

文章编号: 1671-251X(2023)10-0133-09

DOI: 10.13272/j.issn.1671-251x.2022120047

## 煤巷支护参数预测研究

陈攀<sup>1,2</sup>, 马鑫民<sup>1</sup>, 向俊杰<sup>1</sup>, 陈莉影<sup>1</sup>, 梁厅皓<sup>1</sup>

(1. 中国矿业大学(北京)力学与土木工程学院, 北京 100083;

2. 云南省水利水电勘测设计院有限公司, 云南 昆明 650021)

**摘要:** 目前支持向量机(SVM)和随机森林(RF)等算法在煤矿巷道支护领域应用较少。研究了不同的机器学习模型进行支护参数设计的适用性,以建立一个更高性能模型来实现锚杆支护的合理、科学设计。首先建立煤巷支护智能预测数据库:采用现场调研、问卷调查和文献检索等方式收集煤矿巷道样本;采用缺失值填补、箱形图修改离群点和局部异常因子剔除等方式对数据进行处理,建立煤巷支护数据库。提出一种基于合成少数类过采样(SMOTE)-遗传算法(GA)-SVM的煤巷支护参数预测模型:将数据库中的数据分成训练集与测试集,采用SMOTE技术平衡训练样本,提高模型对少数类样本的拟合能力;训练过程采用GA对SVM的超参数进行全局寻优,进一步提高模型整体性能。测试结果表明,SMOTE-GA-SVM模型分类精度达到83.8%,比传统的SVM模型提高了21.8%。将SVM、人工神经网络(ANN)、RF、AdaBoost(ADA)和朴素贝叶斯分类器(NBC)等机器学习方法引入到煤巷锚杆支护参数预测中,建立对应的支护参数预测模型,比较结果表明:从最优到最差的预测模型排序分别为SMOTE-GA-SVM、RF、GA-ANN、SVM、NBC和ADA,6种模型的平均分类精度达69.9%,验证了机器学习方法在煤巷锚杆支护参数预测方面的可行性。在山西霍宝干河煤矿有限公司对SMOTE-GA-SVM模型进行了应用,模型预测准确率达87.5%,具有较强的适用性和可靠性。

**关键词:** 煤矿巷道; 机器学习; 锚杆支护参数; 合成少数类过采样; 遗传算法优化支持向量机

中图分类号: TD353

文献标志码: A

### Research on prediction of support parameters for coal roadways

CHEN Pan<sup>1,2</sup>, MA Xinmin<sup>1</sup>, XIANG Junjie<sup>1</sup>, CHEN Liying<sup>1</sup>, LIANG Tinghao<sup>1</sup>

(1. School of Mechanics and Civil Engineering, China University of Mining and Technology-Beijing,

Beijing 100083, China; 2. Yunnan Institute Co., Ltd. of Water & Hydropower Engineering Investigation Design, Kunming 650021, China)

**Abstract:** Currently, algorithms such as support vector machine (SVM) and random forest (RF) are less applied in the field of coal mine roadway support. The paper studies the applicability of different machine learning models for support parameter design. Thus a higher performance model would be established to achieve reasonable and scientific design of anchor bolt support. Firstly, it is suggested to establish an intelligent prediction database for coal mine roadway support. Through on-site research, questionnaire survey, and literature search, the coal mine roadway samples are collected. The data is processed using methods such as filling in missing values, modifying outliers in box charts, and removing local abnormal factors to establish a coal roadway support database. The paper proposes a coal roadway support parameter prediction model based on synthetic minority oversampling technique (SMOTE) - genetic algorithm (GA) - SVM. The data in the database is divided into

收稿日期: 2022-12-15; 修回日期: 2023-09-20; 责任编辑: 王晖, 郑海霞。

基金项目: 国家自然科学基金资助项目(52074301)。

作者简介: 陈攀(1998—), 男, 云南曲靖人, 硕士, 主要从事巷道支护和水利勘察工作, E-mail: 18811432245@163.com。通信作者: 马鑫民(1979—), 男, 山东菏泽人, 副教授, 主要研究方向为矿山工程爆破和巷道支护智能化技术, E-mail: mxm@cumt.edu.cn。

引用格式: 陈攀, 马鑫民, 向俊杰, 等. 煤巷支护参数预测研究[J]. 工矿自动化, 2023, 49(10): 133-141.

CHEN Pan, MA Xinmin, XIANG Junjie, et al. Research on prediction of support parameters for coal roadways[J]. Journal of Mine Automation, 2023, 49(10): 133-141.



扫码移动阅读

training and testing sets. The SMOTE technology is used to balance training samples, and improve the model's fitting capability for minority class samples. The training process uses GA to globally optimize the hyperparameters of SVM, further improving the overall performance of the model. The test results show that the classification precision of the SMOTE-GA-SVM model reaches 83.8%, which is 21.8% higher than the traditional SVM model. The machine learning methods such as SVM, artificial neural network (ANN), RF, AdaBoost (ADA), and naive Bayesian classifier (NBC) are introduced into the prediction of coal roadway anchor support parameters. The corresponding support parameter prediction models are established. The comparison results showed that the best to worst prediction models are ranked as SMOTE-GA-SVM, RF, GA-ANN, SVM, NBC, and ADA, with an average classification precision of 69.9%. The result verifies the feasibility of machine learning methods in predicting the parameters of coal roadway bolt support. The SMOTE-GA-SVM model is applied in Shanxi Huobaoganhe Coal Mine Co., Ltd., with a precision of 87.5% and strong applicability and reliability.

