

<https://doi.org/10.3799/dqkx.2024.114>



基于 AI 技术的滑坡易发性制图研究进展

刘 聪^{1,2,3}, 陈永吉¹, 张 条¹, 卢全中^{1,2,3*}

1. 长安大学地质工程与测绘学院, 陕西西安 710054
2. 自然资源部地裂缝地质灾害重点实验室, 江苏南京 210018
3. 自然资源部陕西西安地裂缝与地面沉降野外科学观测研究站, 陕西西安 710054

摘 要: 基于 AI 技术的滑坡易发性制图具有高效、准确的优点, 为推进其在滑坡灾害防治中的应用, 在介绍和总结机器学习、深度学习以及集成学习模型的原理及特点的基础上, 选择支持向量机、深度随机森林和随机森林等代表性模型在陕西省略阳县进行了应用分析; 探讨了 AI 技术在滑坡易发性领域的应用及发展方向。研究结果表明: 基于决策树的集成学习模型相比于逻辑回归、支持向量机等, 表现出更高的效能, AUC 值在 0.90 以上; LSM (landslide susceptibility mapping) 中常用的类不平衡采样策略下, 基于 Boosting 的集成模型具有优势, 并且其受采样比的影响相对较小; 对抗生成网络可以提高在深度学习模型在数据限制情况下的性能, 本文中 AUC 值从 0.77 提升至 0.82; 滑坡理论模型与 AI 数据模型相结合, 具有巨大的潜力; 通过充分利用时序数据的 AI 模型可以提升模型的性能, 并有助于揭示滑坡的链式灾害效应和时空演化特征; 进行各种学习模型的系统性研究, 对于 AI 技术在滑坡易发性制图中的应用具有重要的意义。

关键词: 滑坡易发性制图; 深度学习; 机器学习; 集成学习; 工程地质。

中图分类号: P694

文章编号: 1000-2383(2025)06-2270-14

收稿日期: 2024-04-26

Landslide Susceptibility Mapping Based on AI Technology

Liu Cong^{1,2,3}, Chen Yongji¹, Zhang Tiao¹, Lu Quanzhong^{1,2,3*}

1. School of Geological Engineering and Geomatics, Chang'an University, Xi'an 710054, China
2. Key Laboratory of Earth Fissures Geological Disaster, Ministry of Natural Resources, Nanjing 210018, China
3. Field Scientific Observation and Research Station of Ground Fissure and Land Subsidence in Xi'an of Shaanxi, Ministry of Natural Resources, Xi'an 710054, China

Abstract: Utilizing AI technology for landslide susceptibility mapping offers the advantages of efficiency and accuracy. To promote its application in landslide disaster prevention and control, in this article it introduces and summarizes the principles and characteristics of machine learning, deep learning, and ensemble learning models. Representative models such as Support Vector Machines, Deep Random Forests, and Random Forests were applied for analysis in Lueyang County, Shaanxi Province. It discusses the application and development directions of AI technology in the field of landslide susceptibility. The results indicate that ensemble learning models based on decision trees, compared to logistic regression and support vector machines, demonstrate higher efficacy with AUC values above 0.90. Under the commonly used class imbalance sampling strategy in LSM (landslide

基金项目: 自然资源部地裂缝地质灾害重点实验室开放基金(No.EFGD2021-05-01); 国家自然科学基金项目(No.42341101), 陕西省自然科学基金(No.2023-JC-YB-231)。

作者简介: 刘聪 (1979-), 男, 博士, 副教授, 研究生导师, 主要从事工程地质与地质灾害防治方面的教学与研究工作。ORCID: 0009-0000-0339-0874. E-mail: liucong@chd.edu.cn

*** 通讯作者:** 卢全中, 教授, 博士研究生导师, 主要从事地质工程与岩土工程方面的研究与教学工作。E-mail: dcdgx14@chd.edu.cn

引用格式: 刘聪, 陈永吉, 张条, 卢全中, 2025. 基于 AI 技术的滑坡易发性制图研究进展. 地球科学, 50(6): 2270-2283.

Citation: Liu Cong, Chen Yongji, Zhang Tiao, Lu Quanzhong, 2025. Landslide Susceptibility Mapping Based on AI Technology. *Earth Science*, 50(6): 2270-2283.

susceptibility mapping), ensemble models based on Boosting show advantages and are relatively less affected by sampling ratios. Generative Adversarial Networks can enhance the performance of deep learning models under data constraints, where in this study, the AUC value increased from 0.77 to 0.82. Combining landslide theoretical models with AI data models has great potential; leveraging AI models that fully utilize time-series data can improve model performance and help reveal the chain disaster effects and spatio-temporal evolution characteristics of landslides; conducting systematic studies of various learning models is of significant importance for the application of AI technology in landslide susceptibility mapping.

Key words: landslide susceptibility mapping; deep learning; machine learning; ensemble learning; engineering geology.

0 引言

全球自然灾害频发并且还在不断增加.自1990年以来,已报告的灾害已在全球造成160多万人死亡,经济损失估计为10万亿美元,每年2 600~3 100亿美元,防灾减灾是人类发展过程中的永恒课题.而滑坡作为自然界中分布最广泛的灾害,由于其破坏性和广泛性越来越受到世界各地的重视.目前,气候变化极大地增加了洪水和暴雨等特定极端天气和气候事件的频率(Stott, 2016),从而加速了全球滑坡的发生.为了应对这一挑战,LSM(landslide susceptibility mapping)被广泛应用于滑坡易发性评价中.LSM是指根据当地地形条件,说明一个地区发生山体滑坡的可能性.它是基于在发生滑坡的相似条件下会发生滑坡的假设,将某一地区不同滑坡易发性区域的空间分布可视化.

LSM的生成方法主要分为定性方法和定量方法.定性方法主要是根据灾害区的工程地质条件结合专家的个人经验进行评价,这种评价方法紧密结合了人类对滑坡灾害的前沿研究成果,具有极强的可解释性并且模型容易建立,有学者利用启发式决策树建立了全球滑坡灾害预测系统,在一些地区取得了成功(Kirschbaum and Stanley, 2018).但是这种方法不可避免地具有主观性(van Westen *et al.*, 2006).定量评价主要包括机制驱动模型和数据驱动模型.机制驱动是指基于数学物理理论和一定假设构架的计算模型,主要基于极限平衡法和有限元数值模拟.以极限平衡法为基础发展了很多的区域性滑坡评价计算模型,包括无限边坡(Sinmap)模型、Trigrs模型及Scoops3D模型等.但滑坡易发性受到多种因素的影响,包括地形特征(如海拔、岩性)、场地条件(如土壤湿度、地震活动性)以及潜在的触发因素(如极端降雨或地震).由于这些因素的复杂性和多样性,开发一个统一的物理模拟方法来预测滑坡易发性是非常困难的(Ward *et al.*, 2020).数据驱动模型包含传统数学统计方法,如逻辑回归、权重

证据法和频率比(李郎平等, 2017),还包括近年来越来越流行的机器学习方法,例如随机森林(random forest, RF)、支持向量机(support vector machine, SVM)、人工神经网络(artificial neural network, ANN)和深度学习算法.

图1是近年机器学习在LSM中的发文趋势, Liu *et al.* (2023)的研究揭示了在Web of Science (WOS)核心数据库中机器学习学习概念的流行,越来越多的学者聚焦于采用这些方法来挖掘数据特征进行LSM.

