

机器学习模型在滑坡易发性评价中的应用

刘福臻, 王 灵, 肖东升

Application of machine learning model in landslide susceptibility evaluation

LIU Fuzhen, WANG Ling, and XIAO Dongsheng

在线阅读 View online: <https://doi.org/10.16031/j.cnki.issn.1003-8035.2021.06-12>

您可能感兴趣的其他文章

Articles you may be interested in

基于机器学习的区域滑坡危险性评价方法综述

A review of the methods of regional landslide hazard assessment based on machine learning

方然可, 刘艳辉, 黄志全 中国地质灾害与防治学报. 2021, 32(4): 1-8

基于遥感影像多尺度分割与地质因子评价的滑坡易发性区划

Landslide susceptibility assessment based on multi-scale segmentation of remote sensing and geological factor evaluation

李文娟, 邵海 中国地质灾害与防治学报. 2021, 32(2): 94-99

基于机器学习的滑坡崩塌地质灾害气象风险预警研究

Exploring early warning and forecasting of meteorological risk of landslide and rockfall induced by meteorological factors by the approach of machine learning

李阳春, 刘黔云, 李潇, 顾天红, 张楠 中国地质灾害与防治学报. 2021, 32(3): 118-123

北京山区突发性地质灾害易发性评价

Assessment on the susceptibility of sudden geological hazards in mountainous areas of Beijing

罗守敬, 王珊珊, 付德荃 中国地质灾害与防治学报. 2021, 32(4): 126-133

国产GB-InSAR在特大型水库滑坡变形监测中的应用

Application of GB-InSAR in deformation monitoring of huge landslide in reservoir area

郭延辉, 杨溢, 杨志全, 高才坤, 田卫明, 何玉童 中国地质灾害与防治学报. 2021, 32(2): 66-72

地理探测器在判别滑坡稳定性影响因素中的应用

Application of geographic detector in identifying influencing factors of landslide stability: A case study of the Jiangda County, Tibet

支泽民, 陈琼, 张强, 周强, 刘峰贵, 赵富昌, 陈永萍 中国地质灾害与防治学报. 2021, 32(2): 19-26



关注微信公众号, 获得更多资讯信息

DOI: 10.16031/j.cnki.issn.1003-8035.2021.06-12

机器学习模型在滑坡易发性评价中的应用

刘福臻, 王 灵, 肖东升

(西南石油大学土木工程与测绘学院, 四川 成都 610500)

摘要:机器学习在滑坡的易发性评价中面临两个难点,一是评价指标的客观量化,二是训练样本的选择。鉴于此,采用频率比法实现了评价指标的客观量化,利用k均值聚类算法实现了非滑坡样本数据的筛选。结果表明,以k均值聚类算法筛选非滑坡为前提,神经网络的训练精度由73%提升到了97%,支持向量机的训练精度由75%提升到了96%。基于GIS平台,将神经网络和支持向量机模型计算的全区易发性指数按自然断点法分为五个区域,分区图与历史灾害点的叠加分析统计结果显示,神经网络在全局范围内的评价结果优于支持向量机模型,全局精度分别为76%和74%。研究结果可为南江县的防灾减灾工作提供参考。

关键词:滑坡易发性;频率比量化;k均值聚类;神经网络;支持向量机

中图分类号: P642.22 **文献标志码:** A **文章编号:** 1003-8035(2021)06-0098-09

Application of machine learning model in landslide susceptibility evaluation

LIU Fuzhen, WANG Ling, XIAO Dongsheng

(School of Civil Engineering and Surveying, Southwest Petroleum University, Chengdu, Sichuan 610500, China)

Abstract: Machine learning faces two difficulties in the evaluation of landslide susceptibility. One is the objective quantification of evaluation index, and the other is the selection of training sample-0.5pts. For that reason, the frequency ratio method is used to achieve the objective quantification of evaluation index, and the k-means clustering algorithm is used to achieve the selection of non-landslide sample data. The results show that based on the premise that the k-means clustering algorithm selects non-landslides, the training accuracy of the neural network has increased from 73% to 97%, and the training accuracy of the support vector machine has increased from 75% to 96%. Based on the GIS platform, the susceptibility index calculated by the neural network and support vector machine model is divided into five regions according to the natural break point method. The statistical results of the overlay analysis of the zoning map and the historical disaster points show that the evaluation result of the neural network is better than the support vector machine model in the global scope, and the global accuracy is 76% and 74%, respectively. The research results can provide reference for disaster prevention and mitigation in Nanjiang County of China.

Keywords: landslide susceptibility; frequency ratio quantization; k-means clustering; neural networks; support Vector Machines

0 引言

滑坡是我国西南地区最为常见的地质灾害,由于滑

坡发育的复杂性和非线性,目前还没有一套完全成熟的理论能对其进行有效的防治。而我国西南片区受地形地貌、地质构造等复杂环境的影响,更是为滑坡的发育

收稿日期: 2020-11-03; 修订日期: 2021-05-18

基金项目: 国家自然科学基金项目(51774250)

第一作者: 刘福臻(1973-)男,副教授,主要研究方向为地质灾害防治。E-mail:2233896@qq.com

通讯作者: 王 灵(1996-),男,硕士,主要研究方向为地质灾害防治。E-mail: 635370097@qq.com

提供了有利条件。为加强对西南片区滑坡的防治和管理, 对该区域进行精确的滑坡易发性评价具有重要的现实意义。

滑坡易发性评价主要研究内容是: 在区域范围内, 某一确定位置在确定条件下发生滑坡的空间概率^[1]。作为判断滑坡危险性和风险性的基础, 易发性评价是防灾减灾中不可缺少的工作^[2]。国内外学者对滑坡的易发性评价开展了一定程度的研究, 这些研究大多集中在易发性模型的选择以及量化方法上, 而少有学者对其非滑坡样本的选择进行研究。对于回归性和预测性模型而言, 样本的选择通常决定了整个模型的精确性和稳定性, 若将一些高易发区的点位作为了非滑坡的样本, 模型的适用性会大打折扣。为了更加精准的表达滑坡的易发性分区, 有必要对样本的选择进行一定程度的研究。

