

DOI: 10.16030/j.cnki.issn.1000-3665.201911034

# 基于逻辑回归的四川青川县区域滑坡灾害预警模型

方然可<sup>1,2</sup>, 刘艳辉<sup>2</sup>, 苏永超<sup>2</sup>, 黄志全<sup>3</sup>

(1. 华北水利水电大学, 河南 郑州 450045; 2. 中国地质环境监测院(自然资源部地质灾害技术指导中心), 北京 100081; 3. 洛阳理工学院, 河南 洛阳 471023)

**摘要:** 四川省青川县滑坡灾害群发, 点多面广, 区域滑坡灾害预警是有效防灾减灾的重要手段, 预警模型是成功预警的核心。由于研究区滑坡诱发机理复杂、调查监测大数据及分析方法不足等原因, 传统区域地质灾害预警模型存在预警精度有限、精细化不足等问题。文章在青川县地质灾害调查监测和降水监测成果集成整理与数据清洗基础上, 构建了青川县区域滑坡灾害训练样本集, 样本集包括地质环境、降雨等 27 个输入特征属性和 1 个输出特征属性, 涵盖了青川县近 9 年(2010—2018 年)全部样本, 数量达 1 826 个(其中, 正样本 613 个, 负样本 1 213 个)。基于逻辑回归算法, 对样本集进行 5 折交叉验证学习训练, 采用贝叶斯优化算法进行模型优化, 采用精确度、ROC 曲线和 AUC 值等指标校验模型准确度和模型泛化能力。其中, ROC 曲线也称为“受试者工作特征”曲线; AUC 值表示 ROC 曲线下的面积。校验结果显示, 基于逻辑回归算法的模型训练结果准确率和泛化能力均较好(准确率 94.3%, AUC 为 0.980)。开展区域滑坡实际预警时, 按训练样本特征属性格式, 输入研究区各预警单元 27 个特征属性, 调用预先学习训练好的模型, 输出滑坡灾害发生概率, 根据输出概率分段确定滑坡灾害预警等级。当输出概率  $P \geq 40\%$  且  $P < 60\%$  时, 发布黄色预警; 当输出概率  $P \geq 60\%$  且  $P < 80\%$  时, 发布橙色预警; 当输出概率  $P \geq 80\%$  时, 发布红色预警。

**关键词:** 滑坡灾害; 预警模型; 逻辑回归; 模型构建; 预警等级

中图分类号: P642.22

文献标志码: A

文章编号: 1000-3665(2021)01-0181-07

## A early warning model of regional landslide in Qingchuan County, Sichuan Province based on logistic regression

FANG Ranke<sup>1,2</sup>, LIU Yanhui<sup>2</sup>, SU Yongchao<sup>2</sup>, HUANG Zhiquan<sup>3</sup>

(1. North China University of Water Resources and Electric Power, Zhengzhou, Henan 450045, China;  
2. China institute of geological environment monitoring (Technical Guidance Center for Geological Hazards, Ministry of Natural Resources), Beijing 100081, China; 3. Luoyang Institute of Science and Technology, Luoyang, Henan 471023, China)

**Abstract:** In Qingchuan County of Sichuan Province, landslide disasters occur in a large number of places and cover a wide range of areas. Early warning of regional landslide disaster is an important means of effective disaster prevention and mitigation, and an early warning model is the core of successful early warning. The traditional regional geological disaster warning model is limited by the lack of big data and analysis methods of the complicated investigation and monitoring mechanism of the landslide in the study areas, and it has some problems, such as limited warning precision and insufficient refinement. In this paper, the training sample set of landslide disaster in Qingchuan County is constructed on the basis of the integrated collation and data cleaning of

收稿日期: 2019-11-12; 修订日期: 2020-01-12

基金项目: 国家重点研发计划(2018YFC1505503); 国家自然科学基金项目(42077440; 41202217); 国家科技支撑计划子课题(2015BAK10B021)

