**基于浅层机器学习和深度学习的滑坡易发性评价比较研究**

**许石罗1；宋英旭2,\*；XXX**

**摘要：**

本文的主要目的是对比浅层机器学习方法（随机森林、逻辑回归）和深度学习方法（LSTM）在滑坡易发性评价中的表现。本文以三峡库区秭归到巴东段为研究区域，选取坡度、坡向等因子作为滑坡因子，通过对比可以发现，深度学习方法在滑坡易发性评价当中表现出更好的性能，AUC值更大。特别地，本文对比了考虑类别平衡情况下各个机器学习方法的优劣，发现，逻辑回归方法最差；本文也考虑样本维度对滑坡易发性评价的影响，使用主成分分析方法进行因子约减，发现\*\*\*\*\*\*

关键词：深度学习；机器学习；滑坡易发性；类别不平滑；样本维度

# 0背景

滑坡是我国危害最严重的地质灾害之一，在多山地区、库岸地区广泛发育，严重威胁人民群众的生命财产安全。滑坡易发性评价根据影响滑坡发生的一系列评价因子，利用评估模型从空间上预测滑坡发生的概率。滑坡易发性评价模型大致可以分为两类，分别为模型驱动和数据驱动。早期以模型驱动的评价方法为主。后来随着技术发展各类观测和调查数据逐渐丰富，研究者开始从多源数据中寻找滑坡发生的规律，称之为数据驱动的方法。代表性的方法有信息量法[1]、逻辑回归[2]、支持向量机（Support Vector Machine，SVM）[3]、随机森林[4]、BP神经网络[5]、层次分析法[6]、模糊逻辑[7]。这些方法的共同点是网络层次较少，随着积累的各类滑坡数据越来越大，这些浅层机器学习方法已经不能充分地提取出数据隐含的内部特征，制约了滑坡灾害的空间评价的发展。深度学习以深度神经网络为代表，其具有更深的网络层次，拥有更多的内部参数，数据拟合和特征提取能力更强，在滑坡易发性评价中得到了越来越多的应用。如利用RNN和CNN对伊朗作滑坡易发性评价，Yang YI等人将CNN应用于四川九寨沟地区进行了实验应用，验证了CNN相对于多层感知机（MLP）具有更好的数据拟合优势[8]。如Sameen, M I等(2020)基于CNN、ANN和SVM等方法，通过评价滑坡灾害易发性来间接检测滑坡[9]。陈涛等利用深度信念网络，对滑坡的易发性评价作了探索性研究，并将其与浅层神经网络和逻辑回归对比，验证了DBN具有更好的预测能力[10]。可以看出，深度学习在滑坡易发性中具备一定的应用价值。

本文考虑多种影响滑坡发生的诱发因素和控制因素，比对常见的传统浅层机器学习方法和深度学习方法在滑坡易发性评价中的性能表现，探索深度学习方法和浅层机器学习方法各自的优势和缺点。

# 1评价方法

## 1.1逻辑回归

逻辑回归（Logistic Regression, LR）在滑坡易发性评价中被广泛使用。一般地，使用y=w1x1+w2x2+…wn+b对数据作线性拟合，但是滑坡易发性评价是非线性的二分类问题，需要在此基础上对y施加一个非线性的映射，一般使用Sigmoid函数添加非线性特性。Sigmoid函数如下：

上式会使模型输出范围映射值[0,1]，输出值接近0判定归属于非滑坡；输出值接近1判定归属于滑坡。为便于计算，一般会将上式取对数。

## 1.2随机森林

随机森林（Random Forest，RF）[11]是一种集成学习模型，通过集成多棵弱决策树模型形成强分类模型。由于RF随机对输入样本和特征采样，因此可以一定程度上避免过拟合，并且有较好的抗噪能力。RF的构建过程如下：

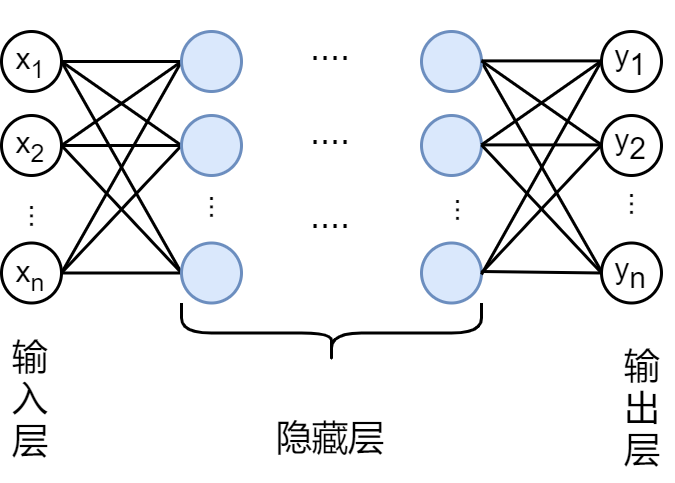
（1）样本随机采样。假定总样本数为N，RF中有m棵决策树；每次从N个样本中有放回的抽取s个样本，建立样本子集Ns；共抽样m次，形成m个样本子集。