Liu *et al.* (2023)的综述文章很好回顾了2023年以前的相关研究,但忽视了Transformer在图像视觉领域的性能优势以及时间序列数据在滑坡易发性制图中的作用.事实上,随着数据采集技术的爆炸式增长,深度学习方法在该领域越来越受欢迎,将映射精度和泛化提升到一个新的水平(Ngo *et al.*, 2021).在LSM中,AI技术的应用已经促成了多种模型的开发,这些模型大致可分为两大类:一类是基础机器学习模型,如逻辑回归(logistic regression, LR)、决策树(decision tree, DT)和支持向量机等;另一类是深度学习模型,包括卷积神经网络(convolutional neural network, CNN)和长短期记忆

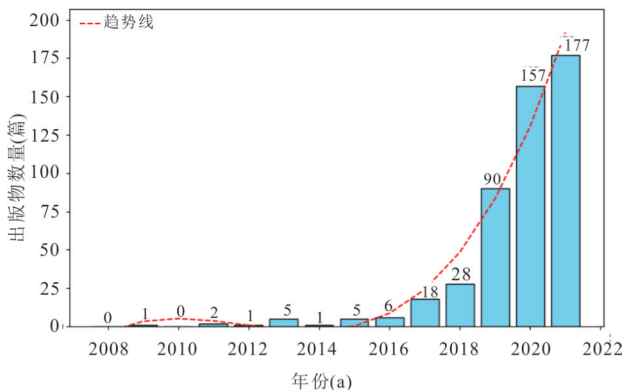


图1 机器学习的发文趋势(据 Liu *et al.*, 2023)

Fig.1 Publication trends on machine learning (Liu *et al.*, 2023)

网络(long short-term memory network,LSTM).此外,集成学习模型如随机森林和深度随机森林等,是对基础模型采用集成方法进一步的改进.由于LSM的研究区域差异,尚未有哪一个模型被证明是最有效的,但有大量学者得出了在特定研究区域的最优模型,同时应用于LSM中的模型也不断朝着高级、更复杂的趋势发展.Fang *et al.* (2020)证明了集成CNN可以提高机器学习模型在LSM中的性能.Lyu *et al.* (2023) 利用机器学习、深度学习模型和密集连接的卷积网络进行了湖北省秭归县、巴东县的滑坡易发性分区,发现深度学习模型和密集网络具有明显优势.本文分析了这些模型的原理和适用性,并进行了案例研究,比较了不同采样策略下多种学习模型的性能,通过分析2023年至今WOS核心数据库中的相关文献,展望了未来研究趋势.

1 人工智能方法基本原理与适用性分析

1.1 机器学习基本原理与适用分析

1.1.1 机器学习 机器学习是一种从数据中提取知识的方法,是统计学、人工智能和计算机科学交叉的研究领域.与传统编程中的规则和决策逻辑由人类开发者明确编写的不同,机器学习使模型能够

自动从大量数据中学习这些规则,从而对新的、未见过的数据做出准确预测或决策.在LSM任务中,常用的机器学习模型有逻辑回归、决策树和支持向量机.表1简单介绍了这几种模型的基本原理,对相关研究进行了分析,总结了不同模型在LSM中应用的优劣势.

1.1.2 集成模型(ensemble learning) 集成学习是指构建并结合多个学习器来完成任务,这一方法是为了结合多种学习方法的优势而弥补它们的劣势而设计的.集成模型主要类型有:Bagging、Boosting and Stacking.根据集成学习方法可以对多种基础机器学习模型进行改进,例如极端梯度增强(XGBoost)和随机森林等集成模型在LSM中有着广泛的应用,表2展示了两种集成方法的基本原理及一些相关研究,并分析了其优劣势.

1.2 深度学习基本原理

1.2.1 深度学习的基础 深度学习是机器学习的一个分支,是从数据中学习的一种新的方式,强调的是从连续的层中进行学习.决策树只进行了一层的学习,并且这个过程是人为编程来实现的.但在深度学习中存在数十个甚至上百个表示层,并且学习的过程是完全自动的.张量被用作数据容器,0D张量指的是1个数,1D张量指的是向量即一个数组,2D张量则是矩阵,当矩阵向更高维度推广,也就

表1 机器学习基本原理及适用性分析
Table 1 Basic principles and applicability analysis of machine learning

模型	基本原理	在LSM中的应用与研究	优劣势分析
逻辑回归	线性模型广泛应用于二元分类问题,通过输入变量的线性函数(决策边界)将研究对象分为两个类别.	逻辑回归的目标是找到最合理的模型来表述滑坡与非滑坡的不同条件因素(如坡度、高程和坡向)之间的关系.	逻辑回归是线性模型,能够清晰解释每个特征的影响方向和大小,但是其不能很好地处理特征与滑坡之间的非线性关系.
决策树	决策树是一种基本的分类工具,通过不断将空间中的对象分组,逐步提高分类精度.它会根据不同特征逐层分类,直到达到所需的精确度.	Saito <i>et al.</i> (2009)将决策树用于分析日本明石山大范围滑坡易发性,揭示滑坡发生与地形、地质条件的关系,指出决策树的优势是可解释性强.Alkhasawneh <i>et al.</i> (2014)用决策树模型确定造成马来西亚槟城岛大范围滑坡的24个因素的重要性,并通过集成学习方法对决策树的增强带来了更强的性能.	决策树和基于树的模型有较高的性能且有强的可解释性,但存在过拟合的现象.
支持向量机	支持向量机通过找到一个超平面来分类数据,目标是使数据点到超平面的间隔最大.关键作用的是距离超平面最近的“支持向量”点.核技巧是支持向量机的重要方法,它允许将低维无法线性分割的数据映射到更高维空间,使得数据可以被线性分割.这一过程依赖于核函数.	支持向量机的在LSM中的应用非常广泛(徐胜华等, 2020).有学者对支持向量机在LSM中的应用做了详细的综述并指出将该模型与其他机器学习或深度学习集成将带来更强大的性能(Huang and Zhao, 2018).	SVM的显著优势是泛化错误率低,并且在样本数据不足或数据异常的情况下能保持高性能.但它比较依赖超参数的选择,且当特征维度高时表现不稳定.

表 2 集成学习基本原理及适用性分析

Table 2 Basic principles and applicability analysis of ensemble learning

集成模型	基本原理	在 LSM 中的应用与研究	优劣势分析
Boosting	Boosting 是多个基学习器串行的集成方法. 这种算法是先由初始训练数据训练一个基学习器, 再根据基学习器的测试表现对样本数据分布进行调整.	Sahin (2022) 比较了梯度增强机(GBM)、分类增强器(CatBoost)、极端梯度增强(XGBoost)和轻度增强机(LightGBM)在 LSM 中应用的性能, 研究结果表明 Boosting 在 LSM 中可以构建更加有效的模型.	Boosting 算法可以显著提升模型的预测准确性, 能够很好地处理数据中复杂的非线性关系, 适用于滑坡预测等复杂问题. 但对于 Boosting 算法容易被数据集中的噪声和异常值影响, 对类不平衡数据需要进行额外处理.
Bagging (Breiman, 1996)	Bagging 是多个基学习器并行的集成方法, 每个学习器的预测结果通过简单投票得出最终结果.	随机森林(Breiman, 2001)是 Bagging 的拓展, 它在以决策树为基学习器的训练中加入随机特征选择. Zhang <i>et al.</i> (2023) 发现 RF 在重庆市奉节县的 LSM 应用性能优于 XGBoost.	随机森林具有计算成本低、便于超参数整定以及在求解复杂非线性问题时具有鲁棒性等优点 (Taalab <i>et al.</i> , 2018). Bagging 将多个模型的结果进行平均或投票, 形成的最终模型较难解释.

产生了更高维度的张量. 对于图像数据则有 3 个维度即高度、宽度和颜色, 而对于滑坡来说, 滑坡数据含有更多的维度 (Reichenbach *et al.*, 2018).

1.2.2 感知机及人工神经网络 人工神经网络是深度学习的基石, 神经网络包含输入层、隐藏层和输出层. 神经网络单个神经元采用了连续函数作为激活函数且可以自动学习权重和偏置. 是 Sigmoid 函数和 ReLU 函数被广泛使用的激活函数, 相比较于阶跃函数的剧烈突变, 它们都具有连续的特性, 也就是说在神经网络中可以传递 0 到 1 之间的任意实数. 以在滑坡分析中的应用为例, 如果输入神经元的信号对边坡稳定有利 (负值信号), 那么 Sigmoid 函数会输出一个接近于 0 的小值, 而 ReLU 函数则会给出一个负值. 从这一角度来看, ReLU 函数在 LSM 的应用中可能会优于 Sigmoid 函数.

人工神经网络近年被学者广泛应用于 LSM 中. 多位学者对比了人工神经网络与机器学习方法的对比, 说明了人工神经网络在 LSM 中的性能优势 (田乃满等, 2020).

1.2.3 卷积神经网络、对抗神经网络 (generative adversarial networks, GANs) 和迁移学习 (transfer learning, TL) 卷积神经网络包括卷积层、池化层和全连接层. 它通过使用卷积层自动提取图像数据为三维特征, 池化层的作用是通过减少每个特征映射的维度来降低数据的空间大小, 减少后续层的计算负担. 在 LSM 中, 可以将遥感影像、DEM 数据、降水数据、地质图等多种图层结合成为多通道的输入, 卷积神经网络可以自动提取滑坡易发区的特

征, 如地形倾斜度、植被覆盖率和岩石类型等. 这种方法显著提升了模型的预测精度和泛化能力 (Krizhevsky *et al.*, 2012). 因此, CNN 近年来开始被用于 LSM 中. Wang *et al.* (2019) 首次将卷积神经网络结构用于滑坡易发性制图, 取得了优于已有方法的性能. 卷积神经网络在 LSM 中表现出了更高的性能, 但是这种性能的优越性是建立在数据量足够大时. 事实上, 在很多地区获取足够多的样本数据并不容易, 为了应对数据量不足的挑战, Liu *et al.* (2023) 建议了加强迁移学习的研究. 迁移学习可以充分的利用已有的类似数据, 因为迁移学习允许将一个领域的知识迁移到另一个领域. 在 LSM 中, 通过迁移学习, 可以利用在其他地区或条件下训练的模型, 从而在样本不足的情况下提升模型准确性 (Wang *et al.*, 2019).