第一作者: 方然可(1996-), 男, 硕士研究生, 主要从事地质灾害预警相关研究工作。E-mail: 1361853780@qq.com

通讯作者: 刘艳辉(1978-), 女, 博士, 教授级高级工程师, 主要从事地质灾害预警与防治等方面的研究工作。E-mail: liuyh@cigem.cn

the results of geological disaster investigation and monitoring and precipitation monitoring. The sample set includes 27 input feature attributes such as geological environment rainfall and 1 output feature attribute, covering the total number of the samples in Qingchuan County in the past 9 years (2010—2018) up to 1 826 (613 positive samples, 1 213 negative samples). Based on the logistic regression algorithm, the study and training of the sample set is carried out with a 50%-fold cross validation. The Bayesian optimization algorithm is used for model optimization, and the accuracy and model generalization ability of the model are verified by such indicators as accuracy, ROC curve and AUC value. The ROC curve is also known as the “Receiver Operating Characteristic” curve. AUC value represents the area under the ROC curve. The verification results show that the training result model based on logistic regression algorithm is of good accuracy and generalization ability (accuracy 94.3% and AUC 0.980). Finally, it is proposed that in the actual warning of regional landslide, 27 characteristic attributes of each warning unit in the research area are input according to the format of characteristic attributes of training samples, and the pre-learned and trained model is called to output the probability of occurrence of landslide disaster, and the warning level of landslide disaster is segmented according to the output probability. A yellow alert is issued when the output probability  $P$  is greater than or equal to 40% and  $P$  is less than 60%. An orange alert is issued when the output probability  $P$  is greater than or equal to 60% and  $P$  is less than 80%. A red alert is issued when the output probability  $P$  is greater than or equal to 80%. In the next step, the accuracy of the model will be further verified in the landslide disaster early warning business in Qingchuan county.

**Keywords:** landslide hazard; Early warning model; logistic regression; model building; warning level

四川省青川县位于四川盆地的北部边缘, 龙门山断裂带的地震多发带, 辖区内山高坡陡, 地形切割强烈, 地质构造复杂, 是滑坡等地质灾害高发区。近年来, 特别是“5.12”汶川地震后, 多次区域强降水作用下, 滑坡灾害频发, 点多面广的防灾形势日益严峻, 居民正常生产生活受到严重威胁。在区域尺度上, 开展降水诱发的区域滑坡灾害预警十分关键。自 2003 年以来, 中国广泛开展各级地质灾害气象预警, 在提高社会公众的防灾减灾意识和有效减轻地质灾害造成人员伤亡和财产损失等方面起到了积极作用<sup>[1-3]</sup>。

预警模型是成功开展地质灾害气象预警的关键, 为此, 大量学者开展了不懈的努力和探索。区域地质灾害预警模型主要分为统计预警模型和动力预警模型两大类, 能在区域尺度得到实际应用的主要是统计预警模型, 特别是基于统计原理的临界降水阈值模型<sup>[4-6]</sup>, 在美国、中国香港、日本等滑坡早期预警系统中都得到了广泛应用。2009 年以来, 刘传正等<sup>[2-3]</sup>提出了显式统计预警原理, 多位学者针对不同地区的特点, 构建了相应的统计预警模型<sup>[7-13]</sup>, 有效支撑了中国大陆各级地质灾害气象预警工作。但预警模型研究受制于研究区滑坡诱发机理复杂、调查监测数据不足等限制, 仍存在区域预警精度有限、精细化不足等问题。

2015 年, 原国土资源部和中国气象局在四川省青

川县建立了第一个国家级地质灾害气象预警试验区, 针对地质灾害预警的地质环境专项调查监测、降雨加密站点布设与监测等工作逐步展开并取得了丰硕成果<sup>[14-15]</sup>。多年的数据积累和系统的预警专项调查, 为区域滑坡灾害预警模型研究奠定了雄厚的数据基础; 人工智能和大数据的蓬勃发展, 为预警模型发展提供了新的方法。

本文以四川省青川县为研究区, 开展了基于逻辑回归算法的区域滑坡灾害预警模型研究并校验, 推动了人工智能在滑坡灾害预警领域中的应用, 有效支撑了我国正在开展的区域地质灾害气象预警业务, 具有重要理论意义和应用价值。

## 1 数据获取与清洗

本研究区为四川省青川县, 位于川、甘、陕交界处, 辖 36 个乡镇, 总人口约 22 万。研究数据主要来源于四川省青川县 1:5 万地质灾害与地质环境调查成果、青竹江流域地质灾害调查成果、地质灾害灾情直报系统、青川县加密雨量监测数据以及青川预警试验区补充调查监测成果等<sup>[14-16]</sup>。

### 1.1 滑坡和非滑坡(正负样本)数据

据调查成果<sup>[16]</sup>, 截至 2015 年底, 青川县共发生崩塌、滑坡、泥石流和不稳定斜坡灾害 1 672 处, 其中崩塌 262 处、滑坡 643 处、泥石流 45 处和不稳定斜坡

722处。灾害类型以不稳定斜坡最多,其次为滑坡、崩塌,泥石流数量相对少。本文研究对象为广义滑坡,包含了崩塌、滑坡和不稳定斜坡。青川县广义滑坡占总数的97.3%,是青川县地质灾害的最主要灾种。

