

文章编号: 1671-251X(2025)06-0021-08

DOI: 10.13272/j.issn.1671-251x.2025060027

基于 XGBoost-SHAP 的综采工作面上隅角瓦斯溯源模型

盛武^{1,2}, 王灵子¹

(1. 安徽理工大学 经济与管理学院, 安徽 淮南 232001;

2. 安徽理工大学 煤炭无人化开采数智技术全国重点实验室, 安徽 淮南 232001)

摘要: 针对目前综采工作面上隅角瓦斯浓度预测模型由于“黑盒”结构导致内部运行逻辑未知、预测结果可解释性弱的问题, 提出一种基于 XGBoost-SHAP 的综采工作面上隅角瓦斯溯源模型。对综采工作面瓦斯涌出浓度关联监测数据进行相关分析, 筛选出特征变量; 基于 XGBoost 搭建上隅角瓦斯浓度预测模型, 引入 SHAP 算法计算每个特征变量对预测结果的贡献值, 增强模型透明度, 为 XGBoost 提供全局性解释; 最后利用现场多源传感监测数据对模型性能进行验证。实例分析结果表明: ① XGBoost 模型的决定系数 R^2 、平均绝对误差(MAE)、均方根误差(RMSE)分别为 0.93, 0.007, 0.008, 相较于随机森林(RF)、支持向量回归(SVR)和梯度提升决策树(GBDT), 拟合优度最高, 误差最低。② XGBoost 模型的平均相对误差为 4.478%, 相较于对比模型, 具有较高的精度与较好的泛化性能。③ 依据各输入特征的平均绝对 SHAP 值, 工作面 T_1 瓦斯浓度对上隅角瓦斯浓度影响最大, 工作面上隅角瓦斯抽采管道内瓦斯浓度次之, 回采煤层瓦斯含量、回采煤层顶板压力等紧随其后, 说明 XGBoost 能捕捉变量间的非线性关系和交互作用, SHAP 算法可为 XGBoost 模型提供全局性解释。

关键词: 瓦斯浓度预测; 上隅角瓦斯溯源; XGBoost 模型; SHAP; 可解释性

中图分类号: TD712

文献标志码: A

A traceability model for upper corner gas in fully mechanized mining faces based on XGBoost-SHAP

SHENG Wu^{1,2}, WANG Lingzi¹

(1. College of Economic and Management, Anhui University of Science and Technology, Huainan 232001, China;

2. State Key Laboratory of Digital Intelligent Technology for Unmanned Coal Mining, Anhui University of Science and Technology, Huainan 232001, China)

Abstract: To address the weak interpretability caused by the "black-box" structure of current gas concentration prediction models in the upper corner of fully mechanized mining faces, a gas traceability model based on XGBoost-SHAP was proposed for the upper corner of fully mechanized mining faces. Correlation analysis was conducted on the monitoring data of gas emission from fully mechanized mining faces to select feature variables. An upper corner gas concentration prediction model was constructed based on XGBoost, and the SHAP algorithm was introduced to calculate the contribution of each feature variable to the prediction results, thereby enhancing the model's transparency and providing a global interpretation for the XGBoost model. Finally, the model performance was evaluated using multi-source sensor monitoring data from the field. Case analysis results showed that: ① the coefficient of determination (R^2), mean absolute error (MAE), and root mean square

收稿日期: 2025-06-08; 修回日期: 2025-06-25; 责任编辑: 胡娴。

基金项目: 国家自然科学基金项目(71971003); 安徽省优秀创新团队项目(2022AH010054); 淮北矿业股份公司委托项目(淮矿技合2024(17))。

作者简介: 盛武(1969—), 男, 安徽淮南人, 副教授, 博士, 研究方向为智能风险预警、安全管理, E-mail: wsheng116@163.com。通信作者: 王灵子(2001—), 女, 安徽安庆人, 硕士研究生, 研究方向为数据分析、风险预测, E-mail: 2893392512@qq.com。

引用格式: 盛武, 王灵子. 基于 XGBoost-SHAP 的综采工作面上隅角瓦斯溯源模型[J]. 工矿自动化, 2025, 51(6): 21-27, 140.

SHENG Wu, WANG Lingzi. A traceability model for upper corner gas in fully mechanized mining faces based on XGBoost-SHAP[J]. Journal of Mine Automation, 2025, 51(6): 21-27, 140.



扫码移动阅读

error (RMSE) of the XGBoost model were 0.93, 0.007, and 0.008, respectively, indicating the highest goodness of fit and the lowest errors compared with random forest (RF), support vector regression (SVR), and gradient boosting decision tree (GBDT). ② The mean relative error of the XGBoost model was 4.478%, demonstrating higher accuracy and better generalization performance compared with the other models. ③ Based on the mean absolute SHAP values of input features, the gas concentration at T1 on the working face had the greatest influence on the gas concentration in the upper corner, followed by the gas concentration in the upper corner extraction pipeline, with the gas content and roof pressure of the mining coal seam following closely. These findings indicate that XGBoost can capture the nonlinear relationships and interactions between variables, and that the SHAP algorithm can provide global interpretability for the XGBoost model.