**Key words:** coal mine roadway; machine learning; anchor rod support parameters; oversampling of synthesized minority classes; genetic algorithm optimizing support vector machine

## 0 引言

巷道是煤矿井工开采的脉络,科学合理的煤巷围岩稳定性控制是保障煤矿安全高效开采的关键<sup>[1]</sup>。锚杆支护自 1956 年引入我国至今,得到了广泛的应用,已成为最主要的煤矿巷道支护方式<sup>[2]</sup>。随着煤矿开采深度、范围和强度的不断增加,巷道支护面临的条件逐渐复杂化,巷道支护设计繁琐、效果差等问题日益突出<sup>[3]</sup>。

近年来,越来越多的专家学者将计算机智能设计方法应用到煤巷支护领域,利用智能算法来实现巷道支护的合理、科学设计<sup>[4]</sup>。谢广祥等<sup>[5]</sup>提出了通过构建多级人工神经网络来确定锚杆支护的方式,优化了支护参数。王茂源<sup>[6]</sup>采用模糊聚类对围岩稳定性进行分类,采用神经网络算法实现了锚杆支护参数设计。王哲哲等<sup>[7]</sup>结合模糊理论与人工神经网络构建评价模型,对围岩稳定性进行分级,从而进行巷道支护方案优选。Xu Qingyun 等<sup>[8]</sup>基于人工神经网络预测模型对围岩进行分类,构建了决策系统的支持网络模型,通过数值模拟和现场测试验证了该模型的可行性。Ren Heng 等<sup>[9]</sup>为了评价神东矿区的围岩稳定性,添加动量项来修正 BP 神经网络的权值,进一步提升了模型精度。Zhang Xiliang 等<sup>[10]</sup>提出粒子群优化(Particle Swarm Optimization, PSO)算法与人工神经网络(Artificial Neural Network, ANN)相结合的 PSO-ANN 模型来评价和预测巷道稳定性,评价结果表明 PSO-ANN 模型可准确评估巷道的稳定性。

通过查阅大量参考文献,发现目前支持向量机(Support Vector Machine, SVM)、随机森林(Random Forest, RF)等算法在岩土工程领域取得了很好的应

用效果<sup>[11-14]</sup>,但在煤矿巷道支护领域应用较少。为了全面研究不同的机器学习模型进行支护参数设计的适用性,笔者建立了煤巷支护智能预测数据库,将 SVM、ANN、RF、AdaBoost(ADA)和朴素贝叶斯分类器(Naive Bayes Classifier, NBC)5 种机器学习方法引入煤巷锚杆支护参数预测中,建立评价体系对模型的性能进行评价,验证机器学习方法在煤巷锚杆支护参数预测方面的可行性。

## 1 智能算法原理

### 1.1 SVM

SVM 是一种针对小样本数据的机器学习模型<sup>[15]</sup>。以风险最小化原则将误差风险控制到最小,通过核函数将在低维线性不可分的数据映射到高维空间,使线性内积运算非线性化,从而在高维空间寻找最优分类超平面实现数据线性可分<sup>[16]</sup>。SVM 寻找最优超平面的过程实际上就是支持向量到超平面的间隔  $D$  最大化问题。

$$D = \min \frac{1}{2} \|\omega\|^2 + c \sum_{i=1}^n \xi_i \quad (1)$$

式中:  $\omega$  为特征空间中的权向量;  $c$  为惩罚系数;  $i$  为样本个数,  $i=1, 2, \dots, n$ ,  $n$  为样本总个数;  $\xi_i$  为损失函数。

SVM 通过核函数将数据从低维空间转换至高维空间,高斯核函数为空间中任一点  $x_i$  到某一中心  $x_j$  ( $j=1, 2, \dots, n, j \neq i$ ) 之间欧氏距离的单调函数,可记作  $K(x_i, x_j)$ ,常被用于高维度、线性不可分的数据。

$$K(x_i, x_j) = \exp\left(-\frac{\|x_i - x_j\|^2}{2g^2}\right) \quad g > 0 \quad (2)$$

式中  $g$  为高斯核函数参数。

惩罚系数  $c$  表征对离群点的重视程度,即模型对

错误分类样本的惩罚力度,  $c$  越大, 惩罚力度越大, 容易使模型过拟合;  $c$  取值太小, 会使模型过于简化, 对错误分类的样本学习训练不够, 导致模型欠拟合。高斯核函数参数  $g$  决定了数据映射到新特征空间后的分布, 通过函数的径向范围影响模型的计算速度。因此, 惩罚系数  $c$  和高斯核函数参数  $g$  决定了高斯核 SVM 模型的整体性能。

## 1.2 遗传算法(GA)对 SVM 超参数寻优

传统的 SVM 只能通过经验或试错法来进行惩罚系数  $c$  和高斯核函数参数  $g$  的选择, 只能寻找到局部最优参数, 不能完全将 SVM 的性能体现出来。遗传算法(Genetic Algorithm, GA)是通过计算机模拟自然界生物进化过程的一种随机化搜索算法, 通过对群体中具有某种特征的个体进行选择、交叉和变异操作, 生成新的群体, 逐渐逼近最优解<sup>[17]</sup>。

采用 GA 对 SVM 的惩罚系数  $c$  和高斯核函数参数  $g$  进行全局寻优。优化的主要流程如图 1 所示。GA 先对惩罚系数  $c$  和高斯核函数参数  $g$  进行编码, 解空间向编码空间映射; 生成初始种群, 定义适应度函数并计算个体初始适应度; 进行选择、交叉和变异操作, 形成新的个体, 并计算个体适应度, 重复进行这一操作直到满足终止条件; 完成解码, 获取最佳参数  $c_{best}$  和  $g_{best}$ 。

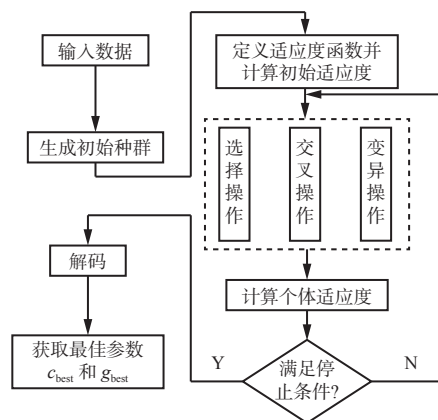


图 1 GA 对 SVM 超参数寻优流程

Fig. 1 GA optimization process for super parameters of SVM

## 2 煤巷支护数据库建立

### 2.1 评价指标的选取

根据巷道支护设计的需要, 选择顶板锚杆和帮部锚杆的直径、长度、间距、排距和数量, 将顶板锚索的直径、长度、数量和布置方式作为输出参数, 其中布置方式是锚索排距对顶板锚杆排距的倍数。煤矿巷道锚杆支护的影响参数涉及广泛且复杂, 将所有参数都输入到机器学习模型中容易加大模型的复杂度, 导致模型过度拟合而失去泛化性, 因此, 对

