

王锦红, 蒋海昆. 2023. 基于地震观测数据的机器学习地震预测研究综述[J]. 地震研究, 46(2): 173-187, doi:10.20015/j.cnki.ISSN1000-0666.2023.0022.

Wang J H, Jiang H K. 2023. Research progress in field of earthquake prediction by machine learning based on seismic data[J]. *Journal of Seismological Research*, 46(2): 173-187, doi:10.20015/j.cnki.ISSN1000-0666.2023.0022.

# 基于地震观测数据的机器学习地震预测研究综述\*

王锦红<sup>1</sup>, 蒋海昆<sup>2\*</sup>

(1. 中国地震局地震预测研究所, 北京 100036; 2. 中国地震台网中心, 北京 100045)

**摘要:** 机器学习突出的隐式特征提取和复杂任务处理能力正推动着地震预测科学的发展, 为系统了解机器学习技术在地震预测领域的发展现状, 从指定时空窗的地震震级预测、发震位置和发震时间估计三方面, 综述了国内外机器学习在地震预测领域中的应用, 其中在震级预测问题上 AI 应用最为广泛; 总结了机器学习地震预测的主要特征参数、模型和评价相关问题, 从多种评价机制中探索地震活动性参数对地震预测结果的影响, 并对地震预测领域存在的问题进行初步讨论和展望。在可预见的未来, AI 技术的引入和应用领域的拓展, 有可能引领地震预测领域的持续发展。

**关键词:** 地震预测; 机器学习; 特征提取; 模型评价

中图分类号: P315.7 文献标识码: A 文章编号: 1000-0666(2023)02-0173-15  
doi:10.20015/j.cnki.ISSN1000-0666.2023.0022

## 0 引言

地震预测是世界性的科学难题。在庆祝 Science 创刊 125 周年之际, 该刊杂志社公布了 125 个最具挑战性的科学问题 (Kennedy, Norman, 2005), 其中第 55 个问题明确提出: 是否存在有助于预报地震的前兆? 地震预测困难之处在于一方面涉及地震前兆的地球物理观测基本属于远离震源地方的地表观测, 所观测到的物理量除包含可能的地震孕育信息外, 更多的是与环境、路径、干扰等有关的信息; 另一方面是研究人员对地震孕震机理尚无清晰的认识, 即地震预测的理论问题远未解决。因此, 目前的地震预测基本停留在基于观测资料的现象类比和统计分析层面。

近几十年来, 人工智能 (Artificial Intelligence, 简称 AI) 技术在计算机视觉和图像识别、时间序

列分析和异常检测、生物医学等许多行业应用的成功经验表明, 在理论问题尚未解决之前, AI 技术对识别未知特性、发现系统隐含规律、解决实际问题等方面作用巨大, 显著推动了相关行业的发展。AI 是用于模拟和扩展人类智能的理论、方法和技术, 开发相应应用系统的一门技术科学, 而机器学习是 AI 的最重要组成部分, 是一门研究如何实现 AI 的技术性学科, 涉及统计理论、系统辨识、逼近理论、神经网络、优化理论、计算机科学、脑科学等领域。即机器学习是从历史数据中学习特征和规律, 然后将其应用到未知事物探索中的一门科学 (中国电子技术标准化研究院, 2018)。2006 年“深度神经网络”概念被提出 (Hinton, Salakhutdinov, 2006), 该方法具有能够在超大数据集中发现隐藏模式的性能, 使其成为一种机器学习技术的新方法。

由于地震孕育的地球物理过程存在非线性

\* 收稿日期: 2022-08-22.

基金项目: 地震动力学国家重点实验室开放基金 (LED2022B05)。

第一作者简介: 王锦红 (1999-), 硕士研究生在读, 主要从事人工智能强余震预测研究。  
E-mail: 17866618823@163.com.

✦ 通讯作者简介: 蒋海昆 (1964-), 研究员, 主要从事余震统计、余震机理及余震预测研究。  
E-mail: jianghaikun@seis.ac.cn.

复杂性等问题，任何可记录的物理参数与未来地震的时间、震级或空间位置之间缺乏确定的数学或经验关系 (Panakkat, Adeli, 2007)。但随着 AI 技术的发展，基于已有地震观测数据的 AI 技术在地震预测领域得到越来越广泛的应用。学习大量观测数据可以发现其隐含的特征或规律，一方面有利于深化对地震机理的认识，更重要的是可以利用这些数据建立/训练模型，对未来数据进行拟合进而开展地震预测，这可以在机理和内在规律尚不清楚的情况下改善地震预测的准确性。

本文介绍了机器学习预测地震的一般流程，并从数据特征、模型算法和评价指标等方面，简要综述机器学习在地震预测领域的最新研究进展。

## 1 机器学习地震预测简介

### 1.1 机器学习

根据学习模式，机器学习主要分为有监督学习、无监督学习以及强化学习 (图 1)。有监督学习是指利用已标记的有限训练数据集，通过某种学习方法建立模型，实现对新输入数据的标记，要求训练样本的分类标签已知，分类标签精确度越高，样本越具有代表性，学习模型的准确度越高。最典型的监督学习算法包括回归 (标签为连续变量) 和分类 (标签为分类变量)。无监督学习是利用无标记的有限数据描述隐藏在未标记数据

中的结构/规律，其典型算法包括降维、聚类等。强化学习是通过交互和反馈指导学习器进行决策，不同于以上两种学习方式，强化学习训练需要环境给予反馈，基于这种反馈不断强化学习器性能以完成特定的目的，其在机器人控制、自动驾驶、游戏等领域已有许多成功应用。

常见的机器学习算法还包括半监督学习、迁移学习等，其中半监督学习介于无监督学习和监督学习之间。经常遇到的实际情况是，部分数据有标签，部分数据无标签，无标签的数据数量远远大于有标签的数据。未标记的样本虽未直接包含标记信息，但若它们与完全标记的样本是从相同数据源独立分布采样得来，则它们所包含的关于数据分布的信息对建立学习模型将大有裨益，这也是半监督学习日益受到重视的重要原因。迁移学习是指当在某些领域无法取得足够多的数据进行模型训练时，利用另一领域数据获得的关系进行学习，可以把已训练好的模型参数迁移到新的模型上来指导新模型的训练，从而更有效地学习底层规则、减少数据量 (中国电子技术标准化研究院, 2018)。目前地震预测领域的机器学习基本上使用有监督学习类算法。

