doi: 10.3969/j.issn.1005-7854.2026.01.016
引用格式：张化进，宋方家，孙妍，李兵磊，蔡和，龙翼. 基于CatBoost算法的矿井风速传感器布设优化[J]. 矿冶，2026，35(1)：122-129.
ZHANG Huajin, SONG Fangjia, SUN Yan, LI Binglei, CAI He, LONG Yi. Optimization of mine wind speed sensors layout based on CatBoost algorithm[J]. Mining and Metallurgy, 2026, 35(1): 122-129.

基于CatBoost算法的矿井风速传感器布设优化

张化进[1] 宋方家[1] 孙 妍[1] 李兵磊[1] 蔡 和[2] 龙 翼[3]
（1. 福州大学 紫金地质与矿业学院，福州 350108；
2. 贵州紫金矿业股份有限公司，贵州 贞丰 562200；
3. 紫金矿业集团股份有限公司，福建 上杭 364200）

摘 要：传感器可为矿井通风系统的安全运行和管理提供有效参考数据。为解决矿井通风故障诊断中的传感器布设问题，基于集成学习理论，将CatBoost算法引入至矿井通风故障诊断中，提出基于CatBoost算法的矿井通风故障诊断与传感器布设优化方法。以矿井通风网络分支风量和故障位置分别作为输入和目标标签，利用基于决策树的集成学习算法CatBoost进行通风网络故障诊断和特征选择，在保证模型准确率的基础上，合理选择重要网络分支布设风速传感器，减少传感器的布设数量以降低通风系统监测成本。分别以经典T型角联（算例1）和基于矿井实际布局的室内实验（算例2）通风网络为例，对比分析了3种常见机器学习算法（逻辑斯蒂回归、支持向量机、人工神经网络），验证基于CatBoost集成学习方法的可行性和优越性。两个算例的研究结果表明：随着训练分支数量的增多，通风网络故障准确率先增大后保持不变，由此可见选择重要分支布设风速传感器便可有效监控整个通风系统。对比常见机器学习模型，CatBoost模型预测效果明显优于其它模型，可采用最少的传感器数量获得最高的故障诊断准确率。因此，基于CatBoost算法的风速传感器布设优化方法，能够合理确定传感器布设数量及其位置，有效节约矿井通风网络监测成本，符合矿井绿色集约式发展趋势。

关键词：矿井通风；机器学习；特征选择；CatBoost；集成学习
中图分类号：TD72 **文献标志码：**A **文章编号：**1005-7854（2026）01-0122-08

Optimization of mine wind speed sensors layout based on CatBoost algorithm

ZHANG Huajin[1] SONG Fangjia[1] SUN Yan[1] LI Binglei[1] CAI He[2] LONG Yi[3]
(1. Zijin College of Geology and Mining, Fuzhou University, Fuzhou 350108, China;
2. Guizhou Zijin Mining Co. Ltd., Zhenfeng 562200, Guizhou, China;
3. Zijin Mining Group Co. Ltd., Shanghang 364200, Fujian, China)

Abstract: Sensors can provide effective reference data for the safe operation and management of mine ventilation systems. In order to solve the problem of sensor layout in mine ventilation fault diagnosis, based on the ensemble learning theory, the CatBoost algorithm is introduced into the mine ventilation fault diagnosis, and a mine ventilation fault diagnosis and sensor layout optimization method based on CatBoost algorithm is proposed. Taking the air volume of the branch of the mine ventilation network and the location of the fault as the input and target labels, respectively, the fault diagnosis and feature selection of the ventilation network are carried out by using the ensemble learning algorithm CatBoost based on the decision tree, and on the basis of ensuring the accuracy of the model, the important network branches are reasonably selected to deploy wind speed sensors, and the number of sensors is reduced to lower the monitoring cost of the ventilation system. Taking the classical angular (Case 1) and the indoor

收稿日期：2024-12-10
基金项目：福建省自然科学基金资助项目(2022J01567)
第一作者：张化进，博士，讲师，主要从事岩石力学、矿井通风方向研究。E-mail：zhanghuajin@fzu.edu.cn
通信作者：龙翼，博士，正高级工程师；E-mail：1158247306@qq.com

experiment (Case 2) based on the actual layout of the mine as examples, three common machine learning algorithms (Logistic Regression, Support Vector Machine, and Artificial Neural Network) are compared and analyzed to verify the feasibility and superiority of the CatBoost ensemble learning method. The results of the two examples show that with the increase of the number of training branches, the accuracy of the ventilation network fault increases at first and then remains constant, which shows that the wind speed sensor can be effectively monitored by selecting important branches. Compared with common machine learning models, the CatBoost model has a significantly better prediction performance than other models, and can use the minimum number of sensors to obtain the highest fault diagnosis accuracy. Therefore, the optimization method of wind speed sensors deployment based on CatBoost algorithm can reasonably determine the number and location of sensor deployment, effectively save the cost of mine ventilation network monitoring, and conform to the development trend of green intensive mine development.