Key words: gas concentration prediction; upper corner gas traceability; XGBoost model; SHAP; interpretability

0 引言

采掘工作面瓦斯异常涌出是瓦斯灾害风险的直接特征表现,极易诱发瓦斯重特大事故,严重威胁安全生产^[1]。瓦斯异常涌出受地质构造、煤层赋存、生产技术及大气压等影响,在多灾种耦合作用下,很难精准识别瓦斯异常涌出致因。依据瓦斯分源理论,采煤工作面区域瓦斯主要来源于开采煤层、采空区及邻近煤层。瓦斯溯源基于分源理论,对瓦斯涌出来源进行精准识别,锁定核心致因,从而为精准防控提供科学依据^[2]。

目前采煤工作面瓦斯异常涌出来源大多是通过浓度预测及分源来识别。Peng Yujie 等^[3]采用指标动态优化和双向长短时记忆网络(Bidirectional Long Short-Term Memory, Bi-LSTM)组合模型预测气体浓度,可提前发现气体异常。梁运培等^[4]构建了基于频率增强分解的上隅角瓦斯浓度预测模型,具有较高的长时预测精度,泛化能力较好;Du Feng 等^[5]对多煤层混合气体进行溯源,通过数值模拟实现了每个煤层瓦斯量的准确测量。Liu Jun 等^[6]建立源分离的三维瓦斯开采系统,实现了大型采掘工作面不同来源瓦斯的三维提取。周福宝等^[7]采用稳定碳氢同位素法对上隅角、采空区、回风巷瓦斯溯源,得到工作面 4 个测点位置瓦斯来源的占比及规律。

多场耦合作用下传统方法在捕捉数据中的非线性关系存在不足,难以获得精准瓦斯浓度及溯源。随机森林(Random Forest, RF)^[8]、支持向量回归(Support Vector Regression, SVR)^[9]、梯度提升决策树^[10](Gradient Boosting Decision Tree, GBDT)、极限梯度提升算法^[11](eXtreme Gradient Boosting, XGBoost)等方法能发现特征间复杂的依赖关系,建立高精度、高稳健模型,得到学者青睐。但由于模型“黑盒”结构,内部运行逻辑未知,预测结果可解释性弱。对

此,SHAP^[12](SHapley Additive exPlanations)、Hybrid-Grid^[13]等解释方法得到了广泛运用。

可解释性预测模型在医疗^[14]、航空^[15]等其他行业已有成功应用,但在煤矿领域具有可解释性的模型很少。本文创新性地提出一种基于 XGBoost-SHAP 的综采工作面上隅角瓦斯溯源模型,对多传感器数据进行关联分析,将筛选出的特征数据作为输入,通过 SHAP 算法增强模型透明度,为 XGBoost 提供全局性解释,进而实现瓦斯溯源。

1 模型构建

1.1 XGBoost 模型

XGBoost 是基于 GBDT 改进的模型,通过迭代训练多棵决策树,根据特征筛选逐步优化目标函数,具有高精度、强灵活性及对缺失值的鲁棒性,善于捕捉复杂的非线性关系^[16]。模型目标函数由损失函数和正则化项 2 个部分组成。损失函数表示模型拟合数据的程度,通过引入二阶导数加快梯度收敛速度,精准逼近真实的损失函数。正则化项控制模型的复杂度,防止过拟合。XGBoost 目标函数为

$$o^{(t)} = \sum_{i=1}^N L(y_i, \hat{y}_i^{(t)}) + \Omega(f_t) + c \quad (1)$$

式中: t 为所有树的集合; N 为样本总量; L 为损失函数,表示真实值 y_i 与预测值 $\hat{y}_i^{(t)}$ 的损失和; $\Omega(f_t)$ 为正则化项,由叶子节点总数和节点权重共同决定, f_t 为树模型; c 为常数。

$$\Omega(f_t) = \gamma Z + \frac{1}{2} \lambda \sum_{j=1}^Y w_j^2 \quad (2)$$

式中: γ 和 λ 为自定义参数,用于避免过拟合; Z 为树中叶子节点的总数; w_j 为第 j 个叶子节点的权重。

1.2 SHAP 算法

针对 XGBoost 模型难以解释的问题,引入

SHAP 算法, 为特征与模型预测结果间的关系提供解释, 增强模型的可靠性。SHAP 算法是一种基于博弈论思想的方法, 依赖于 Shapley 值, 通过加权平均所有可能特征子集的边际贡献, 公平分配每个特征的贡献值, 以此衡量每个特征对预测结果的贡献程度^[17]。Shapley 值计算公式为

$$\phi_i = \sum_S \frac{(M-S-1)!|S|!}{M!} [F(S \cup \{I\}) - F(S)] \quad (3)$$

式中: S 为不包含特征 I 的所有可能子集; M 为特征总数; $|S|$ 为集合 S 中的特征数; $F(S \cup \{I\})$ 为包含特征 I 时 S 的预测值; $F(S)$ 为 S 的预测值。