### 1.2 机器学习地震预测的一般流程

机器学习地震预测主要基于地震观测数据某些特征来进行，如地震活动性参数、地震前兆、地震波以及卫星监测数据等 (Asencio - Cortés *et al*,

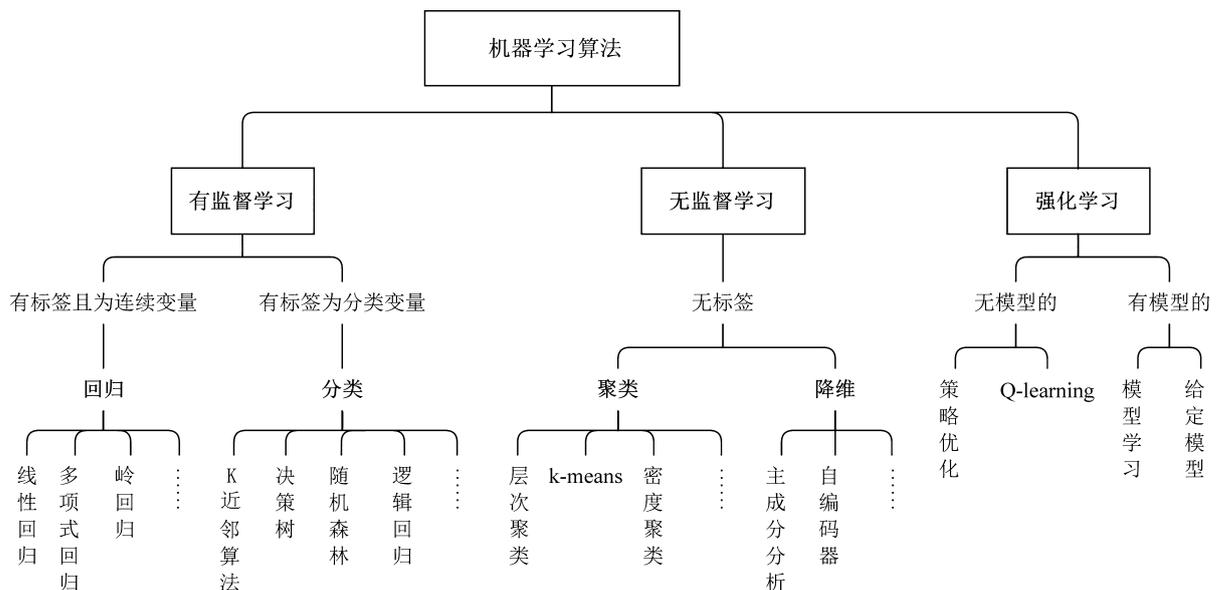


图 1 机器学习算法分类

Fig. 1 Classification of machine learning algorithms

2016; Asim *et al.*, 2016; Niksarlioglu, Kulahci, 2013; Wang *et al.*, 2022)。通过数据预处理, 将其转换为适合分类或回归算法的形式, 找到数据中隐藏的模式, 以达到预测地震强度、发生时间和可能地点的目的。

目前的机器学习地震预测模型大多为降维预测, 即大多只对未来地震的发震时间、地点和震级三要素进行单要素或其中两个要素进行预测。在实际应

用中, 还需对发震时间或震级进一步进行离散化或进行空间网格化, 从而使预测过程简化为分类或回归问题。对于二元分类 (如预测震级是否高于或低于阈值) 或回归 (如可能的震级估计), 训练和预测的输出结果通常表达为概率形式。一般的机器学习地震预测模式和流程如图 2 所示, 包括数据集建立、数据预处理、分类/回归模型构建、外推预测、结果或模型评估等 5 个环节。

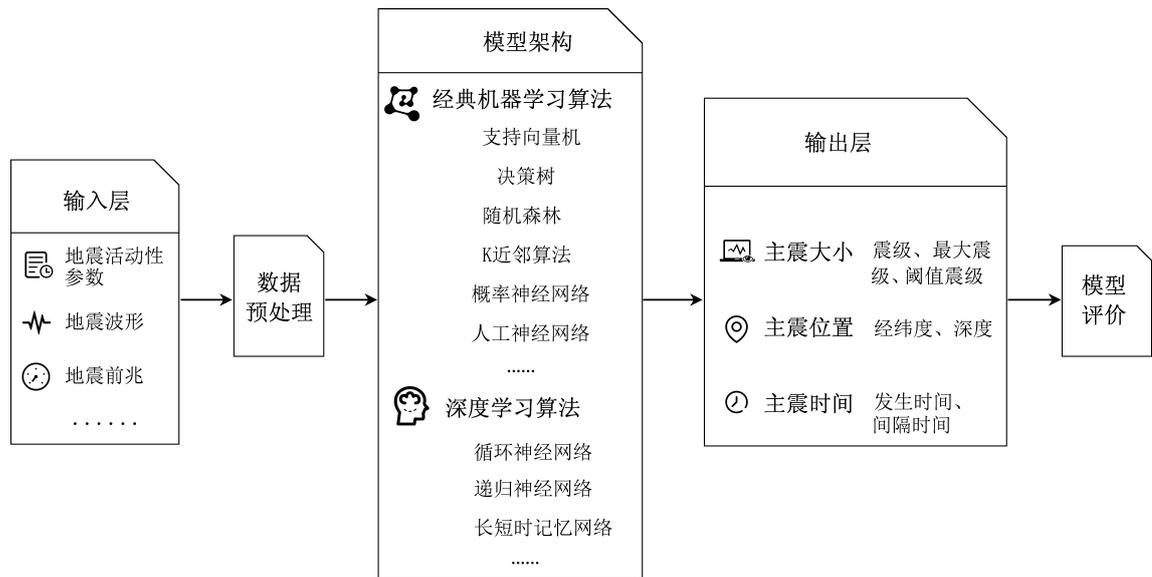


图 2 机器学习地震预测的基本环节

Fig. 2 The data flow of machine learning for earthquake forecasting

## 2 主要特征参数

在目前已开展的研究中, 以地震目录为主的地震观测数据, 由于其具有易获取、较好的完备性、无干扰、较长时期的数据积累等特性, 成为机器学习地震预测最为常用的输入特征数据。通过简单的预处理可将地震目录转换成适合模型输入要求的形式, 作为机器学习的输入来训练模型和开展预测, 如直接将地震目录参数进行简单转换, 将地震发生时间转换为序号 (Zhang *et al.*, 2019)、将地震位置 (经、纬度和深度) 转换为预先划分好的网格 (Wang *et al.*, 2017; Kail *et al.*, 2021) 等。

Panakkat 和 Adeli (2007) 基于地震统计参数及地震学中的基本统计特征, 依据 G-R 关系和特征地震两个先验统计模型, 将地震目录转换为 8 个地震活动性参数作为特征参数构建机器学习

