

# 基于QDA-RF-MLP的矿井突水水源辨识模型

张晨亮, 姚彬, 许明镜, 薛俊辰

河北工程大学地球科学与工程学院, 河北 邯郸

收稿日期: 2026年1月29日; 录用日期: 2026年2月23日; 发布日期: 2026年3月3日

## 摘要

受采动破坏与断裂构造影响, 深部煤矿突水水源的水化学特征难以解释, 具有混合性和重叠性, 导致传统单一模型识别精度受限。本文以安徽祁南煤矿为研究背景, 提出一种基于二次判别分析(QDA)、随机森林(RF)与多层感知机(MLP)的堆叠集成水源辨识模型。在系统解析矿井主要含水层水化学演化规律的基础上, 利用SMOTE技术重构非平衡数据集, 并引入水化学离子比率等高阶特征。通过五折交叉验证优选基学习器参数, 构建双层异构融合模型。实验结果表明: 该集成模型在独立测试集上的识别准确率达95.83%, F1分数达0.96, 显著优于单一最优基模型RF (92.7%), 有效克服了小样本条件下的过拟合问题, 对复杂混合水源具有较强的鲁棒性。

## 关键词

突水水源识别, 集成学习, 二次判别分析, 随机森林, 多层感知机

# Water Inrush Source Identification Model Based on QDA-RF-MLP

Chenliang Zhang, Bin Yao, Mingjing Xu, Junchen Xue

School of Earth Science and Engineering, Hebei University of Engineering, Handan Hebei

Received: January 29, 2026; accepted: February 23, 2026; published: March 3, 2026

## Abstract

Under the influence of mining disturbances and geological fault structures, the hydrochemical signatures of water inrush sources in deep coal mines often exhibit complex nonlinear mixing and overlapping characteristics. Consequently, traditional single-model approaches frequently suffer from limited identification accuracy. Taking the Qinan Coal Mine in Anhui Province as a case study, this paper proposes a stacking ensemble learning model for water source identification, integrating Quadratic Discriminant Analysis (QDA), Random Forest (RF), and Multilayer Perceptron (MLP). First, the hydrochemical evolution mechanisms of the primary aquifers were systematically analyzed. Subsequently,

the Synthetic Minority Over-sampling Technique (SMOTE) was applied to address dataset imbalance, and high-order features—specifically hydrochemical ion ratios, were introduced to enhance discriminative capability. A two-layer heterogeneous fusion model was then constructed, with base learner parameters optimized via five-fold cross-validation. Experimental results demonstrate that the proposed ensemble model achieved an accuracy of 95.83% and an F1-score of 0.96 on an independent test set. This performance significantly outperforms the best single base model (RF, 92.7%). The study confirms that the proposed method effectively mitigates overfitting risks associated with small sample sizes and exhibits strong robustness in identifying complex, mixed water sources.

## Keywords

Water Inrush Source Identification, Ensemble Learning, Quadratic Discriminant Analysis (QDA), Random Forest (RF), Multilayer Perceptron (MLP)

Copyright © 2026 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## 1. 引言

煤矿水害是制约我国煤炭资源安全开发的重大灾害之一，其突发性强、破坏力大，严重威胁矿井安全生产[1][2]。随着开采深度向深部延伸，含水层系统呈现多层叠置特征，断裂构造导水使得突水水源往往不是单一含水层，而是表现为多源混合及水质特征的非线性变化。准确、快速地判识突水水源，是制定科学合理的堵水方案、保障矿井安全生产的前提[3][4]。

长期以来，国内外学者在矿井水源识别领域开展了大量研究。早期的判识方法主要依赖水文地质专家的经验定性判断[5]或常规水化学图解分析[6][7]。随后，基于水温、水位等物理场的分析方法[8]及同位素示踪技术[9]也得到了一定应用。为实现水源的定量识别，Bayes 判别[10][11]、Fisher 判别[12][13]等数理统计方法被广泛引入。于小鸽等[14]对比了 Bayes 和 Fisher 判别法，发现 Bayes 法在多分类问题上具有优势；潘军等[15]则结合层次聚类与灰色关联度进行了水源识别研究；马雷等[16]基于 GIS 技术实现了水源的快速判别。

近年来，随着人工智能技术的发展，机器学习算法逐渐成为主流。BP 神经网络[17][18]、极限学习机[19][20]等浅层神经网络模型因其非线性映射能力而被广泛应用。同时，随机森林[21][22]、支持向量机[23][24]及 LightGBM [25]等算法在处理小样本数据方面表现出良好的鲁棒性。季媛[26]、董东林等[27]进一步探讨了不同机器学习模型在复杂水文地质条件下的适用性。然而，单一模型往往存在局限性：统计模型对数据分布假设严格，难以适应复杂水化学数据；浅层机器学习模型在样本不平衡时易陷入过拟合；而深度学习模型[28]虽具备强大的特征提取能力，但对训练数据量要求较高。

针对上述问题，集成学习[29]通过组合多个基学习器的预测结果，能显著降低单一模型的方差与偏差。针对皖北矿区深部含水层特征[30]，本文选取在统计学、集成树及神经网络领域具有代表性的 QDA、RF 及 MLP 作为基模型，构建 QDA-RF-MLP 堆叠集成识别模型，旨在通过挖掘水化学数据的高阶特征，提高矿井突水水源的判识精度。

## 2. 研究区概况

祁南煤矿位于淮北煤田宿县矿区，地质构造呈现“东西分区、南北分带”特征。井田被巨厚新生界

松散层覆盖，主要含水层自上而下包括：

(1) 新生界松散层第四含水层

简称“四含”，直接覆盖于煤系地层之上，厚度变化大。西北部以砾岩为主，渗透性强；中东部以粘土质砂为主，富水性较弱。其水质受大气降水与氧化环境影响， $SO_4^{2-}$ 含量较高。