机器学习模型的输入参数进行精选十分必要。选择影响参数时遵循获取方便、可量化、物理意义明确、独立性、普适性等要求, 综合我国目前的煤巷生产情况和锚杆支理论, 从围岩参数、围岩节理裂隙发育程度和巷道埋深及断面尺寸 3 个方面选出 12 个参数作为机器学习的输入参数。

1) 围岩参数。对煤矿巷道进行支护的根本目的是保持巷道畅通和围岩稳定, 降低围岩移动量, 改善围岩力学性质。由此可见, 围岩参数对巷道变形与破坏的影响十分显著。本文选择煤层、基本顶、直接顶及直接底的厚度和强度表征巷道围岩特性。

2) 围岩节理裂隙发育程度。在围岩内部存在的节理、裂隙、层理等构造对岩体的整体强度存在不同程度的影响, 一般情况下, 节理、裂隙多的岩体完整性较差, 锚杆、锚索对岩体稳定性的影响很大。

3) 巷道埋深及断面尺寸。巷道埋深和断面尺寸对围岩支护有重要影响。垂直应力随着巷道开采深度的增加而逐渐增大, 高地应力情况下围岩更易发生塑性破坏, 保持围岩稳定性更加困难。另外, 巷道的宽度和高度对围岩稳定性的影响也不可忽视。

### 2.2 数据采集与处理

1) 数据采集。为了保证煤巷支护数据的丰富性和数据来源的可靠性, 采用现场调研、问卷调查和文献检索等方式收集 2010—2022 年的典型煤巷支护数据, 共 157 条。

2) 缺失值填补。在数据收集的过程中, 不可避免地形成少量的数据缺失, 在进行数据分析前参照相同矿场类似的巷道对缺失值进行填补。

3) 离群点处理。为了让样本满足机器学习模型训练和测试要求, 需统一数据的整体分布。通过数据箱形图找出在样本中出现明显偏离大多数观测值的个别值, 如图 2 所示。上四分位数加上 1.5 倍的四分位距为上限, 下四分位数减去 1.5 倍的四分位距为下限, 在上下限之外的点判定为离群点, 采用三角形将离群点标出来, 并采用极值化处理离群点: 偏大值修改为箱形图极大值, 偏小值修改为箱形图极小值。

4) 异常样本剔除。采用局部异常因子(Local Outlier Factor, LOF)对数据离群样本进行检测<sup>[18]</sup>。LOF 是一种基于样本局部密度检测识别离群样本的经典算法, 通过计算样本点的局部密度量化每一个样本点的异常程度, 样本点的异常程度取决于样本点与周围点的局部密度比较。LOF 算法步骤: ① 计算样本点  $p$  的第  $k$  距离邻域, 即与点  $p$  相距最近的  $k$  个点的集合。② 计算点  $p$  与点  $p$  的第  $k$  距离邻域内各点的可达距离, 其中, 点  $p$  到点  $o$  的可达距离为

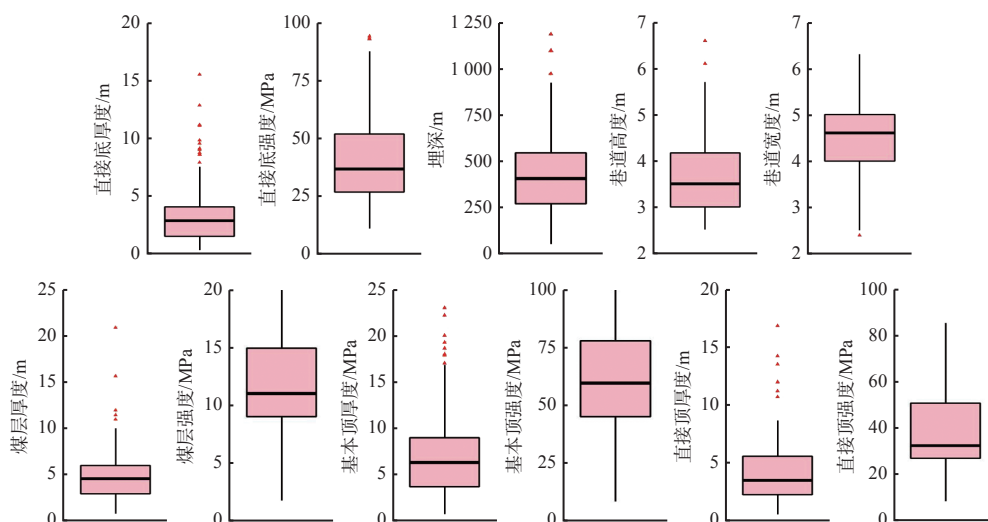


图 2 原始数据的箱形图

Fig. 2 Box diagram of the original data

点  $o$  的  $k$  邻近距离和点  $p$  与点  $o$  之间的直接距离的最大值。③ 计算点  $p$  的局部可达密度, 即样本点  $p$  与它的第  $k$  距离领域内各点的平均可达距离的倒数。④ 对每个点的第  $k$  局部可达密度进行比较并排序。样本点的第  $k$  局部可达密度越大, 表明它的异常程度越小, 反之, 异常程度越大。

在不同  $k$  值下使用 LOF 算法进行 7 次检测, 取每个  $k$  值下异常程度最高的 10 个样本, 结果见表 1。7 次检测中出现次数不少于 4 的样本被确定为异常样本。可看出 10 个异常样本为 33, 84, 105, 124, 151, 75, 14, 129, 25, 109, 将 10 个异常样本被剔除, 剩余 147 个样本组成煤巷支护智能设计数据库。