地震预测模型。8 个预测特征参数中 (表 1 中特征 1~8),  $T$  为地震序列时间分布,  $M_{\text{mean}}$  为地震的平均活动强度,  $dE^{1/2}$  为地震应变释放的快慢,  $\eta$  和  $\Delta M$  为 G-R 关系  $b$  值及数据总体对 G-R 关系的偏离,  $\mu$  和  $c$  为特征地震的平均复发周期及其离散程度,  $b$  值被用来描述研究区域的地震震级-频度特征与地震活动水平, 是地震预测的一个常用指标参数。Reyes 等 (2013) 基于 ANN 模型预测智利和伊比利亚半岛的地震, 使用了 7 个特征参数 (表 1 特征 9~15), 其中 5 个为等地震个数计算的 G-R 关系  $b$  值及等地震序号间隔的  $b$  值增量, 目的是间接考察  $b$  值随时间的变化; 考虑最近的地震活动水平及预测区 6 级地震的背景危险性, 将研究区近 1 周以来的最大地震震级及发生 6 级地震的背景概率也作为特征使用。Martínez - Álvarez 等 (2013) 把上述所有特征合并用于机器学习建模。

表 1 用于机器学习的地震特征参数  
Tab. 1 Seismic characteristic parameters used for machine learning

特征序号	特征	物理含义	计算公式	参考文献
1	$T$	最近 $n$ 个地震事件的时间跨度	$T = t_n - t_1$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez - Alvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Asim <i>et al.</i> , 2016, 2017a, 2018a, b, 2020
2	$M_{\text{mean}}$	最近 $n$ 个地震事件的平均震级	$M_{\text{mean}} = \frac{\sum M_i}{n}$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez - Alvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Asim <i>et al.</i> , 2016, 2017a, 2018a, b, 2020
3	$dE^{1/2}$	地震应变释放速率	$dE^{1/2} = \frac{\sum dE^{1/2}}{T}$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez - Alvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Asim <i>et al.</i> , 2017a, 2018a, b
4	$\eta$	均方偏差	$\eta = \frac{\sum [\log_{10} N_i - (a - bM_i)]^2}{n - 1}$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez - Alvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Asim <i>et al.</i> , 2017a, 2018b, 2020
5	$\Delta M$	实际最大震级与预期最大震级之差	$\Delta M = M_{\text{max, observed}} - M_{\text{max, expected}}$ 其中 $M_{\text{max, expected}} = a/b$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez - Alvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Asim <i>et al.</i> , 2017a, 2020
6	$\mu$	特征地震平均复发时间	$\mu = \sum \frac{t_i^{\text{characteristic}}}{n^{\text{characteristic}}}$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez - Alvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Asim <i>et al.</i> , 2017a
7	$c$	平均复发时间变异系数	$c = \frac{STD \text{ of the observed times}}{\mu}$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez - Alvarez <i>et al.</i> , 2013; Asim <i>et al.</i> , 2017a; Asencio - Cortés <i>et al.</i> , 2018
8	$b$	G - R 关系中的 $b$ 值	$b_i = \frac{\log(\epsilon)}{(1/50) \sum_{j=0}^{49} M_{i-j} - 3}$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez - Alvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Shodiq <i>et al.</i> , 2017, 2018; Asim <i>et al.</i> , 2017a, 2018a, b, 2020
9	$b_1$		$x_{1i} = \Delta b_{1i} = b_i - b_{i-4}$	Reyes <i>et al.</i> , 2013; Martinez - Alvarez <i>et al.</i> , 2013; Morales - Esteban <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Shodiq <i>et al.</i> , 2017, 2018
10	$b_2$	G - R 关系中的 $b$ 值增量	$x_{2i} = \Delta b_{2i} = b_{i-4} - b_{i-8}$	Reyes <i>et al.</i> , 2013; Martinez - Alvarez <i>et al.</i> , 2013; Morales - Esteban <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Shodiq <i>et al.</i> , 2017, 2018
11	$b_3$		$x_{3i} = \Delta b_{3i} = b_{i-8} - b_{i-12}$	Reyes <i>et al.</i> , 2013; Martinez - Alvarez <i>et al.</i> , 2013; Morales - Esteban <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016, 2018; Shodiq <i>et al.</i> , 2017, 2018

续表 1

特征序号	特征	物理含义	计算公式	参考文献
12	$b_4$		$x_{4i} = \Delta b_{4i} = b_{i-12} - b_{i-16}$	Reyes <i>et al.</i> , 2013; Martinez – Alvarez <i>et al.</i> , 2013; Morales – Esteban <i>et al.</i> , 2013; Asencio – Cortés <i>et al.</i> , 2016, 2018; Shodiq <i>et al.</i> , 2017, 2018
G – R 关系中的 $b$ 值增量				
13	$b_5$		$x_{5i} = \Delta b_{5i} = b_{i-16} - b_{i-20}$	Reyes <i>et al.</i> , 2013; Martinez – Alvarez <i>et al.</i> , 2013; Morales – Esteban <i>et al.</i> , 2013; Asencio – Cortés <i>et al.</i> , 2016, 2018; Shodiq <i>et al.</i> , 2017, 2018
14	$P(M_s \geq 6.0)$	震级大于或等于 6.0 级地震的概率	$P(M_s \geq 6.0) = e^{-3b_i/\log(c)} = 10^{-3b_i}$	Martinez – Alvarez <i>et al.</i> , 2013; Morales – Esteban <i>et al.</i> , 2013; Asencio – Cortés <i>et al.</i> , 2016, 2018; Shodiq <i>et al.</i> , 2017, 2018; Asim <i>et al.</i> , 2018b, 2020
15	$M_{\max}$	一周内的最大震级	$M_{\max} = \text{Max} \{M_s\}, t \in [-7, 0)$	Panakkat, Adeli, 2007; Adeli, Panakkat, 2009; Martinez – Alvarez <i>et al.</i> , 2013; Asencio – Cortés <i>et al.</i> , 2016, 2018; Asim <i>et al.</i> , 2017a, 2020
16	$T_r$	概率复发时间	$T_r = \frac{T}{10^{a-bM}}$	Asim <i>et al.</i> , 2016, 2018b, 2020
17	$a$	G – R 关系中的 $a$ 值	$a = \frac{\sum (\log_{10} N_i + bM_i)}{n}$	Martinez – Alvarez <i>et al.</i> , 2013; Asencio – Cortés <i>et al.</i> , 2016; Asim <i>et al.</i> , 2018a, b, 2020
18	$z$	地震率变化	$z = \frac{R_1 - R_2}{\frac{S_1 - S_2}{n_1 - n_2}}$ 式中: $R_{1,2}$ 表示平均地震活动率; $S_{1,2}$ 表示地震活动率的标准偏差	Asim <i>et al.</i> , 2016, 2018a, b, 2020
19	$\beta$	地震率变化	$\beta = \frac{M(t, \delta) - n\delta}{n\delta(1 - \delta)}$	Asim <i>et al.</i> , 2016, 2018a, b, 2020
20	$\sigma_b$	$b$ 值的标准偏差	$\sigma_b = 2.3\delta^2 \sqrt{\frac{\sum_{i=1}^n (M_i - \bar{M})^2}{n(n-1)}}$	Asim <i>et al.</i> , 2016, 2018b, 2020