(2) 二叠系煤系砂岩裂隙含水层

简称“裂隙水”，为煤层顶底板直接充水含水层，以静储量为主，富水性弱，但在构造破碎带处可能富集。

(3) 石炭系太原组灰岩含水层

简称“太灰”，岩溶裂隙发育不均，富水性中等，是下组煤开采的主要威胁。

(4) 奥陶系灰岩含水层

简称“奥灰”，岩溶极发育，富水性强，是区域性强含水基底。

各含水层在天然状态下水力联系较弱，但在采动与断层影响下，不同水源间易发生水力联系，形成成分复杂的水化学场。研究区含水层与隔水层结构如图1所示。

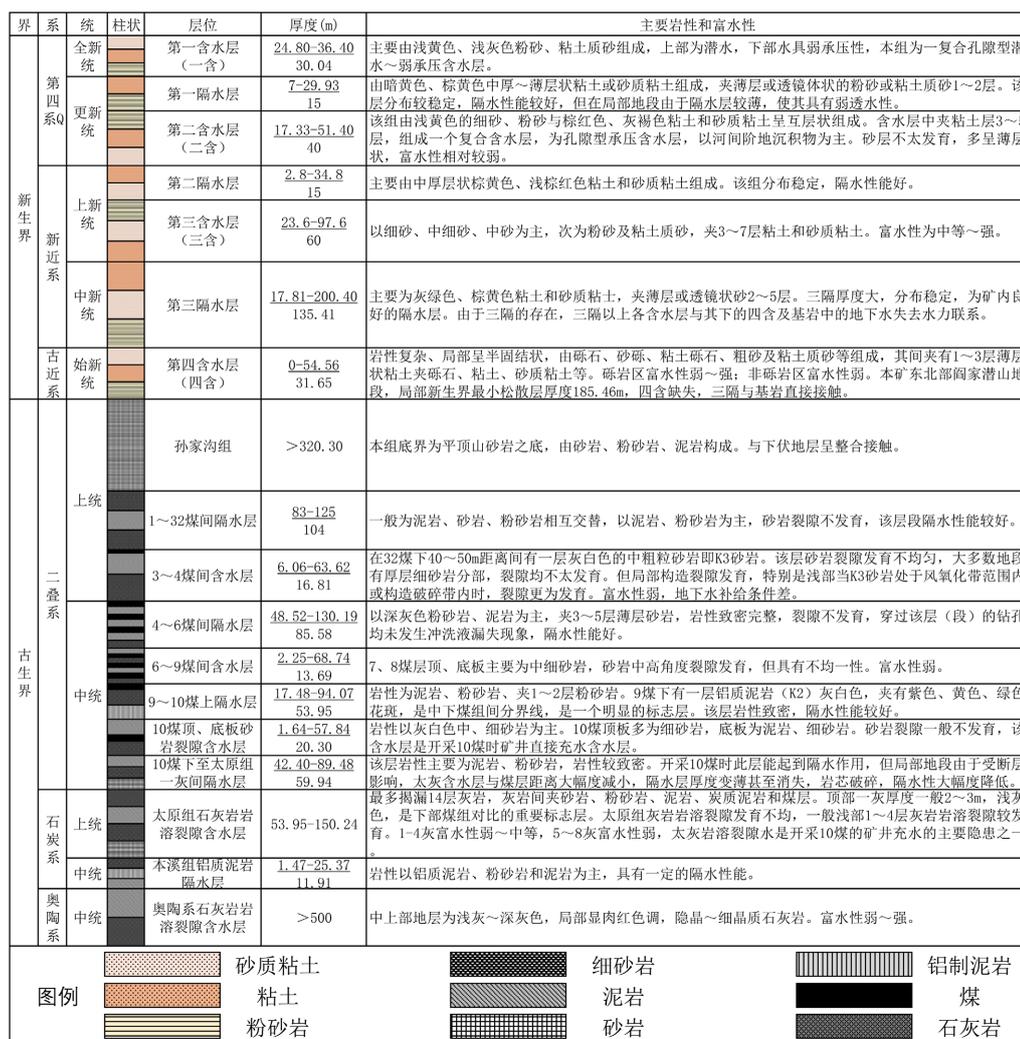


Figure 1. Schematic diagram of aquifers and aquicludes in the study area

图 1. 研究区含水层与隔水层结构示意图

### 3. 模型原理与架构

针对矿井水化学数据特征维度高、样本量少且分布复杂的特点，本研究选取三种差异化显著的模型作为基学习器，并基于 Stacking 策略构建集成模型。

#### 3.1. 二次判别分析模型(QDA)原理

二次判别分析(Quadratic Discriminant Analysis, QDA)是一种基于贝叶斯决策理论的统计分类算法。与线性判别分析(Linear Discriminant Analysis, LDA)假设所有类别共享同一协方差矩阵不同，QDA 允许不同水源类别拥有独立的协方差矩阵。假设第  $k$  类水源样本服从多变量高斯分布，其均值向量为  $\mu_k$ ，协方差矩阵为  $\Sigma_k$ 。QDA 的判别函数  $\delta_k$  定义为：

$$\delta_k(x) = -\frac{1}{2} \log |\Sigma_k| - \frac{1}{2} (x - \mu_k)^T \Sigma_k^{-1} (x - \mu_k) + \log \pi_k$$

式中， $\pi_k$  为第  $k$  类的先验概率。由于协方差矩阵  $\Sigma_k$  各不相同，判别函数中的二次项不会被抵消，因此 QDA 生成的决策边界为二次曲面。这种特性使其在处理不同含水层水质离散程度差异较大(例如，松散层水的方差显著大于深部灰岩水)的数据时，比线性模型具有更强的适应性。

#### 3.2. 随机森林模型(RF)原理

随机森林(Random Forest, RF)是基于 Bagging 策略的集成树算法。该算法通过“样本扰动”与“属性扰动”构建多棵决策树(如图 2 所示)。在训练过程中，采用 Bootstrap 自助法随机抽取样本子集，并在节点分裂时随机选择特征子集。最终分类结果由所有决策树投票决定。对于矿井水源识别任务，RF 对噪声和异常值具有鲁棒性，且能处理高维特征，不易过拟合，同时可输出特征重要性排序，辅助水文地质成因分析。