（2）特征随机采样。假定样本的总特征数目为F，从总特征数F中抽取h个特征形成特征子集，每次决策树分裂时的最优特征从特征子集中产生。

（3）分类器投票。m棵决策树共产生m个分类结果，以投票数多的类别作为最终的预测类别。

## 1.3深度全连接神经网络

传统的浅层神经网络（如BP神经网络[12]）一般由三层组成，分别是输入层、隐藏层、输出层。前一层每个节点与后一层的每个节点之间都存在着关联关系。三层的神经网络由于网络层次较浅，拟合能力受到一定制约。深度全连接神经网络（Deep Fully Connected Neural Network，DFCNN）在原始三层结构的基础上，增加若干个隐藏层，形成层次更丰富的深度网络。其结构图如下：



网络的输入为，一般会给输入施加权重w以及偏置b，为使得神经元具有非线性映射特性，通常会让该输出通过一个激活函数，过程可以用下式表达：

常见的激活函数有Sigmoid函数、Tanh函数和ReLU函数。由于Sigmoid函数和Tanh函数存在梯度消失的问题，本文采用ReLU函数，ReLU函数在输入大于0时，导数为1，可以确保梯度不衰减，因此ReLU函数一定程度上可以缓解梯度消失问题。ReLU函数的定义如下：

由于滑坡易发性评价实际上是二分类问题，在经过多次上述计算之后，最后的输出层后连接一个Sigmoid函数，输出每个类别的概率值。

## 1.4长短期记忆神经网络

长短期记忆神经网络（Long Short-Term Memory Neural Network，LSTM）是一种特殊的循环神经网络（Recurrent Neural Network，RNN）。循环神经网络在隐藏层中它存在一个定向循环，将上一个隐藏层中的信息输入至当前隐藏层。不同时间点的输入数据之间就存在一定的联系，而不是彼此割裂。但是对于比较长的序列而言，RNN仍然存在着梯度消失的问题。为解决该问题，LSTM采用了一种门控机制来控制信息流的状态。记忆块中存在着三种门控单元，分别是记忆门、遗忘门和输出门。记忆门选择性的保留重要的关键信息，遗忘门过滤不重要或者干扰信息，输出门选择对外输出的信息。门控单元的示意图如下：



图 1 记忆块中的门控示意图（t时刻）

图中xt表示t时刻的原始输入，ht-1为t-1时刻传递的隐藏层状态，Ct-1为t-1时刻传递的记忆块状态。经过门控单元之后产生三个输出yt, ht, Ct，分别是当前时刻t的目标输出、当前时刻t的隐藏层输出、当前时刻t的记忆块状态输出。记忆块中同样使用sigmoid和tanh激活函数建立非线性映射能力。

## 1.5精度评价

由于滑坡面积占比相对于整个研究区是比较小的，因此滑坡易发性评价实际上是不平衡类别的分类问题：即非滑坡区多，滑坡区少。根据数据统计，研究区内的非滑坡与滑坡比高达18:1。因此，简单地使用准确度评估易发性预测的精度会导致模型倾向于将分类结果偏向样本较多的类别（非滑坡区）。将非滑坡定义为负类（Negative），滑坡定义为正类（Positive）。那么滑坡和非滑坡有四种预测结果：TP（True Positive）是预测为滑坡，实际也为滑坡，预测正确；FP（False Positive）是预测为滑坡，实际为非滑坡，预测错误；TN（True Negative）是预测为非滑坡，实际为非滑坡，预测正确；FN（False Negative）预测为非滑坡，实际为滑坡，预测错误。

（1）Recall与F1得分

Recall和F1 得分定义如下：

其中Recall为召回率，Precision为精确率。Recall值为实际滑坡数据中，被正确预测的比例，也称为查全率；Precision值为所有预测为滑坡的数据中预测正确的比例大小。一般而言，发生滑坡造成的损失比较大，需要采用Recall值衡量滑坡预测是否全面；但是只采用Recall作为评价标准又会使得模型倾向于更多地将样本预测为滑坡，造成虚警率比较高。F1得分是Recall和Precision的调和平均，避免模型完全倾向于某个类别。因此，本文使用Recall和F1得分作为衡量滑坡易发性评价的精度判定标准。

（2）AUC值

受试者工作曲线（Receiver Operating Characteristic Curve，ROC）的横轴为FPR（False Positive Rate），纵轴为TPR（True Positive Rate），两者的定义如下：

ROC曲线不随样本类别分布的变化而变化，因此用于衡量类别不均衡且样本分布变化较大的滑坡易发性评价问题较为客观。为量化ROC曲线，采用ROC曲线下的面积AUC（Area Under Curve）评价模型的性能。AUC的取值范围介于[0,1]之间。AUC值越大，模型性能越好，反之模型性能越差。一般情况下模型的AUC值会大于0.5，若AUC值小于0.5，模型预测结果比随机猜测的结果还要差。