1.3 瓦斯溯源流程

基于 XGBoost-SHAP 的综采工作面上隅角瓦斯溯源流程如图 1 所示。采集采煤工作面各测点的瓦斯浓度及温度、风速、压力等关联传感器数据, 对传感器数据进行关联分析, 筛选特征, 并进行对数处理、数据集划分; 基于 XGBoost 搭建预测模型, 利用 SHAP 算法计算每个特征变量对预测结果的贡献值, 可视化揭示“黑箱”决策逻辑。

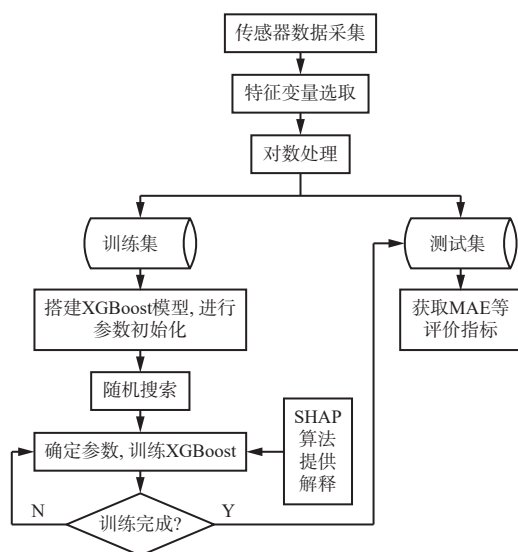


图 1 基于 XGBoost-SHAP 的综采工作面上隅角瓦斯溯源流程

Fig. 1 Source tracing process of upper corner gas in fully mechanized mining face based on XGBoost-SHAP

1.4 模型评价指标

为评价模型预测结果的准确程度, 采用平均绝对误差 (Mean Absolute Error, MAE)、均方根误差 (Root Mean Square Error, RMSE) 和决定系数 R^2 作为评价指标。MAE 为模型预测的偏差平均值, RMSE 为预测值与真实值间的平均偏差, R^2 为模型的拟合程度。MAE 和 RMSE 值越小, 说明模型预测性能越好, R^2 越趋近于 1, 说明模型拟合效果好, 结果预测越精准。

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}_i| \quad (4)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2} \quad (5)$$

$$R^2 = 1 - \frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - \bar{y})^2} \quad (6)$$

式中 $\bar{y}$ 为样本中所有真实值的均值。

2 实例分析

2.1 数据集构建

本文实验数据来源于中煤新集能源股份有限公司某智能化煤矿 171105 综采工作面。该综采工作面为 U 型布置, 单一煤层开采, 上隅角上方、开采煤层回采方向、回风巷布置瓦斯抽采管道, 多传感器在线实时监测数据, 现场布置如图 2 所示。采集综采工作面区域瓦斯浓度关联数据: 工作面 T_1 瓦斯浓度 X_1 、工作面回风巷 T_2 瓦斯浓度 X_2 、工作面上隅角瓦斯抽采管道内瓦斯浓度 X_3 、回风巷抽采管道内瓦斯浓度 X_4 、回采煤层瓦斯抽排巷抽采管道内瓦斯浓度 X_5 、温度 X_6 、风速 X_7 、一氧化碳 X_8 、氧气 X_9 、回采煤层顶板压力 X_{10} 、回采煤层瓦斯含量 X_{11} 、煤层埋藏深度 X_{12} 、煤层倾角 X_{13} 、煤层厚度 X_{14} 及预测目标工作面上隅角 T_0 瓦斯浓度 Y 等 15 个变量。将监测数据统一到同一时空区间, 得到 720 组数据。部分原始数据见表 1。

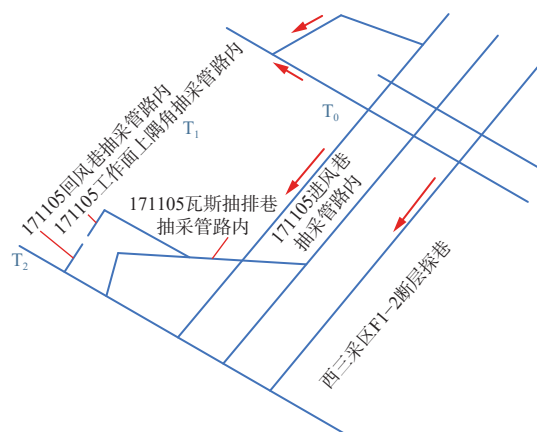


图 2 171105 综采工作面现场布置(上隅角区域)