一般认为, 滑坡的易发性评价体系包含评价指标的量化和评价模型的选择这两个过程。常用的量化方法有专家打分法^[3]、信息量法^[4]、证据权法^[5]、确定性系数法^[6]和频率比法^[7]等。在评价模型上, 常用的模型有层次分析法^[8]、模糊综合评判法^[9]、灰色理论^[10]、粗集理论^[11]、分形理论^[12]以及近年来发展迅速的机器学习

法^[13]。实验选择频率比法作为滑坡易发性的指标量化方法, 该方法从数据本身的结构信息出发, 打破了人为主观给分的限制, 从而实现了滑坡因子二级属性的客观量化。评价模型上, 由于机器学习能够更为准确的反应滑坡易发性与各评价因子之间的非线性关系^[14], 实验选择了机器学习中的聚类模型和回归模型, 利用两个模型的有机结合, 不仅解决了回归模型中训练样本的选择问题, 同时实现了滑坡易发性指数与各评价因子之间的非线性回归。

1 研究区概况及数据源

1.1 研究区概况

南江县位于我国四川省巴中市西北部, 是巴中连接汉中市的重要交通枢纽。全县幅员辽阔, 总面积 3 382.8 km², 是巴中市海拔最高的一个县, 巴中市海拔高度大于 1 800 m 的山体 90% 以上分布在南江县, 且地形起伏大, 最高地段海拔达 2 493 m, 而最低为 332 m。复杂多变的地质、地貌条件为南江县带来了丰富的矿产资源, 同时也为地质灾害的孕育提供了有利条件, 其中地质灾害以滑坡为主。南江县在巴中市的地理位置及其 1958—2009 年的历史滑坡点位置如图 1 所示。

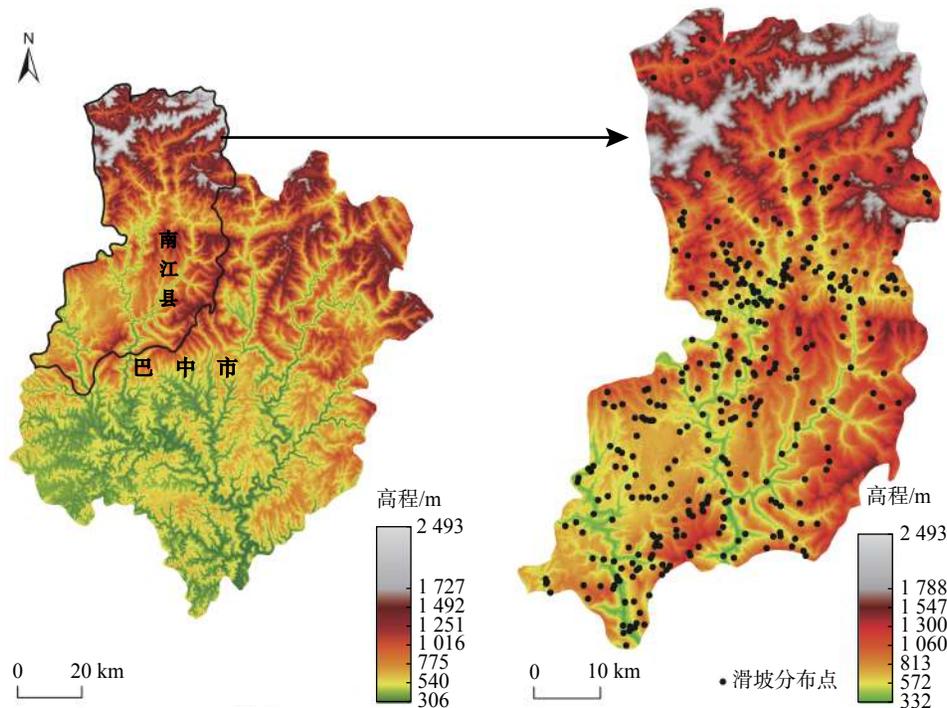


图 1 研究区地理位置及历史滑坡点分布图

Fig. 1 Geographical location of the study area and historical landslide distribution map

1.2 数据源

实验用到的数据有南江县 1958—2009 年地质灾害

详查汇总表、南江县 DEM、南江县 1 : 25 万地质图 (表 1)。通过 DEM 提取出坡度、坡向、坡型、水系和高

程的初始状态因子。将地质图矢量化后转为栅格数据,从而获取地层岩组因子。以上因子全部投影到同一坐标系下,并重采样为 30 m×30 m 的栅格,全区共分为 3 784 091 个栅格,其中滑坡占了 359 个。

表 1 数据源
Table 1 Data source

数据名称	数据类型	数据来源
滑坡灾害点	excel	南江县 1 : 5 万地质灾害详查
DEM	栅格	地理空间数据云
1 : 25 万地质图	栅格	91 卫图

2 模型介绍

2.1 频率比法

评价因子的量化一直以来都是评价模型中最为重要的一环,通常情况下因子的量化值都是采用专家经验打分,该方法虽然操作简单快捷,但存在的主观性太大。文章选择频率比法作为量化模型,该模型从统计学原理出发,将历史滑坡点数据和评价因子二级属性进行叠加分析,从而实现了评价因子的客观量化,见式(1)。

$$X_i = \frac{n_i/s_i}{N/S} \quad (1)$$

式中: X_i ——因子 X 在二级属性 i 下的频率比值;

n_i ——因子 X 在二级属性 i 下的滑坡个数;

s_i ——研究区内因子 X 在二级属性 i 下的栅格个数;