正样本是指已经发生滑坡的点,正样本的采样一般以历史滑坡编目数据为依据进行筛选。筛选标准为具有确定的空间地理坐标和时间坐标(一般精确到日)。负样本是指没有发生滑坡的点,无法直接获取。本研究以正样本为基础,通过时空采样确定负样本<sup>[17]</sup>。空间上,在正样本120 m(3倍的预警网格单元大小)缓冲区外空间随机采样(正负样本采样比例1:2),确定负样本的空间位置;然后在汛期范围内随机时间,确定负样本的时间属性。最终确定2010—2018年滑坡灾害正负样本1 826个:613个滑坡点(正样本)和1 213个非滑坡点(负样本)。

## 1.2 地质环境数据

孕育滑坡灾害的地质环境数据主要包括地形地貌、地层岩性、地质构造、沟谷水系,以及人类工程活动等。据相关分析结果<sup>[16]</sup>,地形地貌是滑坡发育的重要条件,地层岩性是滑坡发育的物质基础条件,控制了青川县滑坡灾害的空间分布;地质构造特别是断裂带的分布对滑坡灾害发育影响明显;沟谷水系广泛分布,直接影响着斜坡稳定性;人类工程活动,特别是道路切坡、房屋建筑等,对自然斜坡进行改造,诱发或加剧了滑坡灾害的发生发展。因此,在前期研究成果基础上,本文选取了坡度、地貌、地层岩性、断裂、房屋分布等11个地质环境因子参与模型构建。

## 1.3 降雨诱发数据

降雨是研究区滑坡灾害发生的主要诱发因素。本研究降雨数据主要来源于气象部门,收集整理了青川县域43个加密雨量站点逐日雨量数据,通过空间插值,以公里网格为单元构建雨量数据库,保证了降雨数据的站点密度和监测精度。选取当日雨量,前1~15日逐日雨量等16个降雨因子参与模型构建。

## 1.4 训练样本集

以200 m×200 m为单元对研究区进行网格剖分(约82 000个网格单元),分别匹配27个影响因子(包括坡度等11个地质环境因子,当日雨量等16个降雨因子)作为输入特征参数,滑坡是否发生(滑坡正样本为1,非滑坡负样本为0)作为输出特征参数。构建了青川县区域滑坡灾害预警训练样本集,样本个数1 826个(图1),样本集输入特征及参数见表1。

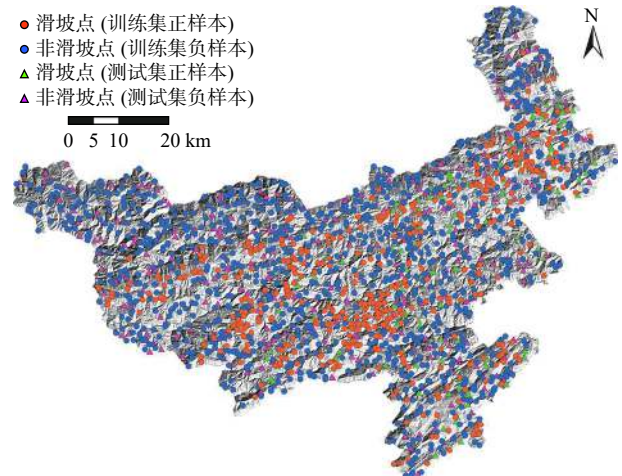


图1 青川县区域滑坡灾害预警训练样本集

Fig. 1 Training sample set of regional landslide disasters in Qingchuan County

表1 训练样本输入特征及参数

Table 1 Input characteristics and parameters of the training samples

| 序号  | 输入特征        | 输入特征参数  |
|-----|-------------|---|
| 1   | 坡度/(°)      | ①0~10; ②10~20; ③20~30; ④30~40; ⑤≥40                           |
| 2   | 坡向/(°)      | ①0~90; ②90~180; ③180~270; ④270~360                            |
| 3   | 高程/m        | ①0~800; ②800~1 200; ③1 200~1 600; ④1 600~2 000; ⑤≥2 000       |
| 4   | 地貌类型        | ①中低山; ②中山; ③高中山   |
| 5   | 地层岩性        | ①松散堆积层; ②软弱-半坚硬薄-中层状岩组; ③半坚硬-坚硬薄-中层状岩组; ④坚硬-半坚硬中-厚层状岩组; ⑤未知岩性 |
| 6   | 距断裂距离/m     | ①0~500; ②500~1 000; ③1 000~1 500; ④1 500~2 000; ⑤≥2 000       |
| 7   | 年雨量/mm      | ①0~500; ②500~800; ③800~1 000; ④1 000~1 200; ⑤≥1 200           |
| 8   | 距房屋距离/m     | ①0~200; ②≥200   |
| 9   | 距道路距离/m     | ①0~200; ②≥200   |
| 10  | 距沟谷距离/m     | ①0~200; ②≥200   |
| 11  | 网格单元历史灾点数/个 | ①0; ②1; ③2; ④3~4; ⑤5~7  |
| 12  | 当日雨量/mm     | ①<10; ②10~25; ③25~50; ④50~100; ⑤>100                          |
| 13  | 前1日雨量/mm    | ①<10; ②10~25; ③25~50; ④50~100; ⑤>100                          |
| 14  | 前2日雨量/mm    | ①<10; ②10~25; ③25~50; ④50~100; ⑤>100                          |
| ... | ...         | ...   |
| 27  | 前15日雨量/mm   | ①<10; ②10~25; ③25~50; ④50~100; ⑤>100                          |