之后,以这些指标为基础又演化出更多的特征参数,如 Asencio - Cortés 等 (2016) 考察了 19 个地震活动性相关参数在机器学习地震预测中的灵敏性,所使用的预测指标主要取自 Panakkat 和 Adeli (2007)、Reyes 等 (2013) 确定的地震参数,增加了不同震级间隔范围内地震计算所得的  $G-R$  关系  $a$  值、基于概率模型的复发时间  $T_r$  等参数及其组合 (表 1 中特征 16、17)。Asim 等 (2018a) 利用机器学习对东都库什、智利和南加州地区进行地震预测,把基于目录的预测指标全部汇集在一起,利用最小二乘和最大似然两种方法计算特征,将地震相关的特征参数扩充到 60 个,包括  $G-R$  关系系数、地震发生率、前震频次、地震应变释放等参数,基于最大相关和最小冗余 (mRMR) 准则来确定具体的人选特征。需要指出的是,有些特征的应用不但于预测效率提升无益,甚至可能会降低预测效果。因此,一些学者将特征选择技术应用于地震预测,发现  $G-R$  关系相关参数以及描述地震频次变化、发生时间、地震应变释放速率等参数,对机器学习地震预测具有较大的贡献度 (Martínez - Álvarez *et al.*, 2013; Asim *et al.*, 2018a), 地震率变化  $z$ 、 $\beta$  (Asim *et al.*, 2016, 2020, 2018a, b),  $b$  值的标准偏差  $\sigma_b$  (Asim *et al.*, 2016, 2018b, 2020) 等指标也被用于机器学习地震预测模型训练,具体见表 1 的特征 18 ~ 20。

除基于地震目录构建的各类特征外,在地震预测中还用到一些其它类型的数据,如静态库仑应力变化 (DeVries *et al.*, 2018)、地脉动 (马士振等, 2020)、氡浓度 (Külahcı *et al.*, 2009; Niksarlioglu, Külahcı, 2013)、地电信号 (Suratgar *et al.*, 2008; Moustra *et al.*, 2011)、强度函数 (Nicolis *et al.*, 2021)、模型表面变形特征 (Corbi *et al.*, 2019) 等。除以上使用的地震特征外,还有研究通过深度学习方法从数据集中提取特征,再与传统地震特征相融合来开展地震预测 (Li *et al.*, 2020)。

### 3 模型评价

AI 领域存在较多的外推预测算法,与之对应的评价机制对于模型的训练和优化显得尤为重要。地震预测模型评估指标中,最简单的方法是通过混淆矩阵来进行评测,这也是机器学习中最常见的模型评估方法,主要用于地震预测中诸如地震

事件是否发生的二分类问题,其中混淆矩阵的 4 个指标 (表 2) 描述了所有可能预测的结果。True Positive (简称 TP): 预测有震且实际有震的数量,即有震报准的数量。False Positive (简称 FP): 预测有震但实际无震的数量,即虚报的数量。True Negative (简称 TN): 预测无震且实际无震的数量,即无震报准的数量。False Negative (简称 FN): 预测无震但实际有震的数量,即漏报的数量。

表 2 地震预测评估模型中的混淆矩阵

Tab. 2 Confusion matrix in earthquake prediction evaluation model

混淆矩阵		实际情况	
		有震	无震
预测结果	有震	TP	FP
	无震	FN	TN

混淆矩阵在机器学习地震预测评价领域得到较多的应用,如 Gulia 和 Wiemer (2019) 用混淆矩阵评估了“红绿灯”形式的分类器在地震发生过程中对地震序列属于前震还是余震进行实时识别的能力; Mignan 和 Broccardo (2020) 利用混淆矩阵分析了 ANN 模型与其它机器学习简易模型在预测增益上的不同。但总的来看,由于混淆矩阵统计的只是地震的次数,单一的统计值很难全面衡量模型的优劣,因此在混淆矩阵的基本统计结果上又延伸出许多检验指标,见表 3。与混淆矩阵有关的模型评估曲线可以更直观地显示预测效能评价结果,例如 Precision - Recall (简称 P - R) 曲线 (Kail *et al.*, 2021)、Receiver Operating Characteristic (简称 ROC) 曲线 (Kanarachos *et al.*, 2017)、Area Under Curve (简称 AUC) 曲线 (DeVries *et al.*, 2018; Aslam *et al.*, 2021) 等,这些曲线可直接用于二分类问题,但无法直接应用于多分类问题。在一些地震预测回归算法中,还使用了平均绝对误差 Mean Absolute Error (简称 MAE) (Alarifi *et al.*, 2012; Nicolis *et al.*, 2021)、均方误差 Mean Squared Error (简称 MSE) (Alarifi *et al.*, 2012; Nicolis *et al.*, 2021)、均方根误差 Root Mean Squared Error (简称 RMSE) (Hu *et al.*, 2012)、相对误差 Relative Error (简称 RE) (Asencio - Cortés *et al.*, 2018) 等指标来表征预测值和观测值之间的差异。由于不同地震预测模型的要求不同,需根据具体的预测模型选择合适的评价指标,才能得到有效的评估结果,进而对模型进行反馈和优化。

表 3 地震预测评估模型中的评估指标  
Tab. 3 Evaluation indicators in earthquake prediction assessment models