表 1 基于 LOF 的异常样本检测结果

Table 1 Test results of abnormal samples based on local outlier factor(LOF)

$k=4$	$k=5$	$k=6$	$k=7$	$k=8$	$k=9$	$k=10$
105	105	105	105	33	14	14
84	33	33	33	105	33	25
129	84	84	84	14	84	49
33	129	154	124	84	25	33
124	124	129	154	154	49	84
154	154	124	14	124	105	105
36	75	90	129	25	154	154
90	36	17	25	49	124	124
75	44	44	109	109	75	75
70	14	109	75	129	109	59

### 2.3 合成少数类过采样技术(SMOTE)平衡样本

为便于模型的学习训练和测试, 从数据库中随机抽取 80%(117 组)的数据作为训练集, 剩余 20%(30 组)的数据作为测试集。训练集中各支护参

数的频数统计结果见表 2、表 3 和表 4, 各输入参数的分布统计情况如图 3 所示。

表 2 顶板锚杆支护参数统计

Table 2 Statistics of roof anchor bolt support parameters

参数名称	参数值	频数	参数名称	参数值	频数
直径	18 mm	11	间距	700 mm	5
	20 mm	46		800 mm	39
	22 mm	60		900 mm	36
长度	2 000 mm	14		1 000 mm	16
	2 200 mm	17		1 100 mm	9
	2 400 mm	57		1 200 mm	12
	2 500 mm	20	排距	700 mm	10
	2 600 mm	9		800 mm	33
数量	4 根	13		900 mm	16
	5 根	28		1 000 mm	38
	6 根	53		1 100 mm	7
	7 根	23		1 200 mm	13

表 3 顶板锚索支护参数统计

Table 3 Statistics of roof anchor cable support parameters

参数名称	参数值	频数	参数名称	参数值	频数
直径	15.24 mm	14	长度	5 300 mm	20
	17.89 mm	36		6 300 mm	39
	18.7 mm	18		7 300 mm	22
	21.6 mm	17		8 300 mm	32
	22 mm	32		9 300 mm	4
数量	1 根	4	布置方式	1	24
	2 根	68		2	71
	3 根	39		3	22
	4 根	6			



表4 帮部支护参数统计

Table 4 Side support parameter statistics

参数名称	参数值	频数	参数名称	参数值	频数
直径	16 mm	7	间距	700 mm	10
	18 mm	20		800 mm	41
	20 mm	41		900 mm	21
	22 mm	49		1 000 mm	27
长度	1 800 mm	16	排距	1 100 mm	4
	2 000 mm	25		1 200 mm	14
	2 200 mm	12		700 mm	10
	2 400 mm	43	700 mm	800 mm	33
	2 500 mm	21		900 mm	15
数量	2 根	7		1 000 mm	39
	3 根	29		1 100 mm	7
	4 根	52		1 200 mm	13
	5 根	29			

由表2、表3和表4可知,支护参数是不均衡的类分布,存在某一类或某几类样本数量显著少于其他类别的情况,这会降低模型对少数样本的拟合。采用SMOTE对数据进行类平衡处理。SMOTE是基于随机过采样算法的一种改进方案,若2个同类样本间的欧氏距离足够近,则SMOTE假设这2个样本之间的样本与这2个样本同类<sup>[19]</sup>。SMOTE平衡样本流程如图4所示。根据样本不平衡比例确定采样倍率 $b$ ;随机选取一个少数类样本 $X$ ,计算与其他样本间的欧氏距离,并找出样本 $X$ 的 $k$ 个同类近邻样本;在每个少数类样本 $X$ 的 $k$ 个同类近邻样本中随机选取一个样本 $m$ ;在 $m$ 和 $X$ 之间的连线上随机选取一个点作为新的少数类样本 $X'$ ,重复以上步骤,直到少数类样本满足采样倍率 $b$ ,输出最终的数据集。