Fig. 2 Site layout of fully mechanized mining face 171105 (upper corner area)

综采工作面上隅角瓦斯溯源模型开发环境为 Pycharm, 编程语言为 Python3.12, 实验中数据预处理、模型构建、“黑箱”解释等部分基于 sklearn,

表 1 部分原始数据
Table 1 Partial raw data

样本	$X_1/\%$	$X_2/\%$	$X_3/\%$	$X_4/\%$	$X_5/\%$	$X_6/^\circ\text{C}$	$X_7/(\text{m}\cdot\text{s}^{-1})$	$X_8/\%$	$X_9/\%$	X_{10}/MPa	$X_{11}/(\text{m}^3\cdot\text{t}^{-1})$	X_{12}/m	$X_{13}/(^{\circ})$	X_{14}/m	$Y/\%$
1	0.13	0.17	0.61	14.1	5.6	26.8	2.3	0	20.3	31.4	3.49	-705.2	6.8	4.5	0.14
2	0.12	0.15	0.67	13.9	5.61	26.8	2.4	0	20.3	23.5	3.5	-683.8	1.1	3.6	0.13
3	0.1	0.13	0.62	14	5.6	26.8	2.3	0	20.2	23.2	3.49	-670.2	4.9	3.8	0.13
4	0.11	0.12	0.65	14	5.61	26.8	2.2	0	20.2	23.2	3.47	-663.2	5	3.8	0.16
5	0.11	0.16	0.7	14	5.6	26.8	2.2	0	20.3	23	3.49	-667.5	7.7	3.3	0.14
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
719	0.1	0.12	0.73	11.9	5.78	26.8	2.3	0	20.3	15.9	3.23	-664.5	3.2	4.2	0.04
720	0.11	0.12	0.73	11.9	5.79	26.8	2.2	0	20.3	15.7	3.23	-672	2.2	4.9	0.06

xgboost, shap, tensorflow 等库实现。

2.2 特征变量选取

参与模型训练的输入特征数量增加,并不代表训练后模型性能会提升,相反,过多的输入特征会导致模型复杂度上升,训练时间更长^[18]。表 1 中影响采空区上隅角瓦斯浓度的因子较多,特征维度较大,因此,采用特征关联法筛选特征。

将数据导入 SPSS26.0 版本,先进行标准化处理,再进行相关性检验。对数据进行 Spearman 相关系数检验,结果显示 $X_1, X_2, X_3, X_5, X_7, X_{10}, X_{11}$ 与工作面上隅角瓦斯浓度 Y 相关性显著,因此筛选出这 7 个特征作为模型输入。

2.3 数据预处理

井下环境复杂,在数据监测、传输及存储过程中易发生数据偏差及缺失,且瓦斯浓度、温度等数据常呈现右偏分布,无法满足机器学习模型内置的数据服从正态分布的假设,不利于建模分析。因此,本文采用对数处理方法对原始样本数据进行预处理^[19]。对数处理公式为处理后的数据更接近适合机器学习算法的正态分布,降低了噪声影响,提高了数据的解释性和鲁棒性,为后续建模提供了支撑。

2.4 模型拟合效果

将预处理后的数据按照 8:2 的比例划分为训练集和测试集,导入 XGBoost 模型中。在训练过程中,考虑到模型参数繁多,若优化所有参数会降低效率,因此根据文献^[20]和实际情况,选取 5 个主要的超参数,使用 5 折交叉验证的随机搜索方法寻找最佳参数组合,其余超参数保持默认值。模型参数设置及寻优结果见表 2。

为了验证所建模型的性能,将其与 RF, SVR, GBDT 模型对比,结果如图 3 所示,具体评价指标见表 3。

由图 3 可知,本文模型预测的工作面上隅角瓦

表 2 模型参数设置及寻优结果

Table 2 Model parameter settings and optimization results

参数	含义	搜索范围	步长	寻优结果
learning_rate	学习率	0.01~0.3	0.005	0.08
max_depth	树的最大深度	3~15	1	9
reg_alpha	L1正则化项系数	0~1	0.01	0.4
subsample	随机采样比例	0.5~1	0.1	0.6
min_child_weight	子树最小观测权重和	1~10	1	9

斯浓度相比于其他模型更接近实际值。由表 3 可知, XGBoost 模型的误差最低, MAE 为 0.007、RMSE 为 0.008,拟合优度最高, R^2 为 0.93。上述结果表明本文构建的工作面上隅角瓦斯浓度预测模型具有更高的预测精度和有效性。