对抗神经网络由一个生成网络和一个判别网络组成, 通过二者之间的对抗过程来生成新的数据样本. 在 LSM 中, 对抗神经网络可以用于生成更多的训练样本, 特别是在数据稀缺的情况下, 可用于增强现有数据集, 提高模型训练的效果和准确性.

1.2.4 长短时记忆网络和自注意力模型 (Transformer) 融合递归神经网络 (recurrent neural network, RNN) 的核心是隐藏状态的概念, 它作为记忆的一种形式, 图 2 展示了 RNN 的基本架构, 网络状态会随着网络处理序列的每个元素而更新. 融合递归神经网络设计用于处理序列数据, 能够捕捉时间序列中的长期依赖关系. 然而, 传统 RNN 在实践中面临一些问题 (Hochreiter and Schmidhuber, 1997).

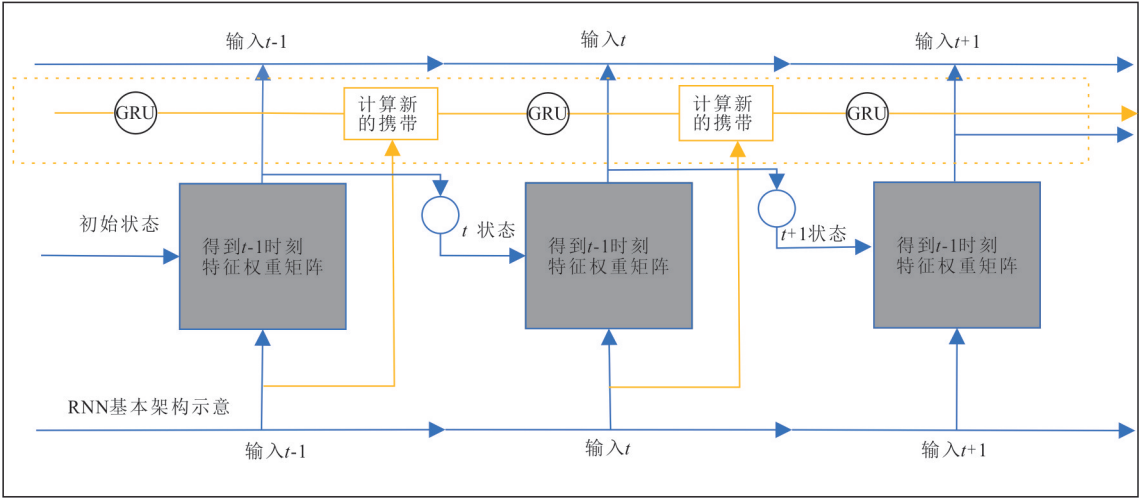


图2 RNN及LSTM结构示意图
Fig.2 RNN and LSTM structure diagram

例如,梯度消失问题即反向传播过程中使用的梯度变得非常小,以至于它们不会对模型权重产生显著变化,网络会停止学习.并且融合递归神经网络难以捕捉长期依赖关系,长期依赖关系是指在时序数据中极早的时间点的数据对后边时间点起到的影响.长短时记忆网络作为融合递归神经网络的改进版本,通过引入携带轨道和门控机制(gated recurrent unit,GRU层)解决了这个问题,使模型能够学习长期依赖.图2示意了长短时记忆网络对融合递归神经网络的改进,其中携带轨道保证序列中任意时刻的信息可以被传送到更晚的时间点,GRU层则可以控制这些信息的隐藏和重新出现.

在LSM中,长短时记忆网络可以用于分析时间序列数据,如降雨量、温度变化以及地表形变等,以预测滑坡的时空发生概率.Transformer依赖于自注意力机制(self-attention)来处理序列数据,不同于融合递归神经网络和长短时记忆网络的循环结构,其能够并行处理序列中的元素,这使得它在处理长序列数据时更加高效.Transformer的核心自注意力机制,它能够让模型在处理每个序列元素时考虑到序列中的所有元素,从而捕获它们之间的全局依赖关系,这对于处理降雨及地表形变这种时间序列数据是很有利的.

在已有的研究中,很少有研究者致力于捕捉降雨或地表形变这类的时间序列数据用于LSM.Wang *et al.* (2021)首次实现了CNN-LSTM算法在香港的滑坡易发性分析中的应用,在分析中采用了4 h降雨和24 h降雨的时间序列数据、通过与其他算

法的对比显示出了最优的性能.在此之后几乎没有相关的研究进展,但时间序列被用于滑坡监测预测中.Wang *et al.* (2022,2023)对比了长短时记忆网络与其他算法在三峡库区滑坡位移的预测性能.

虽然Transformer是为了处理时间序列数据而设计的,但是其也可以捕捉空间特征.在图像识别和处理领域,卷积神经网络长期以来一直是主导技术,但近年来,一些研究表明Transformer也能在这些任务上取得相似甚至更好的性能(Dosovitskiy *et al.*, 2020).

1.3 模型可解释性算法

在AI技术的应用中,许多学者已经关注到模型可解释性研究的必要性(Dikshit *et al.*, 2021).模型可解释性算法主要可分为3类:基于模型的可解释性、基于后处理的可解释性技术和可视化工具.

基于模型的可解释性是指采用简单的模型如决策树或逻辑回归模型,这些方法可以直接输出特征的权重.*t*-SNE/PCA等可视化工具在深度学习模型中被用于可视化网络激活,能帮助理解网络是如何响应不同的特征.基于后处理的可解释技术适用于任何已经训练好的模型,尤其适用于深度学习和复杂的集成学习模型.包括,局部可解释模型—不透明模型(local interpretable model-agnostic explanations,LIME):通过在模型的预测结果附近随机扰动输入并观察输出的变化,生成一个局部简单模型(如线性模型),来近似解释复杂模型的局部行为;SHAP(shapley additive explanations):基于博弈论的Shapley值,量化每个特征对模型预测的贡献;

集成梯度(integrated gradients):用于深度学习模型,特别是在图像和神经语言学的应用中,通过计算模型输入与基线输入之间的积分梯度来确定每个特征的重要性。

1.4 文献分析

2023年至今Web of Science核心库302篇文献的核心关注点如图3所示。这些出现频率最高的关键词显示多方面的信息:(1)“China”显示了我国学者在AI技术应用于LSM的相关研究众多,同时“3 gorges reservoir”说明我国在三峡地区有大量研究。(2)关键词“spatial prediction”、“remote sensing”和“GIS”表明了遥感技术和地理信息系统在数据收集和分析中的关键作用,这包括利用卫星图像、空中摄影和地形测量来评估滑坡风险。同时,也一定程度上表明了LSM中主要以空间数据为主导,忽视了坡体变形破坏的过程机理,存在一定的缺陷(刘凡等, 2023)。(3)学者对以支持向量机和人工神经网络为代表的机器学习模型进行了大量研究,以卷积神经网络为主的深度学习模型也被关注,以随机森林、GBDT、XGBoost为代表的集成学习模型是应用最多。这揭示了学习模型朝更复杂、性能更高的方向探索的趋势。

1.5 研究趋势与挑战

基于机器学习和深度学习算法的滑坡易发性制图表现出了极高的准确度和效能,尤其是深度学

习方法在数据样本充足的情况下表现出高度的性能,但仍面临多方面的挑战。

(1)技术应用的挑战。集成学习和深度学习可以带来更强的性能,探索这些模型的应用是明显的趋势(Liu *et al.*, 2023)。但这些模型对数据量的需求非常高,在数据量不足的地区这些更高级的技术应用面临困难。迁移学习和对抗生成网络提供了可能的解决方案。迁移学习需要有效地选择合适的源域和调整模型以适应目标域,而对抗生成网络如何生成真实且多样化的训练数据仍亟待研究。未来研究应进一步探索如何优化这两种技术并探索他们的结合,提升模型在数据匮乏情况下的训练效果和预测性能。