## 1.5 数据清洗

数据清洗是指对数据进行缺失值插补和异常值(噪音值)识别处理<sup>[18]</sup>。在机器学习模型构建过程中,训练样本集的好坏直接决定了模型准确性和泛化能力。原始数据集中常存在人工错误、数据传输误差、

设备故障和地质信息模糊等问题,严重影响着训练样本集的质量。因此,对原始数据集进行数据清洗,是有效提升模型精度的必要基础。

数据清洗方法要根据数据实际情况进行选择。例如,若空值较少可直接删除,若空值所占比例较大,则不可直接删除,一般可用均值代替补全。对于离群值问题,因离群值不一定是错值,一般通过降低权重或是增加此类样本方式进行清洗<sup>[19-20]</sup>。

## 2 模型构建与检验

### 2.1 算法与工具

逻辑回归算法是一种常用的非线性二分类因变量回归统计模型,在机器学习领域也得到了广泛使用。该算法通过极大似然估计法估计参数,具有一致的渐进正态性<sup>[21-22]</sup>。与一般线性回归算法的区别在于逻辑回归算法通过 Sigmoid 函数可以把输出结果约束在 [0, 1] 之内:

$$Y = \frac{1}{1 + e^{-x}} \quad (1)$$

由图 2 所示,逻辑回归算法输出 Y 值介于 0~1 之间,因此可以引入阈值的概念,从而对输出 Y 进行二分类。例如设置阈值为 0.5,算法输出大于 0.5 则判定为 1,反之则判定为 0。

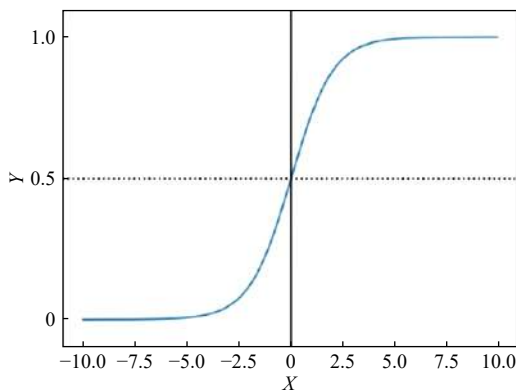


图 2 Sigmoid 函数  
Fig. 2 Sigmoid function

滑坡灾害预警时,影响滑坡是否发生的地质环境、降雨等因素可以作为自变量,滑坡发生或不发生可以作为分类因变量。设 P 为滑坡发生的概率,取值范围为 [0, 1], 1-P 即为滑坡不发生的概率。P/(1-P) 为滑坡发生与不发生的概率比值,对其取自然对数 ln[P/(1-P)]:

$$\ln \frac{P}{1-P} = \beta_0 + \beta_1 x_1 + \dots + \beta_m x_m \quad (2)$$

$$P = \frac{e^{\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m}}{1 + e^{\beta_0 + \beta_1 x_1 + \dots + \beta_m x_m}} \quad (3)$$

式中: P——滑坡发生概率;

$x_1, x_2, \dots, x_m$ ——影响因子;

$\beta_0, \beta_1, \dots, \beta_m$ ——逻辑回归算法的回归系数。

国内外学者在该领域也开展了相应探索, Lee S 等<sup>[23]</sup>和 Ohlmacher G C 等<sup>[24]</sup>分别在韩国龙仁地区和美国堪萨斯州东北地区,选择诱发滑坡的致灾影响因子,采用了逻辑回归算法开展滑坡预测。李铁锋等<sup>[25]</sup>以长江三峡为研究区,把有效雨量与 Logistic 回归模型结合,建立了预警模型。孙德亮<sup>[26]</sup>应用逻辑回归算法,设置了不同阈值开展对比分析,建立了滑坡灾害易发性分区模型等。