为进一步验证模型泛化性能,随机取未参与模型拟合的 20 个样本点进行预测,并与 RF, SVR, GBDT 模型的预测效果对比,结果见表 4。可看出 XGBoost 模型的平均相对误差为 4.478%,远低于对比模型,验证了 XGBoost 具有较高的精度与较好的泛化性能。

2.5 工作面上隅角瓦斯溯源模型解释

2.5.1 上隅角瓦斯浓度特征贡献度

XGBoost 模型预测性能优异,但其内部决策逻辑难以理解,模型难以解释输入特征变量对预测结果的影响程度。SHAP 算法通过量化每一个输入特征对预测结果的贡献并以可视化形式呈现,极大增加了模型的可信度^[21]。

本文处理数据时导入 shap 库,计算每个特征的 SHAP 值(对输出的影响),按贡献度值大小得到特征重要性排序,如图 4 所示。可看出工作面 T_1 瓦斯浓度 X_1 对上隅角瓦斯浓度 Y 的影响最大,平均绝对 SHAP 值为 0.31;其次是工作面上隅角瓦斯抽采管道内瓦斯浓度 X_3 、回采煤层瓦斯含量 X_{11} 、回采煤层顶

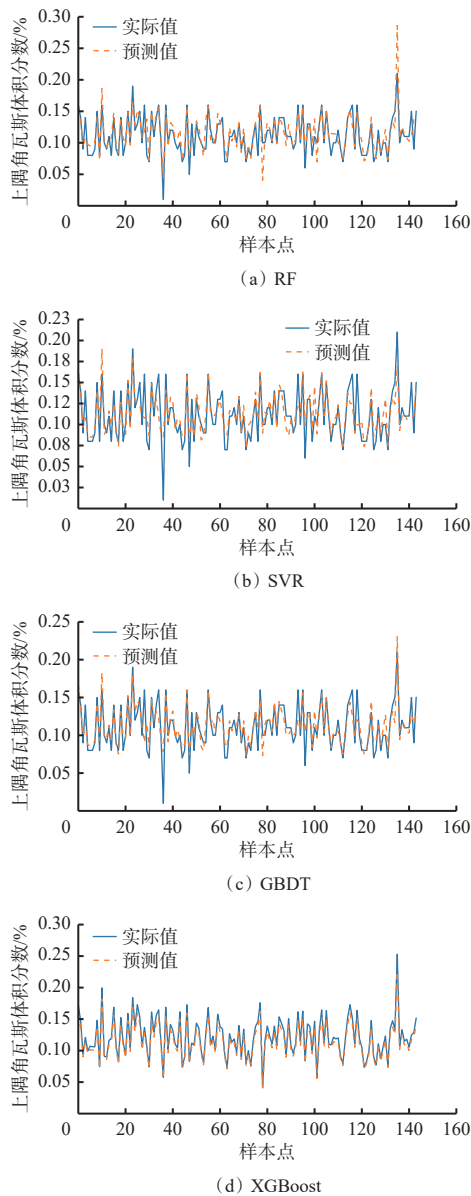


图 3 不同模型预测结果对比

Fig. 3 Comparison of prediction results of different models

表 3 不同模型评价指标对比

Table 3 Comparison of evaluation indicators for different models

模型	MAE	RMSE	R^2
RF	0.013	0.018	0.64
SVR	0.013	0.018	0.66
GBDT	0.011	0.016	0.72
XGBoost	0.007	0.008	0.93

板压力 X_{10} , 平均绝对 SHAP 值分别为 0.23, 0.21, 0.17; 风速 X_7 、回采煤层瓦斯抽排巷抽采管道内瓦斯浓度 X_5 、工作面回风巷 T_2 瓦斯浓度 X_2 对 Y 的影响相对较弱, 平均绝对 SHAP 值依次为 0.1, 0.09, 0.06。

为进一步分析输入特征对上隅角瓦斯涌出浓度的影响, 输出 SHAP 蜂巢图, 如图 5 所示, 图中每个

表 4 不同模型实际值与预测值比较
Table 4 Comparison between actual and predicted values of different models %