# 2.研究区概况

研究区为三峡库区秭归县-巴东段，经度介于110°15′51″E与110°52′33″E之间，纬度介于30°51′21″N与31°5′1″N之间（图 2）。研究区内多山地、河谷和丘陵，总面积662.671平方公里，最高海拔2004 m，长江干流流经整个区域。研究区内滑坡极为发育，主要分布在长江干流及支流两侧。根据实地调查数据，包含已查明滑坡（稳定与不稳定）332个，总面积约4210 m2，其中土质滑坡占比53%、岩质滑坡占37%，岩土混合类型的占比2%，其他类型的占8%。研究区降雨充沛，月平均降雨量可高达1000 mm以上，强降雨导致土壤含水量大幅度上升，是诱发滑坡的重要因素。同时，研究区内有仙女山断裂、牛口断裂等著名断裂，沿着断裂带岩层受到挤压和拉伸，地层不稳定性增加，控制着断裂带周边滑坡的形成。



图 2 研究区地理位置示意图

## 2.1数据来源

本文的数据源主要包括2013年Landsat 8 OLI遥感影像，1：5万地质图，数字高程模型（DEM）采用的是ASTER GDEM V2影像，研究区内各个雨量站获取的雨量数据（2013年）；地震台网数据（1978年至2013年）；另外还有实地调查的滑坡面数据。在原始数据的基础上，通过矢量化、栅格化，得到各个影响因素的栅格图层，栅格图层的空间分辨率为30m，与DEM和Landsat数据分辨率一致。

## 2.2技术路线

在获取不同评价因子数据之后，通过计算和转化获得评价因子和滑坡调查数据样本。随机选择70%的栅格单元作为训练数据集，30%作为测试数据集。选择18个评价因子作为输入特征，实地调查的滑坡面数据转化为栅格单元作为目标特征。输入特征（评价因子）在输入前作标准化处理，避免特征的量纲差异过大导致模型预测能力降低。本文采用两大类模型：一类是传统浅层机器学习模型，采用逻辑回归（LR）、随机森林（RF）；一类是深度神经网络模型，采用深度全连接神经网络（DFCNN）、长短期记忆神经网络（LSTM）对研究区进行易发性评价，在测试数据集上比较不同模型的预测能力、计算时间。整体的技术流程如下：