本文选取逻辑回归算法开展滑坡灾害预警模型训练。模型训练过程通过 Python 语言实现,调用了 sklearn 库里的 LogisticRegression 模型。

### 2.2 模型构建与优化

按照 4 : 1 的比例,将青川县训练样本集(图 1, 表 1)划分为训练集和测试集,进行训练和校验。采用贝叶斯优化算法、五折交叉验证开展模型参数优化。目前最常用的模型参数优化方法包括传统方法和超参数优化算法。传统方法,也称为网格搜索法,方法优化精度和速度成反比。为了参数优化更高效,出现了超参数优化算法——贝叶斯优化算法<sup>[27]</sup>。贝叶斯优化算法采用了高斯过程,通过增加样本数量来拟合目标函数分布,目标函数通过交叉验证精度来进行优化,每次迭代都会输出一次超参数,在寻找最优值的过程中优化超参数。通过贝叶斯优化算法,对逻辑回归模型的正则化指数 C 值进行参数优化,最终得到模型最优参数, C=2。

基于优化后的 Logistic 模型,分别设置 0.25, 0.5, 0.75 三种不同的阈值对模型结果进行二分类,得到混淆矩阵,见表 2。

根据表 2,对比分析三种阈值分类结果的精度,可见,当阈值取 0.5 时,模型总精度最高。

### 2.3 模型检验

模型检验从 3 个方面来进行:

(1) 准确率 (Accuracy, ACC), 表达的是模型的精度。模型准确率,是判断模型预测分类正确的样本数和总样本数的比值。另外,还有精确率 (Precision)、召回率 (Recall) 和 F1 值等指标。

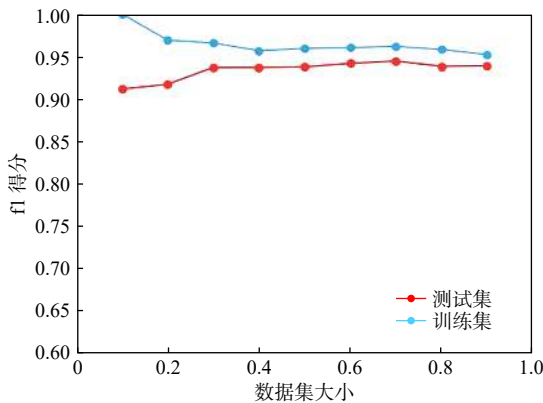
(2) ROC 曲线和 AUC 值,表达的是模型泛化能

表 2 不同阈值下的 Logistic 回归分类结果混淆矩阵

Table 2 Confuse matrix of the result of the logistic regression classification under different thresholds

| 阈值   | 实际值 |            |            |            |            |
|------|-----|------------|------------|------------|------------|
|      |     | 滑坡         | 非滑坡        |            |            |
| 0.25 | 预测值 | 滑坡         | 468        | 58         | 准确率: 0.890 |
|      |     | 非滑坡        | 25         | 910        | 准确率: 0.973 |
|      |     | 召回率: 0.949 | 召回率: 0.940 | 总精度: 0.943 |            |
| 0.5  | 预测值 | 滑坡         | 451        | 26         | 准确率: 0.945 |
|      |     | 非滑坡        | 42         | 942        | 准确率: 0.957 |
|      |     | 召回率: 0.915 | 召回率: 0.973 | 总精度: 0.953 |            |
| 0.75 | 预测值 | 滑坡         | 423        | 15         | 准确率: 0.966 |
|      |     | 非滑坡        | 70         | 953        | 准确率: 0.932 |
|      |     | 召回率: 0.858 | 召回率: 0.985 | 总精度: 0.942 |            |

力。ROC(Receiver Operating Characteristic)曲线是一个画在二维平面上的曲线; AUC(Area Under Curve), 即 ROC 曲线下的面积。通常, AUC 的值介于 0.5 到 1.0



之间, AUC 值越大, 说明模型表现越好。

(3)学习曲线(Learning Curve), 描述模型拟合程度, 判断模型是否存在过拟合或欠拟合。

使用测试集对 Logistic 回归模型的准确率和模型泛化能力指标进行评估。评估结果见表 3 和图 3。可见, 模型准确率为 0.943, AUC 值为 0.980, 模型准确率和泛化能力均较好。

表 3 Logistic 回归模型分类

Table 3 Logistic regression model classification report

|      | 精确率   | 召回率   | f1得分  |
|------|-------|-------|-------|
| 0    | 0.949 | 0.975 | 0.957 |
| 1    | 0.950 | 0.883 | 0.915 |
| 准确率  |       |       | 0.943 |
| 宏平均  | 0.944 | 0.929 | 0.936 |
| 加权平均 | 0.943 | 0.943 | 0.942 |