真实值	RF		SVR		GBDT		XGBoost	
	预测值	相对误差	预测值	相对误差	预测值	相对误差	预测值	相对误差
0.16	0.159	0.480	0.164	2.200	0.168	4.717	0.150	6.855
0.14	0.138	1.651	0.139	0.512	0.138	1.115	0.131	6.396
0.09	0.097	7.939	0.097	7.737	0.098	9.105	0.087	4.261
0.14	0.115	17.86	0.108	22.84	0.117	16.79	0.137	4.011
0.08	0.093	16.26	0.090	12.86	0.088	9.448	0.077	3.775
0.08	0.086	7.674	0.085	6.527	0.086	7.290	0.077	3.785
0.08	0.086	8.091	0.085	6.109	0.088	10.47	0.076	3.816
0.09	0.082	9.143	0.093	3.560	0.092	2.441	0.095	4.661
0.15	0.140	6.762	0.140	6.404	0.142	5.106	0.146	5.576
0.08	0.087	9.121	0.100	24.88	0.094	16.94	0.086	4.222
0.16	0.186	16.40	0.190	18.94	0.183	14.18	0.154	3.758
0.10	0.095	4.994	0.088	11.67	0.096	4.209	0.095	4.692
0.09	0.100	11.51	0.095	5.194	0.097	7.776	0.095	4.670
0.11	0.112	1.892	0.116	5.559	0.113	3.024	0.104	5.089
0.08	0.095	18.61	0.099	23.25	0.095	18.16	0.086	4.239
0.14	0.136	3.141	0.126	10.23	0.134	4.636	0.146	4.286
0.09	0.094	4.927	0.095	5.112	0.088	2.457	0.095	4.676
0.08	0.081	0.927	0.073	8.273	0.075	5.878	0.086	4.241
0.14	0.132	5.909	0.126	9.922	0.131	6.208	0.137	2.295
0.08	0.096	20.62	0.098	22.05	0.097	20.69	0.086	4.258
平均相对误差	8.696		10.69		8.532		4.478	

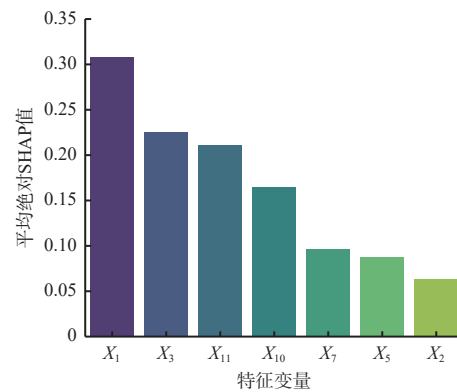


图 4 特征重要性排序

Fig. 4 Feature importance ranking

点代表 1 个特定样本, 不同颜色表示原始值大小。

由图 5 可知, 工作面 T_1 瓦斯浓度 X_1 对 Y 有较为明显的正向关系; 工作面上隅角瓦斯抽采管道内瓦斯浓度 X_3 的 SHAP 值分布范围为 $(-0.6, 0.9)$, 与 Y 有复杂的非线性关系; 随着回采煤层瓦斯含量 X_{11} 的上

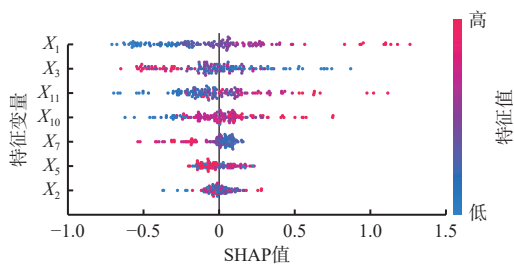


图 5 SHAP 蜂巢图

Fig. 5 SHAP beeswarm plot

升, SHAP 值呈上升趋势; 回采煤层顶板压力 X_{10} 的 SHAP 值随着 X_{10} 的增大而增大, 表明对 Y 起正向反馈作用; 风速 X_7 、回采煤层瓦斯抽排巷抽采管道内瓦斯浓度 X_5 与工作面回风巷 T_2 瓦斯浓度 X_2 的 SHAP 值大部分集中在 0 附近, 对 Y 的影响较小。

2.5.2 上隅角瓦斯浓度特征依赖关系

单个特征依赖图能深层次解释单个特征独立作用及 2 个密切特征间的交互作用。本文选取重要性排序在前 4 名的特征进行分析, 绘制 SHAP 值特征依赖图, 如图 6 所示。

由图 6(a)可知, 工作面 T_1 瓦斯浓度 X_1 在一定区间内增加时, 割煤过程中煤体原生裂隙被破坏, 可能会大量释放吸附态瓦斯, 若此时风速 X_7 稀释不足, 无法运移瞬时涌出的游离瓦斯, 便会使密度较大的瓦斯积聚在通风死角——上隅角, 从而导致 Y 上升。工作面上隅角瓦斯抽采管道内瓦斯浓度 X_3 的上升表明采空区邻近层卸压瓦斯渗流强度增大, 由图 6(b)可知, X_3 到一定阈值时, 工作面 T_1 瓦斯浓度 X_1 也上升, X_1 能够反映回采煤层煤体的瓦斯解吸强度, 解释了中期工作面上隅角瓦斯涌出浓度 Y 呈上升趋势的原因。从图 6(c)可看出, 在回采过程中, 机械扰动可能破坏吸附平衡, 煤体单位时间内解吸的瓦斯量变大, 且随着回采煤层瓦斯含量 X_{11} 的增加, 煤体可释放瓦斯潜能线性增长, Y 整体呈上升趋势。图 6(d)中, 回采煤层顶板压力 X_{10} 的增大可能触发垂直应力增高, 导致煤层裂隙扩展、煤岩渗透性加大, 而煤体骨架压缩使孔隙压力增大, 瓦斯吸附势必降低, 瓦斯解吸加快, 二者共同作用, 对 Y 呈现正向反馈。