序号	评价指标	意义	公式	参考文献
1	准确率 (Accuracy)	预测与实际相一致的比率	$Acc = \frac{TP + TN}{TP + TN + FP + FN}$	Asim <i>et al.</i> , 2017b, 2018b, 2020; Rouet - Leduc <i>et al.</i> , 2020; Aslam <i>et al.</i> , 2021
2	有震报准率、查准率 / 精确度 (Precision) 检测概率 (POD)	预测有震中实际有震所占的比例	$P_1 = \frac{TP}{TP + FP}$	Reyes <i>et al.</i> , 2013; Martínez - Álvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016; Asim <i>et al.</i> , 2017b, 2018b, 2020; Shodiq <i>et al.</i> , 2015, 2018; Aslam <i>et al.</i> , 2021
3	召回率 / 查全率 (Recall) 敏感度 (Sensitivity)	实际有震的前提下, 预测有震所占的比例	$S_n = \frac{TP}{TP + FN}$	Reyes <i>et al.</i> , 2013; Martínez - Álvarez <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016; Asim <i>et al.</i> , 2017b, 2018b, 2020; Shodiq <i>et al.</i> , 2015, 2018; Aslam <i>et al.</i> , 2021
4	$F_1$	精确度和敏感度的调和平均值	$F_1 - Score_1 = \frac{2 \times precision \times recall}{precision + recall} = \frac{2 \times TP}{2 \times TP + FP + FN}$	Aslam <i>et al.</i> , 2021
5	无震报准率	预测无震中实际无震所占的比例	$P_0 = \frac{TN}{TN + FN}$	Adeli, Panakkat, 2009; Martínez - Álvarez <i>et al.</i> , 2013; Reyes <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016; Asim <i>et al.</i> , 2017b, 2018b, 2020; Shodiq <i>et al.</i> , 2015, 2018; Aslam <i>et al.</i> , 2021
6	特异度 (Specificity)	实际无震的前提下, 预测无震所占的比例	$S_p = \frac{TN}{TN + FP}$	Adelil, Panakkat, 2009; Martínez - Álvarez <i>et al.</i> , 2013; Reyes <i>et al.</i> , 2013; Asencio - Cortés <i>et al.</i> , 2016; Asim <i>et al.</i> , 2017b, 2018b, 2020; Shodiq <i>et al.</i> , 2015, 2018; Aslam <i>et al.</i> , 2021
7	虚警率 / 误报率 (FAR)	预测有震中实际无震所占的比例	$FAR = 1 - P_1 = \frac{FP}{TP + FP}$	Panakkat, Adeil, 2007; Adeil, Panakkat, 2009
8	频率偏差 (FB)	—	$FB = \frac{TP + FP}{TP + FN}$	Panakkat, Adeil, 2007
9	R - 评分	—	$R = S_n - FAR = \frac{TP}{TP + FN} - \frac{FP}{TN + FP}$	Panakkat, Adeil, 2007; Adeil, Panakkat, 2009; Huang <i>et al.</i> , 2018; Asim <i>et al.</i> , 2018b, 2020; Aslam <i>et al.</i> , 2021
10	马修斯相关系数 (MCC)	实际结果与预测结果之间的相关系数	$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$	Asim <i>et al.</i> , 2018b, 2020; Aslam <i>et al.</i> , 2021

## 4 主要算法及预测效果

机器学习思想在地震预测中的应用最早起源于我国 20 世纪 80 年代, 如石绍先和范杨 (1990), 石绍先和沈斌 (1986) 研究了多级判别、聚类分析等在地震预测中的应用, 后续这种思想延伸到了神经网络中。由于神经网络在处理复杂的非线性问题时具有明显的优点, 不需要对象有精确的数学模型, 可通过神经网络结构的可变性, 逐步适应外部环境的各种因素作用 (王炜等, 2000a), 许多学者将其应用于地震综合预测 (蔡煜东等, 1993; 李荣峰, 2000; 王炜等, 2000; 王炜, 吴耿峰, 2000, 2006a, b; 陈一超等, 2008; 韩晓飞等, 2012)。蔡煜东等 (1993) 运用神经网络对我国西南地区 17 个震例进行分析, 建立了地震综合预报专家系统, 对确定地区、确定时间范围内的地震强度进行预测; 李荣峰 (2000) 建立人工神经网络模型预测福建及其周边地区的年度最大地震震级, 该模型的输入值为 4 个地震参数 ( $b$  值、 $M_L \geq 2.5$  地震的频次、每年地震释放总能量  $\sum E$  和空间集中度  $C$  值), 兼顾地震三要素, 预测准确率可达 90%; 王炜等 (2000b) 进一步探索 Back Propagation (简称 BP) 神经网络在地震中期预报中的应用, 结果表明中强地震前 1~3 a, 未来震中周围一般都开始出现明显的异常区, BP 神经网络显示出一定的中期预报效果。但由于 BP 神经网络存在收敛速度慢、易陷入局部最小点等不足, 陈一超等 (2008) 提出利用具有全局搜索能力的遗传算法, 建立遗传算法和神经网络相结合的震级预测模型, 震级预测误差小于传统 BP 神经网络; 韩晓飞等 (2012) 提出基于遗传算法结合广义回归神经网络的地震预测算法, 震级预测误差也小于传统的 BP 神经网络算法。

尽管国内开展了许多探索性研究, 但总的来看国外在人工智能地震技术研究方面远远走在前列。地震预测的重点是震级、时间、地点三要素, 但目前大多数研究主要关注震级预测问题, 即在固定区域、固定时间窗前提下, 评估某级以上地震的发震可能性, 即大多数研究实际上是将地震预测问题进行降维处理, 转化为在固定区域、固定时间段内“有”或“无”某级以上地震发生的分类问题。

### 4.1 指定时空窗的地震震级预测

#### 4.1.1 神经网络

人工神经网络 (Artificial Neural Network, 简称 ANN) 较早被应用于地震预测研究。Alves (2006) 针对地震时间序列的混沌特性, 借鉴“金融振荡器 (financial oscillators)”工具构建地震相关数据集, 利用 ANN 对 1998 年 7 月和 2004 年 1 月 2 次 5 级以上地震进行预测, 虽然预测时空、窗口较大, 但在一定程度上证明了 ANN 模型的地震预测潜力。为进一步探索 ANN 的地震预测效能, Reyes 等 (2013) 采用 ANN 模型对指定区域未来 5 天内地震趋势进行预测, 得到 ANN 预测效能受区域差异影响较大。而对于相同区域的预测, 许多研究显示 ANN 似乎优于决策树、支持向量机、朴素贝叶斯、K-最近邻、ANFIS 等分类器 (Reyes *et al*, 2013; Morales - Esteban *et al*, 2013; Pandit, Pandal, 2021)。也有研究指出, 由于地震目录的结构化和表格化性质及有限的特征数量, 比 ANN 更简单、更透明的机器学习模型似乎更为可取 (Mignan, Broccardo, 2019)。