Key words: mine ventilation; machine learning; feature selection; CatBoost; ensemble learning

矿井通风系统是地下开采活动中不可或缺的部分，承担着关键职责，为井下提供新鲜风流的同时，稀释有毒气体并排出被污染的空气，它的安全与稳定是井下生产施工的前提[1-5]。风速传感器布设是矿井通风环境与矿山安全运营的基础，但地下矿山井巷数量庞大，若在所有井巷部署传感器，则会造成传感器的冗余，导致高额投资成本和大量维护工作，不符合集约、高效、可持续矿山发展趋势[6]。为了更加科学合理地获取通风系统监测数据，对风速传感器位置和数量进行布设优化势在必行。

在传感器布设优化方面，国内外学者对风速传感器的优化布置已经进行了一些研究。COHEN 等[7]提出通过对井下监测数据的分析，实时调整传感器位置布设，动态调节优化传感器配置；STEWART 等[8]提出了一种瞬态模拟计算机模型技术，通过预测传感器下游通风状态与时间的关系，最大限度地减少所需的传感器数量，反映整个矿山模型所有位置的通风状况；刘剑等[9]提出利用支持向量机（SVM）和遗传算法（GA）等机器学习算法诊断矿井通风系统阻变型故障位置及故障量，采用邻域粗糙集方法优化风速传感器的布局，实验证明了上述风速传感器布设优化方法的有效性；潘竞涛等[10]对风速传感器的数量及位置问题进行研究并将其应用到大明矿的传感器优化布置中，给出了大明矿风速传感器的布置方案；倪景峰等[11]以决策树基尼系数为传感器优化布置选择标准，解决了风速传感器布设优化与诊断模型不匹配的问题；黄德等[12]提出基于多目标优化的观测特征选择方法，有效地剔除冗余或不相关传感器，提高矿井通风网络故障诊断的模型性能。这些研究成果有效降低了传感器的布设数量，节省矿井通风网络监测的成本，在风速传感器布设优化中具有较高的研究和应用价值。

传感器布设优化本质是进行特征选择，即选择有效网络分支进行通风系统故障诊断。基于集成学习模型的特征选择，将通风网络故障预测准确率作为优化目标，降低维度的同时，提高模型的泛化能力和计算效率。CatBoost 是由 Yandex 开发的一种开源集成学习算法，可以处理多种数据类型，帮助解决各种问题，是一种高效的梯度提升算法，CatBoost 采用特殊的对称树结构及 Ordered Boosting 技术，简化模型的同时提高预测速度以及模型的鲁棒性[13,14]，是一个强大的机器学习算法，在易用性、准确性、高效性等方面表现出色，现已成为重要且应用广泛的人工智能技术。但迄今为止，CatBoost 方法在传感器布设优化领域未见报道与应用。鉴于此，本文以矿井通风网络分支风量和网络故障位置作为输入特征和目标标签，利用基于对称决策树为基学习器的集成学习算法 CatBoost 进行通风网络故障诊断和特征选择，在保证模型准确率的基础上，合理选择重要的网络分支布设风速传感器，减少传感器的布设数量以降低通风系统监测成本。

1 CatBoost 算法与原理

1.1 集成学习算法

集成学习通过调用一些简单的机器学习算法[15,16]，构建出大量具有差异性的基学习器，然后采用某种策略将多个弱基学习器结合起来组成一个强学习器，形成更准确、稳健的集成模型。集成学习[17]具有以下三方面优点：1）提高预测准确性和鲁棒性，减少单个模型的过拟合风险；2）可处理复杂任务，通过结合多个模型来处理复杂的任务，提高整体性能；3）可提供对模型预测的解释，增加了对模型预测结

果的信任度和可解释性。

假设基学习器的错误率为ε，即对应每个基学习器h_i有：

$$P(h_i(x)\neq f(x))=\varepsilon \tag{1}$$

假设构建了T个基学习器，若有超过半数的基学习器正确则集成学习器正确：

$$H(x)=sign(\sum_{i=1}^{T}h_i(x)) \tag{2}$$

假设基学习器的错误率独立，那么由Hoeffding不等式，集成学习的错误率为：

$$P(H(x)\neq f(x))=\sum_{k=0}^{[T/2]}\binom{T}{k}(1-\varepsilon)^k\varepsilon^{T-k}\leqslant \exp(-\frac{1}{2}T(1-2\varepsilon)^2) \tag{3}$$