(2)模型可解释性的挑战。基于AI的LSM是典型的黑箱,并且滑坡灾害存在复杂的机理,滑坡存在“一区一症”甚至“一坡一症”的特点。不能有效的对模型的结果做出解释,将成为技术推广的一大阻碍(窦杰等, 2023)。AI模型与传统物理模型的有效结合是另一研究趋势,这种跨学科的融合有助于提高模型的解释性和预测准确性(Pradhan *et al.*, 2023)。Pradhan *et al.* (2023)的研究表明,这种集成方法不仅能提高预测的准确度,还能有助于认识滑坡这一问题, AI模型可以发现理论建模难以捕捉到的新模式。

(3)缺乏对时序数据的关注。在已有的研究中,对于时序数据的分析是欠缺的(Reichenbach *et al.*, 2018)。在未来的研究中有必要更多的关注时序数据,通过时序数据可以精确的进行易发性评价,也可以更进一步的探索灾害的链式效应(Zuccaro *et al.*, 2018)。降雨和破坏这一类时序数据对滑坡识别是极为重要的。一方面,滑坡易发性制图传统上被认为是静态的;另一方面,这可能是由于这类数据难以获取并且精度难以保证。但实际上,In-SAR和GNSS技术都提供了可靠的地表形变数据,并且滑坡监测与信息采集系统也在迅速发展。而LSTM和Transformer在处理时间序列数据上有着极强的性能,探索动态的滑坡易发性制图系统有望在滑坡灾害领域取得新进展。

(4)由于不同LSM任务之间差异巨大,因此很难生成一系列标准,以确定在特定任务上应使用哪种类型的AI模型(例如,传统机器学习、集成模型或深度学习)。值得在今后的研究中进行系统分析,以指导方法选择(Liu *et al.*, 2023)。已有的研究几乎都

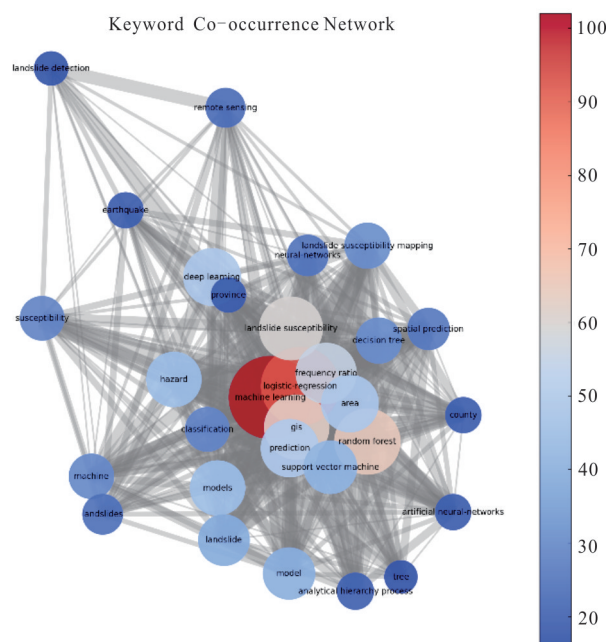


图3 关键词网络图

Fig.3 Keyword co-occurrence network

是关于某一地区某几种方法的对比,总结出了一定的规律,但涉及到的模型和地区有限.事实上,已有研究表明模型在不同区域的性能存在差异(Miao *et al.*, 2023).已经有学者注意到系统性研究的必要性(Zhu *et al.*, 2023),但目前仍然没有关于如何构建学习模型的系统性规律.可能需要在不同地区(地貌、地质构造等情况存在差异)进行大量方法的多方面(样本数据需求、预测精度等)对比,以期得出更加系统的成果.这可以极大程度的指导从业者快速选择适合的学习方法,也有助于形成行业标准.

2 基于AI技术的LSM流程

基于AI技术的LSM主要包含4大步骤:数据准备、特征工程、模型构建与易发性制图.数据准备是指收集历史滑坡灾害点和制图中选取的对滑坡的影响因子,特征工程指选取用于模型训练的影响因子,模型构建则包含数据预处理及模型的选择和训练,最后将模型预测结果进行制图,图4展示了一般流程.

2.1 数据及评价单元选择

数据及预处理对模型的性能有重要的影响(Zhang *et al.*, 2023).所需的数据有滑坡清单和与滑坡相关的影响因子,关于这些数据的来源和预处理问题已经有多位学者做了详细的综述(Liu *et al.*, 2023).在数据处理中,由于采样率对模型的准确性有显著影响被广泛关注(Yang *et al.*, 2023).采样率是指选取的滑坡正类数据和非滑坡点负类数据的比例.用于LSM的评价单元主要有栅格单元和斜坡单元,根据已有研究的结果,由于栅格单元忽略

了滑坡与周围地形地貌的联系,在性能上弱于基于斜坡单元的模型(Chang *et al.*, 2023).

2.2 特征工程

特征工程(feature engineering)是指从原始数据中提取特征并将其转换为适合机器学习的格式,从而改善机器学习性能.关于格式的转化问题,即预处理问题,已在前一节提到,本节主要关注特征因子的提取.

滑坡易发性评价中有众多的特征因子,特征因子的选取体现了对滑坡发生机理的研究,合理选取特征因子可以有效提升模型性能.特征选择方法有以工程地质类比为代表的定性方法、以信息量法为代表的统计学方法和利用机器学习的方法.

工程地质类比法是工程地质领域的一种定性方法,在特征因子的选择中,是指参考已有的类似地质条件地区的研究所采用的特征因子来选取因子的方法.由于方法简便被大量使用,但难以保证特征因子的有效性.

信息量法基于信息论原理,用于评估每个特征提供多少信息以减少结果的不确定性.在LSM中,这意味着衡量每个特征对于区分滑坡发生与否的贡献大小.信息量法能够准确识别出对预测目标有显著影响的特征,实际上这一选择过程体现了对滑坡发生机理的研究.Li *et al.* (2022)通过对比贝叶斯信息准则法和信息增益法,发现信息量法可以有效筛选用于滑坡易发性评价的特征因子,但可能偏向于选择具有更多类别的特征.

机器学习可以通过模型本身的评估特征来提取重要的特征因子.常用的算法包括随机森林和极端梯度提升(XGBoost).这些算法在训练过程中评估各个特征对模型预测性能的贡献,从而识别出重要的特征.例如,随机森林通过平均多棵决策树中的特征重要性评分,来确定特征的重要性.这种方法能够捕捉特征之间的非线性关系和相互作用,适用于复杂数据模型,特别是在特征维度高和特征间相互依赖的情况下.Catani *et al.* (2013)探讨了随机森林技术在估计滑坡易发性方面的敏感性和尺度问题,表明随机森林是一种有效的特征选择工具. Pham *et al.* (2019)展示了XGBoost在特征重要性评估中的优越性.

2.3 性能指标

在机器学习和深度学习领域,性能指标是用来评估模型表现的关键工具.在LSM这样的地理空

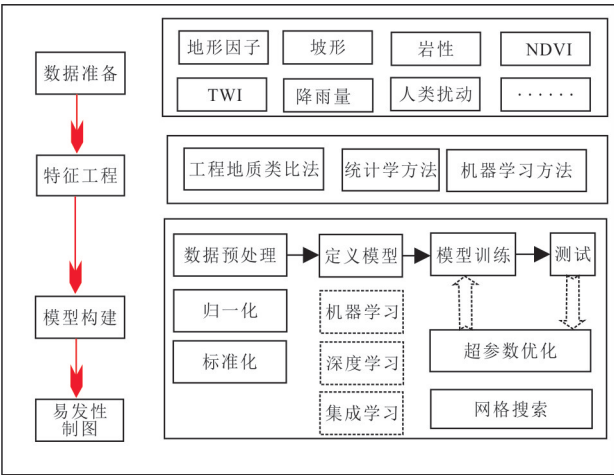


图4 AI技术应用于LSM一般流程

Fig.4 General process of applying AI technology to LSM

间预测任务中,选择合适的性能指标可以帮助我们准确评价模型的预测能力.常用的性能指标有:正确率(accuracy)、精密度(precision)、召回率(recall)、F1分数(F1 score)、AUC-ROC and Kappa系数(kappa coefficient).

正确率是最直观的性能指标,它衡量了模型正确分类的样本占总样本的比例.在均衡的数据集中,即正负样本数相近的情况下,准确率是一个有效的性能衡量标准.然而,在类别不均衡的数据集中,正确率可能会产生误导,因为模型可能仅仅通过预测数量较多的类别来提高准确率,而忽略了少数类的预测.