3 结论

1) 提出一种基于 XGBoost-SHAP 的综采工作面上隅角瓦斯溯源模型。采用关联分析法筛选主要特征, 引入 SHAP 解释特征变量对输出的影响, 构建上隅角瓦斯溯源模型。

2) 对比 RF, SVR, GBDT, 本文模型的拟合优度

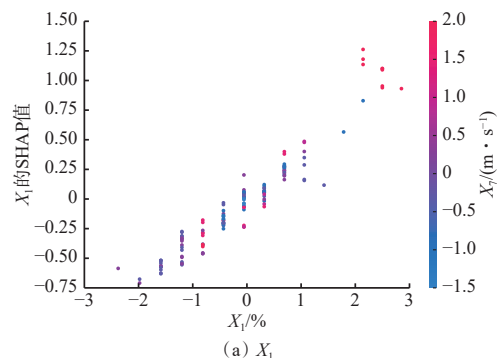
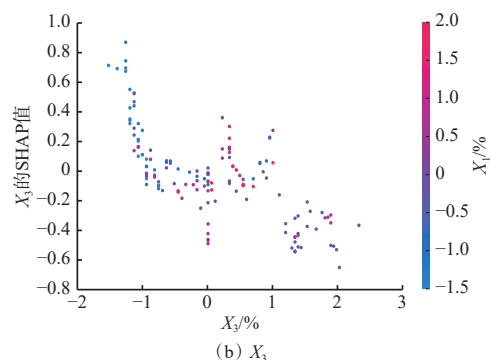
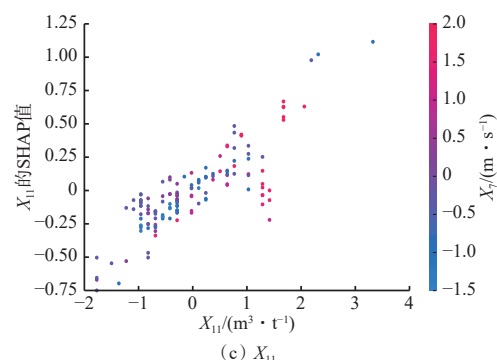
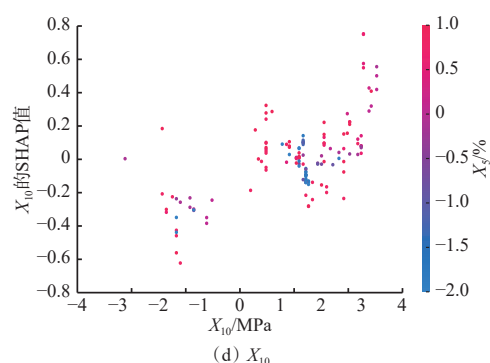
(a) X_1 (b) X_3 (c) X_{11} (d) X_{10}

图 6 上隅角瓦斯浓度 SHAP 值特征依赖

Fig. 6 SHAP dependence plot of gas concentration in the upper corner

最高, R^2 达 0.93; MAE 和 RMSE 均最低, 分别为 0.007, 0.008; 平均相对误差为 4.478%, 说明所建模型具有较高的精度与较好的泛化性能。

3) 通过计算各输入特征的 SHAP 值, 得出工作面 T_1 瓦斯浓度对上隅角瓦斯浓度影响最大, 工作面上隅角瓦斯抽采管道内瓦斯浓度次之, 且对上隅角瓦斯涌出浓度影响复杂, 回采煤层瓦斯含量、回采煤

层顶板压力等紧随其后,说明 XGBoost 能捕捉变量间的非线性关系和交互作用,SHAP 算法可为“黑箱”模型提供全局性解释。

4) 后续研究将扩大综采工作面瓦斯关联监测数据来源,优化 XGBoost 超参数,提高 SHAP 解释关联特征对工作面上隅角瓦斯浓度的影响,实现更精准预测与溯源。

参考文献(References):