由此可见，集成学习的错误率随着基学习器数目T的增大呈指数下降，当基学习器误差相互独立时，错误率最终趋向于0。

1.2 CatBoost 算法

CatBoost算法是一种改进的梯度提升树（GBDT）实现，GBDT框架的基本算法是分类回归树，它假设前1轮基学习器训练结果为$F^{t-1}(x)$，并针对损失函数$L(y, F^{t-1}(x))$进行优化。每一轮中，算法旨在找到一个新的分类回归树强学习器，目标是最小化当前轮次的损失函数h^t。不断迭代此过程，直至生成一个性能优越的强学习器为止[17]。

$$h^t=\arg\min_{h\in H}EL((y,F^{t-1}(x)+h(x))) \tag{4}$$

与GBDT相比，CatBoost在处理标称属性和预测偏移问题方面进行了重要改进，从而减少过拟合的可能性，提升算法的准确性和泛化能力。标称属性通常采用独热编码将类别特征转化为数值特征，但独热编码易导致过拟合。为降低信息损失，CatBoost引入了一种新的类别特征处理方法，基于目标变量的统计进行分组，有效替代了原有的类别变量。为更好的处理类别特征，CatBoost在目标变量统计法的基础上增加了先验分布项p，形成贪婪目标变量统计法[18]，这种方法帮助减少噪声和低频率类别型数据对数据分布的干扰，进而防止过拟合的发生[19]。在贪婪目标变量统计法中，数值型特征被定义为目标统计量$\hat{x}_k^i$：

$$\hat{x}_k^i=\frac{\sum_{j=1}^{n}[x_j^i=x_k^i]\cdot y_j+ap}{\sum_{j=1}^{n}[x_j^i=x_k^i]+a} \tag{5}$$

式中，n为训练集样本数量；a为权重系数，a>0。

CatBoost算法相较于传统的GBDT算法，专为解决多分类问题开发，特别适用于处理类别型特征，并有效应对梯度偏差及预测偏移问题，通过迭代训练多个决策树，逐步提升模型的性能，从而增强算法的准确性和泛化效果。

2 基于CatBoost的矿井风速传感器布设优化方法

基于CatBoost算法的矿井风速传感器布设优化方法，流程如图1所示。具体步骤为：

1）利用矿井通风模拟软件或室内试验等手段构建相应通风系统模型，通过改变风阻等在井巷（分支）中设置故障，监测并记录井巷中的风量数据，构建通风网络故障诊断数据集。

2）以通风网络分支的监测风量作为输入特征，以故障位置为目标标签，将数据划分为训练集与测试集，训练集用于构建模型，测试集用于模型性能评估。

3）依托于训练集数据，利用CatBoost集成学习算法和网格搜索超参数优化，建立矿井通风网络故障诊断模型，并对分支进行重要性排序。

4）依据通风网络分支重要性顺序，依次增加CatBoost算法模型的分支数量，重新训练模型，获得不同分支数量下通风网络模型故障诊断准确率。

5）确定模型和传感器布设数量及其位置，监测监控整个矿井通风系统。

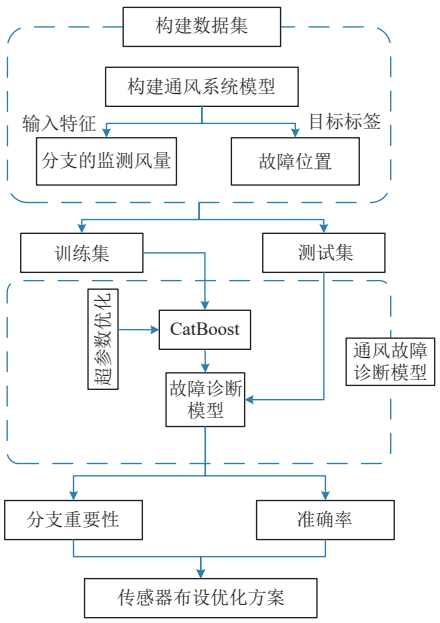


图1 矿井通风网络布设优化流程

Fig. 1 Optimization process for mine ventilation network layout

3 算例验证

为直观展示CatBoost集成学习方法在矿井通风网络故障诊断中的优势与特点，基于两个通风网络故障诊断算例，通过Python程序下的Scikit-learn、Pandas等工具搭建基于CatBoost集成学习的矿井通风网络故障诊断模型，进而对风速传感器布设进行优化。同时，对比分析了3种常见机器学习算法（逻辑斯蒂回归LR、SVM、人工神经网络ANN）的通风网络故障诊断和传感器布设优化结果。