$$X' = X + r(0, 1)(X - m) \quad (3)$$

式中 $r(\cdot)$ 为随机函数。

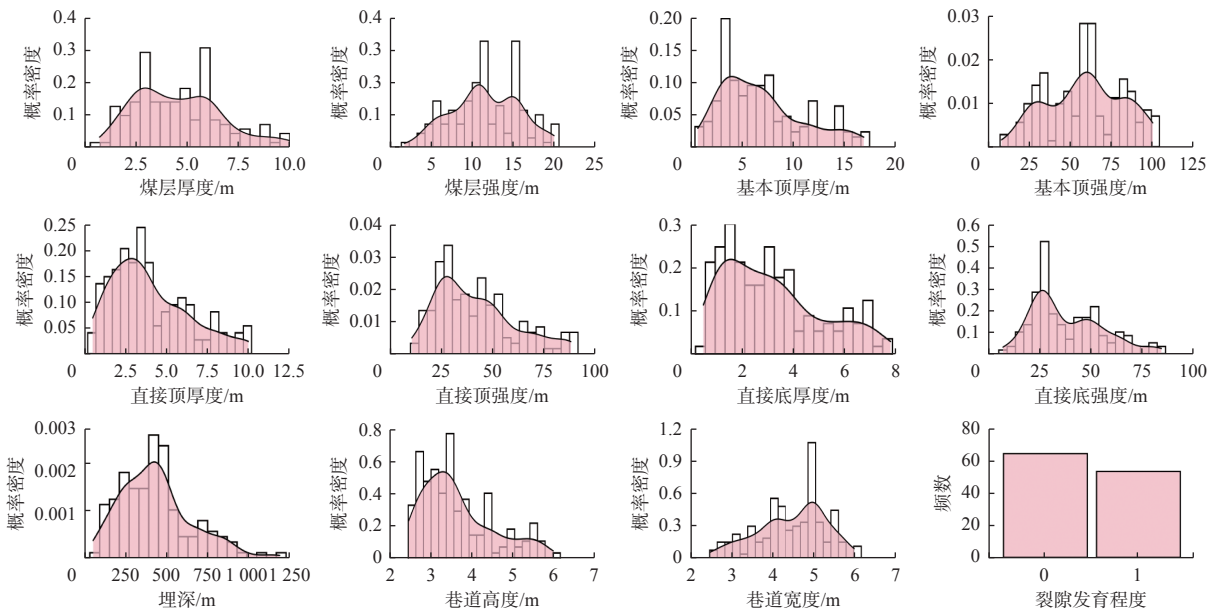


图3 训练集输入参数分布统计

Fig. 3 Distribution statistics of input parameters of training set

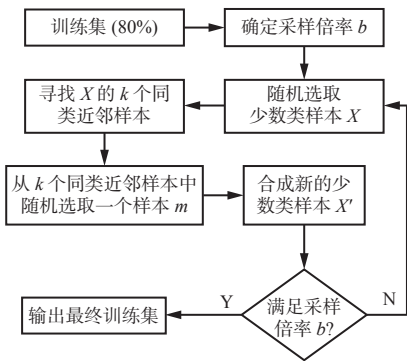


图4 SMOTE 平衡样本流程

Fig. 4 Sample balancing flow by SMOTE

## 2.4 数据标准化

最终训练集中的各个特征向量具有不同的量纲和单位,采用Z-score 标准化方法将数据转换为均值为0、标准差为1的数据。采用式(4)对最终训练集进行处理。

$$y = \frac{a_i - \bar{a}}{\sqrt{\frac{1}{n-1} \sum_{i=1}^n (a_i - \bar{a})^2}} \quad (4)$$

式中: $y$ 为标准化后的值; $a_i$ 为特征值; $\bar{a}$ 为特征均值。

煤巷支护数据库的整体流程如图5所示。首先

通过现场调研、问卷调查和文献检索等方式收集煤巷支护技术资料;然后采用缺失值填补、离群点处理及异常样本剔除等方式对数据进行清洗处理,建立煤巷支护数据库;最后按照 8:2 的比例将数据库中的数据分成训练集与测试集,并采用 SMOTE 平衡训练样本,经标准化处理后即可用来训练和测试模型。

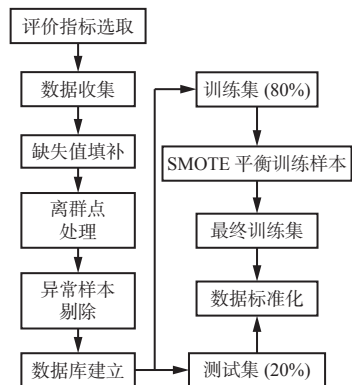


图 5 煤巷支护数据库的整体构建流程

Fig. 5 The overall building process of the coal roadway support database

### 3 煤巷支护参数预测模型建立与评价

#### 3.1 SMOTE-GA-SVM 模型

选用 GA 算法对 SVM 参数进行全局寻优<sup>[20]</sup>。GA 参数设置:种群规模为 50,直接复制到下一代的染色体数量为 20%,交叉概率为 0.8,变异概率为 0.01,繁殖次数为 300,采用精确度作为模型适应度函数。设置惩罚系数  $c$  和高斯核函数参数  $g$  的寻优范围分别为  $[0, 10]$  和  $[0, 200]$ 。GA 生成初始群体并计算初始适应度,然后通过选择、交叉、变异等操作,经过设定繁殖次数或个体适应度满足要求后终止计算,将得到的最大适应度作为全局最优解输出。通过 GA 优化后,确定各模型的最佳超参数组合  $c_{best}$  和  $g_{best}$ ,结果见表 5。

表 5 GA 全局寻优结果

Table 5 Global optimization results of GA

支护特征	$c_{best}$	$g_{best}$	支护特征	$c_{best}$	$g_{best}$
顶板锚杆直径	85.26	2.32	帮部锚杆间距	33.65	5.30
顶板锚杆长度	6.75	0.63	帮部锚杆排距	54.23	2.65
顶板锚杆间距	70.76	7.44	帮部锚杆数量	30.23	4.64
顶板锚杆排距	81.04	2.07	锚索直径	74.62	7.55
顶板锚杆数量	32.97	3.71	锚索长度	61.50	1.51
帮部锚杆直径	73.86	8.96	锚索数量	87.00	2.65
帮部锚杆长度	10.53	5.07	锚索布置	70.72	5.30

SMOTE-GA-SVM 支护参数预测模型建立流程

如图 6 所示。

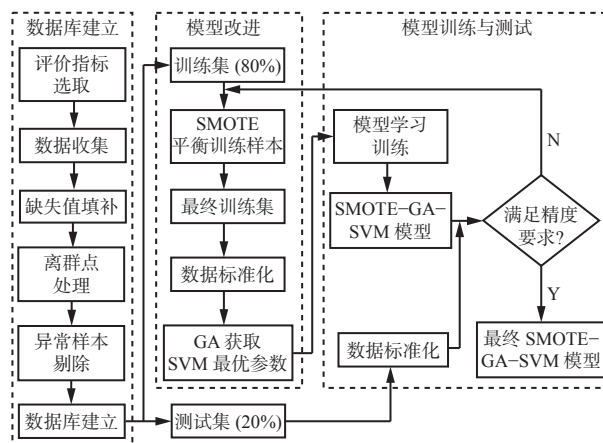


图 6 SMOTE-GA-SVM 支护参数预测模型建立流程

Fig. 6 SMOTE-GA-SVM supporting parameter prediction model establishment process

首先建立煤巷锚杆支护数据库;然后通过 SMOTE 平衡样本、数据标准化和 GA 超参数寻优等步骤优化改进 SVM 支护参数预测模型的性能;最后经过模型训练和测试建立满足精度要求的 SMOTE-GA-SVM 支护参数预测模型。