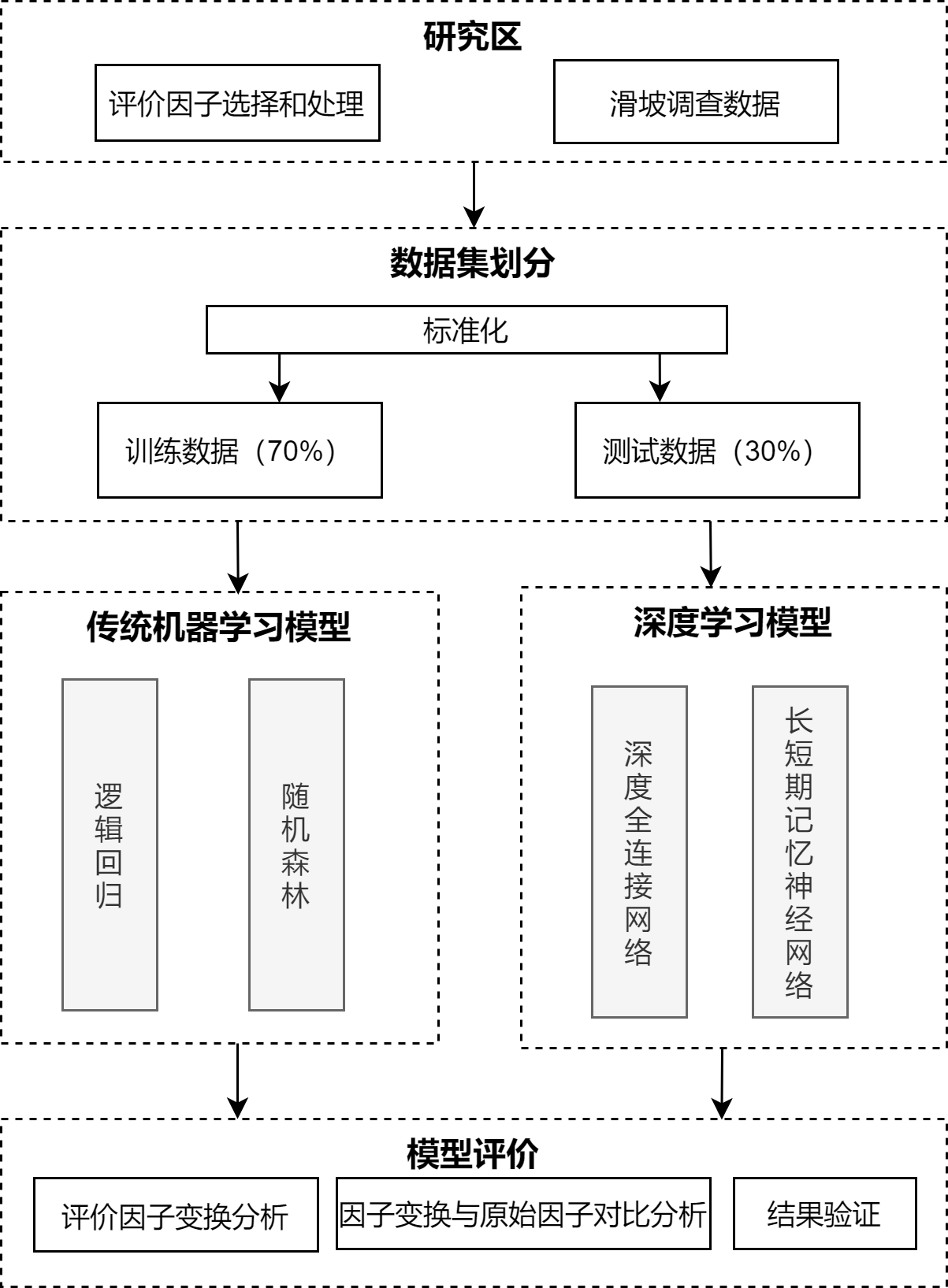


图 3 滑坡易发性评价流程图

## 2.3评价因子处理

根据前人研究成果和经验，本文采用六大类滑坡评价因子。分别是地表覆盖、地形地貌、水文条件、地质条件、地震、降雨。地表覆盖包含土地利用和植被覆盖度。植被的根系可以固定土壤和岩石，同时能够吸水土壤中的水分，对滑坡的影响一般是正向的。土地利用则可以反映出人类工程活动的情况，人类工程活动剧烈的区域，斜坡不稳定性增强。地形地貌是控制滑坡发育的重要条件。在一定范围内的坡度、坡向、斜坡形态、曲率、表面纹理、粗糙指数均会影响边坡的稳定性。水文条件是研究区内滑坡发生的重要影响因素。研究区内的滑坡主要分布在水域的两侧。地质条件对滑坡的发育起着决定性的作用，是重要的内部控制因素。断裂会影响岩层的稳定，距离断层越近，岩层受到的影响越大。地震能够使的岩层发生较大的移位、会挤压拉伸岩层，造成地质环境发生剧烈变化，影响边坡稳定性。降雨能够影响土壤中的含水量，短时强降雨可能造成土壤侵蚀、地表径流增大，造成土壤和岩层吸附力减弱；同时，降雨引起库水位升降，静动水压力随之变化，进一步加剧斜坡的不稳定性。

表 1滑坡评价因子数据来源与数据处理方法

|  |  |  |  |
| --- | --- | --- | --- |
| **原始数据来源** | **因子类别** | **处理方法** | **滑坡影响因子** |
| 遥感图像 | 地表覆盖 | 预处理、模型计算 | 植被覆盖度 |
| 预处理、像素分类 | 土地利用 |
| 数字高程模型 | 地形地貌 | 直接使用 | 高程 |
| 模型计算 | 坡度 |
| 模型计算 | 坡向 |
| 模型计算 | 斜坡形态 |
| 模型计算 | 地形表面纹理 |
| 模型计算 | 地形粗糙指数 |
| 模型计算 | 地形曲率 |
| 数字高程模型 | 水文条件 | 模型计算 | 流域面积 |
| 模型计算 | 流路长度 |
| 模型计算 | 河流强度指数 |
| 水系面  数字高程模型 | 模型计算 | 水系距离 |
| 地质图件 | 地质条件 | 矢量化、栅格化 | 工程岩组 |
| 矢量化、距离计算、栅格化 | 断层距离 |
| 矢量化、模型计算、栅格化 | 斜坡结构（联合坡度和坡向计算） |
| 地震台网 | 地震 | 插值 | 地震震级 |
| 雨量站数据 | 降雨 | 插值 | 降雨因子 |
| 调查数据 | 滑坡面 | 栅格化 | 滑坡数据 |

地质图件通过矢量化的形式转为数字图形，主要用于提取地层岩性、断层距离；遥感图像通过一系列的预处理之后提取植被覆盖度信息和土地利用信息；降雨则通过克里金插值得到整个研究区域的栅格图像。地震数据通过对峰值加速度作三角插值得到。高程数据直接使用，除裁剪之外不需要额外处理。斜坡结构、斜坡形态、坡度、坡向、地形表面纹理、地形粗糙指数、剖面曲率、地形曲率则由高程数据通过一定的算法计算获得。不同的因子不论原始数据来源是栅格还是矢量，最后都会转化为30×30m的栅格图层。最终形成的滑坡控制因素包括地层岩性、斜坡结构、斜坡形态、坡度、坡向、地形表面纹理、地形粗糙指数、地形曲率、高程、峡谷深度、流域面积、水系距离、流路长度。滑坡影响因素包括植被覆盖度、土地利用、降雨、地震等级。文中采用的滑坡影响因子数据处理清单如表 1所示（模型计算指的是利用ArcGIS模型工具或建模工具计算）。