N ——研究区总滑坡个数;

S ——研究区总栅格数。

2.2 机器学习

机器学习的应用主要包含回归、分类、聚类和数据降维四大领域。各模块的相互结合更能突出机器学习的优势。本次实验将聚类思想和回归思想相结合,用于南江县的滑坡易发性评价。实验的算法和数据分析通过 Python 编程平台实现,空间分析模型和出图模块由 ArcGIS 平台完成。通过 Python 和 ArcGIS 的结合,可以方便且快速的对地理空间数据进行分析。

2.2.1 k 均值聚类

k 均值聚类算法能在没有先验数据的情况下,对原始数据进行初步分类,分类结果通过后续的信息补充加以验证。其思想是,随机选择 k 个样本数据作为聚类中心,计算出每个样本与聚类中心的距离,并把相近的样本作为一类。每分配一个样本时,聚类中心会根据当前类中出现的样本重新计算,反复迭代这一过程,直到聚类中心不再发生变化为止^[15]。

2.2.2 神经网络

神经网络算法通过模仿生物神经网络的结构和功能,创建出用于连接输入端和输出端的神经元。每个神经元节点作为一个激励函数,传达不同数据之间的交流信息。节点之间的权重可类比于生物的记忆功能,通过不断的学习进而不断的更新,从而实现类似人一样的判断能力^[16]。在样本数据足够完善的情况下,通过神经网络的训练学习可以很好的表达出一些复杂的、非线性的模型,因此可将神经网络模型用于滑坡易发性指数的回归。

2.2.3 支持向量机

支持向量机模型的思想是:将输入向量根据一定的法则映射到更高维数的特征空间,并在该特征空间构造一个最优分类面,利用最优分类面对数据进行分类或者回归,从而实现学习和预测的功能^[17]。将支持向量机模型作为神经网络的对照模型,对比两个模型在相同训练集情况下的学习效果,筛选出合适的模型并计算评价结果。

3 滑坡评价因子量化结果

滑坡是一个复杂且非线性的系统,实验从滑坡的易发性角度出发^[1],结合研究区概况和专家建议,选择了坡度、坡向、坡型、水系、岩组、高程和地形起伏共 7 个因子。对于连续数据,首先通过自然断点法将数据分级,再利用频率比实现量化,而对于分类数据则直接利用频率比实现量化。量化结果如表 2。

考虑到各因子之间可能存在的相关性会对结果产生不利影响,在评价之前对量化的因子进行相关性分析,结果见表 3。其中地形起伏因子和坡度、岩组、高程的相关性分别达到了 0.61、0.25 和 0.27,相关性过高。而剩余的六个因子之间相关性都低于 0.2,所以实验选择剔除地形起伏因子,保留剩余的 6 个因子作为南江县滑坡易发性的评价因子。

为方便 ArcGIS 成图,实验将每个因子的量化值扩大 1 000 倍后取整,为每个等级因子配上不同的颜色以区分,结果见图 2。

4 机器学习的应用

4.1 数据预处理

为消除数据之间的量纲不一致性,利用标准化操作对数据进行特征缩放,使其值落在 [0,1] 之间。实践证明,处理后的数据不仅能加快机器学习的收敛速度,同时还能提高模型整体精度,保障模型的稳定性和可行

表 2 因子量化结果

Table 2 Results of factor quantification

因子	因子二级属性	s_i	n_i	x_i
坡度/(°)	0 ~ 10	650 642	67	1.085 425 029
	10 ~ 16	803 628	107	1.403 446 432
	16 ~ 22	805 749	96	1.255 852 346
	22 ~ 29	763 630	53	0.731 576 946
	29 ~ 34	375 968	20	0.560 720 245
	34 ~ 43	311 624	15	0.507 373 154
	>43	72 850	1	0.144 689 684
坡向	北	438 393	28	0.673 227 029
	东北	416 896	37	0.935 494 243
	东	450 786	51	1.192 523 317
	东南	479 436	49	1.077 289 835
	南	510 916	50	1.031 543 684
	西南	491 470	60	1.286 830 544
	西	499 440	39	0.823 092 053
	西北	496 754	45	0.954 856 841
坡型	<-1	131 994	1	0.079 856 989
	-1 ~ -0.3	832 505	66	0.835 649 597
	-0.3 ~ 0.1	1 141 012	127	1.173 223 173
	0.1 ~ 0.8	1 491 162	153	1.081 517 936
	>0.8	187 418	12	0.674 896 336
水系/m	<200	331 210	40	1.272 986 136
	200 ~ 500	479 405	64	1.407 163 424
	500 ~ 800	455 006	48	1.111 965 305
	800 ~ 1 200	558 891	65	1.225 895 254
	1 200 ~ 1 500	381 262	34	0.939 988 453
	1 500 ~ 2 000	548 717	51	0.979 690 471
	2 000 ~ 5 000	1 023 859	57	0.586 815 838
	>5 000	5 741	0	0.000 000 000
岩组	K ₁	882 610	101	1.206 200 914
	J ₃	1 024 962	120	1.234 072 302
	J ₁₋₂	266 052	77	3.050 642 528
	T ₁₋₂	223 391	22	1.038 064 004
	P ₂₋₃	94 493	2	0.223 098 927
	Pz ₁	360 578	21	0.613 885 241
	Z	426 090	14	0.346 332 954
	ξγNh	326 208	2	0.064 625 291
	Pt ₂	148 769	0	0.000 000 000
	S ₁₋₂	30 938	0	0.000 000 000
高程/m	332 ~ 604	424 239	84	2.087 064 249
	604 ~ 768	713 900	121	1.786 549 738
	768 ~ 924	647 443	77	1.253 592 279
	924 ~ 1 094	538 137	44	0.861 840 594
	1 094 ~ 1 273	451 639	23	0.536 788 894
	1 273 ~ 1 464	377 209	7	0.195 606 426
	1 464 ~ 1 670	276 415	3	0.114 400 197
	1 670 ~ 1 897	236 877	0	0.000 000 000
地形起伏/m	1897 ~ 2 493	118 232	0	0.000 000 000
	9 ~ 98	543 341	64	1.241 579 747
	98 ~ 151	997 907	120	1.267 530 155
	151 ~ 202	929 250	105	1.191 033 159
	202 ~ 258	720 598	53	0.775 264 576
	258 ~ 329	447 958	13	0.305 895 564
329 ~ 655	145 037	4	0.290 702 192	