精确率是正类(滑坡易发区)预测正确的样本占有所有预测为正类样本的比例.在LSM中,高精确率意味着较少的非易发区被错误地分类为易发区,这对于资源分配和防灾减灾计划尤为重要.召回率是正类预测正确的样本占实际正类样本的比例下,高召回率表示大多数真实的滑坡易发区都被模型成功预测,有助于减少漏报的风险.F1分数是精确率和召回率的调和平均值,是衡量模型精确率和召

回率平衡性的指标.在处理LSM中类别不平衡的问题时,F1分数提供了一个更全面的性能评价.

AUC-ROC是一个性能度量,用于评价分类问题的预测能力.ROC曲线表示在不同阈值下模型的真正类率又称敏感性(True Positive Rate, TPR)和假正类率又称特异性(False Positive Rate, FPR)的关系.AUC值越高,表示模型的区分能力越好.AUC-ROC可以评估模型在预测滑坡易发性上的准确性.

Kappa系数是一个用于一致性检验的指标,也可以用于衡量分类的效果.因为对于分类问题,所谓一致性就是模型预测结果和实际分类结果是否一致.

2.4 超参数优化

超参数是指在开始学习过程之前设置的参数,这些参数控制着学习过程本身(如学习率、树的数量、树的深度、卷积层的数量等).超参数优化基于试错的原理,通过系统地探索参数空间,评估不同超参数下模型的性能,选择出最佳配置.这一过程需要一个明确的性能指标来评价不同超参数配置下

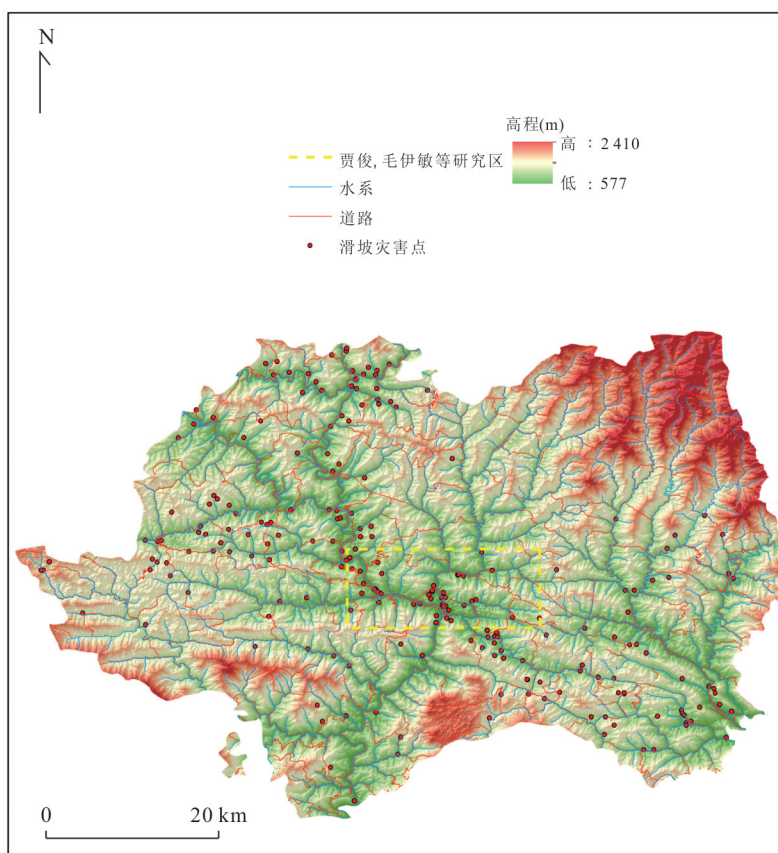


图5 案例研究区位置

Fig.5 Location map of the case study area

模型的表现。

Grid Search and Random Search: 这是一种穷举搜索方法,通过指定一个超参数的网格,系统地遍历所有可能的参数组合,找到最佳的配置。虽然这种方法可以找到全局最优解,但当参数空间大时,计算成本非常高。相对于网格搜索,随机搜索随机选择参数组合进行试验,这种方法在大参数空间中更为高效。虽然它不保证找到全局最优解,但在实践中常常能找到足够好的解。

Bayesian Optimization: 贝叶斯优化是一种基于贝叶斯统计的优化方法,它利用历史的评价结果来推测哪些参数组合可能会带来更好的性能。这种方法旨在减少所需的评估次数,快速收敛到最优解。

Genetic Algorithms: 遗传算法是一种受自然选择启发的优化技术,它通过模拟自然进化过程(如选择、交叉、变异)来搜索最优解。这种方法在处理复杂的、非线性的优化问题时尤为有效。数据时间序列和遗传算法(GA-SVR)模型被用于预测滑坡位移(Miao *et al.*, 2018)。

2.5 案例分析

为分析数据量、采样比以及各学习模型在LSM中对性能的影响,选择陕西省略阳县作为研究区域进行案例分析。贾俊等(2023)以1:1的采样比对略阳县中心城区进行了随机森林和深度随机森林的研究,图5说明了研究区位置以及中心城区的范围。因此,本文设置了8:2的采样比对略阳全域进行对

比研究。

本文所采用的滑坡详细调查数据来源于秦岭南部地区灾害地质调查项目2010年开展的略阳县幅1:50 000地质灾害调查成果与谷歌地球影像解译。针对该区域进行易发性评价所采用的DEM数据来源于地理空间数据云,用于提取坡向、坡度、曲率等信息;地层岩性及断裂分布等图件来源于中国地质调查局西安地质调查中心;植被归一化指数的数据来源于中国科学院资源环境科学数据中心。数据处理基于ArcGIS软件进行,研究单元选取为栅格单元,选择的特征因子与贾俊等(2023)的因子保持一致,图6是所选因子的相关性矩阵热图,其中坡度和起伏度表现出强相关性,但为了与已有研究对比,在本文研究中没有去掉其中之一。图7展示了数据准备的过程。模型训练是基于python语言实现的,对于数据型特征的缺失值替换为相应数据的均值;针对支持向量机对特征进行了标准化处理,将所有特征转化为具有零均值和单位方差的形式;针对深度随机森林通过主成分分析保留解释了95%的方差的主成分,以减少复杂度和避免过拟合。

图8是随机森林、支持向量机、逻辑回归、XGBoost和梯度提升机模型在采样比1:1和采样比8:2情况下的ROC曲线。在贾俊等(2023)的研究中采样比为1:1,随机森林的AUC值为0.863,深度随机森林为0.912。在本文中随机森林的AUC值提升到了0.9以上,这是采用贝叶斯优化带来的提升;基

