,相较于原始数据模型预测准确率提高近 25%。

参 考 文 献

- [1] 冯夏庭, 肖亚勋, 丰光亮. 岩爆孕育过程研究[J]. 岩石力学与工程学报, 2019,38(4): 650-673.
FENG Xiating, XIAO Yaxun, FENG Guangliang. Study on the development process of rockbursts [J]. Chinese Journal of Rock Mechanics and Engineering, 2019,38 (4): 650-673.
- [2] 江飞飞, 周辉, 刘畅, 等. 地下金属矿山岩爆研究进展及预测与防治[J]. 岩石力学与工程学报, 2019,38(5): 956-972.
JIANG Feifei, ZHOU Hui, LIU Chang, et al. Research progress, prediction and prevention of rock bursts in underground metal mines [J]. Chinese Journal of Rock Mechanics and Engineering, 2019,38 (5): 956-972.
- [3] ZHANG Junfei, WANG Yuhang, SUN Yuantian, et al. Strength of ensemble learning in multiclass classification of rockburst intensity[J]. International Journal for Numerical and Analytical Methods in Geomechanics, 2020,44(13): 1 833-1 853.
- [4] CHEN Guangyao, WANG Mingwu, YAN Jiahui, et al. A connection cloud model coupled with improved conflict evidence fusion method for prediction of rockburst intensity[J]. IEEE Access, 2021,9: 113 535-113 549.
- [5] WOJTECKI A, IWASZENKO S, APEL D B, et al. An attempt to use machine learning algorithms to estimate the rockburst hazard in underground excavations of hard coal mine[J]. Energies, 2021,14(21): 6 920-6 928.
- [6] 谭文侃, 胡南燕, 叶义成, 等. 基于四大集成学习的岩爆烈度分级预测[J]. 岩石力学与工程学报, 2022, 41(增 2): 3 250-3 259.
TAN Wenkan, HU Nanyan, YE Yicheng, et al. Rockburst intensity classification prediction based on four ensemble learning[J]. Chinese Journal of Rock Mechanics and Engineering, 2022,41 (S2): 3 250-3 259.
- [7] 夏元友, 张宏伟, 吝曼卿, 等. 基于数据预处理技术并考虑围岩应力梯度影响的隧洞岩爆预测[J]. 岩土工程学报, 2023,45(10): 1 987-1 994.
XIA Yuanyou, ZHANG Hongwei, LIN Manqing, et al. Prediction of tunnel rockbursts based on data preprocessing technology considering influences of stress gradient of surrounding rock[J]. Chinese Journal of Geotechnical Engineering, 2023,45(10): 1 987-1 994.
- [8] YIN Xin, LIU Quansheng, PAN Yucong, et al. Strength of stacking technique of ensemble learning in rockburst prediction with imbalanced data: comparison of eight single and ensemble models[J]. Natural Resources Research, 2021,30(2): 1 795-1 815.
- [9] GONG Fengqiang, WANG Yunliang, LUO Song. Rockburst proneness criteria for rock materials: review and new insights[J]. Journal of Central South University, 2020,27(10): 2 793-2 821.
- [10] 满轲, 武立文, 刘晓丽, 等. 基于灰色关联分析和 SSA-RF 模型的岩爆等级预测[J]. 金属矿山, 2023,34(5): 202-212.
MAN Ke, WU Liwen, LIU Xiaoli, et al. Rockburst grade prediction based on grey correlation analysis and SSA-RF model[J]. Metal Mines, 2023,34(5): 202-212.
- [11] GUO Deping, CHEN Hemaoyang, TANG Libin, et al. Assessment of rockburst risk using multivariate adaptive regression splines and deep forest model[J]. Acta Geotechnica, 2022,17(4): 1 183-1 205.
- [12] 李明亮, 李克钢, 秦庆词, 等. 岩爆烈度等级预测的机器学习算法模型探讨及选择[J]. 岩石力学与工程学报, 2021,40(增 1): 2 806-2 816.
LI Mingliang, LI Kegang, QIN Qingci, et al. Discussion and selection of machine learning algorithm model for rockburst intensity grade prediction[J]. Chinese Journal of Rock Mechanics and Engineering, 2021,40 (S1): 2 806-2 816.
- [13] 侯克鹏, 孙华芬, 包广拓. 改进的 MVO-GRNN 神经网络岩爆预测模型研究[J]. 安全与环境学报, 2023,41(3): 1-12.
HOU Kepeng, SUN Huafen, BAO Guangtuo. Research on improved MVO-GRNN neural network rockburst prediction model [J]. Journal of Safety and Environment, 2023,41(3): 1-12.
- [14] 谢学斌, 李德玄, 孔令燕, 等. 基于 CRITIC-XGB 算法的岩爆倾向等级预测模型[J]. 岩石力学与工程学报, 2020,39(10): 1 975-1 982.
XIE Xuebin, LI Dexuan, KONG Lingyan, et al. Rockburst propensity prediction model based on CRITIC-XGB algorithm[J]. Chinese Journal of Rock Mechanics and Engineering, 2020,39 (10): 1 975-1 982.
- [15] 陈则黄, 李克钢, 李明亮, 等. 基于 PCA-SOFM 模型的岩爆烈度等级预测[J]. 地下空间与工程学报, 2022, 18(增 2): 934-942.