评价因子栅格化之后如图 4所示。栅格因子图层的水系部分不参与模型的构建，采用掩膜方式从图层中剔除。多个图层叠加形成评价模型的输入特征。实地调查的滑坡面栅格化之后，形成滑坡样本点与非滑坡样本点，图层范围与评价因子图层的范围一致。评价因子图层与滑坡栅格图层在GIS软件中叠加，形成评价因子特征与目标属性的一一对应关系。

|  |  |  |
| --- | --- | --- |
| （a）植被覆盖度 | （b）土地利用 | （c）高程 |
| （d）坡度 | （e）坡向 | （f）斜坡形态 |
| （g）地形表面纹理 | （h）地形粗糙指数 | （i）地形曲率 |
| （j）流域面积 | （k）流路长度 | （l）河流强度指数 |
| (m)水系距离 | （n）工程岩组 | （o） |
| （p）斜坡结构 | （q）地震震级 | （r）降雨 |

图 4 滑坡评价因子图层（a-r）

## 2.4评价单元

常见的滑坡易发性评价单元有栅格单元、网格单元和斜坡单元，不同的划分方式各有优缺点。本文采用栅格评价单元。栅格单元即每个像素点为一个评价单元。每个像素点对应位置的栅格因子图层的像素值为模型的输入特征；滑坡栅格图层的像素值为模型的目标属性，取值为0或1。由于栅格图层不便于输入模型，将栅格因子图层和滑坡栅格图层转化为二维表。转化为二维表的过程如下：（1）建立和研究区范围相同的矢量面，并将其转化为研究区栅格图层；（2）将研究区栅格图层转为矢量点；（3）利用ArcGIS的多值提取至点工具，将和矢量点对应位置的栅格图层中的像素值提取至矢量点的属性表。该属性表即为转化后的二维表。表中的行即为每个栅格单元样本，表的列为不同的滑坡评价因子。

## 2.5类别权重策略

由于整个数据集中非滑坡样本与滑坡样本比值非常高，四种模型采用两种训练策略：①一种是类别平衡策略，即多样本类别（非滑坡）权重低，少样本类别（滑坡）权重高。滑坡与非滑坡的权重比为设置为5：1，模型会倾向于将样本预测为滑坡，避免滑坡漏检。②一种是类别不平衡策略，即不对模型和训练数据集作处理，滑坡和非滑坡的权重相同；模型会倾向于将样本预测为非滑坡，易出现滑坡漏检。

# 3易发性评价

## 3.1实验设置

研究区共有636190个栅格单元，实验采用类别平衡与类别不平衡两种策略。LSTM和DFCNN分别训练200次、LR最大迭代200次、RF不需要设置训练次数。LR、RF利用CPU进行训练，训练过程中使用并行设置。DFCNN、LSTM除使用CPU之外，还会利用GPU加速运算。实验软硬件配置为：AMD 5800 CPU，32GB内存，Gefore RTX3060 12GB GPU，Windows 10 64bit操作系统，Python开发语言。浅层机器学习模型调用scikit-learn 0.24.2库中相关的功能实现，深度学习模型采用Tensorflow 2.6.0中自带的Keras包实现。实验评估不同模型在两种类别权重策略下的性能表现。模型先在训练数据集上进行充分的训练，在测试数据集上进行预测。

LR采用scikit-learn自带的LogisticRegressionCV实现，优化算法参数使用newton-cg。RF采用scikit-learn自带的RandomForestClassifier实现，森林决策树的数目设置为100，其他值保持默认。DFCNN采用Keras实现，中间6个隐藏层，隐藏层节点数目为64，隐层激活函数采用Relu。LSTM采用Keras实现，采用7个LSTM层接1个全连接层，隐层激活函数采用Relu。

## 3.2评价因子变换

主成分分析（Principal Component Analysis，PCA）是一种常见的特征筛选方法，本文通过对滑坡评价因子作PCA分析，分别选择不同数量的PCA分量作为新的滑坡评价输入特征，探索不同主成分分量所包含的信息量和对滑坡易发性评价精度的影响，确定PCA变换在滑坡易发性评价中的适用性。PCA本质上是一种线性变换，它将原始滑坡评价因子转换到其他空间来最大化因子之间的方差，转化后的各分量之间的可区分度增加。研究区的原始评价因子在PCA变换之后的信息量比例和累计信息量比例如图 5所示。