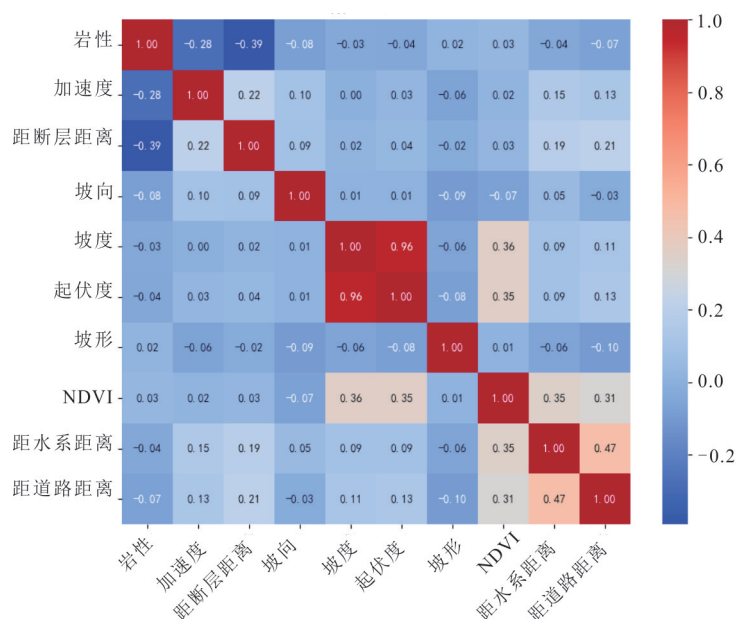


图6 因子相关性矩阵热图

Fig.6 Factor correlation matrix heatmap

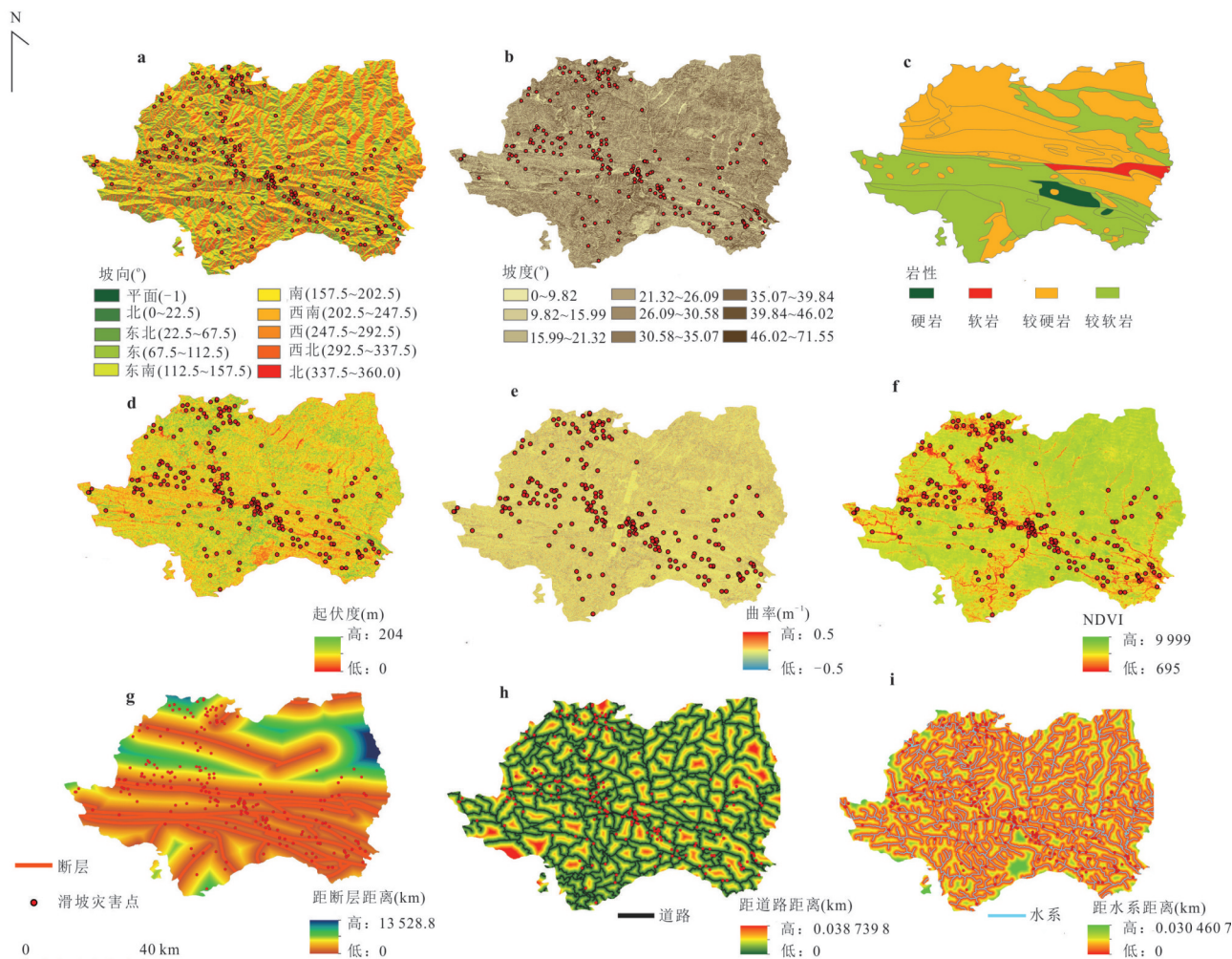


图7 滑坡易发性制图因子图

Fig.7 Landslide susceptibility mapping factor map

a. 坡向图; b. 坡度图; c. 岩性图; d. 起伏度图; e. 曲率图; f. NDVI图; g. 断层距离图; h. 道路距离图; i. 水系距离图

于决策树的集成模型,包括随机森林、XGBoost和梯度提升机的表现优于逻辑回归和支持向量机等机器学习模型.采样比由1:1到8:2的情况下,基于Boosting的集成模型XGBoost和梯度提升机的性能下降幅度小,其余模型性能均有5%以上的下降,其中逻辑回归模型下降幅度最大以致于该模型失去预测能力.虽然采样比1:1情况下模型具有更强的性能,但在实际情况中,由于非滑坡点远远多于滑坡点,1:1的采样比无法反映滑坡与非滑坡的真实分布,这会导致模型实际预测能力不足.因此,在LSM任务中通常采取采样比7:3或8:2的类不平衡策略.可见,基于Boosting的模型在采样比8:2这一类不平衡采样策略下具有优势.

略阳地区的滑坡灾害清单集中于中心城区,

216个滑坡点中有136处位于中心城区,将研究区扩大到整个县区后,面积扩大了约10倍或更大,滑坡灾点数量仅增加1.6倍,数据量可能存在不足的情况.因此本文还训练了深度随机森林和集成了神经网络对抗网络的深度随机森林模型,用于探索神经网络在数据量不足情况下的应用.当研究区由略阳城中心扩展到略阳全境受到数据量的限制模型性能出现下降,本文的深度随机森林模型在采样比1:1的情况下精度为0.90(图9a),集成对抗神经网络后,AUC值为0.88(图10a),表明1:1的采样比不利于对抗神经网络生成符合实际的数据,导致了模型性能下降.当采用8:2采样比精度为0.77(图9b),集成对抗神经网络后,AUC值为0.82(图10b),在类不平衡和数据量限制的情况下仍然可以表现出较优秀

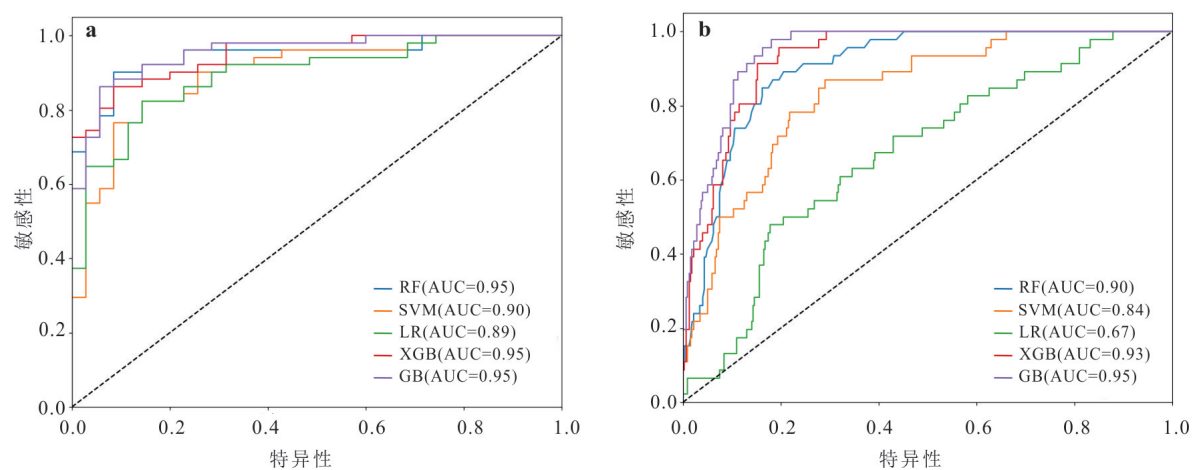


图 8 多种 AI 模型 ROC 曲线
Fig.8 ROC curves of various AI models
a. 采样比 1:1 多种 AI 模型 ROC 曲线; b. 采样比 8:2 多种 AI 模型 ROC 曲线

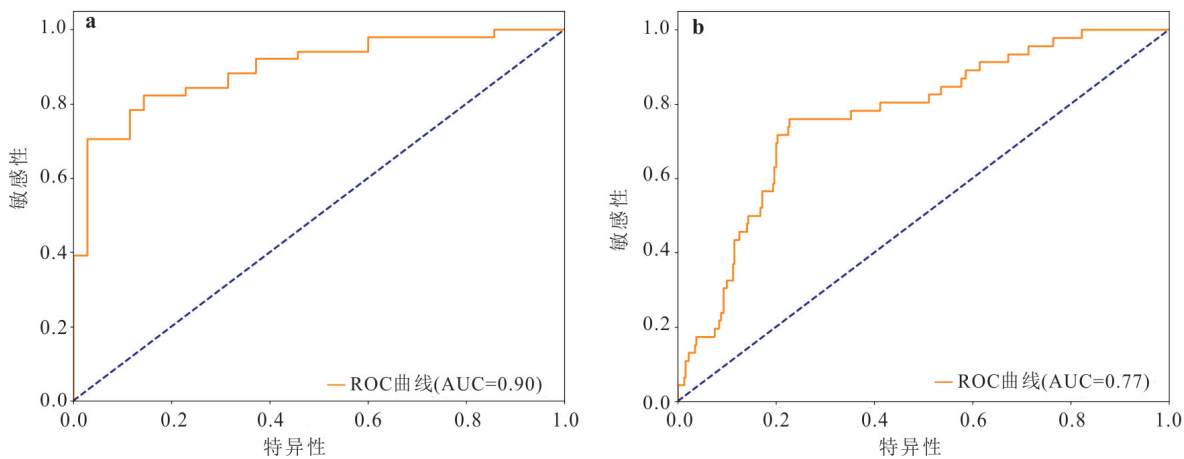


图 9 深度随机森林模型 ROC 曲线
Fig.9 ROC curve of deep random forest model
a. 采样比 1:1 深度随机森林模型 ROC 曲线; b. 采样比 8:2 深度随机森林模型 ROC 曲线