#### 3.2 RF 模型

RF 是一种集成学习算法,因其优越的性能成为一种流行的分类算法<sup>[21]</sup>。RF 的树构建过程允许特征之间的交互作用和高度相关性,可量化输入变量对于模型的重要程度。各指标对支护设计的重要程度排序如图 7 所示。可看出裂隙发育程度是最重要的指标,直接底厚度对模型的贡献度最小,这为支护设计时的变量选择提供了参考。

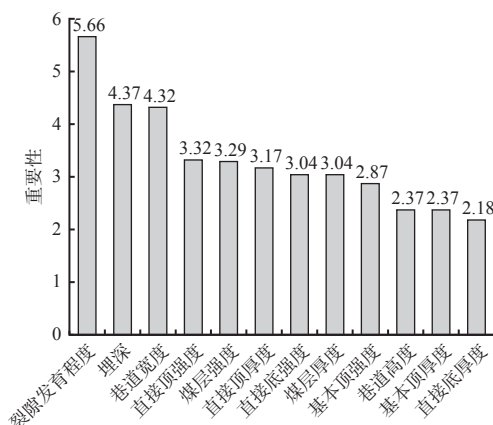


图 7 输入变量在支护参数预测模型上的重要性

Fig. 7 Importance of input variables in support parameter prediction model

#### 3.3 GA-ANN 模型

ANN 是应用最为广泛的机器学习算法之一,输入层、隐藏层和输出层共同组成了 ANN 结构,其中,输入层和输出层的结构由模型的输入参数和输出参

数决定,不可改变,调整隐藏层的结构能显著改变模型的性能。通过试错法确定了隐藏层结构为9-7时模型性能最好。为了提高网络性能,采用GA对ANN的权重和偏差进行优化。其中,GA参数设置与SMOTE-GA-SVM模型一致。另外,ANN模型采用支持回溯的弹性反向传播算法,设置误差学习率为0.01,以Sigmoid函数作为激活函数,模型最大迭

代次数为1 000。同时,对模型进行10倍交叉验证,每次模型的初始权重都不一样,一定程度上避免出现局部最优解。构建的顶板锚杆间距GA-ANN预测模型的网络拓扑结构如图8所示,其中, $I_1$ - $I_{12}$ 为输入指标, $H_1$ - $H_9$ 为隐藏层神经元, $O$ 为输出层, $B$ 为偏差,线条颜色和粗细分别代表不同的权值和偏差。

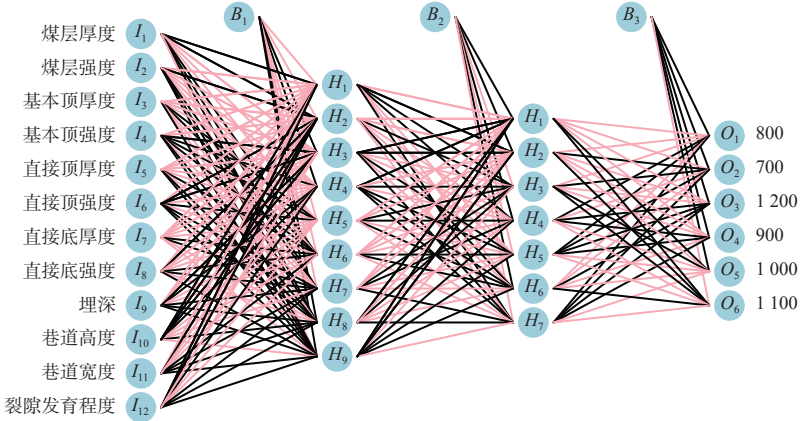


图8 顶板锚杆间距 GA-ANN 预测模型的网络拓扑结构

Fig. 8 Network topology of GA-ANN prediction model of roof bolt spacing

3.4 ADA模型和NBC模型

ADA是在Boosting基础上的一种优化算法。ADA在面对多分类问题时,通过不断调整错误样本的权重迭代升级,一定程度上避免了模型过拟合。

NBC算法是一种采用最大似然估计对样本进行概率分类的算法,通过求解样本在各个类别下的概率进行求解。

3.5 模型性能评估

采用分类精度来评价支护参数预测模型在测试集上的性能,分类精度为模型在测试集上正确分类样本数与总样本数的比值。建立的所有机器学习模型在测试集上的分类精度见表6。可看出SMOTE-GA-SVM模型的表现最佳,每一个支护参数的分类精度均超过75%,模型平均分类精度达到83.8%。ADA模型在预测帮部锚杆间距时表现最差,分类精度只有37.6%;SMOTE-GA-SVM模型预测锚索长度时表

现最佳,分类精度达到93.5%。SMOTE-GA-SVM模型在处理样本类不均衡数据集的问题上性能更加强大,平均分类精度较传统的SVM模型提高了21.8%。在SMOTE和GA的优化改进下,SVM模型整体性能有了较大提高。可见,SMOTE方法可作为处理样本类不均衡的有效方法,GA对SVM模型的超参数有很好的全局寻优能力。且SMOTE-GA-SVM模型优于其他模型,平均分类精度达83.8%;ADA在测试集上的性能最低,精度为60.5%。预测模型分类精度排序为SMOTE-GA-SVM、RF、GA-ANN、SVM、NBC、ADA,6种模型的平均分类精度达69.9%。

4 工程应用

为了验证基于SMOTE-GA-SVM的煤巷支护参数预测模型在工程实践中的适用性和可靠性,在山西霍宝干河煤矿的4条巷道进行了实际工程验证,

表6 机器学习模型在测试集上的分类精度

Table 6 Classification precision of machine learning model on test set %

模型	顶板锚杆精度					帮部锚杆精度					顶板锚索精度				平均精度
	直径	长度	间距	排距	数量	直径	长度	间距	排距	数量	直径	长度	数量	布置	
SVM	66.0	83.6	66.0	79.2	74.8	54.0	70.4	70.4	64.8	57.2	54.0	79.2	74.8	69.2	68.8
SMOTE-GA-SVM	86.0	89.8	82.4	84.3	82.1	75.3	79.2	88.4	77.1	77.2	83.3	93.5	90.6	84.0	83.8
RF	79.2	74.8	52.8	83.6	66.0	70.4	61.6	61.6	74.8	52.8	57.2	92.4	88.0	79.2	71.0
ADA	72.6	64.5	53.8	57.8	57.8	48.4	51.1	37.6	59.2	61.8	48.4	78.7	80.7	75.3	60.5
GA-ANN	93.4	63.4	64.4	55.0	66.0	62.7	66.0	57.2	71.5	55.0	74.8	89.8	79.8	82.5	70.1
NBC	66.0	66.0	44.0	44.0	79.2	70.4	57.2	57.2	57.2	74.8	48.4	81.8	88.0	74.8	64.9