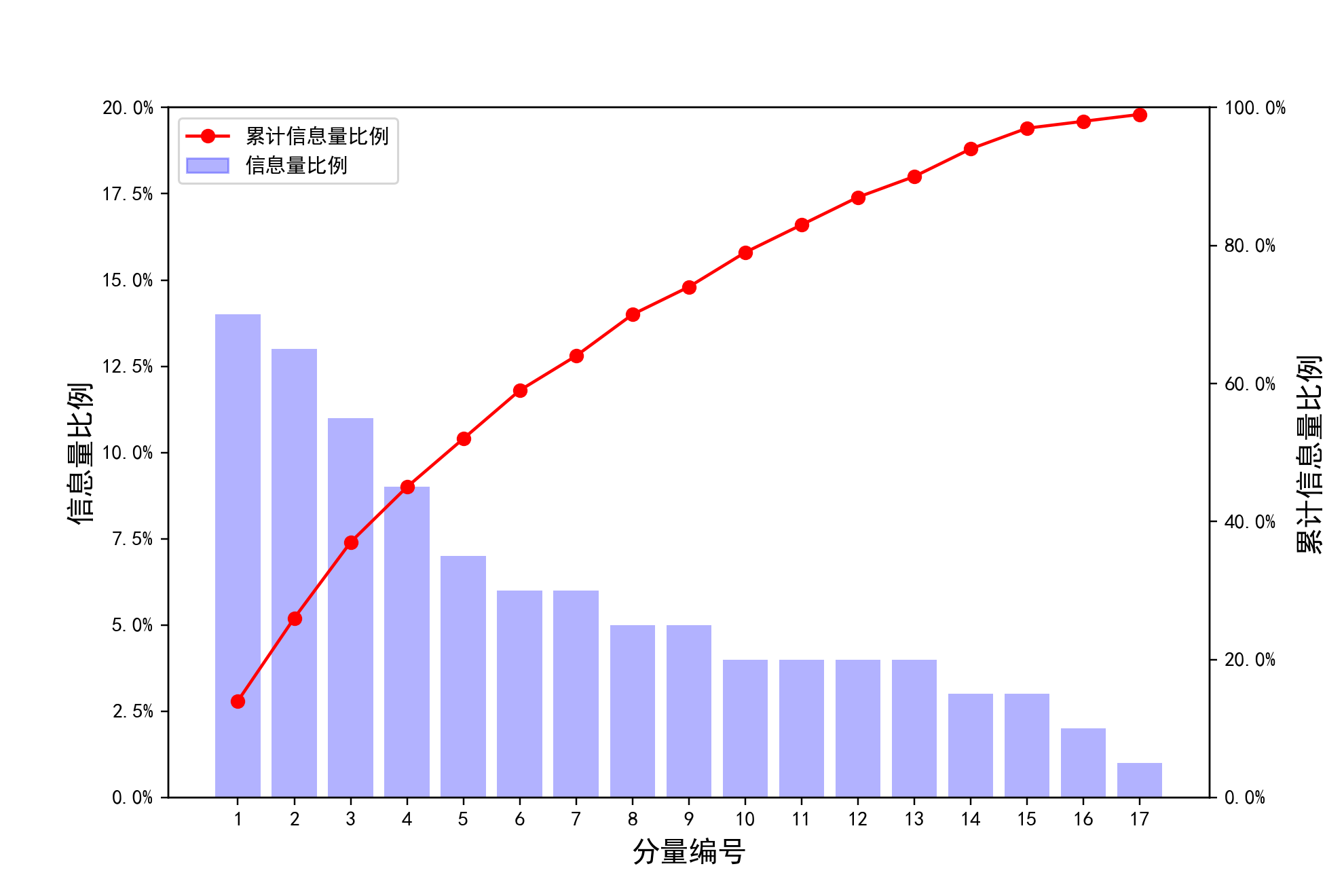


图 5 滑坡评价因子PCA变换后的信息量占比

可以看出PCA变换后的每个分量的信息量占比均不超过15%，前3个分量的累计信息量不超过40%，前13个分量的累计信息量恰好临近90%，前17个分量信息量临近99%。这表明PCA变换后滑坡评价因子的信息量并不集中，比较分散。因此，采用变换后的部分分量作滑坡易发性评价可能会导致精度下降（相对于原始评价因子）。为了验证这一点，分别取PCA变换后的前1~17个分量作为评价模型的评价因子，计算得到四种方法在不同PCA分量时的精度表现，如图 6所示。

|  |  |
| --- | --- |
| （a） | (b) |
| (c) | (d) |
| (e) | (f) |

图 6 不同PCA分量与模型的精度变化

图 6表明，随着模型输入的PCA分量增多，三个精度评价指标均稳步上升。这说明滑坡评价因子在作PCA变换之后，PCA的前若干个分量并不能很好地表达滑坡评价因子的内在特征，即信息量不够，这与图 5的表现是一致的。