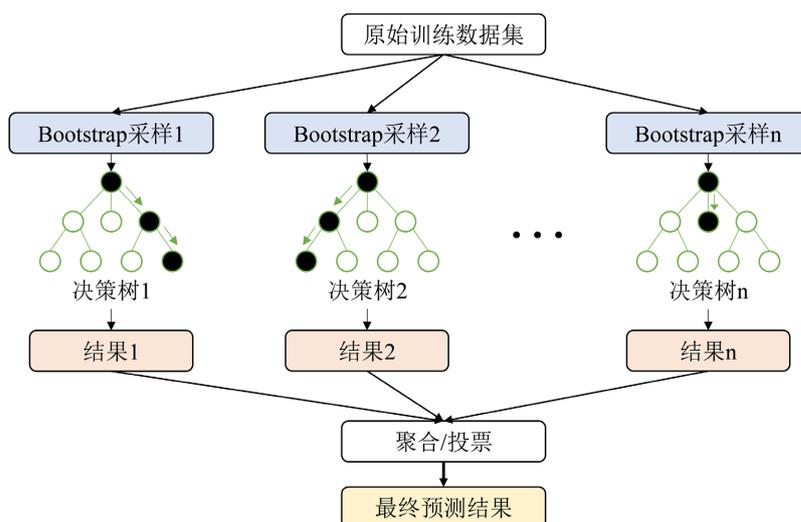


Figure 2. Schematic structure of Random Forest (RF) model  
图 2. 随机森林(RF)模型结构示意图

#### 3.3. 多层感知机(MLP)原理

多层感知机(Multilayer Perceptron, MLP)是一种前馈人工神经网络，由输入层、多个隐藏层和输出层组成(如图 3 所示)。各神经元之间通过加权连接，并引入非线性激活函数。MLP 具有通用函数逼近能力，

通过反向传播算法调整权重和偏置，能够从原始离子浓度数据中学习高阶抽象特征，挖掘离子间隐含的非线性耦合关系，适用于处理成因复杂、非线性强的水化学数据。

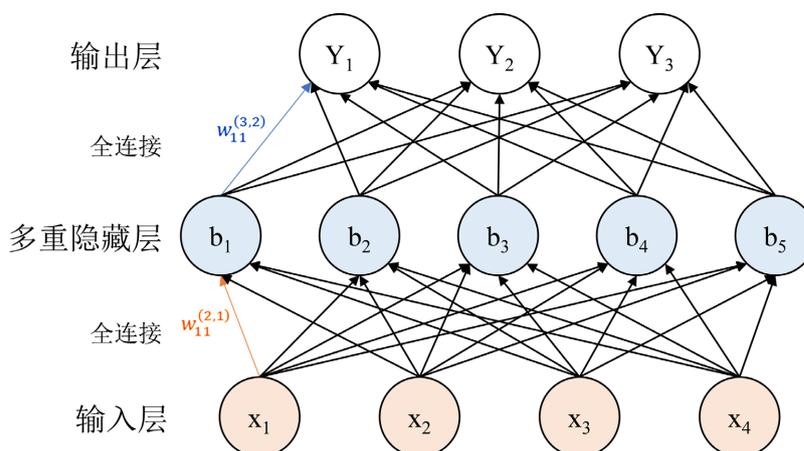


Figure 3. Topology schematic of Multilayer Perceptron (MLP)

图 3. 多层感知机(MLP)拓扑结构示意图

### 3.4. 模型训练流程

集成模型训练流程如图 4 所示。首先对原始水化学数据进行清洗与标准化处理，利用合成少数类过采样技术(Synthetic Minority Over-sampling Technique, SMOTE)平衡样本，SMOTE 方法只在训练集上进行使用，不涉及后续测试集的评估和判断；随后，并行训练 QDA、RF 和 MLP 三个基模型，并通过网格搜索优选超参数；最后，采用双层 Stacking 策略，将基模型的预测结果输入元学习器进行二次融合，输出最终的水源类别。

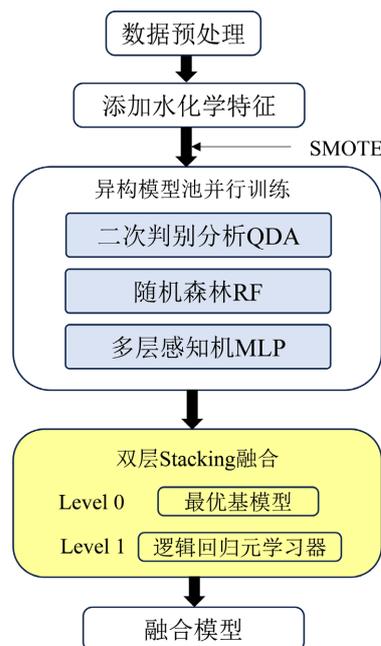


Figure 4. Training flowchart of the ensemble model based on QDA-RF-MLP

图 4. 基于 QDA-RF-MLP 的集成模型训练流程图

## 4. 模型应用与分析

### 4.1. 数据来源与预处理

数据来源于祁南煤矿井下实测水样及自建矿以来的水文地球化学观测台账。采样范围覆盖了矿井主要的充水含水层，包括松散层含水层、二叠系煤系砂岩裂隙水、石炭系太原组灰岩水及奥陶系灰岩水，并补充了部分采空区积水(老空水)样本。所有采集水样均严格按照《水质 32 种元素的测定电感耦合等离子体发射光谱法 HJ 776-2015》及离子色谱法标准进行化验，共获取包含  $\text{Na}^+\text{K}^+$ 、 $\text{Ca}^{2+}$ 、 $\text{Mg}^{2+}$ 、 $\text{Cl}^-$ 、 $\text{SO}_4^{2-}$ 、 $\text{HCO}_3^-$ 、 $\text{CO}_3^{2-}$  浓度及 pH 值在内的原始数据 228 组。