3.1 算例 1—经典角联通风网络

以文献[20]中的T型经典角联通风网络为例，如图2所示。该通风网络包含8个节点（①~⑧），10个分支（e_1~e_{10}）和3条角联分支，e_1为进风井，e_{10}为回风井。主扇所在分支编号为e_{10}，对应的节点编号为8。

基于上述T型角联通风网络，保持网络拓扑关系和风机动力不变，假设通风网络中的某一分支出现了阻变型故障（该分支风阻发生了增减变化），发生降阻变化，分支风阻变为$r_i'=r_i-k\times\Delta r_i$；发生增阻变化，

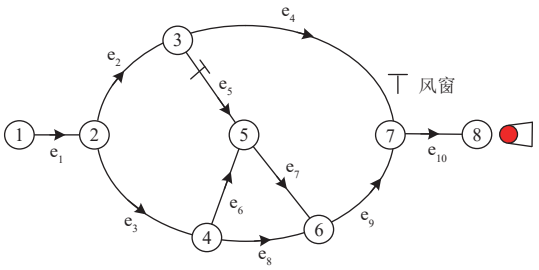


图2 算例1通风网络

Fig. 2 Ventilation network of Case 1

分支风阻为$r_i'=r_i+j\times\Delta r_i$（$\Delta r_i$为阻变量，$j$、$k$为任意自然数）。利用矿井通风仿真程序对通风网络进行风量解算，生成新的通风网络数据，并记录各个分支的风量，从而形成一个样本。对通风系统内的每个分支依次执行以上步骤，形成对应分支的故障样本集。值得注意的是，本例中未对总进风e_1分支、总出风e_{10}分支生成对应的故障样本，分支按照Δr_i=1发生故障，增阻样本j=（1，2，3，4，5，6），降阻样本r=（1，2，3，4），8个分支共计生成52个样本，同理以阻变量插值方式增补了95个样本，合计147个样本。受篇幅限制，本文仅列出其中8个样本数据，见表1。

表1 经典角联通风网络数据集

Table 1 Classical angular ventilation network dataset

序号	故障位置	风量/($m^3\cdot s^{-1}$)								
		e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	e_9
1	e_2	86.69	14.93	71.76	13.96	0.96	29.93	30.90	41.83	72.73
2	e_4	89.36	34.39	54.97	13.64	20.75	14.60	35.35	40.37	75.72
3	e_5	88.77	24.39	64.38	16.27	8.12	24.05	32.17	40.33	72.50
4	e_6	89.48	38.42	51.06	14.76	23.66	7.51	31.17	43.55	74.72
5	e_7	87.40	33.14	54.26	16.09	17.05	2.08	19.12	52.18	71.31
6	e_8	85.28	37.04	48.24	16.47	20.57	31.31	51.88	16.93	68.81
7	e_9	53.92	22.84	31.08	22.72	0.11	13.08	13.19	18.01	31.20
8	e_9	62.83	26.18	36.66	21.86	4.31	13.82	18.13	22.84	40.97

然后，根据第2节的流程与步骤，首先将通风网络风量和故障位置分别作为模型的输入和输出标签，将数据集样本按照4∶1随机划分成训练集和测试集。训练集用来训练CatBoost集成学习模型，结合网格搜索算法对模型进行超参数优化。测试集用来评估CatBoost集成学习模型的性能效果。根据模型训练结果，基于CatBoost集成学习算法的故障诊断准确率为1，可有效预测矿井通风网络故障位置。在此基础上，根据模型特征重要性对各分支的贡献率进行排序，结果如图3所示。分支重要性顺序为e_8、e_3、e_6、e_2、e_4、e_7、e_1、e_5、e_9。其中分支e_3与e_8的特征重要性分