- CHEN Zehuang, LI Kegang, LI Mingliang, et al. Prediction of rock burst intensity level based on PCA-SOFM model[J]. Chinese Journal of Underground Space and Engineering, 2022, 18 (S2): 934-942.
- [16] 杨小彬, 裴艳宇, 程虹铭, 等. 基于 SOFM 神经网络模型的岩爆烈度等级预测方法[J]. 岩石力学与工程学报, 2021, 40(增 1): 2 708-2 715.
- YANG Xiaobin, PEI Yanyu, CHENG Hongming, et al. Prediction method of rockburst intensity grade based on SOFM neural network model[J]. Chinese Journal of Rock Mechanics and Engineering, 2021, 40 (S1): 2 708-2 715.
- [17] 张传庆, 卢景景, 陈琚, 等. 岩爆倾向性指标及其相互关系探讨[J]. 岩土力学, 2017, 38(5): 1 397-1 404.
- ZHANG Chuanqing, LU Jingjing, CHEN Jun, et al. Discussion on rock burst proneness indexes and their relation[J]. Rock and Soil Mechanics, 2017, 38 (5): 1 397-1 404.
- [18] 夏元友, 刘昌昊, 刘夕奇, 等. 均布与梯度应力加载路径下岩爆破坏特征试验[J]. 中国安全科学学报, 2020, 30(5): 149-155.
- XIA Yuanyou, LIU Changhao, LIU Xiqi, et al. Experimental study on rockburst characteristics under uniform and gradient stress loading paths[J]. China Safety Science Journal, 2020, 30 (5): 149-155.
- [19] 宋英华, 庞昭胜, 李墨潇, 等. 基于反赋权与 MBCT-SR 多维云模型算法岩爆预测研究[J]. 中国安全生产科学技术, 2022, 18(3): 40-46.
- SONG Yinghua, PANG Zhaosheng, LI Moxiao, et al. Research on rock burst prediction based on anti weighting and MBCT-SR multidimensional cloud model algorithm [J]. Journal of Safety Science and Technology, 2022, 18 (3): 40-46.
- [20] 黄建, 夏元友, 吝曼卿. 基于改进组合赋权的岩爆多维云模型预测研究[J]. 中国安全科学学报, 2019, 29(7): 26-32.
- HUANG Jian, XIA Yuanyou, LIN Manqing. Study on prediction of rock burst by multi-dimensional cloud model based on improved combined weight [J]. China Safety Science Journal, 2019, 29 (7): 26-32.
- [21] 胡建华, 黄鹏莅, 周坦, 等. 岩爆倾向性的改进有限云评价模型与工程应用[J]. 中国安全科学学报, 2022, 32(2): 90-98.
- HU Jianhua, HUANG Pengli, ZHOU Tan, et al. Improved finite cloud evaluation model for rock burst tendency and its engineering application [J]. China Safety Science Journal, 2022, 32 (2): 90-98.
- [22] 白明洲, 王连俊, 许兆义. 岩爆危险性预测的神经网络模型及应用研究[J]. 中国安全科学学报, 2002, 12(4): 65-69.
- BAI Mingzhou, WANG Lianjun, XU Zhaoyi. Neural network model and application research for predicting rock burst risk [J]. China Safety Science Journal, 2002, 12 (4): 65-69.
- [23] 何佳其, 吝曼卿, 刘夕奇, 等. 引入梯度应力的岩爆预测方法[J]. 岩土工程学报, 2020, 42(11): 2 098-2 105.
- HE Jiaqi, LIN Manqing, LIU Xiqi, et al. New method for introducing gradient stress into rock-burst prediction[J]. Chinese Journal of Geotechnical Engineering, 2020, 42(11): 2 098-2 105.
- [24] ZHANG Hongwei, XIA Yuanyou, LIN Manqin, et al. A three-step rockburst prediction model based on data preprocessing combined with clustering and classification algorithms [J]. Bulletin of Engineering Geology and the Environment, 2024, 83: 1-19.
- [25] 王艳芳. 基于微震监测数据的岩爆智能预警方法研究[D]. 沈阳: 东北大学, 2019.
- WANG Yanfang. Research on intelligent early warning method for rock burst based on microseismic monitoring data [D]. Shenyang: Northeastern University, 2019.

作者简介: 范成强 (1997—), 男, 安徽合肥人, 硕士研究生, 主要研究方向为岩爆预测。
E-mail: 1909636694@qq.com。