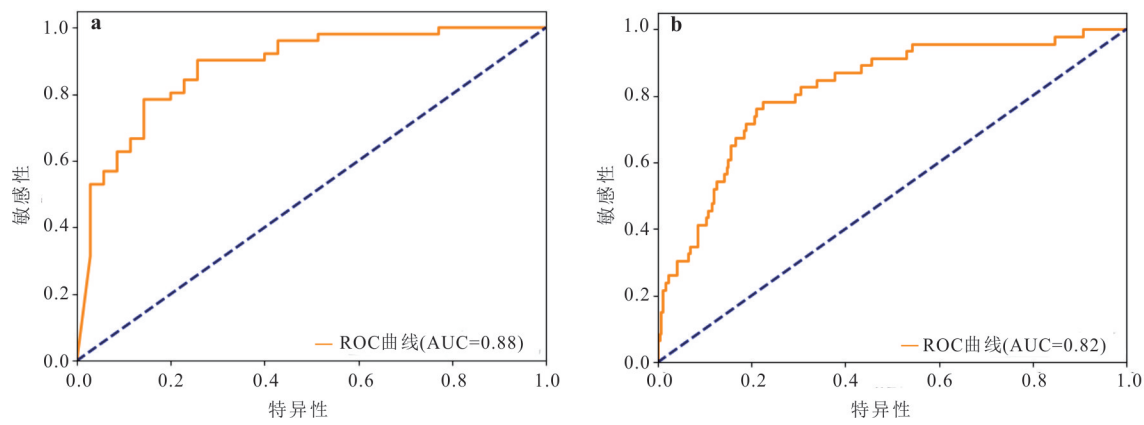


图 10 集成对抗神经网络的深度随机森林模型 ROC 曲线
Fig.10 ROC curve of deep random forest model with integrated generative adversarial networks
a. 采样比 1:1 对抗神经网络与深度随机森林集成模型 ROC 曲线; b. 采样比 8:2 对抗神经网络与深度随机森林集成模型 ROC 曲线

的性能,这表明了对抗神经网络在深度学习模型受到数据量限制的情况下的确可以带来一定的性能提升。

3 结论与展望

随着气候变化和人类活动的影响,全球滑坡灾害的风险持续上升,因此对滑坡易发性的准确预测变得尤为重要。AI技术用于滑坡易发性制图是一大研究热点,可以高效、准确地进行滑坡的识别。本研究分析了多种AI模型的基本原理和适用性,进行了案例研究,发现:(1)基于决策树的集成模型、深度学习模型相比于其他机器学习模型有更高的性能,在本文案例中AUC值均在0.9以上;(2)基于Boosting的集成模型在类不平衡采样策略下表现更好,且受采样比影响小;(3)集成了对抗神经网络的模型可以在数据匮乏的情况下提升深度学习模型的性能,本文深度学习模型集成对抗神经网络后AUC值由0.77提升到了0.82;(4)在LSM中类平衡的采样策略不利于对抗神经网络生成有效的数据。

然而,在实际应用中,AI技术面临一些的挑战,如缺乏模型的选取标准、AI模型与已有的滑坡灾害机理研究结合程度低、数据需求高等。鉴于此,建议未来研究应集中在以下4个方面:(1)探索如何优化数据采集和处理流程,降低对大规模数据集的依赖,以便在资源有限的环境中也能有效使用;(2)进行AI技术与传统物理模型的融合,进一步提升模型的可解释性和准确性;(3)结合新算法来处理时间序列数据,以提高对滑坡事件时空动态的预测能力;(4)研究如何将这些技术应用于实际的灾害响应和规划中,包括开发可操作的政策建议,确保技术成果能够被实际采纳和应用。

References

- Alkhasawneh, M.S., Ngah, U.K., Tay, L.T., et al., 2014. Modeling and Testing Landslide Hazard Using Decision Tree. *Journal of Applied Mathematics*, (1):929768. <https://doi.org/10.1155/2014/929768>
- Breiman, L., 1996. Bagging Predictors. *Machine Learning*, 24: 123—140.
- Breiman, L., 2001. Random Forests. *Mach. Learn.*, 45(1): 5—32. <https://doi.org/10.1023/a:1010933404324>
- Catani, F., Lagomarsino, D., Segoni, S., et al., 2013. Landslide Susceptibility Estimation by Random Forests Technique: Sensitivity and Scaling Issues. *Natural Hazards and Earth System Sciences*, 13(11): 2815—2831. <https://doi.org/10.5194/nhess-13-2815-2013>
- Chang, Z.L., Catani, F., Huang, F.M., et al., 2023. Landslide Susceptibility Prediction Using Sope Unit - Based Machine Learning Models Considering the Heterogeneity of Conditioning Factors. *Journal of Rock Mechanics and Geotechnical Engineering*, 15(5): 1127—1143. <https://doi.org/10.1016/j.jrmge.2022.07.009>
- Dikshit, A., Pradhan, B., Alamri, A.M., 2021. Pathways and Challenges of the Application of Artificial Intelligence to Geohazards Modelling. *Gondwana Research*, 100: 290—301. <https://doi.org/10.1016/j.gr.2020.08.007>
- Dosovitskiy, A., Beyer, L., Kolesnikov, A., et al., 2020. An Image is Worth 16×16 Words: Transformers for Image Recognition at Scale. *ArXiv e-Prints: arXiv:2010.11929*. <https://doi.org/10.48550/arXiv.2010.11929>
- Dou, J., Xiang, Z.L., Xu, Q., et al., 2023. Application and Development Trend of Machine Learning in Landslide Intelligent Disaster Prevention and Mitigation. *Earth Science*, 48(5): 1657—1674(in Chinese with English abstract).
- Fang, Z. C., Wang, Y., Peng, L., et al., 2020. Integration of Convolutional Neural Network and Conventional Machine Learning Classifiers for Landslide Susceptibility Mapping. *Computers & Geosciences*, 139:104470. <https://doi.org/10.1016/j.cageo.2020.104470>
- Hochreiter, S., Schmidhuber, J., 1997. Long Short - Term Memory. *Neural Computation*, 9(8):1735—1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- Huang, Y., Zhao, L., 2018. Review on Landslide Susceptibility Mapping Using Support Vector Machines. *Catena*, 165: 520—529. <https://doi.org/10.1016/j.catena.2018.03.003>
- Jia, J., Mao, Y.M., Meng, X.J., et al., 2023. Comparison of Landslide Susceptibility Evaluation by Deep Random Forest and Random Forest Model: A Case Study of Luyang County, Hanzhong City. *Northwestern Geology*, 56(3):239—249(in Chinese with English abstract).
- Kirschbaum, D., Stanley, T., 2018. Satellite - Based Assessment of Rainfall - Triggered Landslide Hazard for Situational Awareness. *Earth's Future*, 6(3): 505—523. <https://doi.org/10.1002/2017ef000715>
- Krizhevsky, A., Sutskever, I., Hinton, G., 2012. ImageNet Classification with Deep Convolutional Neural Networks. *Communications of the ACM*, 60(6):84—90. <https://doi.org/10.1145/3065386>
- Li, L.P., Lan, H.X., Guo, C.B., et al., 2017. Geohazard Susceptibility Assessment along the Sichuan-Tibet Railway and Its Adjacent Area Using an Improved Frequency Ra-