Panakkat 和 Adeli (2007) 采用前馈反向传播神经网络 Levenberg - Marquardt (简称 LMBP)、递归神经网络 Recurrent Neural Network (简称 RNN) 和径向基函数神经网络 Radial Basis Function (简称 RBF) 3 种模型预测南加州和旧金山地区 1991 年 1 月至 2005 年 9 月后续单位时间 (月) 的地震强度, 模型输入即表 1 中 8 个地震相关特征参数, 并使用 4 种不同的统计指标 (检测概率 POD、虚警率 FAR、频率偏差 FB 和 R 分数) 评估模型的预测精度, 结果表明与 LMBP 和 RBF 网络相比, RNN 具有更高的预测精度, 尤其在中强度地震 ( $5.5 \leq M < 6.5$ ) 预测时, 不同模型之间的差距最明显; Adeli 和 Panakkat (2009) 使用特征作为概率神经网络 (Probabilistic Neural Network, 简称 PNN) 的输入来预测相同地区的地震, 结果显示 PNN 对  $M \leq 6.0$  地震的预测结果比 RNN 更好。上述 8 个地震参数还被用于预测兴都库什和巴基斯坦北部的地震, 在兴都库什地区, 线性规划增强集成分类器 (Linear Programming Boosting, 简称 LPBoost) 在预测灵敏度方面比较好的结果, 而神经网络则倾向于产生更低的误报率 (Asim *et al*, 2016); 在巴基斯坦北部地区, 前馈神经网络则表现更好, 其准确率可达 75% (Asim *et al*, 2017b)。从上述

研究可见，即使用相同的地震活动性特征和相同的机器学习算法，在不同地区也会产生不同的预测效果，可能是由每个区域构造及地震活动特点存在差异所致。尽管机器学习提供了特征或模型选取算法，但主要为针对特定数据的模型与数据之间吻合程度的一系列统计学指标，不同特点的地震活动区域适用哪些特征和算法并没有明确的认识，还是只能使用不停的尝试、检验、调参等方法来进行选择和优化。

随观测数据的急剧增加，ANN 模型不断复杂化，深度神经网络 Deep Neural Networks (简称 DNN)、RNN 等学习模型也不断推出。Li 等 (2020) 提出一种可以有效融合显性和隐性特征的深度学习模型 A Deep Learning Model for Earthquake Prediction (简称 DLEP)，在 8 个具有不同特征的典型地震带上进行测试，认为 DLEP 在地震震级预测方面具有更好的泛化能力和更高的准确性。RNN 的优势在于对序列数据具有记忆，能够更好地反映地震活动的时间依赖特性，因而也被应用于地震预测研究。长短期记忆网络 Long Short Term Memory (简称 LSTM) 作为一种特殊的 RNN 能够更好地解决长序列训练过程中的梯度消失和梯度爆炸问题，因而被广泛应用，如 Wang 等 (2017) 从地震活动空间相关的角度出发，利用 LSTM 学习不同位置地震之间的时空关系，即使在地震监测能力低的地区，仍然能够使用该方法判断是否有地震发生；Kail 等 (2021) 同样基于 LSTM，研究预测震级高于给定阈值的地震是否会从选定时刻开始的 10 ~ 50 天在给定区域发生，并进而解释日本地区历史地震之间的时间依赖特性。由于卷积神经网络 (Convolutional Neural Network, 简称 CNN) 和 RNN 等震相自动识别技术的出现，基于深度学习的地震监测技术快速发展 (Perol *et al*, 2018; Mousavi, Beroza, 2020)，使得地震监测台网可检测出更多的地震事件，其空间分辨率和震级完备性也显著提高，这不但可以改进传统的统计和基于物理的地震预测方法，对基于 AI 的预测方法、特征构建以及模型训练也大有裨益 (Mousavi, Gregory, 2022)。

#### 4.1.2 多种机器学习模型集成

同一种数据、多种机器学习模型的联合使用和相互印证可能是当前可以考虑的一种途径，如 Asim 等 (2018b) 将地震特征与基于遗传规划 Ge-

netic Programming (简称 GP) 和自适应增强 Adaptive Boosting (简称 AdaBoost) 集成分类方法相结合，定义了一个地震预测系统 (Enhanced Particle Swarm Optimization, 简称 EPSO)，用于预测未来 15 天  $M \geq 5.0$  地震发生的可能性，对兴都库什、智利和南加州地区进行的地震预测，预测精度  $P_1$  分别为 74%、80% 和 84%，与 Asim 等 (2017b) 预测结果相比，提高了约 15%；在巴基斯坦北部也印证了 EPSO 系统的低误报率 (Aslam *et al*, 2021)。Asim 等 (2018a) 利用基于支持向量回归器 (Support Vector Regression, 简称 SVR) 和混合神经网络 (Hybrid neural network, 简称 HNN) 的混合系统再次对以上 3 个地区进行地震预测检验，与灵敏度高达 90% 但精度  $P_1$  较小的 SVM - HNN 预测模型相比较，该分类预测模型灵敏度相对低但精度  $P_1$  较高，从实际应用的角度更为可取 (Asim *et al*, 2018b)。Asencio - Cortés 等 (2018) 利用 1970—2017 年南加州地区约 140 万多条地震目录数据，采用 16 种地震特征参数，探索了 4 种回归算法 (广义线性模型、梯度推进机、深度学习和随机森林) 与集成学习相结合用于时间窗为 7 天的中强地震震级预测尝试，不同算法预测震级与实际震级的对比结果显示非常离散。史翔宇 (2021) 采用相同的 16 个地震活动性特征参数，对上述 4 种模型对预测结果的单独贡献进行分析，发现 G - R 类特征参数贡献度较大，地震能量类特征参数次之，不同模型和特征参数在不同地震区 (带) 的地震预测贡献度有较大的差异。

#### 4.1.3 聚类和地震活动图像识别

纯粹的分类/聚类算法也被用于地震预测研究，如 Florido 等 (2015) 将原始数据转换成包含地震活动性参数  $b$  值的数据集，应用聚类算法对转换后的数据集进行离散化，构建地震活动的参数“序列”，将指定区域  $M \geq 4.4$  地震作为标签，对  $M \geq 4.4$  地震前的所有“序列”进行搜索，找出可能的地震前兆“序列”用于地震预测，在指定的评价语境下，其结果的准确率接近 70%，肯定了所提出方法的有效性；Shodiq 等 (2015) 提出一种空间分析和自动聚类技术来判断后续是否有大地震发生，使用山谷追踪技术确定最佳分类数为 6 类，在质心  $K$  均值优于其它  $K$  均值的情况下进行分层聚类，其结果显示对印尼地区未来 1 ~ 6 年的地震预测具有较高的准确性；Shodiq 等 (2017)

进一步将自动聚类 and 人工神经网络相结合, 预测印尼地区发生 5.1 级地震后 5 天内发生 5.5 级或 6 级以上地震的可能性, 结果表明该模型能较好地预测 6 级以上地震,  $P_0$ 、 $P_1$ 、 $S_N$ 、 $S_p$  四项指标平均值达到 75%, 其中  $P_0$  高达 97%; Shodiq 等 (2018) 使用同样的方法对该区域的余震进行预测, 发现对  $M \geq 6$  地震的余震具有较好的预测效果, 特异度  $S_p$  为 99.16%。