数据预处理首先利用阴阳离子电荷平衡原理校验，剔除相对误差绝对值大于 5% 的样本；随后采用四分位距法(Interquartile Range, IQR)剔除超出 1.5 倍 IQR 范围的离群点，最终保留 202 组有效样本。为消除量纲影响，对数据进行归一化处理。各含水层水样的统计特征如表 1 所示。针对奥灰水与老空水样本量较少导致的类别不平衡问题，本研究采用合成少数类过采样技术对训练集进行处理。需明确指出，SMOTE 仅在训练集上应用，用于生成合成样本来平衡各类别权重，而后续用于评估模型泛化性能的测试集，则完全由未经过任何采样处理的原始独立样本构成，以保证性能评估的客观性。

**Table 1.** Statistical characteristics of hydrochemical parameters of groundwater samples

**表 1.** 矿井各含水层水样水化学参数统计特征表

类型	数量	pH	$\text{Na}^+\text{K}^+$	$\text{Ca}^{2+}$	$\text{Mg}^{2+}$	$\text{Cl}^-$	$\text{SO}_4^{2-}$	$\text{HCO}_3^-$	$\text{CO}_3^{2-}$
二含水	5	7.3~8.06	46.44~314.59	0.00~78.16	26.34~69.41	0.00~227.44	10.70~309.52	32.10~645.18	0.00~457.89
		7.63	144.83	51.41	42.85	69.62	107.11	392.92	91.58
三含水	5	7.48~8.36	189.82~344.63	0.00~143.81	9.72~102.90	0.00~203.25	75.73~740.06	289.01~507.44	0.00~359.77
		7.97	264.27	72.79	58.78	113.27	388.29	410.09	74.81
四含水	57	5.83~13.32	0.25~977.91	0.00~210.95	3.68~144.24	0.00~887.23	19.76~2266.68	0.00~602.29	0.00~292.90
		8.31	355.34	73.41	58.77	131.91	689.15	323.38	20.79
裂隙水	60	7.3~10.78	3.10~860.78	0.00~104.20	0.24~69.45	0.00~623.55	0.41~947.50	1.60~1554.30	0.00~1553.00
		8.47	430.4	16.44	10.62	116.8	180.35	500.74	239.19
太灰水	54	7.01~8.63	28.02~483.25	3.96~251.23	0.00~175.32	0.00~262.29	4.13~693.55	130.49~793.82	0.00~466.19
		7.51	238.68	138.24	71.37	178.89	411.65	471.3	84.27
奥灰水	7	7.11~7.97	64.58~432.95	17.17~238.10	6.45~156.41	108.68~279.58	218.56~645.39	333.67~476.61	0.00~0.00
		7.35	214.49	144.72	75.17	206.05	507.88	384.23	0
老空水	14	7.46~9.54	417.27~983.04	0.59~17.40	0.99~13.43	92.64~368.53	2.47~563.89	599.88~1698.55	0.00~176.77
		8.49	583.32	7.71	4.06	147.82	105.84	1087.74	58.16

为进一步增强模型对水化学差异的识别能力，在 pH 值及 7 项主要离子浓度这 8 项原始特征的基础上，本研究构建了 24 项衍生特征，主要可分为四类：

(1) 直接离子比值：包括反映碳酸盐岩溶解平衡的  $\text{Ca}^{2+}/\text{Mg}^{2+}$  比，指示水体演化路径的  $\text{Cl}^-/\text{SO}_4^{2-}$  比。

(2) 毫克当量百分比特征：计算了  $\text{Ca}^{2+}$ 、 $\text{Mg}^{2+}$ 、 $\text{Na}^+\text{K}^+$  的阳离子百分比，以及  $\text{Cl}^-$ 、 $\text{SO}_4^{2-}$ 、 $\text{HCO}_3^-$  的阴离子百分比，用以精确量化 Piper 图中的点位信息。

(3) 综合与多样性指标：包括基于主要离子总和和估算的总溶解固体，以及阴离子浓度标准差。

(4) 统计变换特征：如对原始浓度进行对数变换、Z 值标准化及归一化，以探索数据的不同尺度。

综上，模型输入的特征总数共计 32 项。这些衍生特征旨在从不同维度刻画水化学组成，为模型提供超越常规图解法的判别信息。

基于处理后的数据绘制 Piper 三线图(如图 5 所示)。图中显示，太灰水与奥灰水投点集中于上方，均为 Ca-Mg-SO<sub>4</sub> 型，两者重叠严重，难以通过常规图解法区分；裂隙水与老空水集中于右下角，为 Na-HCO<sub>3</sub> 型，同样存在部分重叠；四含水分布离散，主要为 Na-SO<sub>4</sub> 型，但部分样本向裂隙水区域延伸。含水层水样点在图上的混杂与重叠，表明各水源间存在复杂的水力联系或混合作用，单纯依赖图形特征难以实现高精度的水源判别，需借助非线性模型挖掘深层特征。