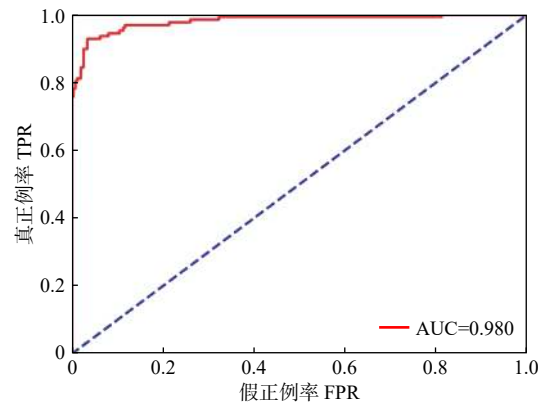


图 3 逻辑回归模型学习曲线和 ROC 曲线

Fig. 3 Learning curve and ROC curve of the logistic regression model

开展区域滑坡实际预警时, 按训练样本特征属性格式, 输入研究区各预警网格单元 27 个特征属性, 调用基于逻辑回归算法训练保存好的模型开展概率预测, 依据模型输出概率  $P$  划分滑坡灾害预警等级。输出概率的预警等级, 可参考地质灾害气象风险预警标准<sup>[28]</sup>中的预警等级划分, 也可根据研究区具体情况微调确定。考虑到青川县基于逻辑回归算法的预警模型训练中阈值设定为 0.5, 结合研究区具体情况, 将地质灾害气象预警概率等级划分标准调整如下, 当输出概率  $P \geq 40\%$  且  $P < 60\%$  时, 发布黄色预警; 当输出概率  $P \geq 60\%$  且  $P < 80\%$  时, 发布橙色预警; 当输出概率  $P \geq 80\%$  时, 发布红色预警, 见表 4。

表 4 预警等级划分

Table 4 Early warning level division

| 预警等级 | 风险等级 | 概率 $P$      |
|------|------|-------------|
| 红色预警 | 风险很高 | [80%, 100%] |
| 橙色预警 | 风险高  | [60%, 80%)  |
| 黄色预警 | 风险较高 | [40%, 60%)  |

### 3 结论

(1)以清洗后的 2010—2018 年滑坡灾害为正样

本, 以 1 : 2 的比例采样负样本, 以 200 m×200 m 为网格单元, 构建了青川县区域滑坡灾害训练样本集。训练样本数量为 1 826 个, 其中正样本为 613 个, 负样本为 1 213 个。各训练样本包含了 27 个输入特征参数(地质环境因子、降雨因子等), 1 个输出特征参数(滑坡正样本为 1, 非滑坡负样本为 0)。

(2)基于 1 826 个训练样本, 采用逻辑回归算法开展滑坡灾害预警模型学习训练。以训练样本集的 80% 作为训练集, 20% 作为测试集, 进行 5 折交叉验

证,采用精确度(Accuracy)、ROC曲线和AUC值校验模型精确度和模型泛化能力。采用贝叶斯优化算法对模型参数进行优化。结果显示,设置阈值为0.5时,混淆矩阵的总精度最高,模型准确率为0.943,AUC值为0.980,模型准确率和模型泛化能力均较好。

(3)开展区域滑坡灾害实际预警时,可调用基于逻辑回归算法训练好的模型输出概率,根据概率分段确定预警等级,输出概率 $P \geq 40\%$ 且 $P < 60\%$ 时,发布黄色预警;输出概率 $P \geq 60\%$ 且 $P < 80\%$ 时,发布橙色预警;当输出概率 $P \geq 80\%$ 时,发布红色预警。