表 3 因子相关性分析结果

Table 3 Results of factor correlation analysis

因子	坡度	坡向	坡型	水系	岩组	高程	地形
坡度	1.00	0.02	0.15	-0.05	0.16	0.16	0.61
坡向	0.02	1.00	0.02	0.01	0.01	0.00	0.00
坡型	0.15	0.02	1.00	0.01	0.09	0.09	0.17
水系	-0.05	0.01	0.01	1.00	0.12	0.13	-0.06
岩组	0.16	0.01	0.09	0.12	1.00	0.19	0.25
高程	0.16	0.00	0.09	0.13	0.19	1.00	0.27
地形	0.61	0.00	0.17	-0.06	0.25	0.27	1.00

性。标准化公式如式(2)。

$$N_i = \frac{X_i - X_{\min}}{X_{\max} - X_{\min}} \quad (2)$$

式中: N_i —— X 标准化后的值;

X_i —— 因子 X 在二级属性 i 下的频率比值;

X_{\min} —— X_i 量化的最小值;

X_{\max} —— X_i 量化的最大值。

4.2 样本选取

训练样本的质量决定着模型整体的稳定性, 因此在滑坡样本的选取上, 实验将 359 个滑坡点全部作为样本数据, 以此来外推出符合滑坡性质的栅格点。在非滑坡样本的选取上, 实验通过随机选择和通过 k 均值聚类两种方法筛选非滑坡点。再将非滑坡点和滑坡点的数据整合到一起, 以此获得完整的样本数据。

4.2.1 随机选取非滑坡点

实验通过随机选择研究区的点位作为非滑坡点, 以此来和聚类的结果作为对照。将随机选择的 1 000 个非滑坡点和 359 个滑坡点整合, 再打乱顺序, 按照 7 : 3 的比例分配训练集和测试集。

4.2.2 K 均值聚类选取非滑坡点

将因子标准化结果作为聚类的输入数据, 采用 k 均值聚类算法将原始数据分为五大类。为了挑选出非滑坡的样本点, 将聚类结果和历史滑坡点进行叠加分析, 结果如表 4。统计聚类结果中滑坡所占的个数和滑坡所占的相对比例。选择滑坡所占个数最少且滑坡所占相对比例最低的聚类结果作为非滑坡的样本, 由表中可知, 聚类结果为 2 的栅格满足要求, 因此从该类中随机采样 1 000 个点作为非滑坡点。同样将筛选出来的 1 000 个非滑坡样本和 359 个滑坡样本整合, 随机打乱顺序, 按照 7 : 3 的比例分配训练集和测试集。