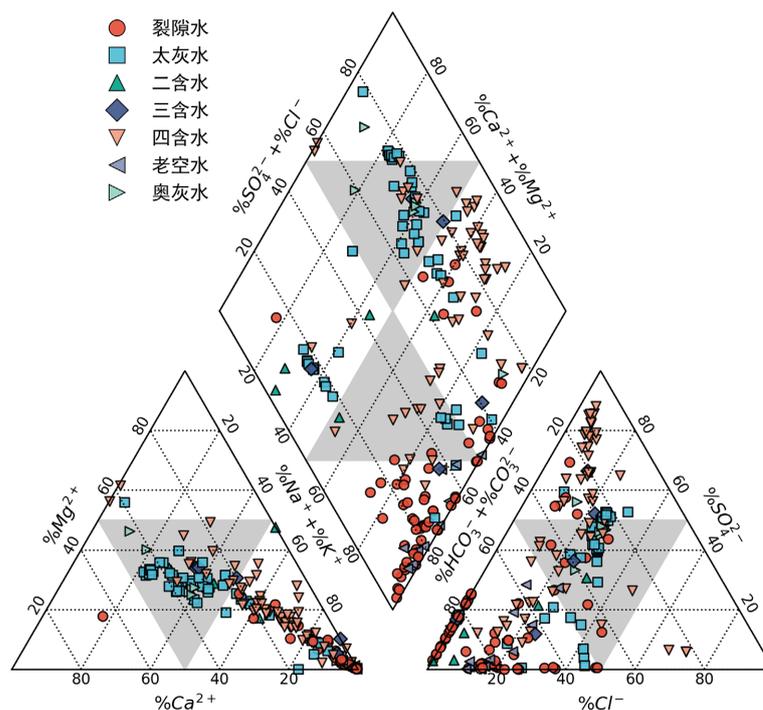


Figure 5. Piper trilinear diagram of groundwater samples in the study area  
图 5. 研究区地下水样 Piper 三线图

## 4.2. 模型参数设置与融合策略

基于 Python 的 Scikit-learn 框架搭建模型。为了最大化模型性能，对各基模型进行了精细的超参数调优。各模型的最优参数设置如表 2 所示。

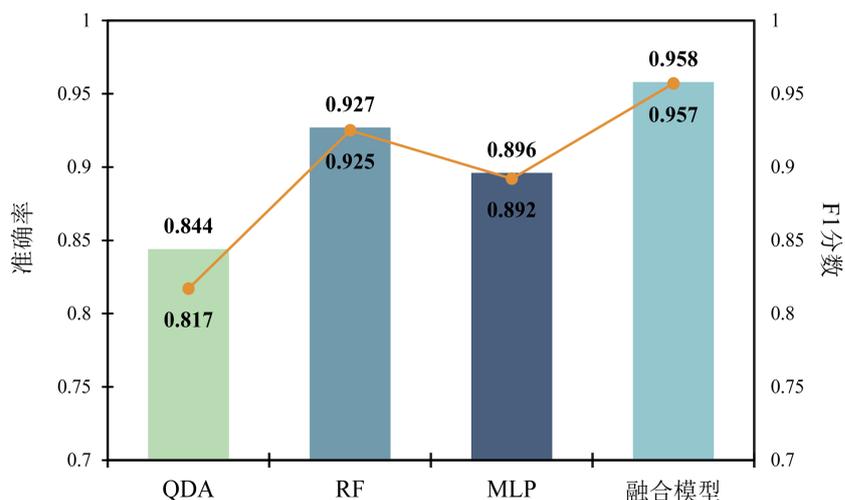
为实现各基模型的互补，采用双层 Stacking 集成架构。Layer-0 为基学习层，包含优化后的 QDA、RF 和 MLP 模型；Layer-1 为元学习层，选用具有概率解释性的逻辑回归(Logistic Regression)模型。训练过程中，在 Layer-0 阶段采用 5 折交叉验证(5-Fold Cross Validation)生成元特征。即将训练集分为 5 份，轮流用其中 4 份训练基模型，预测剩余 1 份，拼接成与原始训练集等长的“预测概率矩阵”作为 Layer-1 的输入。这种机制使元学习器能学习不同基模型在特定样本上的可信度，纠正单一模型的偏差。

## 4.3. 模型对比与结果分析

将数据集按 8:2 随机划分为训练集和独立测试集。为验证融合策略的有效性，将 QDA-RF-MLP 融合模型与三个单一基模型在独立测试集上的性能进行对比(如图 6 所示)。

**Table 2.** Optimal parameter settings for base models**表 2.** 基模型最优参数设置表

模型类型	关键参数	参数值说明
QDA	reg_param	0.1 (正则化参数, 用于解决小样本协方差矩阵估计不稳定问题)
	n_estimators	150 (决策树数量)
RF	max_depth	8 (树的最大深度, 防止过拟合)
	criterion	Gini (基尼系数)
	hidden_layer_sizes	(100, 50, 25) (三层隐藏层结构)
MLP	activation	ReLU (激活函数)
	solver	Adam (优化器, 初始学习率 0.001)

**Figure 6.** Comparison of accuracy and F1 score of different models on test set**图 6.** 不同模型在测试集上的准确率与 F1 分数对比

实验结果显示, 单一模型中, RF 表现最佳, 准确率为 92.7%, F1 分数为 0.925, 体现了集成树模型在处理高维小样本数据时的优势; MLP 次之, 准确率为 89.6%, 显示了其对非线性特征的提取能力; QDA 准确率为 84.4%, 在处理复杂边界时略显不足。融合后, QDA-RF-MLP 融合模型的准确率达到 95.83%, F1 分数为 0.96。与表现最好的单一模型 RF 相比, 准确率提升了 3.13%; 与 MLP 和 QDA 相比, 提升幅度更为显著。这一结果表明, Stacking 策略有效地结合了 QDA 的概率统计特性、RF 的抗噪能力和 MLP 的非线性映射能力。通过元学习器的二次融合, 模型能够修正单一算法在特定样本上的误判, 从而构建出泛化能力更强的分类边界。

为探究模型高精度决策背后的水化学依据, 本研究采用 SHAP (SHapley Additive exPlanations) 值对融合模型进行可解释性分析(结果如图 7 所示, 仅显示前 20 项)。分析显示, 对模型输出贡献最大的前五项特征均为原始离子浓度指标, 按贡献度依次为:  $\text{Na}^+\text{K}^+$ 、 $\text{HCO}_3^-$ 、 $\text{Ca}^{2+}$ 、 $\text{SO}_4^{2-}$  和  $\text{Cl}^-$ 。这一排序具有明确的地质意义:  $\text{Na}^+\text{K}^+$  与  $\text{HCO}_3^-$  的高贡献度, 直接指向对 Na-HCO<sub>3</sub> 型水(裂隙水、老空水)的识别; 而  $\text{Ca}^{2+}$  与  $\text{SO}_4^{2-}$  的重要性, 则对应了对 Ca-Mg-SO<sub>4</sub> 型水(太灰水、奥灰水)的判别。总溶解固体估算值(TDS\_estimate)和阴离子总和(Anion\_Diversity)也位列前茅, 分别作为矿化度和水化学复杂性的综合指标发挥作用。相比