- [1] 袁亮. 我国煤炭主体能源安全高质量发展的理论技术思考[J]. 中国科学院院刊, 2023, 38(1): 11-22.
YUAN Liang. Theory and technology considerations on high-quality development of coal main energy security in China[J]. Bulletin of Chinese Academy of Sciences, 2023, 38(1): 11-22.
- [2] 李国富, 李超, 张碧川, 等. 我国煤矿瓦斯抽采与利用发展历程、技术进展及展望[J]. 煤田地质与勘探, 2025, 53(1): 77-91.
LI Guofu, LI Chao, ZHANG Bichuan, et al. Gas drainage and utilization in coal mines in China: history, technological advances, and prospects[J]. Coal Geology & Exploration, 2025, 53(1): 77-91.
- [3] PENG Yujie, SONG Dazhao, QIU Liming, et al. Combined prediction model of gas concentration based on indicators dynamic optimization and Bi-LSTMs[J]. Sensors, 2023, 23(6). DOI: [10.3390/s23062883](https://doi.org/10.3390/s23062883).
- [4] 梁运培, 李赏, 李全贵, 等. 基于 FEDformer-LGBM-AT 架构的采煤工作面上隅角瓦斯浓度预测[J]. 煤炭学报, 2025, 50(1): 360-378.
LIANG Yunpei, LI Shang, LI Quanguai, et al. Prediction of gas concentration in the upper corner of mining working face based on the FEDformer-LGBM-AT architecture[J]. Journal of China Coal Society, 2025, 50(1): 360-378.
- [5] DU Feng, CUI Weilong, WANG Kai, et al. Research on precise quantitative traceability of combined gas extraction in close-distance coal seam group[J]. Environmental Earth Sciences, 2024, 83(21). DOI: [10.1007/s12665-024-11904-8](https://doi.org/10.1007/s12665-024-11904-8).
- [6] LIU Jun, WANG Zheng, WANG Qian, et al. Study on the application of comprehensive gas control technology for source-separated three-dimensional gas extraction in large mining faces[J]. ACS Omega, 2024, 9(30): 32777-32788.
- [7] 周福宝, 刘宏, 刘应科, 等. 煤层群开采工作面瓦斯精准定量溯源原理与技术[J]. 煤炭科学技术, 2021, 49(5): 11-18.
ZHOU Fubao, LIU Hong, LIU Yingke, et al. Principle and technology of precise and quantitative gas traceability in coal seam group mining face[J]. Coal Science and Technology, 2021, 49(5): 11-18.
- [8] 张增辉, 马文伟. 基于随机森林回归算法的回采工作面瓦斯涌出量预测[J]. 工矿自动化, 2023, 49(12): 33-39.
ZHANG Zenghui, MA Wenwei. Prediction of gas emission in mining face based on random forest regression algorithm[J]. Journal of Mine Automation, 2023, 49(12): 33-39.
- [9] 林海飞, 刘时豪, 周捷, 等. 基于 STL-EEMD-GA-SVR 的采煤工作面瓦斯涌出量预测方法及应用[J]. 煤田地质与勘探, 2022, 50(12): 131-141.
LIN Haifei, LIU Shihao, ZHOU Jie, et al. Prediction method and application of gas emission from mining workplace based on STL-EEMD-GA-SVR[J]. Coal Geology & Exploration, 2022, 50(12): 131-141.
- [10] 郭风景, 王斌, 贾澎涛, 等. 基于梯度提升决策树的瓦斯浓度在线预测[J]. 中国煤炭, 2023, 49(10): 61-67.
GUO Fengjing, WANG Bin, JIA Pengtao, et al. Online prediction of gas concentration based on gradient boosting decision tree[J]. China Coal, 2023, 49(10): 61-67.
- [11] 马恒, 任美学, 高科. 基于随机搜索优化 XGBoost 的瓦斯涌出量预测模型[J]. 中国安全生产科学技术, 2022, 18(5): 129-134.
MA Heng, REN Meixue, GAO Ke. Prediction model of gas emission amount based on XGBoost optimized with random search algorithm[J]. Journal of Safety Science and Technology, 2022, 18(5): 129-134.
- [12] 闵继源, 鲁统宇, 袁伟, 等. 基于 TW-Focal Loss 的债券违约预测及可解释性分析[J]. 系统管理学报, 2025, 34(3): 790-807.
MIN Jiyuan, LU Tongyu, YUAN Wei, et al. Bond default prediction based on TW-Focal Loss and its interpretability analysis[J]. Journal of Systems & Management, 2025, 34(3): 790-807.
- [13] 朱凯雯, 尤亚楠, 曹婧宜, 等. Hybrid-Gird: 遥感图像细粒度分类可解释方法[J]. 遥感学报, 2024, 28(7): 1722-1734.
ZHU Kaiwen, YOU Yanan, CAO Jingyi, et al. Hybrid-Gird: an explainable method for fine-grained classification of remote sensing images[J]. National Remote Sensing Bulletin, 2024, 28(7): 1722-1734.
- [14] 黄艺龙, 秦小林, 陈芋文, 等. 利用 LIME 对脓毒症预测模型进行可解释性分析[J]. 计算机应用, 2021, 41(增刊1): 332-335.
HUANG Yilong, QIN Xiaolin, CHEN Yuwen, et al. Interpretability analysis of sepsis prediction model using LIME[J]. Journal of Computer Applications, 2021, 41(S1): 332-335.