在类别不平衡策略下，AUC表现LSTM > RF > DFCNN > LR；F1得分LSTM > DFCNN > RF > LR；Recall表现LSTM > RF > DFCNN > LR。在平衡策略下，AUC表现LSTM > RF=DFCNN > LR；F1得分LSTM > DFCNN > RF > LR；Recall表现RF > LSTM > DFCNN > LR。同时发现LR对类别权重非常敏感，在类别不平衡模式下，LR的F1和Recall精度非常差，而采用类别平衡策略之后，LR的F1和Recall精度迅速上升。综合看来，LSTM模型最优，DFCNN和RF次之，LR再次。

比较类别平衡和类别不平衡策略可以发现：采用平衡策略之后，四种模型的Recall值均有较大幅度上升，说明平衡策略可以将滑坡区评估得更准确；除LR之外，三种模型的F1得分有小幅度下降，这是由于类别平衡策略倾向于将样本预测为滑坡区，代价则是增大非滑坡样本的预测错误率，导致二者的调和平均值（即F1得分）下降。 两种策略下各个模型的AUC值变化不明显。

综上所述，当采用PCA变换时，LSTM、DFCNN、RF在平衡与类别不平衡策略下均有较好的表现；若采用LR作为滑坡易发性评价模型，则需要滑坡与非滑坡样本的数目相当，或者采用类别平衡策略。PCA变换后的滑坡评价因子信息量不集中，是由于各个滑坡评价因子本身是异构数据，评价因子的特征类别和物理意义差别较大，评价因子间本身有较大的区分度，数据冗余不明显。因此，采用更多的PCA分量可以获得更好的滑坡易发性评价精度。

## 3.3原始评价因子与PCA对比

为了进一步探索PCA是否能够筛选出滑坡评价因子中的主要特征，分别采用PCA变换的前17个分量作为（精度较高）和原始评价因子作对比实验。实验结果如表 3所示。

表 2 类别平衡模式精度对比

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **评价指标** | **AUC** | | **F1** | | **Recall** | |
| 原始 | PCA-17 | 原始 | PCA-17 | 原始 | PCA-17 |
| LR | 0.89 | 0.89 | 0.37 | 0.37 | 0.45 | 0.45 |
| RF | 0.99 | 0.97 | 0.62 | 0.45 | 0.99 | 0.94 |
| DFCNN | 0.99 | 0.98 | 0.82 | 0.62 | 0.92 | 0.87 |
| LSTM | 0.99 | 0.99 | 0.81 | 0.82 | 0.89 | 0.89 |

表 3 类别不平衡模式精度对比

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **评价指标** | **AUC** | | **F1** | | **Recall** | |
| 原始 | PCA-17 | 原始 | PCA-17 | 原始 | PCA-17 |
| LR | 0.89 | 0.89 | 0.00 | 0.00 | 0.00 | 0.00 |
| RF | 0.99 | 0.98 | 0.89 | 0.59 | 0.83 | 0.43 |
| DFCNN | 0.99 | 0.97 | 0.83 | 0.66 | 0.86 | 0.64 |
| LSTM | 0.99 | 0.99 | 0.83 | 0.83 | 0.84 | 0.82 |

上面两表中，“原始”指的是采用原始的滑坡评价因子作为模型的特征输入；PCA-17指的是采用PCA变换后的17个分量作为模型的特征输入。LR在原始评价因子和PCA-17下精度相当，说明LR模型的拟合能力有较大的限制，滑坡评价因子信息量到达一定程度后，精度也随之到达上限。RF和DFCNN在原始评价因子下取得的精度显著高于PCA变换因子取得的精度，表明原始的滑坡评价因子包含的信息量更丰富，特征间的差异明显；RF和DFCNN两种模型能够充分地发现和利用原始评价因子中包含的主要信息。LSTM在PCA-17和原始评价因子下取得的精度差异较小，对数据的表现形式要求更宽泛，相对于其他三种模型具有更好的特征提取和分析能力。

对比平衡与不平衡两种策略模式，深度学习模型的精度差异比较小，而浅层机器学习模型的精度差异较大。表明浅层模型（LR和RF）对类别平衡策略较为敏感，而深度模型（DFCNN和LSTM）则对类别平衡策略不敏感。

浅层机器学习方法和深度学习方法在训练模型时消耗的平均时间如下表 4所示。其中s为秒。

表 4 滑坡易发性评价模型训练耗时

|  |  |  |
| --- | --- | --- |
| **模型** | **训练时间（s）** | |
| 平衡策略 | 不平衡策略 |
| LR | 12 | 11 |
| RF | 5 | 14 |
| DFCNN | 418 | 405 |
| LSTM | 765 | 761 |

可以看出，深度学习方法的训练时间是浅层机器学习训练时间的数倍，最高可达153倍。LSTM精度最高，其时间开销也最大。RF作为浅层模型，其时间开销非常小，而且精度较高。