4.3 神经网络模型的构建

通过反复对比实验, 神经网络模型在构建时选择一个隐含层, 且隐含层中设置 13 个神经节点的效果较

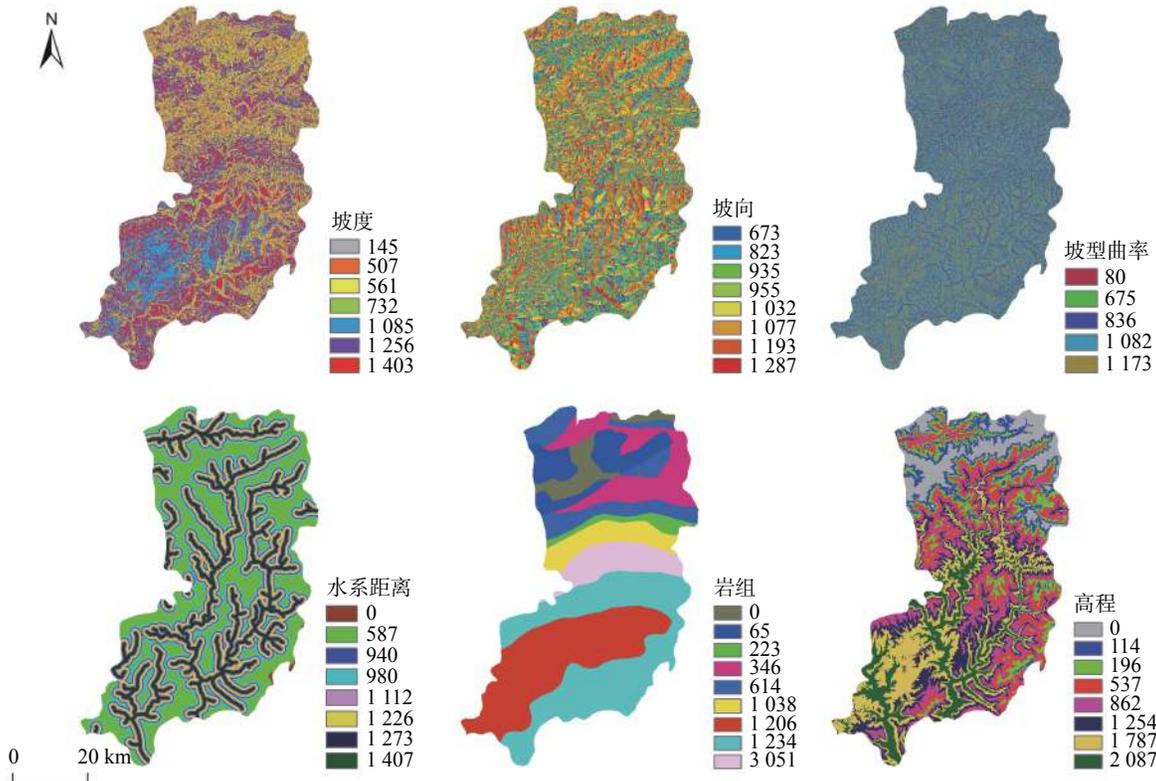


图 2 因子量化结果

Fig. 2 Results of factor quantification

表 4 k 均值聚类统计分析结果

Table 4 Results of k-means clustering statistical analysis

聚类结果	栅格数量	滑坡点数	相对滑坡比
0	909 306	37	0.428 902 710
1	836 114	106	1.336 310 845
2	963 716	28	0.306 249 991
3	265 935	77	3.051 984 680
4	809 020	111	1.446 208 281

好。每个神经元的激活函数选择 Logistic 函数，权重迭代器选择了基于随机梯度的优化器。将随机模型和 k 均值聚类模型生成的数据集分别作为神经网络模型的输入，保存训练好的两个模型以及相应的训练结果参数。

4.4 支持向量机模型的构建

支持向量机在模型构建时，在核函数上选择了能处理非线性特征的多项式核函数。同时将误差项的惩罚参数设置为 1，通过验证该值能很好的提高模型的预测能力和泛化能力。同理，将随机模型和 k 均值聚类模型生成的数据集分别作为支持向量机模型的输入，保存训练好的两个模型以及相应的训练结果参数。

4.5 模型精度验证及模型选择

上述的对照实验，一共保存了 4 个模型，对应 4 组精度验证曲线，分别为：(1)通过 k 均值聚类算法筛选非

滑坡点，以神经网络为训练模型，训练结果的 ROC 曲线 ANN_km。(2)通过 k 均值聚类算法筛选非滑坡点，以支持向量机为训练模型，训练结果的 ROC 曲线 SVR_km。(3)通过随机筛选非滑坡点，以神经网络为训练模型，训练结果的 ROC 曲线 ANN_sj。(4)通过随机筛选非滑坡点，以支持向量为训练机模型，训练结果的 ROC 曲线 SVR_sj(图 3)。