与计算机视觉类似的地震活动图像识别也被用于地震预测, 如 Huang 等 (2018) 提出一种基于图深度学习的机器学习地震预测方法, 将中国台湾地震活动图转换为图像, 共产生 65 536 个特征 (对应  $256 \times 256$  像素图像), 利用 CNN 从标有地震信息的地理图像中提取隐含特征, 用过去 120 天的地震事件预测未来 30 天的地震。如果未来 30 天的最大地震震级大于等于 6 级, 则将该图像标记为 1, 否则标记为 0。该 CNN 模型的 R 值评分约等于 0.303, 在不需要人工设计特征的情况下, 该方法已展现出良好的应用前景。

#### 4.2 发震位置和发震时间预测

目前大多数机器学习地震预测研究均针对指定时空域的未来地震活动水平估计, 对地点及时间预测的研究并不多, 如 Madahizadeh 和 Allame-

hzadeh (2009) 基于神经网络对余震空间分布模式进行了研究, 但研究对象仅为一个地震序列。DeVries 等 (2018) 基于 DNN 对余震空间分布区域进行预测研究, 共使用了一个包含 30 000 多个主震-余震对的独立测试数据集对余震进行二元分类 (余震存在或不存在), 其模型特征主要基于主震破裂模型计算的应力变化张量而设计, 并非常用的地震活动性特征; DNN 模型包含 6 个隐藏层, 每个隐藏层由 50 个节点组成, 总共产生 13 451 个自由参数, DNN 在测试数据集上遍历所有的破裂滑移分布和网格单元得到的累积 AUC 值为 0.849, 大于经典库仑破裂应力准则 AUC 值 (0.583), 与 ROC 评价结果对比来看也是如此 (图 3)。结果表明 DNN 在余震地点预测方面比传统库仑破裂应力变化有更好的优势, 能够更好地解释余震的空间分布。他们进一步分析认为, 对余震地点预测起主要作用的是偏应力张量第二不变量, 因而认为 DNN 给出的预测结果在物理上可解释。该项工作受到广泛关注的同时也引起极大的争议, 如 Mignan 和 Broccardo (2019) 认为, 余震预测属于一种降维预测, 本身就应该比主震预测容易, 仅用较简单的 ANN 就完全可以达到很好的效果, 且无需应力相关参数的计算, 简化了余震预测过程,