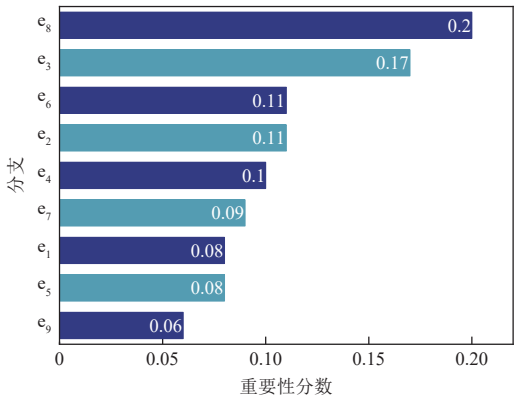


图3 算例1分支重要性

Fig. 3 The branching importance of Case 1

数明显比其余分支大，由此可见，这2个分支对于通风网络故障诊断模型的贡献更大，更为重要，布设传感器时应优先考虑。

根据分支重要性顺序，以最重要的分支e_8为基础，依次增加分支数量，分别构建了LR、SVM、ANN机器学习模型和CatBoost集成学习模型，获得不同分支数量下的模型故障诊断准确率，如图4所示。可知，只采用分支e_8训练模型时，LR、SVM、ANN和CatBoost模型的通风网络故障的准确率分别为0.70、0.80、0.78和0.83，随着参与训练的分支数量增加，模型故障诊断准确率显著提升至1后保持不变；当训练分支数量增加到2时，SVM与CatBoost模型故障诊断准确率达1后不变；当训练分支数量增加到4时，LR与ANN模型故障诊断准确率也达1后不变；相对其余3种机器学习算法，CatBoost集成学习风速传感器布设优化效果显然最好，通风网络故障诊断准确率和传感器布设优化效率均较高。

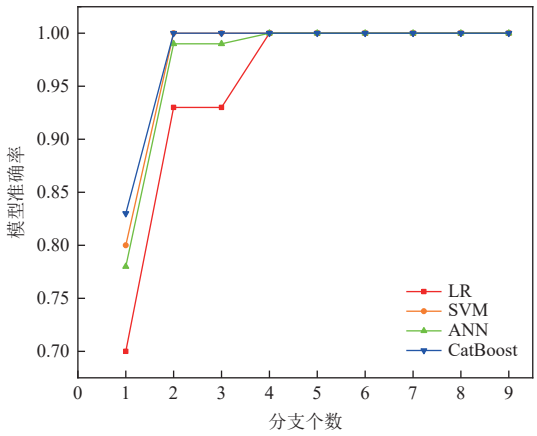


图4　算例1传感器布设数量优化

Fig. 4　Optimization of the number of sensors deployed in the Case 1

3.2　算例2—室内实验通风网络

为进一步探讨本文所提方法的可行性和适用性，采用基于矿井实际布局的室内通风网络实验模型进行验证。结合文献[21]建立的通室内通风网络实验模型，采用中央并列式和抽出式通风方式，主副斜井进风，回风立井布置1台2.5 kW轴流式通风机进行回风，实验模型通风网络如图5所示。其中e_1~e_{62}为网络分支，①~㊳为网络节点。为确保数据的全面性和代表性，风量监测点的选取基于通风网络的整体性和多样性，避免监测数据冗余或偏差。在实验过程中，保持风机频率50 Hz不变，采用蝶阀替代风窗、风门，模拟不同故障场景下的风阻变化，选取e_{55}、e_{56}分支模拟风门异常打开故障，e_{16}、e_{58}分支模拟风门打开情况下工作面风流短路故障，分别从0度到100度调节e_{16}、e_{55}、e_{56}、e_{58}分支蝶阀开度，但不包括其初始角度0度，记录e_1、e_2、e_7、e_9、e_{10}、e_{20}、e_{24}、