今后,将在青川县区域滑坡灾害预警中进一步校验模型的准确性。

### 参考文献 (References) :

- [ 1 ] 刘传正,温铭生,唐灿.中国地质灾害气象预警初步研究[J].地质通报,2004,23(4):303-309. [LIU Chuazheng, WEN Mingsheng, TANG Can. Meteorological early warning of geo-hazards in China based on raining forecast[J]. Geological Bulletin of China, 2004, 23(4): 303-309. (in Chinese with English abstract) ]
- [ 2 ] 刘传正,刘艳辉,温铭生,等.中国地质灾害区域预警方法与应用[M].北京:地质出版社,2009. [LIU Chuazheng, LIU Yanhui, WEN Mingsheng, et al. Method and application of regional warning for geo-hazards in China [M]. Beijing: Geological Publishing. 2009. (in Chinese) ]
- [ 3 ] 刘传正,刘艳辉,温铭生,等.中国地质灾害气象预警实践:2003-2012[J].中国地质灾害与防治学报,2015,26(1):1-8. [LIU Chuazheng, LIU Yanhui, WEN Mingsheng, et al. Early warning for regional geo-hazards during 2003-2012, China[J]. The Chinese Journal of Geological Hazard and Control, 2015, 26(1): 1-8. (in Chinese with English abstract) ]
- [ 4 ] CANNON S H. Rainfall conditions for abundant debris avalanches, San Francisco Bay region, California[J]. California Geology, 1985, 38(12): 267-272.
- [ 5 ] AU S W C. Rain-induced slop instability in Hong Kong[J]. Engineering Geology, 1998, 51(1): 1-36.
- [ 6 ] ALEOTTI P. A warning system for rainfall-induced shallow failures[J]. Engineering geology, 2004, 73: 247-265.
- [ 7 ] 刘艳辉,刘传正,连建发,等.基于显式统计原理的地质灾害区域预警方法初步研究[J].中国地质,2008,35(2):344-350. [LIU Yanhui, LIU Chuazheng, LIAN Jianfa, et al. Preliminary study of geo-hazards regional early warning based on explicit statistical theory[J]. Geology in China, 2008, 35(2): 344-350. (in Chinese with English abstract) ]
- [ 8 ] 刘艳辉,刘传正,温铭生,等.中国地质灾害气象预警模型研究[J].工程地质学报,2015,23(4):738-746. [LIU Yanhui, LIU Chuazheng, WEN Mingsheng, et al. Study of early warning models for regional geo-hazards in China[J]. Journal of Engineering Geology, 2015, 23(4): 738-746. (in Chinese with English abstract) ]
- [ 9 ] 刘艳辉,张振兴,苏永超.地质灾害承灾载体脆弱性评价方法研究[J].工程地质学报,2018,26(5):1121-1130. [LIU Yanhui, ZHANG Zhenxing, SU Yongchao. Case study of vulnerability evaluation for geo-hazards bearing capacity of a region[J]. Journal of Engineering Geology, 2018, 26(5): 1121-1130. (in Chinese with English abstract) ]
- [ 10 ] 刘艳辉,苏永超.四川青川县区域地质灾害气象风险预警模型研究[J].工程地质学报,2019,27(1):134-143. [LIU Yanhui, SU Yongchao. Early-warning model of regional geological disasters based on meteorological factor in Qingchuan County, Sichuan Province[J]. Journal of Engineering Geology, 2019, 27(1): 134-143. (in Chinese with English abstract) ]
- [ 11 ] 侯圣山,李昂,韩冰,等.四川雅安地质灾害预警预报及分析[J].中国地质灾害与防治学报,2014,25(4):134-138. [HOU Shengshan, LI Ang, HAN Bing, et al. An approach of geo-hazard warning system in Ya'an, Sichuan and its analysis[J]. The Chinese Journal of Geological Hazard and Control, 2014, 25(4): 134-138. (in Chinese with English abstract) ]
- [ 12 ] 李守定,白亚恒,姜越,等.基于内外动力耦合成因理论的新疆地质灾害气象预警显式统计模型[J].地球科学与环境学报,2017,39(2):287-300. [LI Shouding, BAI Yaheng, JIANG Yue, et al. An explicit statistical model for meteorological early warning of geological hazards in Xinjiang based on the coupling theory of internal and external dynamics[J]. Journal of Earth Sciences and Environment, 2017, 39(2): 287-300. (in Chinese with English abstract) ]
- [ 13 ] 魏平新,李秀娟.广东省突发性地质灾害气象预警实践[J].中国地质灾害与防治学报,2015,26(1):138-144. [WEI Pingxin, LI Xiujian. The meteorologic early warning research of sudden geo-hazard in Guangdong province[J]. The Chinese Journal of Geological Hazard and Control, 2015, 26(1): 138-144. (in Chinese with English abstract) ]