之下，后续的离子比值及各类统计变换特征的贡献度均显著降低。这表明，模型决策主要依赖于表征水体基本化学类型的宏量离子浓度及其综合性指标，其决策逻辑符合水文地球化学的基本原理，验证了模型的合理性。

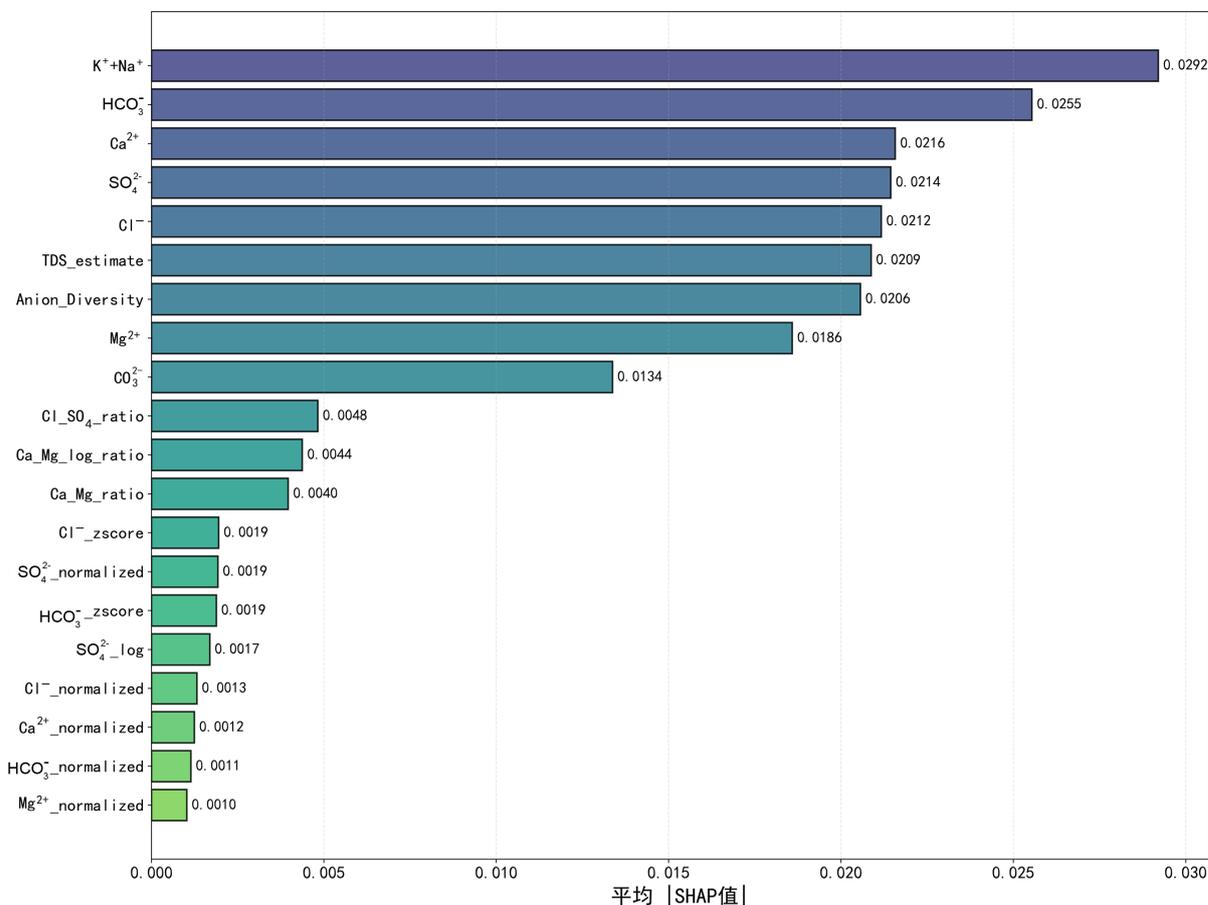


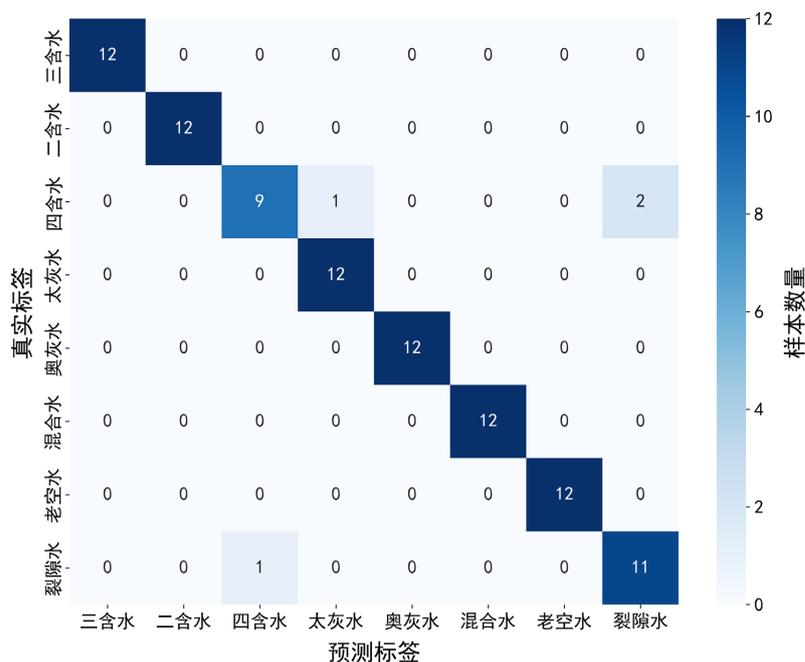
Figure 7. SHAP summary plot for global feature importance ranking