巷道的特征参数见表 7。对巷道特征参数进行 Z-score 标准化,并将其输入已经建立好的基于 SMOTE-

GA-SVM 的煤巷支护参数预测模型中,通过模型计算得到巷道的锚杆、锚索支护参数,结果见表 8。

表 7 霍州矿区干河煤矿的特征参数

Table 7 Characteristic parameters of Ganhe Coal Mine in Huozhou Mining area

序号	巷道名称	煤层厚度/m	煤层强度/MPa	基本顶厚度/m	基本顶强度/MPa	直接顶厚度/m	直接顶强度/MPa	直接底厚度/m	直接底强度/MPa	埋深/m	巷道高度/m	巷道宽度/m
1	2-1161巷	4.20	9.38	1.70	65.06	6.40	51.79	1.10	51.79	450	3.6	5.0
2	2-1261巷	3.75	15.00	4.80	86.09	2.45	65.06	2.90	19.16	420	3.7	5.0
3	三采区辅助运输巷	0.78	9.38	3.12	65.06	4.96	86.09	4.60	45.40	420	3.5	4.8
4	2-1021巷	4.20	14.54	3.10	57.32	3.89	25.00	0.80	37.45	500	3.8	5.0

表 8 SMOTE-GA-SVM 模型应用结果

Table 8 Application result of SMOTE-GA-SVM model

序号	巷道名称		顶板锚杆					帮部锚杆					顶板锚索			
			直径/mm	长度/mm	间距/mm	排距/mm	数量/根	直径/mm	长度/mm	间距/mm	排距/mm	数量/根	直径/mm	长度/mm	布置方式	数量/根
1	2-1161巷	真实值	22	2 500	900	900	6	22	2 500	900	900	4	21.60	8 300	2	3
		测试值	22	2 400	900	900	6	22	2 400	900	900	4	21.60	8 300	2	3
2	2-1261巷	真实值	22	2 500	800	800	7	22	2 500	800	800	5	21.60	8 300	3	3
		测试值	22	2 500	800	800	7	22	2 500	800	800	5	21.60	8 300	3	3
3	三采区辅助运输巷	真实值	20	2000	800	800	7	20	2000	800	800	2	15.24	7 300	2	2
		测试值	20	2000	800	1 000	7	20	2000	800	1 000	2	17.89	7 300	2	2
4	2-1021巷	真实值	22	2 400	900	1 000	6	22	2 400	1 000	1 000	4	21.60	6 300	2	2
		测试值	22	2 400	900	1 000	6	20	2 400	1 000	1 000	4	21.60	7 300	2	2

由表 8 可知,2-1261 巷实际采用的锚杆、锚索支护参数与 SMOTE-GA-SVM 模型计算结果一致,2-1161 巷和 2-1021 巷都有 2 个参数预测错误,三采区辅助运输巷有 3 个参数预测错误,4 条巷道的 56 个支护参数中预测结果与真实值相同的有 49 个,模型预测的正确率为 87.5%。其中,SMOTE-GA-SVM 模型错误预测的 7 个参数都被预测为相邻类别的参数,相对误差较小。由此可见,SMOTE-GA-SVM 模型能够很好地掌握巷道特征参数到锚杆、锚索支护参数的非线性映射能力,具有较强的适用性和可靠性,对煤巷锚杆、锚索支护参数设计具有一定的参考价值。

## 5 结论

1) 建立了煤巷支护智能预测数据库。采用现场调研、问卷调查和文献检索等方式收集煤矿巷道样本;采用缺失值填补、箱形图修改离群点和 LOF 剔除异常样本等方式对数据进行处理后,建立煤巷支护智能预测数据库。

2) 提出一种基于 SMOTE-GA-SVM 的煤巷支护参数预测模型,该模型在训练前采用 SMOTE 方法平衡训练集中类别不平衡的样本,提高模型对少数类样本的拟合能力;训练过程中使用 GA 对超参数进

行全局寻优,进一步提高了模型整体性能。测试结果表明,基于 SMOTE-GA-SVM 的煤巷支护参数预测模型的分类精度达 83.8%,比传统的 SVM 模型提高了 21.8%。

3) 将 ANN、RF、ADA 和 NBC 引入煤巷支护参数预测中,建立对应的支护参数预测模型。结果显示,从最优到最差的预测模型排名分别为 SMOTE-GA-SVM、RF、GA-ANN、SVM、NBC 和 ADA,验证了机器学习方法在煤巷支护参数预测方面的可行性。

4) 在山西霍宝干河煤矿的 4 条巷道对 SMOTE-GA-SVM 模型进行了应用,模型预测准确率达到 87.5%,具有较强的适用性和可靠性。

## 参考文献(References):

- [1] 康红普. 我国煤矿巷道锚杆支护技术发展 60 年及展望[J]. 中国矿业大学学报, 2016, 45(6): 1071-1081.  
KANG Hongpu. Sixty years development and prospects of rock bolting technology for underground coal mine roadways in China[J]. Journal of China University of Mining & Technology, 2016, 45(6): 1071-1081.
- [2] 康红普. 我国煤矿巷道围岩控制技术发展 70 年及展望[J]. 岩石力学与工程学报, 2021, 40(1): 1-30.  
KANG Hongpu. Seventy years development and prospects of strata control technologies for coal mine roadways in China[J]. Chinese Journal of Rock