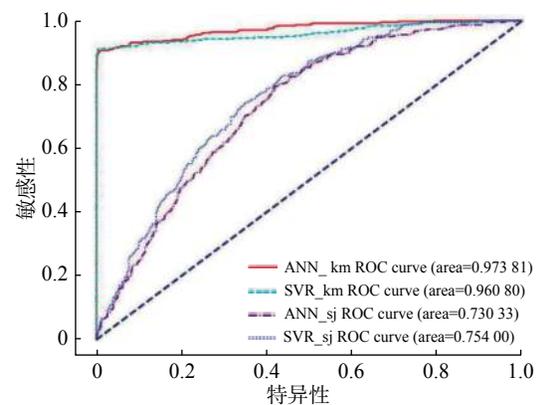


图 3 训练集 ROC

Fig. 3 Training set ROC

为检验模型的稳定性和可靠性，将相应的测试集输入到训练好的 4 个模型，同样也获得了对应的 4 条

ROC 曲线。分别为: ANN_km_t, SVR_km_t, ANN_sj_t, SVR_sj_t(图 4)。

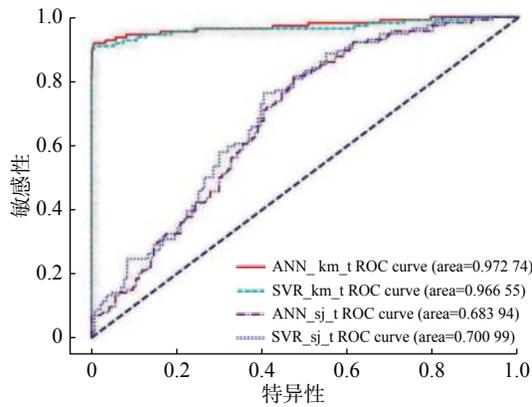


图 4 测试集 ROC
Fig. 4 Test set ROC

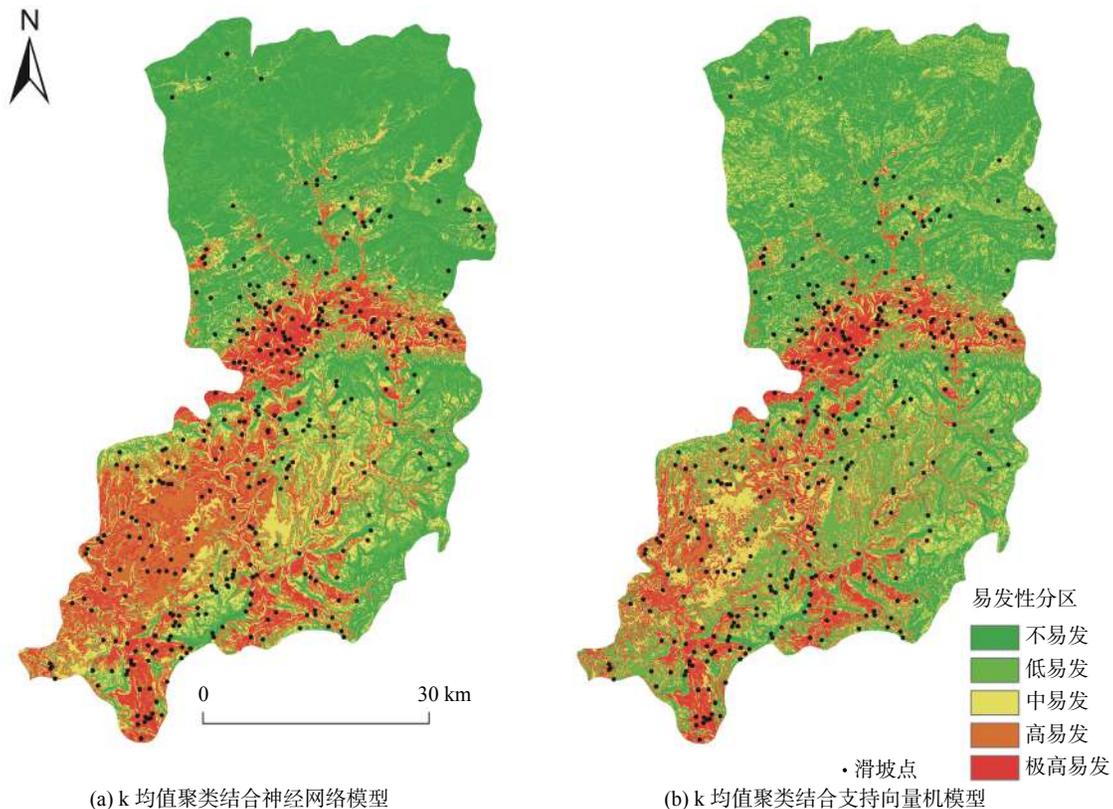
由结果可知,在随机筛选非滑坡样本的前提下,神经网络和支持向量机的训练精度约为 70%,表明两个模型在一定程度上能反映滑坡的易发性指数。在 k 均值聚类筛选非滑坡样本的前提下,神经网络和支持向量机的训练精度约为 97% 和 96%,远高于随机模型下的训练精度,由此可知,k 均值聚类和两个回归模型的结合

能够更好的反映滑坡的易发性指数。

对比训练集和测试集的 ROC 曲线面积,在相同前提和模型下,训练集和测试集的精度基本相同,说明以上 4 个模型都拥有稳定的预测和外推能力。其中在 k 均值聚类的前提下,神经网络及支持向量机的训练集和测试集精度都高于 95%,表明 k 均值聚类和两个回归模型的结合不仅能很好的反映滑坡易发性指数,而且模型具有很好的稳定性、泛化性和外推性。实验最终选择了训练精度和稳定性都较高的 k 均值聚类和神经网络的结合模型及 k 均值聚类和支持向量机的结合模型作为评价的基础。

4.6 易发性分区

将研究区的所有栅格点数据分别输入到保存好的 k 均值聚类—神经网络模型和 k 均值聚类—支持向量机模型,以此获取不同模型下的易发性指数,再利用自然断点法根据易发性指数将南江县分为五个区域,分别为不易发、低易发、中易发、高易发和极高易发(图 5)。其中图 5(a)为在 k 均值聚类结合神经网络模型得到的结果;图 5(b)为 k 均值聚类结合支持向量机模型获得的结果。两个模型所得到的分区结果在大体上相似,其中不易发和低易发区域集中出现在南江县北部区域。



(a) k 均值聚类结合神经网络模型

(b) k 均值聚类结合支持向量机模型

图 5 易发性分区图

Fig. 5 Susceptibility zone map

高易发和极高易发集中出现在南江县中部。

为验证两个模型在南江县整个范围内的适用性,实验将两个模型得到的易发性分区图与历史灾害点叠加分析,通过相对滑坡频率比来验证模型精度,计算公式如式(3)。

$$A_i = \frac{m_i/s_i}{N/S} \quad (3)$$

式中: A_i ——易发性分区*i*的相对滑坡频率比;

m_i ——易发性分区*i*下的滑坡个数;

s_i ——易发性分区*i*的栅格个数;

N ——研究区总滑坡个数;

S ——研究区总栅格数。

神经网络模型的精度验证结果见表 5,支持向量机模型的精度验证结果见表 6。从结果中我们可以看出,不管是神经网络还是支持向量机,两个模型所得结果的滑坡点在不易发到极高易发的个数都是递增的,相应的相对滑坡比也是递增的。说明两个模型在研究区范围内具有稳定的适用性。其中神经网络在极高易发区的相对滑坡比高于支持向量机,在不易发的相对滑坡比低于支持向量机,这表明神经网络在全局上的精度高于支持向量机。

表 5 神经网络分区统计结果

Table 5 Partition statistics results of neural network

易发性等级	栅格数量	分区面积比例/%	滑坡点数	相对滑坡频率比
不易发	1 333 655	35.24	35	0.276 625 155
低易发	654 221	17.29	38	0.612 246 399
中易发	738 454	19.51	70	0.999 175 361
高易发	626 759	16.56	103	1.732 222 873
极高易发	431 002	11.39	113	2.763 543 349