图 7. 基于 SHAP 值的模型特征全局重要性排序图

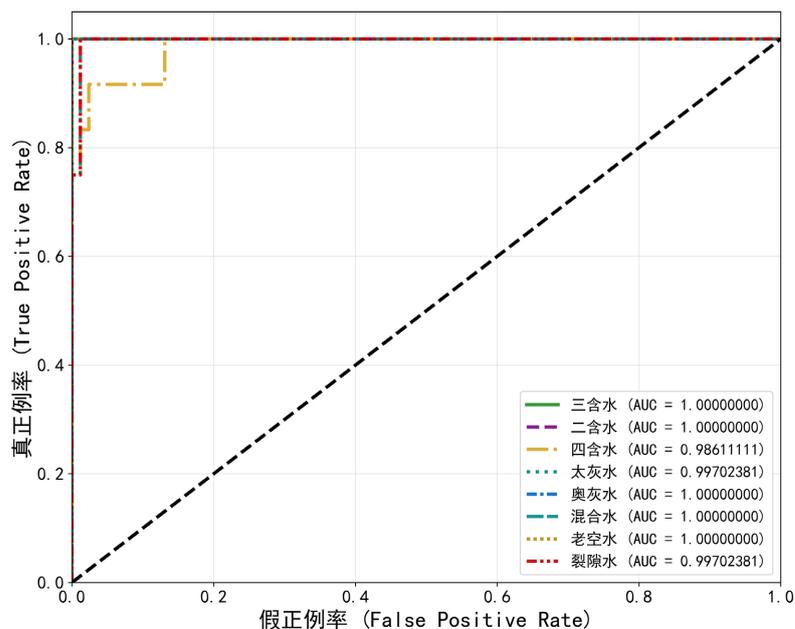
融合模型的混淆矩阵(图 8(a))显示，模型对太灰水、奥灰水及老空水等类别的识别准确率达到 100%。主要的误判发生在“四含水”与“裂隙水”之间，这与 Piper 图中部分四含水点向 Na-HCO<sub>3</sub> 区域延伸的现象相符，可归因于浅部含水层局部发生的混合作用。ROC 曲线(图 8(b))进一步证实了模型的强大分类能力，除太灰水与裂隙水的 AUC 值均为 0.997，四含水的 AUC 值为 0.986，略低于其他类别外，其余水源类别的 AUC 值均达到 1.00，与混淆矩阵中近乎完美的分类表现一致。

测试集上模型表现出极高的识别精度，与 Piper 图中水源投影点严重重叠的现象形成了直观矛盾。这一矛盾揭示了传统图解方法与机器学习模型在分析维度上的根本差异。Piper 三线图的核心信息来源于主要阴阳离子的毫克当量百分比，其重叠仅表明不同水源在离子相对比例上可能接近。然而，本模型识别所依赖的是一个包含 32 项特征的体系，其中不仅包含了比例信息，更关键的是引入了各离子的绝对浓度及多项衍生指标。正是这些多维特征，捕捉并量化了 Piper 图所丢失的关键判别信息。例如，尽管太灰水与奥灰水在 Piper 图上同属 Ca-Mg-SO<sub>4</sub> 型，但模型可利用 Ca<sup>2+</sup>/Mg<sup>2+</sup> 比值来反映方解石与白云石溶解平衡的细微差异，或通过 Na<sup>+</sup>+K<sup>+</sup> 与 HCO<sub>3</sub><sup>-</sup> 的绝对浓度来精准刻画 Na-HCO<sub>3</sub> 型水的特征。因此，模型的高精度

并非推翻了传统图示结果，而是通过更全面的数值化描述，从机理上分辨了那些在二维投影中重叠、实则源于不同水文地球化学过程的本质差异。需要指出的是，奥灰水样本量极少，虽训练中应用 SMOTE 技术进行了样本平衡，但极小的初始样本量可能导致对该类别水化学特征空间边界的估计不够稳健，其结果的统计不确定性较高。测试集上奥灰水 AUC 值为 1.00 的完美表现，有可能是因为测试样本恰好与训练样本的特征高度相似，未来需更多数据验证。



(a)



(b)

Figure 8. Confusion matrix and ROC curve of the fusion model

图 8. 融合模型混淆矩阵与 ROC 曲线

## 5. 结论

(1) 祁南煤矿地下水化学场存在显著分异：深部岩溶水(太灰、奥灰)受溶滤作用控制，呈高矿化度 Ca-Mg-SO<sub>4</sub> 型；煤系水(裂隙、老空)受阳离子交替吸附作用控制，呈 Na-HCO<sub>3</sub> 型；浅部四含水受氧化与混合作用影响，呈 Na-SO<sub>4</sub> 型。但不同含水层在常规图解上存在重叠，界限模糊，需借助非线性模型进行识别。

(2) 构建的 QDA-RF-MLP 堆叠集成模型有效融合了统计模型、集成树与深度神经网络的优势。实验结果表明，该模型在独立测试集上的准确率达 95.83%，优于任一单一基模型(RF: 92.7%, MLP: 89.6%, QDA: 84.4%)，有效克服了小样本、非平衡数据条件下的识别难题。

(3) 该研究提出的模型架构通过双层融合策略，显著提升了水源判识的准确性与鲁棒性。基于该算法可开发智能化的水源识别系统，为矿井突水水源的快速诊断及防治水决策提供科学依据。

## 致 谢

感谢河北工程大学地球科学与工程学院提供的科研平台与学术指导。此外，本文引用了国内外多位学者的研究成果与学术观点，这些文献为本研究提供了坚实的理论基础，在此向所有原作者致以最诚挚的谢意。