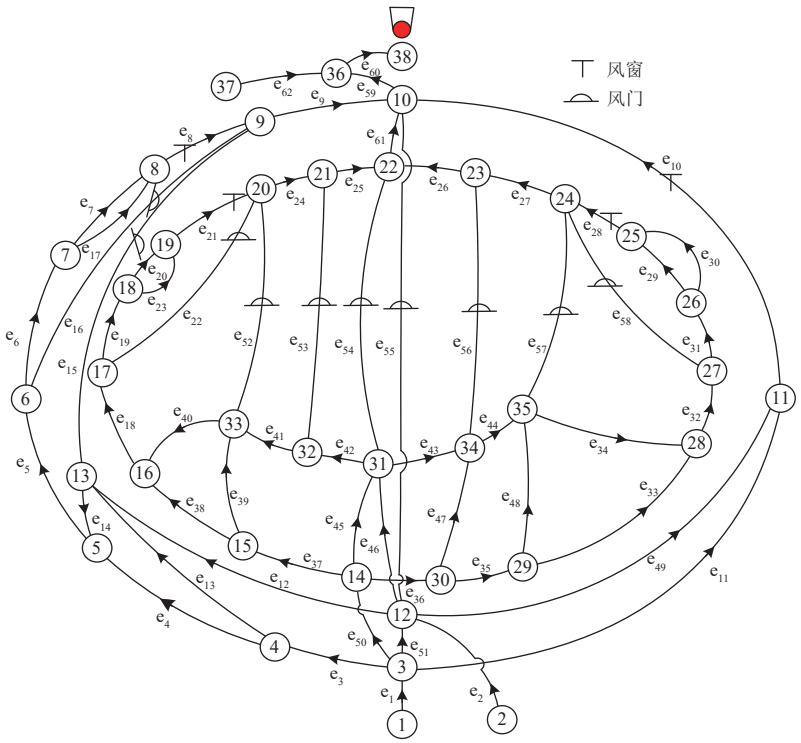


图5　算例2通风网络

Fig. 5　Ventilation network of Case 2

e_{27}、e_{29}、e_{60} 10个分支的风速值。在此基础上，选取e_8、e_{21}分支安装调节风窗，模拟风量不足故障，调节e_8、e_{21}分支蝶阀开度从0度到50度，不包括其角度50度，记录上述10个分支的风速值，并将上述风速数据换算成风量。最终每种故障位置处收集监测分支的风量数据各30组，共收集风量数据180组，构成通风网络故障诊断和传感器布设优化数据集，受篇幅限制，本文仅列出其中8个样本数据，见表2。

表2　室内实验通风网络数据集

Table 2　Indoor experimental ventilation network dataset

序号	故障位置	风量/($m^3 \cdot min^{-1}$)									
		e_1	e_2	e_7	e_9	e_{10}	e_{20}	e_{24}	e_{27}	e_{29}	e_{60}
1	e_8	7.35	7.32	4.06	4.19	3.79	5.40	5.69	3.99	3.65	14.78
2	e_8	7.31	4.43	4.27	4.85	4.43	3.93	4.33	4.33	3.93	14.07
3	e_8	7.34	7.30	4.37	2.88	2.41	5.91	6.25	4.30	3.90	14.78
4	e_{16}	7.31	4.34	4.34	4.62	4.15	3.99	4.40	4.40	3.99	14.48
5	e_{16}	7.31	4.20	4.20	5.12	3.92	3.87	4.26	4.25	3.86	14.64
6	e_{55}	7.31	4.38	3.87	4.06	3.70	3.59	3.93	3.92	3.59	14.57
7	e_{56}	7.32	4.40	3.96	4.18	3.80	3.43	3.73	3.86	3.54	14.68
8	e_{58}	7.31	4.27	4.27	4.50	4.06	3.85	4.73	4.31	3.91	14.75

基于上述数据集，依据第2节的流程与步骤，同样将通风网络风量和故障位置分别作为模型的输入和输出标签，按照4∶1随机划分数据集为训练集和测试集结合网格搜索，训练和测试CatBoost集成学习模型。根据模型训练结果，基于CatBoost集成学习算法的故障诊断准确率为1，可有效预测矿井通风网络故障位置。在此基础上，同算例1，得出各分支重要性，结果如图6所示。分支重要性顺序为e_9、e_{24}、e_{60}、e_2、e_{10}、e_7、e_{20}、e_{29}、e_{27}、e_1，其中e_9、e_{24}明显比其他分支大，布设传感器时应优先考虑。

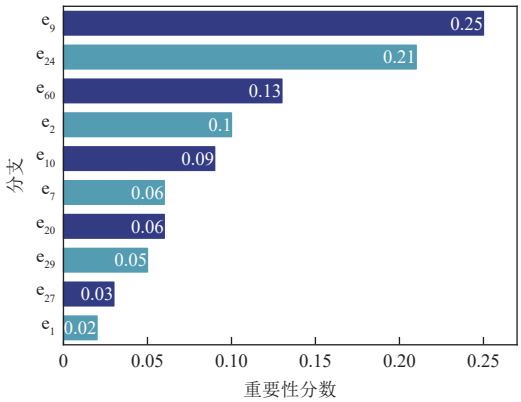


图6　算例2分支重要性

Fig. 6　The branching importance of Case 2

根据分支重要性顺序，以最重要的e_9为基础，依次增加分支数量，分别构建了LR、SVM、ANN机器学习模型和CatBoost集成学习模型，获得不同分支数量下的模型故障诊断准确率，如图7所示。可知，只采用分支e_9训练模型时，LR、SVM、ANN和CatBoost模型的通风网络故障的准确率分别为0.42、0.38、0.49和0.50，随着参与训练的分支数量增加，模型故障诊断准确率逐渐提升至最大值后保持不变，当训练分支数量增加到3时，SVM模型故障诊断准确率达0.89后保持不变；当训练分支数量增加到5时，CatBoost模型故障诊断准确率达1后保持不变；当训练分支数量增加到7时，ANN模型故障诊断准确率达0.73后保持不变；当训练分支增加到9时，LR模型故障诊断准确率达0.75后保持不变。相对其余3种机器学习算法，CatBoost集成学习通风网络故障诊断准确率最高，同时传感器布设优化效率也较为突出。