- Mechanics and Engineering, 2021, 40(1): 1-30.
- [3] 单仁亮, 彭杨皓, 孔祥松, 等. 国内外煤巷支护技术研究进展[J]. 岩石力学与工程学报, 2019, 38(12): 2377-2403.
- SHAN Renliang, PENG Yanghao, KONG Xiangsong, et al. Research progress of coal roadway support technology at home and abroad[J]. Chinese Journal of Rock Mechanics and Engineering, 2019, 38(12): 2377-2403.
- [4] 顾清华, 江松, 李学现, 等. 人工智能背景下采矿系统工程发展现状与展望[J]. 金属矿山, 2022(5): 10-25.
- GU Qinghua, JIANG Song, LI Xuexian, et al. Development status and prospect of mining system engineering under the background of artificial intelligence[J]. Metal Mine, 2022(5): 10-25.
- [5] 谢广祥, 曹伍富, 王德润, 等. 基于人工神经网络的煤巷锚杆支护设计研究[J]. 煤炭学报, 1999(6): 599-604.
- XIE Guangxiang, CAO Wufu, WANG Derun, et al. The study on bolting support design in coal roadway based on artificial neural networks[J]. Journal of China Coal Society, 1999(6): 599-604.
- [6] 王茂源. 煤巷锚杆支护设计混合智能系统研究[D]. 北京: 中国矿业大学(北京), 2016.
- WANG Maoyuan. Hybrid intelligent system on coal roadway bolting design[D]. Beijing: China University of Mining and Technology-Beijing, 2016.
- [7] 王哲哲, 许梦国, 程爱平, 等. 模糊神经网络在巷道支护方案选择中的应用[J]. 化工矿物与加工, 2019, 48(1): 16-19, 23.
- WANG Zhezhe, XU Mengguo, CHENG Aiping, et al. Application of fuzzy neural network in selection of roadway support scheme[J]. Industrial Minerals & Processing, 2019, 48(1): 16-19, 23.
- [8] XU Qingyun, LI Yongming, LU Jie, et al. The use of surrounding rock loosening circle theory combined with elastic-plastic mechanics calculation method and depth learning in roadway support[J]. PLoS ONE, 2020, 15(7). DOI: 10.1371/journal.pone.0234071.
- [9] REN Heng, ZHU Yongjian, WANG Ping, et al. Classification and application of roof stability of bolt supporting coal roadway based on BP neural network[J]. Advances in Civil Engineering, 2020. DOI: 10.1155/2020/8838640.
- [10] ZHANG Xiliang, NGUYEN H, BUI X, et al. Evaluating and predicting the stability of roadways in tunnelling and underground space using artificial neural network-based particle swarm optimization[J]. Tunnelling and Underground Space Technology, 2020, 103. DOI: 10.1016/j.tust.2020.103517.
- [11] PU Yuanyuan, APEL D, HALL R. Using machine learning approach for microseismic events recognition in underground excavations: comparison of ten frequently-used models[J]. Engineering Geology, 2020. DOI: 10.1016/j.enggeo.2020.105519.
- [12] 马鑫民, 范皓宇, 林天舒, 等. 基于 GA-SVM 的煤矿岩巷爆破效果智能预测[J]. 煤炭工程, 2019, 51(5): 148-153.
- MA Xinmin, FAN Haoyu, LIN Tianshu, et al. Intelligent prediction of blasting effect of coal mine roadway based on GA-SVM[J]. Coal Engineering, 2019, 51(5): 148-153.
- [13] MAHDEVARI S, KHODABAKHSHI M B. A hierarchical local-model tree for predicting roof displacement in longwall tailgates[J]. Neural Computing and Applications, 2021, 33(21): 14909-14928.
- [14] 赵汝星. 基于随机森林的回采巷道围岩稳定性分类[J]. 煤矿安全, 2014, 45(11): 200-202, 206.
- ZHAO Ruxing. Classification of roadway surrounding rock stability based on random forest[J]. Safety in Coal Mines, 2014, 45(11): 200-202, 206.
- [15] 汪海燕, 黎建辉, 杨风雷. 支持向量机理论及算法研究综述[J]. 计算机应用研究, 2014, 31(5): 1281-1286.
- WANG Haiyan, LI Jianhui, YANG Fenglei. Overview of support vector machine analysis and algorithm[J]. Application Research of Computers, 2014, 31(5): 1281-1286.
- [16] JU Xuchan, YAN Zhenghao, WANG Tianhe, et al. Overview of optimization algorithms for large-scale support vector machines[C]. IEEE International Conference on Data Mining Workshops, Ningbo, 2021: 909-916.
- [17] CHAHAR V, KATOCH S, CHAUHAN S S. A review on genetic algorithm: past, present, and future[J]. Multimedia Tools and Applications, 2020, 80(5): 8091-8126.
- [18] 谭文侃, 叶义成, 胡南燕, 等. LOF 与改进 SMOTE 算法组合的强烈岩爆预测[J]. 岩石力学与工程学报, 2021, 40(6): 1186-1194.
- TAN Wenkan, YE Yicheng, HU Nanyan, et al. Severe rock burst prediction based on the combination of LOF and improved SMOTE algorithm[J]. Chinese Journal of Rock Mechanics and Engineering, 2021, 40(6): 1186-1194.
- [19] FENG Shuo, KEUNG J, YU Xiao, et al. Investigation on the stability of SMOTE-based oversampling techniques in software defect prediction[J]. Information and Software Technology, 2021, 139(6). DOI: 10.1016/j.infsof.2021.106662.
- [20] CHAO Ying, YIN Kunlong, ZHOU Chao, et al. Establishment of landslide groundwater level prediction model based on GA-SVM and influencing factor analysis[J]. Sensors, 2020, 20(3): 845.
- [21] 吕红燕, 冯倩. 随机森林算法研究综述[J]. 河北省科学院学报, 2019, 36(3): 37-41.
- LYU Hongyan, FENG Qian. A review of random forests algorithm[J]. Journal of the Hebei Academy of Sciences, 2019, 36(3): 37-41.