表 6 支持向量机分区统计结果

Table 6 Partition statistical results of support vector machines

易发性等级	栅格数量	分区面积比例/%	滑坡点数	相对滑坡频率比
不易发	1 125 524	29.74	43	0.402 699 248
低易发	943 452	24.93	49	0.547 448 656
中易发	830 829	21.96	72	0.913 456 715
高易发	463 728	12.25	87	1.977 529 889
极高易发	420 558	11.11	108	2.706 854 924

从数据的直观性上我们可初步判定神经网络的精度优于支持向量机,为了更准确的反应两个模型的精度,实验用成功率验证曲线^[18]来量化两个模型在全局的精度,其中横轴为根据易发性指数从高到低的累计栅格百分比,纵轴为历史滑坡累计发生频率(图 6)。实验表

明神经网络在全局的精度约为 76%,支持向量机的精度约为 74%。这也验证了上述由相对滑坡比所得到的结论。

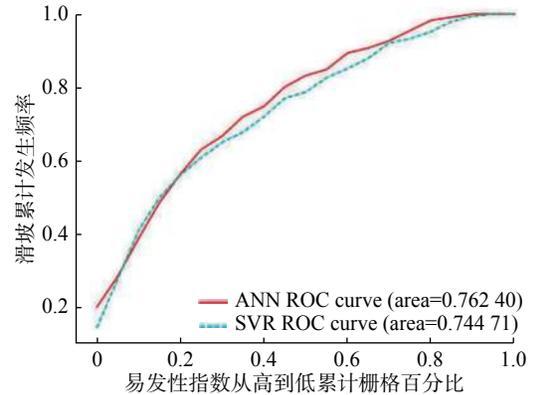


图 6 模型全局精度验证曲线

Fig. 6 Model's global accuracy verification curve

5 结论

(1)利用 k 均值聚类筛选非滑坡样本,得到的结果用于神经网络和支持向量机两个模型的训练,其训练结果精度分别为 97% 和 96%,远高于利用随机模型筛选滑坡时获得的训练精度。

(2)将 k 均值聚类算法同神经网络结合,得到南江县的易发性分区图,其中从不易发到高易发分区所占的比例分别为 35.24%、17.29%、19.51%、16.56%、11.39%。将 k 均值聚类算法同支持向量机结合,得到了另一份南江县易发性分区图,其中从不易发到高易发分区所占的比例分别为 29.74%、24.93%、21.96%、12.56%、11.11%。两个模型所得到的分区结果在大体上相似,其中不易发和低易发区域集中出现在南江县北部区域。高易发和极高易发集中出现在南江县中部。

(3)利用相对滑坡比作为模型的评价指标,其中神经网络在极高易发区的相对滑坡比高于支持向量机,在不易发的相对滑坡比低于支持向量机,这表明神经网络在全局上的精度高于支持向量机。为了量化模型的全局精度,采用成功率验证曲线,结果表明神经网络在全局的精度约为 76%,支持向量机在全局的精度约为 74%。

参考文献 (References) :

[1] 邱海军. 区域滑坡崩塌地质灾害特征分析及其易发性和危险性评价研究:以宁强县为例[D]. 西安:西北大学, 2012. [QIU Haijun. Study on the regional landslide characteristic analysis and hazard assessment: A case study of