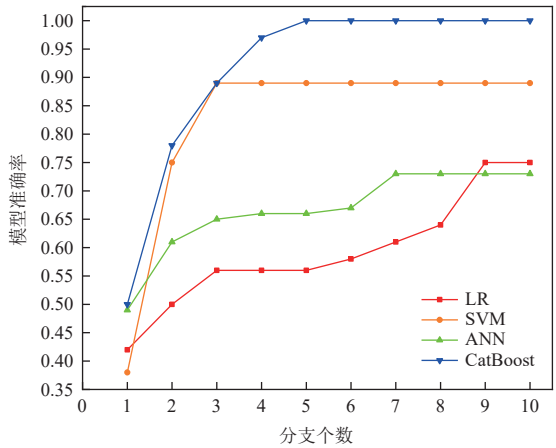


图7　算例2传感器布设数量优化

Fig. 7　Optimization of the number of sensors deployed in the Case 2

综上所述，相较于传统机器学习，CatBoost作为一种先进的集成学习算法，独特的训练与学习策略使其准确性高、性能卓越，能够出色地完成矿井通风系统故障诊断和传感器布设优化。算例1只需在e_3、e_8

各安装1个风速传感器，算例2在e_9、e_{24}、e_{60}、e_2、e_{10}各安装1个风速传感器，辅以CatBoost集成学习算法，便可对整个通风网络进行有效故障诊断，降低传感器的布设数量，降低生产成本。

4 结论

1）矿井风速传感器的布设数量及其位置决定着故障诊断模型的预警效果，基于集成学习理论，将CatBoost算法引入到矿井通风故障诊断中，提出了基于CatBoost集成学习的矿井通风故障诊断与传感器布设优化方法。

2）CatBoost可反映通风网络分支的重要性，进而对风速传感器进行布设优化。在T型经典角联通风网络中，e_3、e_8分支较为重要；在基于矿井实际布局的室内通风网络中，e_9、e_{24}、e_{60}、e_2、e_{10}相对较为重要，表明这些分支对监测监控通风网络具有较大影响。

3）两个算例研究结果表明，参与模型训练的分支数量与通风网络故障诊断准确率成正相关，随着分支数量的增加，故障准确率持续提升后保持不变，说明只需选择重要分支布设风速传感器便可有效监控整个通风系统。

4）通过对比分析3种常见机器学习，CatBoost模型表现最为优越，可采用最少的传感器数量获得最高的故障诊断准确率，能够有效节约矿井通风网络监测成本，符合矿井绿色集约式发展趋势。

参考文献

[1] 王翔飞，谢贤平，许旭．某铁矿通风系统监测点选址布局研究[J]. 有色金属(矿山部分)，2021，73(1)：47-52.
WANG X F, XIE X P, XU X. Research on site selection and layout of monitoring points for a certain iron mine ventilation system[J]. Nonferrous Metals (Mine Section), 2021, 73(1): 47-52.

[2] JIA J Z, JIA P, LI Z. Theoretical study on stability of mine ventilation network based on sensitivity analysis[J]. Energy Science and Engineering, 2020, 8(8): 2823-2830.

[3] 程纬华，乔登攀，张磊，等．BP神经网络在露天矿边坡稳定性分析中的应用[J]. 矿冶，2012，21(2)：10-14.
CHENG W H, QIAO D P, ZHANG L, et al. Application of BP networks in the stability analysis of slopes in the open-pit mine[J]. Mining and Metallurgy, 2012, 21(2): 10-14.

[4] 张庆华，姚亚虎，赵吉玉．我国矿井通风技术现状及智能化发展展望[J]. 煤炭科学技术，2020，48(2)：97-103.
ZHANG Q H, YAO Y H, ZHAO J Y. Status of mine ventilation technology in China and prospects for intelligent development[J]. Coal Science and Technology, 2020, 48(2): 97-103.

[5] 李军，王海军，赵旭林，等．基于Ventsim的超深部多中段同时开采通风系统优化[J]. 矿冶，2024，33(3)：356-361.
LI J, WANG H J, ZHAO X L, et al. Optimization study of the ventilation system for ultra-deep multi-level simultaneous mining based on Ventsim[J]. Mining and Metallurgy, 2024, 33(3): 356-361.

[6] 孙继平，唐亮，李春生，等．风量比例法在甲烷传感器优化布置中的应用[J]. 煤炭学报，2008，33(10)：1126-1130.
SUN J P, TANG L, LI C S, et al. Application of air-volume proportion rule in optimal placement of methane sensor in mine[J]. Journal of China Coal Society, 2008, 33(10): 1126-1130.