## 参考文献

- [1] 黎良杰, 钱鸣高, 李树刚. 断层突水机理分析[J]. 煤炭学报, 1996, 21(2): 119-123.
- [2] 程磊, 孙洁. 2016-2022 年我国煤矿事故统计与规律分析[J]. 煤炭工程, 2023, 55(11): 125-129.
- [3] 崔中良, 黄保胜, 龙向前. 矿井突水水源判别研究现状及发展趋势[J]. 中国钨业, 2018, 33(4): 42-50.
- [4] 靳德武. 煤矿水害防治中的综合水文地质分析方法[J]. 煤田地质与勘探, 1998, 26(2): 54-56.
- [5] 刘文明, 桂和荣, 孙雪芳, 等. 潘谢矿区矿井突水水源的 QLT 法判别[J]. 中国煤炭, 2001, 27(5): 31-34.
- [6] 宋宝文. 根据水化学成份预测矿井突水水源几例[J]. 煤田地质与勘探, 1981, 44(2): 65-69.
- [7] 贺亮亮, 吕广罗, 胡安焱, 等. 基于水化学特征分析的矿井突水水源判别[J]. 中国煤炭地质, 2022, 34(6): 34-39.
- [8] 桑向阳, 林云, 刘保民, 等. 基于水位、水温、突水量和水质的充水水源识别神经网络模型[J]. 河南理工大学学报(自然科学版), 2024, 43(5): 36-42.
- [9] 潘国营, 王素娜, 孙小岩, 等. 同位素技术在判别矿井突水水源中的应用[J]. 矿业安全与环保, 2009, 36(1): 32-34, 90.
- [10] 宋立兵, 董东林, 王振荣, 等. 基于多方法耦合的矿井突水水源识别模型研究[J]. 煤炭工程, 2022, 54(2): 140-146.
- [11] Yan, B., Ren, F., Cai, M. and Qiao, C. (2020) Bayesian Model Based on Markov Chain Monte Carlo for Identifying Mine Water Sources in Submarine Gold Mining. *Journal of Cleaner Production*, **253**, Article ID: 120008. <https://doi.org/10.1016/j.jclepro.2020.120008>
- [12] 申雄, 栗继祖, 赵德康. 基于 PCA 的 Fisher 多元统计方法识别矿井充水水源[J]. 矿业安全与环保, 2024, 51(3): 144-152.
- [13] 黄平华, 陈建生. 基于多元统计分析的矿井突水水源 Fisher 识别及混合模型[J]. 煤炭学报, 2011, 36(S1): 131-136.
- [14] 于小鸽, 裴富华, 刘焱菲, 等. 贝叶斯判别和 Fisher 判别在矿井突水水源判别中的应用比较[J]. 中国科技论文, 2022, 17(1): 9-14.
- [15] 潘军. 基于 Piper-层次聚类-灰色关联度的突水水源识别[J]. 煤炭科学技术, 2024, 52(S1): 221-227.
- [16] 马雷, 钱家忠, 赵卫东. 基于 GIS 和水质水温的矿井突水水源快速判别[J]. 煤田地质与勘探, 2014, 42(2): 49-53.
- [17] 徐星, 王公忠. BP 神经网络在矿井突水水源识别中的应用[J]. 煤炭技术, 2016, 35(7): 144-146.
- [18] 刘伟韬, 李蓓蓓, 杜衍辉, 等. 基于改进的 SSA-BP 神经网络的矿井突水水源识别模型研究[J]. 工矿自动化, 2024, 50(2): 98-105.
- [19] 于小鸽, 刘焱菲, 翟培合. 基于 PCA-AWOA-ELM 模型的矿井突水水源识别[J]. 煤炭科学技术, 2023, 51(3): 182-189.

- [20] 孙文洁, 杨恒, 李祥, 等. PCA 与 ELM 模型相结合的矿井突水水源快速识别方法研究[J]. 煤炭工程, 2020, 52(1): 111-115.
- [21] 黄敏, 毛岸, 路世昌, 等. 矿井突水水源识别的主成分分析-混沌麻雀搜索-RF 模型[J]. 安全与环境学报, 2023, 23(8): 2607-2614.
- [22] 肖观红, 鲁海峰. 基于 PCA-GA-RF 的矿井突水水源快速识别模型[J]. 煤矿安全, 2024, 55(6): 184-191.
- [23] 邵良杉, 李相辰. 基于 MIV-PSO-SVM 模型的矿井突水水源识别[J]. 煤炭科学技术, 2018, 46(8): 183-190.
- [24] Wei, Z., Dong, D., Ji, Y., Ding, J. and Yu, L. (2022) Source Discrimination of Mine Water Inrush Using Multiple Combinations of an Improved Support Vector Machine Model. *Mine Water and the Environment*, **41**, 1106-1117. <https://doi.org/10.1007/s10230-022-00884-5>
- [25] Ji, Y., Dong, D., Mei, A. and Wei, Z. (2022) Study on Key Technology of Identification of Mine Water Inrush Source by PSO-LightGBM. *Water Supply*, **22**, 7416-7429. <https://doi.org/10.2166/ws.2022.323>
- [26] 季媛. 矿井突水水源快速识别关键技术研究[D]: [博士学位论文]. 徐州: 中国矿业大学, 2022.
- [27] 董东林, 张陇强, 张恩雨, 等. 基于 PSO-XGBoost 的矿井突水水源快速判别模型[J]. 煤炭科学技术, 2023, 51(7): 72-82.
- [28] Hinton, G.E. and Salakhutdinov, R.R. (2006) Reducing the Dimensionality of Data with Neural Networks. *Science*, **313**, 504-507. <https://doi.org/10.1126/science.1127647>
- [29] 徐继伟, 杨云. 集成学习方法: 研究综述[J]. 云南大学学报(自然科学版), 2018, 40(6): 1082-1092.
- [30] 孙林华, 桂和荣. 皖北桃源矿深部含水层地下水地球化学数理统计分析[J]. 煤炭学报, 2013, 38(S2): 442-447.