- Ningqiang County[D]. Xi'an: Northwest University, 2012. (in Chinese with English abstract)]
- [2] 唐亚明,张茂省,李林,等. 滑坡易发性危险性风险评价例析 [J] . *水文地质工程地质*, 2011, 38(2): 125 - 129. [TANG Yaming, ZHANG Maosheng, LI Lin, et al. Discrimination to the landslide susceptibility, hazard and risk assessment [J] . *Hydrogeology & Engineering Geology*, 2011, 38(2): 125 - 129. (in Chinese with English abstract)]
- [3] 杨秀梅. 基于GIS的地质灾害危险性评价[D]. 兰州: 兰州大学, 2008. [YANG Xiumei. Geological hazard risk assessment based on GIS[D]. Lanzhou: Lanzhou University, 2008. (in Chinese with English abstract)]
- [4] 胡芹龙,王运生. 基于GIS的川西地貌过渡带滑坡灾害易发性评价 [J] . *成都理工大学学报(自然科学版)*, 2018, 45(6): 746 - 753. [HU Qinlong, WANG Yunsheng. The susceptibility assessment of geological disasters in geomorphic transition zone based on GIS, western Sichuan, China [J] . *Journal of Chengdu University of Technology (Science & Technology Edition)*, 2018, 45(6): 746 - 753. (in Chinese with English abstract)]
- [5] 郭长宝,唐杰,吴瑞安,等. 基于证据权模型的川藏铁路加查—朗县段滑坡易发性评价 [J] . *山地学报*, 2019, 37(2): 240 - 251. [GUO Changbao, TANG Jie, WU Ruian, et al. Landslide susceptibility assessment based on WOE model along Jiacha—Langxian County section of Sichuan—Tibet railway, China [J] . *Mountain Research*, 2019, 37(2): 240 - 251. (in Chinese with English abstract)]
- [6] 兰恒星,伍法权,王思敬. 基于GIS的滑坡CF多元回归模型及其应用 [J] . *山地学报*, 2002, 20(6): 732 - 737. [LAN Hengxing, WU Faquan, WANG Sijing. GIS based landslide CF multi-variable regression model and its application [J] . *Journal of Mountain Research*, 2002, 20(6): 732 - 737. (in Chinese with English abstract)]
- [7] 郭子正,殷坤龙,黄发明,等. 基于滑坡分类和加权频率比模型的滑坡易发性评价 [J] . *岩石力学与工程学报*, 2019, 38(2): 287 - 300. [GUO Zizheng, YIN Kunlong, HUANG Faming, et al. Evaluation of landslide susceptibility based on landslide classification and weighted frequency ratio model [J] . *Chinese Journal of Rock Mechanics and Engineering*, 2019, 38(2): 287 - 300. (in Chinese with English abstract)]
- [8] 许冲,戴福初,姚鑫,等. GIS支持下基于层次分析法的汶川地震区滑坡易发性评价 [J] . *岩石力学与工程学报*, 2009, 28(增刊2): 3978 - 3985. [XU Chong, DAI Fuchu, YAO Xin, et al. Gis-based landslide susceptibility assessment using analytical hierarchy process in Wenchuan earthquake region [J] . *Chinese Journal of Rock Mechanics and Engineering*, 2009, 28(Sup 2): 3978 - 3985. (in Chinese with English abstract)]
- [9] 苏强. 基于DEM的黄土滑坡危险性评价研究[D]. 北京: 中国地质大学(北京), 2006. [SU Qiang. Research on loess landslide hazard zonation based on DEM[D]. Beijing: China University of Geosciences, 2006. (in Chinese with English abstract)]
- [10] 李嘉良,马东辉,王威. 基于证据理论和熵权灰色关联的潜在地震滑坡危险性评价 [J] . *中南大学学报(自然科学版)*, 2016, 47(5): 1730 - 1736. [LI Jiali, MA Donghui, WANG Wei. Assessment of potential seismic landslide hazard based on evidence theory and entropy weight grey incidence [J] . *Journal of Central South University (Science and Technology)*, 2016, 47(5): 1730 - 1736. (in Chinese with English abstract)]
- [11] 牛瑞卿,彭令,叶润青,等. 基于粗糙集的支持向量机滑坡易发性评价 [J] . *吉林大学学报(地球科学版)*, 2012, 42(2): 430 - 439. [NIU Ruiqing, PENG Ling, YE Runqing, et al. Landslide susceptibility assessment based on rough sets and support vector machine [J] . *Journal of Jilin University (Earth Science Edition)*, 2012, 42(2): 430 - 439. (in Chinese with English abstract)]
- [12] 王森,许强,罗博宇,等. 基于分形理论的南江县滑坡敏感性分析与易发性评价 [J] . *水文地质工程地质*, 2017, 44(3): 119 - 126. [WANG Sen, XU Qiang, LUO Boyu, et al. Vulnerability analysis and susceptibility evaluation of landslides based on fractal theory in Nanjiang County [J] . *Hydrogeology & Engineering Geology*, 2017, 44(3): 119 - 126. (in Chinese with English abstract)]
- [13] 何静. 基于机器学习的滑坡灾害空间预测及风险评估[D]. 成都: 电子科技大学, 2019. [HE Jing. Spatial prediction and risk assessment of landslides based on machine learning[D]. Chengdu: University of Electronic Science and Technology of China, 2019. (in Chinese with English abstract)]
- [14] 黄发明,殷坤龙,蒋水华,等. 基于聚类分析和支持向量机的滑坡易发性评价 [J] . *岩石力学与工程学报*, 2018, 37(1): 156 - 167. [HUANG Faming, YIN Kunlong, JIANG Shuihua, et al. Landslide susceptibility assessment based on clustering analysis and support vector machine [J] . *Chinese Journal of Rock Mechanics and Engineering*, 2018, 37(1): 156 - 167. (in Chinese with English abstract)]
- [15] 王敞,陈增强,袁著祉. 基于遗传算法的K均值聚类分析 [J] . *计算机科学*, 2003, 30(2): 163 - 164. [WANG Chang, CHEN Zengqiang, YUAN Zhuzhi. K-means clustering based on genetic algorithm [J] . *Computer Science*, 2003, 30(2): 163 - 164. (in Chinese with English abstract)]
- [16] 陈玉萍,袁志强,周博,等. 遗传算法优化BP网络在滑坡

- 灾害预测中的应用研究 [J] . 水文地质工程地质, 2012, 39(1): 114 - 119. [CHEN Yuping, YUAN Zhiqiang, ZHOU Bo, et al. Application of back propagation neural networks with optimization of genetic algorithms to landslide hazard prediction [J] . Hydrogeology & Engineering Geology, 2012, 39(1): 114 - 119. (in Chinese with English abstract)]
- [17] 武雪玲, 沈少青, 牛瑞卿. GIS支持下应用 PSO-SVM模型预测滑坡易发性 [J] . 武汉大学学报(信息科学版), 2016, 41(5): 665 - 671. [WU Xueling, SHEN Shaoqing, NIU Ruiqing. Landslide susceptibility prediction using GIS and PSO-SVM [J] . Geomatics and Information Science of Wuhan University, 2016, 41(5): 665 - 671. (in Chinese with English abstract)]
- [18] 张俊, 殷坤龙, 王佳佳, 等. 三峡库区万州区滑坡灾害易发性评价研究 [J] . 岩石力学与工程学报, 2016, 35(2): 284 - 296. [ZHANG Jun, YIN Kunlong, WANG Jiajia, et al. Evaluation of landslide susceptibility for Wanzhou district of Three Gorges Reservoir [J] . Chinese Journal of Rock Mechanics and Engineering, 2016, 35(2): 284 - 296. (in Chinese with English abstract)]