[7] COHEN A F, FISHER T J, WATSON R A, et al. Location strategy for methane, air velocity, and carbon monoxide fixed-point mine-monitoring transducers[J]. IEEE Transactions on Industry Applications, 1987, IA-23(2): 375-381.

[8] STEWART C M, AMINOSSADATI S M, KIZIL M S. Use of live sensor data in transient simulations of mine ventilation models[J]. Mining Report, 2017, 153(4): 215-222.

[9] 刘剑，蒋清华，刘丽，等．矿井通风系统阻变型故障诊断及风速传感器位置优化[J]. 煤炭学报，2021，46(6)：1907-1914.
LIU J, JIANG Q H, LIU L, et al. Resistance variant fault diagnosis of mine ventilation system and position optimization of wind speed sensor[J]. Journal of China Coal Society, 2021, 46(6): 1907-1914.

[10] 潘竟涛，赵丹，李宗翔，等．大明矿通风系统故障源诊断及风速传感器的布置[J]. 煤炭学报，2013，38(增刊1)：167-172.
PAN J T, ZHAO D, LI Z X, et al. Fault source diagnosis for ventilation system and air velocity transducer placement in Daming Mine[J]. Journal of China Coal Society, 2013, 38(Sup.1): 167-172.

[11] 倪景峰，乐晓瑞，常立峰，等．基于决策树的矿井通风阻变型故障诊断及传感器优化布置[J]. 中国安全科学学报，2021，17(2)：34-39.
NI J F, LE X R, CHANG L F, et al. Resistance variant fault diagnosis and optimized layout of sensors for mine ventilation based on decision tree[J]. China Safety Science Journal, 2021, 17(2): 34-39.

[12] 黄德，刘剑，刘永，等．矿井通风阻变故障观测特征组合选择试验研究[J]. 煤炭学报，2021，46(12)：3922-3933.
HUANG D, LIU J, LIU Y, et al. Experimental research on combination selection of observation feature of resistance variation fault in mine ventilation[J]. Journal of China Coal Society, 2021, 46(12): 3922-3933.

[13] 苗丰顺，李岩，高岑，等．基于CatBoost算法的糖尿病预测方法[J]. 计算机系统应用，2019，28(9)：215-218.

MIAO F S, LI Y, GAO C, et al. Diabetes prediction method based on CatBoost algorithm[J]. Computer Systems and Application, 2019, 28(9): 215-218.

[14] PROKHORENKOVA L, GUSEV G, VOROBEV A, et al. Catboost: unbiased boosting with categorical features[J]. Advances in Neural Information Processing Systems, 2018, 31: 6638-6648.

[15] 赵锴，叶丹．基于机器学习的矿床描述文本多标签分类[J]. 中国矿业，2024, 33(10): 153-161.
ZHAO K, YE D. Multi-label classification for mineral deposit description texts based on machine learning[J]. China Mining Magazine, 2024, 33(10): 153-161.

[16] 崔丽珍，曹坚，李丹阳，等．融合机器学习算法的煤矿井下信道建模研究[J]. 中国矿业，2021, 30(11): 68-74.
CUI L Z, CAO J, LI D Y, et al. Research on the channel modeling in coal mine based on machine learning algorithm[J]. China Mining Magazine, 2021, 30(11): 68-74.

[17] ZHANG H J, WU S C, ZHANG Z X. Prediction of uniaxial compressive strength of rock via genetic algorithm—selective ensemble learning[J]. Natural Resources Research, 2022, 31(3): 1721-1737.

[18] MICCI-BARRECA D. A preprocessing scheme for high-cardinality categorical attributes in classification and prediction problems[J]. ACM SIGKDD Explorations Newsletter, 2001, 3(1): 27-32.

[19] 谭勇，陈记，杨忠民，等．基于CatBoost集成学习的边坡稳定性预测方法[J]. 科学技术与工程，2024, 24(30): 13153-13160.
TAN Y, CHEN J, YANG Z M, et al. Slope stability prediction method based on CatBoost ensemble learning[J]. Science Technology and Engineering, 2024, 24(30): 13153-13160.

[20] HUANG D, LIU J, DENG L J. A hybrid-encoding adaptive evolutionary strategy algorithm for windage alteration fault diagnosis[J]. Process Safety and Environmental Protection, 2020, 136: 242-252.

[21] 张浪，张迎辉，张逸斌，等．基于机器学习的通风网络故障诊断方法研究[J]. 工矿自动化，2022, 48(3): 91-98.
ZHANG L, ZHANG Y H, ZHANG Y B, et al. Research on fault diagnosis method of ventilation network based on machine learning[J]. Journal of Mine Automation, 2022, 48(3): 91-98.