- [14] 中国地质环境监测院. 城镇化过程中承灾载体脆弱性及脆弱性评价方法研究成果报告[R]. 北京: 中国地质环境监测院, 2018. [ China Institute of Geo-Environmental Monitoring. The research report of vulnerability vector and vulnerability assessment method during urbanization[R]. Beijing: China Institute of Geo-Environmental Monitoring, 2018. (in Chinese) ]
- [15] 四川九零九建设工程有限公司. 四川省青川县地质灾害详细调查报告[R]. 成都: 四川九零九建设工程有限公司, 2015. [ Sichuan 909 Construction Engineering Co. LTD. A detailed survey report on geological hazards in Qingchuan County, Sichuan Province[R]. Chengdu: Sichuan 909 Construction Engineering Co. LTD, 2015. (in Chinese) ]
- [16] 中国地质环境监测院. 四川青竹江流域地质灾害调查报告[R]. 北京: 中国地质环境监测院, 2016. [ China Institute of Geo-Environmental Monitoring. Geo-hazards investigation report of Qingzhujiang River region, Sichuan[R]. Beijing: China Institute of Geo-Environmental Monitoring, 2016. (in Chinese) ]
- [17] 刘艳辉, 肖锐铧, 陈春利, 等. 区域滑坡预警中训练样本集的构建方法、系统及存储介质: 202010829816.0[P]. [ LIU Yanhui, XIAO Ruihua, CHEN Chunli, et al. Construction method system and storage medium of training sample set in regional landslide early warning: 202010829816.0[P].2020-08-18. (in Chinese with English abstract) ]
- [18] 韩帅, 孙乐平, 杨艺云, 等. 基于改进K-Means聚类和误差反馈的数据清洗方法[J]. 电网与清洁能源, 2020, 36(7): 9 - 15. [ HAN Shuai, SUN Leping, YANG Yiyun, et al. Data cleaning method based on improved K-means clustering and error feedback[J]. Power System and Clean Energy, 2020, 36(7): 9 - 15. (in Chinese with English abstract) ]
- [19] 朱力. 决策树算法在山区公路地质灾害风险评估系统中的应用[D]. 重庆: 重庆师范大学, 2019. [ ZHU Li. Application of decision tree algorithm in the risk assessment system of mountain highway geological hazards[D]. Chongqing: Chongqing Normal University, 2019. (in Chinese with English abstract) ]
- [20] BOSCHETTI A, MASSARON L. Python data science essentials[M]. 2nd ed. Packt Publishing Ltd, 2015.
- [21] 赵永红, 王航, 张琼, 等. 滑坡位移监测方法综述[J]. 地球物理学进展, 2018, 33(6): 2606 - 2612. [ ZHAO Yonghong, WANG Hang, ZHANG Qiong, et al. A review of monitoring methods for landslide displacement[J]. Progress in Geophysics, 2018, 33(6): 2606 - 2612. (in Chinese with English abstract) ]
- [22] 张铎, 吴中海, 李家存, 等. 国内外地震滑坡研究综述[J]. 地质力学学报, 2013, 19(3): 225 - 241. [ ZHANG Duo, WU Zhonghai, LI Jiacun, et al. A review of seismic landslide studies at home and abroad[J]. Journal of geomechanics, 2013, 19(3): 225 - 241. (in Chinese with English abstract) ]
- [23] LEE S, RYU J, MIN K, et al. Development and application of landslide susceptibility analysis techniques using geographic information system (GIS)[C]//IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium. Taking the Pulse of the Planet: The Role of Remote Sensing in Managing the Environment. Proceedings (Cat. No. 00CH37120). IEEE, 2000, 1: 319-321.
- [24] OHLMACHER G C, DAVIS J C. Using multiple logistic regression and GIS technology to predict landslide hazard in northeast Kansas, USA[J]. Engineering geology, 2003, 69: 331 - 343.
- [25] 李铁锋, 丛威青. 基于Logistic回归及前期有效雨量的降雨诱发型滑坡预测方法[J]. 中国地质灾害与防治学报, 2006, 17(1): 33 - 35. [ LI Tiefert, CONG Weiqing. Rainfall-induced landslide prediction method based on Logistic regression and early effective rainfall[J]. Chinese Journal of Geological Hazards and Prevention, 2006, 17(1): 33 - 35. (in Chinese with English abstract) ]
- [26] 孙德亮. 基于机器学习的滑坡易发性区划与降雨诱发滑坡预报预警研究[D]. 上海: 华东师范大学, 2019. [ SUN Deliang. Study on landslide susceptibility regionalization and rainfall-Induced landslide prediction and Early Warning based on machine learning [D]. Shanghai: East China Normal University, 2019. (in Chinese with English abstract). ]
- [27] SNOEK J, LAROCHELLE H, ADAMS R P. Practical bayesian optimization of machine learning algorithms[J]. Advances in neural information processing systems, 2012, 25: 2951 - 2959.
- [28] 中国地质灾害防治工程行业协会. 地质灾害区域气象风险预警标准(试行): T/CAGHP 039-2018[S]. 武汉: 中国地质大学出版社, 2018. [ China Geological Disaster Prevention Engineering Association. Regional meteorological risk early warning standard for geological disaster (trial): T/CAGHP 039-2018[S]. Wuhan: China University of Geosciences Press, 2018. ](in Chinese) ]