## 4结果

上述实验表明，直接使用原始评价因子作滑坡易发性评价时的精度表现更好。为了更好地查准滑坡区域，最终采用原始评价因子和类别平衡策略对实验区进行滑坡易发性评价。利用自然断点法将模型的评价结果分为四个易发等级，分别是不易发区、低易发区、中易发区、高易发区。四种模型的实验结果分别与栅格评价单元连接，形成可视化图像并与历史滑坡数据比对验证。四种模型的易发性分区评价结果如图 7所示。

|  |  |
| --- | --- |
| LR | RF |
| DFCNN | LSTM |

图 7 四种模型的评价结果（平衡策略）

LR评价结果较差，高易发区与历史滑坡重叠度不高，总体上中易发区和高易发区集中在河流两岸。RF、DFCNN和LSTM三种模型预测的高易发区均集中在河流两岸，与实际相符，较好地预测了高易发区。浅层模型（LR和RF）和深度模型（DFCNN和LSTM）的评价倾向性有所不同。浅层模型评价结果低易发区和中易发区占有一部分比例，而深度模型评价结果中高易发区周边较为纯净，低易发区和中易发区占比非常小。深度模型对高易发性区的评价更加精准，错误预测的面积更小。在实际选择易发性模型时，需要考虑两类模型的适用场景。浅层模型评价结果层次性更加明显，适用于需要分级响应的场景；深度模型的高易发区评价结果更加精准，适用于需要精确调查和定向治理的场景。

## 5.结论

本文以三峡库区秭归-巴东段子区域作为研究对象，根据历史滑坡调查数据和遥感数据、地质资料，选择18个滑坡评价因子，分别使用逻辑回归（LR）、随机森林（RF）、深度全连接神经网络（DFCNN）、长短期记忆神经网络（LSTM）构建滑坡易发性评价模型，并对模型的表现作对比研究。

总体而言，在类别平衡模式下，四种模型的AUC值均在0.8以上，能够作为可信的滑坡易发性评价模型。在类别不平衡模式下，LR的查全率（Recall）低，不适合构建滑坡易发性评价模型。浅层模型中，RF表现最优，兼顾时间开销和精度；深度模型中，LSTM表现最优，虽然时间开销最大，但是能够取得最优异的精度，并且对数据的鲁棒性更好。浅层模型对类别平衡更敏感，深度模型对类别平衡相对不敏感。滑坡不同评价因子之间差异本身较大，对评价因子的数据变换（PCA）不一定能够起到正向作用。浅层模型评价结果具有更丰富的易发区层次，深度模型评价结果能够更好地查准滑坡高易发区。本文的实验结果可为滑坡易发性评价的模型选择提供参考依据。

**References:**

[1]. 陶舒等, 基于信息量与逻辑回归模型的次生滑坡灾害敏感性评价——以汶川县北部为例. 地理研究, 2010. 29(9): 第1594-1605页.

[2]. 赵良军等, 基于二元逻辑回归模型的新疆果子沟滑坡风险区划. 山地学报, 2017. 35(2): 第203-211页.

[3]. 黄发明等, 基于机器学习的滑坡易发性预测建模及其主控因子识别. 地质科技通报, 2021.

[4]. Chen, W., et al., A comparative study of logistic model tree, random forest, and classification and regression tree models for spatial prediction of landslide susceptibility. CATENA, 2017. 151: p. 147-160.

[5]. Dou, J., et al., An integrated artificial neural network model for the landslide susceptibility assessment of Osado Island, Japan. Natural Hazards, 2015. 78(3): p. 1749-1776.

[6]. Chen, W., et al., GIS-based landslide susceptibility mapping using analytical hierarchy process (AHP) and certainty factor (CF) models for the Baozhong region of Baoji City, China. Environmental Earth Sciences, 2016. 75(1).

[7]. Tazik, E., et al., Landslide susceptibility mapping by combining the three methods Fuzzy Logic, Frequency Ratio and Analytical Hierarchy Process in Dozain basin. ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2014. XL-2/W3(2): p. 267-272.

[8]. Yi, Y., et al., Landslide susceptibility mapping using multiscale sampling strategy and convolutional neural network: A case study in Jiuzhaigou region. CATENA, 2020. 195: p. 104851.

[9]. Sameen, M.I., B. Pradhan and S. Lee, Application of convolutional neural networks featuring Bayesian optimization for landslide susceptibility assessment. CATENA, 2020. 186: p. 104249.

[10]. 陈涛等, 利用深度信念网络进行滑坡易发性评价. 武汉大学学报(信息科学版), 2020. 45(11): 第1809-1817页.

[11]. Hastie, T. and R. Tibshirani, The Elements of Statistical Learning: Data Mining, Inference and Prediction. 2009, New York: Springer Science & Business Media.

[12]. Hecht-Nielsen, R., Theory of the back-propagation of neural network, in International Conference on Joint Neural Network. 1989. p. 593-605.