- tio Method. *Geoscience*, 31(5):911–929 (in Chinese with English abstract).
- Li, X.P., Chong, J.X., Lu, Y.B., et al., 2022. Application of Information Gain in the Selection of Factors for Regional Slope Stability Evaluation. *Bulletin of Engineering Geology and the Environment*, 81(11): 470. <https://doi.org/10.1007/s10064-022-02970-y>
- Liu, F., Deng, Y.H., Mu, H.D., et al., 2023. A Study of the Stability Evaluation Method of Rainfall-Induced Shallow Loess Landslides Based on the Maxent-Sinmap Slope Model. *Hydrogeology & Engineering Geology*, 50(5): 146–158(in Chinese with English abstract).
- Liu, S.L., Wang, L.Q., Zhang, W.G., et al., 2023. A Comprehensive Review of Machine Learning-Based Methods in Landslide Susceptibility Mapping. *Geological Journal*, 58(6):2283–2301. <https://doi.org/10.1002/gj.4666>
- Lyu, L., Chen, T., Li, J., 2023. An Interpretation Study on the ML Models for Landslide Susceptibility Mapping. IGARSS 2023—2023 IEEE International Geoscience and Remote Sensing Symposium. July 16–21, 2023, Pasadena, CA, USA. IEEE, 7241–7244. <https://doi.org/10.1109/igarss52108.2023.10282171>
- Miao, F.S., Wu, Y.P., Xie, Y.H., et al., 2018. Prediction of Landslide Displacement with Step-Like Behavior Based on Multialgorithm Optimization and a Support Vector Regression Model. *Landslides*, 15(3):475–488. <https://doi.org/10.1007/s10346-017-0883-y>
- Miao, F.S., Zhao, F.C., Wu, Y.P., et al., 2023. Landslide Susceptibility Mapping in Three Gorges Reservoir Area Based on GIS and Boosting Decision Tree Model. *Stochastic Environmental Research and Risk Assessment*, 37(6): 2283–2303. <https://doi.org/10.1007/s00477-023-02394-4>
- Ngo, P.T.T., Panahi, M., Khosravi, K., et al., 2021. Evaluation of Deep Learning Algorithms for National Scale Landslide Susceptibility Mapping of Iran. *Geoscience Frontiers*, 12(2): 505–519. <https://doi.org/10.1016/j.gsf.2020.06.013>
- Pham, B. T., Prakash, I., Singh, S.K., et al., 2019. Landslide Susceptibility Modeling Using Reduced Error Pruning Trees and Different Ensemble Techniques: Hybrid Machine Learning Approaches. *Catena*, 175:203–218. <https://doi.org/10.1016/j.catena.2018.12.018>
- Pradhan, B., Dikshit, A., Lee, S., et al., 2023. An Explainable AI (XAI) Model for Landslide Susceptibility Modeling. *Applied Soft Computing*, 142. <https://doi.org/10.1016/j.asoc.2023.110324>
- Reichenbach, P., Rossi, M., Malamud, B.D., et al., 2018. A Review of Statistically - Based Landslide Susceptibility Models. *Earth-Science Reviews*, 180:60–91. <https://doi.org/10.1016/j.earscirev.2018.03.001>
- Sahin, E.K., 2022. Comparative Analysis of Gradient Boosting Algorithms for Landslide Susceptibility Mapping. *Geocarto International*, 37(9): 2441–2465. <https://doi.org/10.1080/10106049.2020.1831623>
- Saito, H., Nakayama, D., Matsuyama, H., 2009. Comparison of Landslide Susceptibility Based on a Decision - Tree Model and Actual Landslide Occurrence: The Akaishi Mountains, Japan. *Geomorphology*, 109(3–4):108–121. <https://doi.org/10.1016/j.geomorph.2009.02.026>
- Stott, P., 2016. How Climate Change Affects Extreme Weather Events. *Science*, 352(6293): 1517–1518. <https://doi.org/10.1126/science.aaf7271>
- Taalab, K., Cheng, T., Zhang, Y., 2018. Mapping Landslide Susceptibility and Types Using Random Forest. *Big Earth Data*, 2(2): 159–178. <https://doi.org/10.1080/20964471.2018.1472392>
- Tian, N.M., Lan, H.X., Wu, Y.M., et al., 2020. Performance Comparison of BP Artificial Neural Network and CART Decision Tree Model in Landslide Susceptibility Prediction. *Journal of Geo-Information Science*, 22(12):2304–2316(in Chinese with English abstract).
- van Westen, C.J., van Asch, T.W.J., Soeters, R., 2006. Landslide Hazard and Risk Zonation: Why is It still so Difficult? *Bulletin of Engineering Geology and the Environment*, 65(2): 167–184. <https://doi.org/10.1007/s10064-005-0023-0>
- Wang, H.J., Wang, L., Zhang, L.M., 2023. Transfer Learning Improves Landslide Susceptibility Assessment. *Gondwana Research*, 123: 238–254. <https://doi.org/10.1016/j.gr.2022.07.008>
- Wang, H.J., Zhang L., Luo H., et al., 2021. AI-Powered Landslide Susceptibility Assessment in Hong Kong. *Engineering Geology*, 288:106103. <https://doi.org/10.1016/j.enggeo.2021.106103>
- Wang, Y., Fang, Z.C., Hong, H.Y., 2019. Comparison of Convolutional Neural Networks for Landslide Susceptibility Mapping in Yanshan County, China. *Science of the Total Environment*, 666:975–993. <https://doi.org/10.1016/j.scitotenv.2019.02.263>
- Wang, Y.K., Tang, H.M., Huang, J.S., et al., 2022. A Comparative Study of Different Machine Learning Methods for Reservoir Landslide Displacement Prediction. *Engineering Geology*, 298:106544. <https://doi.org/10.1016/j.enggeo.2022.106544>
- Ward, P.J., Blauhut, V., Bloemendaal, N., et al., 2020. Review

- Article: Natural Hazard Risk Assessments at the Global Scale. *Natural Hazards and Earth System Sciences*, 20(4):1069–1096. <https://doi.org/10.5194/nhess-20-1069-2020>
- Xu, S.H., Liu, J.P., Wang, X.H., et al., 2020. Landslide Susceptibility Assessment Method Incorporating Index of Entropy Based on Support Vector Machine: A Case Study of Shaanxi Province. *Geomatics and Information Science of Wuhan University*, 45(8):1214–1222(in Chinese with English abstract).
- Yang, C., Liu, L.L., Huang, F.M., et al., 2023. Machine Learning-Based Landslide Susceptibility Assessment with Optimized Ratio of Landslide to Non-Landslide Samples. *Gondwana Research*, 123: 198–216. <https://doi.org/10.1016/j.gr.2022.05.012>
- Zhang, W.G., He, Y.W., Wang, L.Q., et al., 2023. Landslide Susceptibility Mapping Using Random Forest and Extreme Gradient Boosting: A Case Study of Fengjie, Chongqing. *Geological Journal*, 58(6):2372–2387. <https://doi.org/10.1002/gj.4683>
- Zhu, J.J., Yang, M.Q., Ren, Z.J., 2023. Machine Learning in Environmental Research: Common Pitfalls and Best Practices. *Environmental Science & Technology*, 57(46): 17671–17689. <https://doi.org/10.1021/acs.est.3c00026>
- Zuccaro, G., De Gregorio, D., Leone, M.F., 2018. Theoretical Model for Cascading Effects Analyses. *International Journal of Disaster Risk Reduction*, 30: 199–215. <https://doi.org/10.1016/j.ijdrr.2018.04.019>

中文参考文献

- 窦杰, 向子林, 许强, 等, 2023. 机器学习在滑坡智能防灾减灾中的应用与发展趋势. *地球科学*, 48(5):1657–1674.
- 贾俊, 毛伊敏, 孟晓捷, 等, 2023. 深度随机森林和随机森林算法的滑坡易发性评价对比: 以汉中市略阳县为例. *西北地质*, 56(3):239–249.
- 李郎平, 兰恒星, 郭长宝, 等, 2017. 基于改进频率比法的川藏铁路沿线及邻区地质灾害易发性分区评价. *现代地质*, 31(5):911–929.
- 刘凡, 邓亚虹, 慕焕东, 等, 2023. 基于最大熵—无限边坡模型的降雨诱发浅层黄土滑坡稳定性评价方法研究. *水文地质工程地质*, 50(5):146–158.
- 田乃满, 兰恒星, 伍宇明, 等, 2020. 人工神经网络和决策树模型在滑坡易发性分析中的性能对比. *地球信息科学学报*, 22(12):2304–2316.
- 徐胜华, 刘纪平, 王想红, 等, 2020. 熵指数融入支持向量机的滑坡灾害易发性评价方法: 以陕西省为例. *武汉大学学报(信息科学版)*, 45(8):1214–1222.