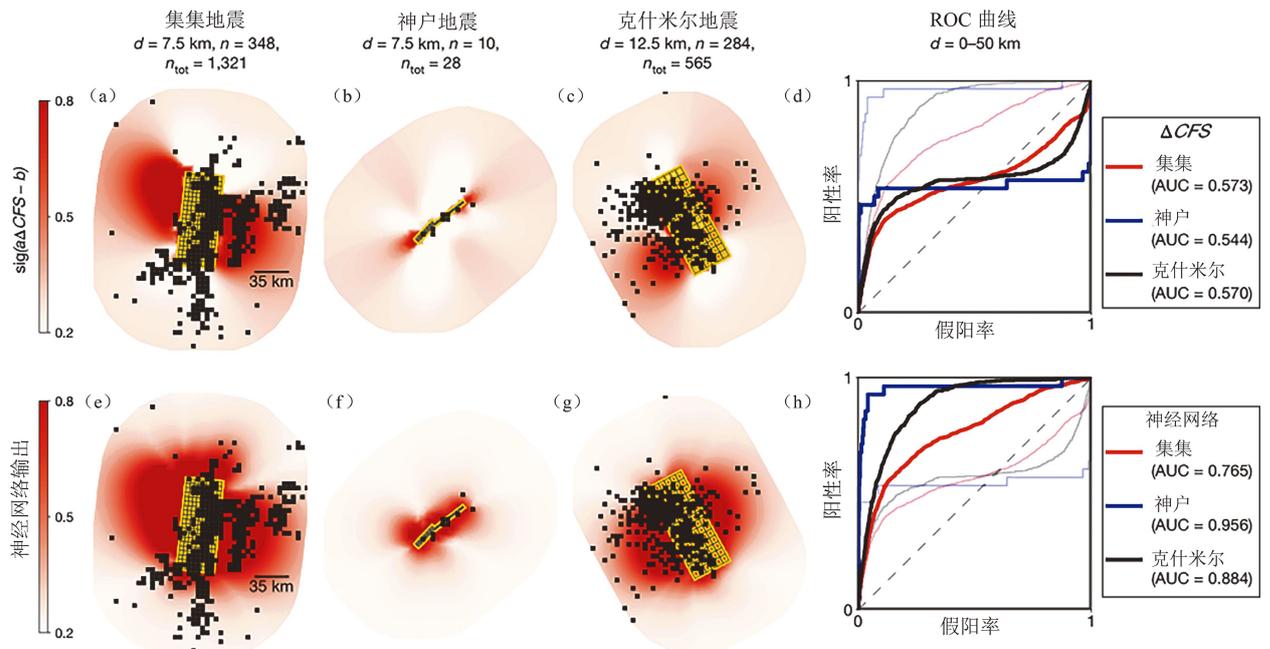


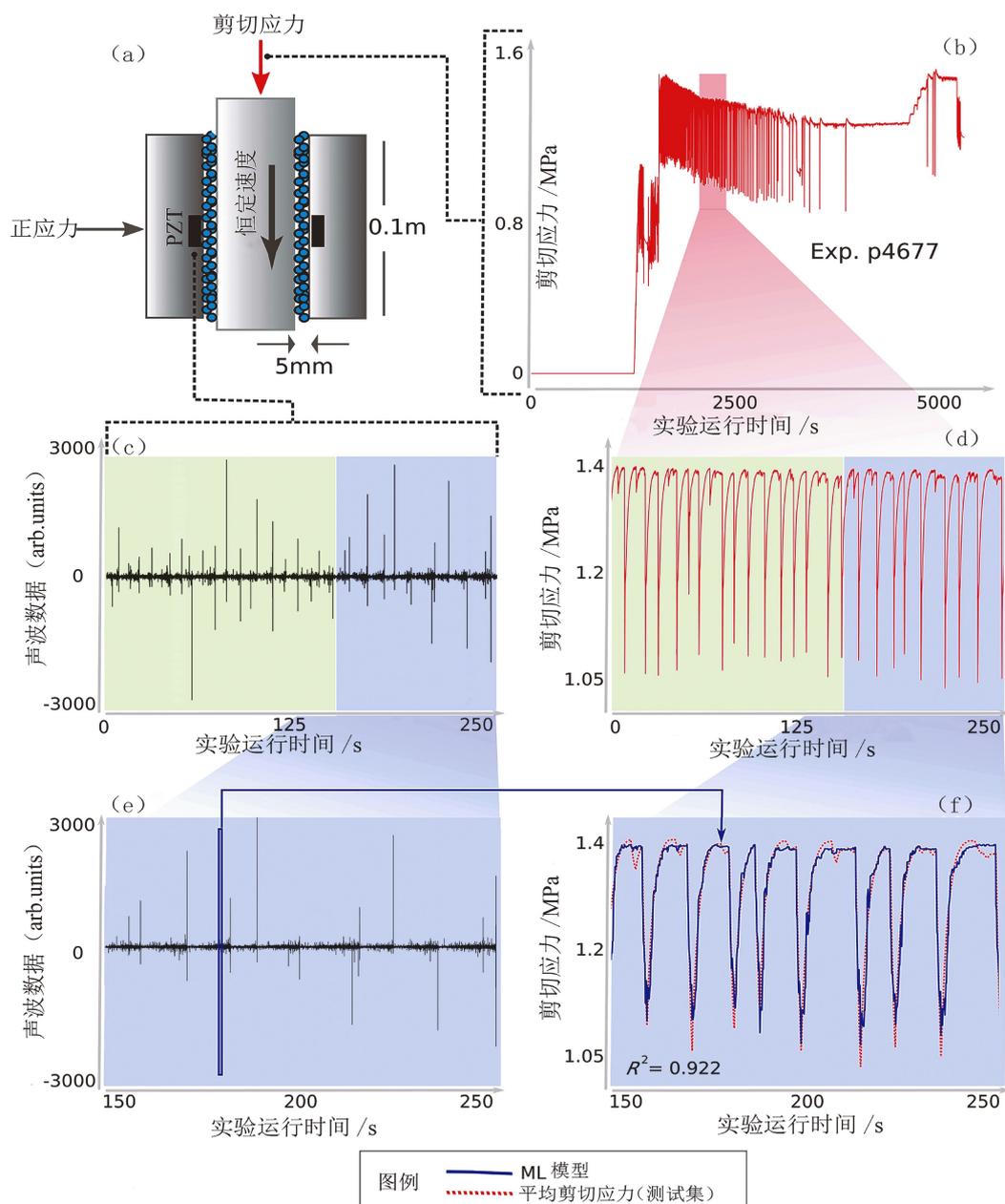
图3 传统库仑破裂应力变化 (a~c) 和神经网络 (e~g) 预测余震位置的方法, 及其对应的 ROC 曲线 (d、h) (据 DeVries *et al.*, 2018)

Fig. 3 Traditional methods for predicting aftershock location by coulombic stress change (a - c) and neural network (e - g), and their corresponding ROC curves (d, h) (according to DeVries *et al.*, 2018)

限制了模型偏差，深化了对余震机理的认识和对预测模型的改善。

机器学习对地震发生时间的预测，目前仍停留在环境条件单一的实验室观测数据处理，而 Rouet - Leduc 等 (2017, 2018) 基于剪切摩擦实

验 AE 连续波形记录 and 同时期的剪应力记录，训练随机森林模型，训练样本约占记录数据的 60%，预测并重现了另外约 40% 时间段的剪应力变化 (图 4)。这意味着可以通过连续波形记录，预测断层失稳 (比拟于大地震) 的时间。



(a) 为断层泥双剪实验装置示意图；(b) 为实验过程的剪应力记录，选择实验过程部分阶段 (b 中红色阴影区域) 放大，显示非周期的粘滑特征如 (d) 所示，与 (d) 时段相对应的声发射波形记录如 (c) 所示；在 (c) 和 (d) 中，绿色阴影区域对应 60% 的训练集，据此建立声发射数据与剪应力之间的相关性模型；(c) 和 (d) 中蓝色阴影区域对应 40% 的测试集，使用声发射数据 (e)，通过 ML 模型预测得到的剪应力时间变化如 (f) 中蓝线所示，与实际观测的剪应力时间变化 (f 中红线) 高度吻合

图 4 利用随机森林从断层摩擦实验声发射记录重构外加荷载应力变化曲线 (据 Rouet - Leduc *et al.*, 2018)

Fig. 4 Reconstructing the applied load stress change curve from the acoustic emission recording of a fault friction experiment using a random forest (according to Rouet - Leduc *et al.*, 2018)

Hulbert 等 (2019) 基于实验室观测数据开展了类似的重构应力变化过程的研究, 并进一步将机器学习用于常规地震与慢地震 (包括震颤和瞬时加速和自驱动传播等其他慢滑模式) 的识别, 其研究肯定了预测各种粘滑和蠕变-滑动失稳模式的可行性, 认为灾难性地震失稳前可能会有一组有组织的、潜在的可预测过程。但总的来看, 迄今为止尚未见到类似方法在实际地震记录中的应用, 其难点可能在于与实验室环境条件相比较, 实际环境过于复杂。Wang 等 (2021) 利用迁移学习算法处理实验室剪切实验数据, 实验数据来自一个双轴剪切装置, 该装置可记录声发射数据以及计算摩擦系数所需的法向应力和剪切应力。首先通过数值方法模拟实验室观测数值, 利用模拟的实验室数据训练机器学习模型, 然后将其应用到实验室实际观测数据中预测地震 (断层滑移), 实现了从数值模拟到实验室剪切实验的迁移学习, 同时引入交叉训练的方法来模拟实际地震观测数据不足的情况。结果表明实验中的失稳时间预测要比强度预测更加准确。通过迁移学习, 研究人员可以从一个模型推广到另一个模型, 从而克服数据的稀疏特性, 但其所采用的数据驱动的方法仅对慢滑事件有效, 而对于大规模粘滑断层却作用有限。

近期有学者对机器学习在地震预测研究中的概况也进行了总结 (Mignan, Broccardo, 2020), 如 Al Banna 等 (2020) 利用 “Earthquake” “Prediction” “Neural Network” “Machine Learning” 等关键词搜寻了 292 篇相关论文, 经过筛选最终选择 84 篇进行进一步分析, 结果显示利用 AI 进行地震预测的文章数量近些年呈逐渐上升的趋势且涵盖了几乎所有的 AI 方法, 其中 ANN 应用最多 (约占 33%)。按不同作者各自的评价规则及数据, 23 种 AI 算法的预测成功率均达到 60%, 其中神经动态分类-优化模型、反向传播自适应模型和改进的人工蜂群-多层感知系统以及随机森林等的预测准确率相对更高。DNN 对复杂问题表现出更优的解决问题能力, 在地震预测研究中应用也较为广泛。

## 5 结论与讨论

地震预测是地震危险性评估、震灾预防和建

筑结构安全设计的关键。随着观测站网密度增加, 地震和地震物理观测技术的进步, 地震观测数据急剧增加, 使 AI 技术可以更好地发挥其在大数据处理、隐式特征提取等方面的优势, 以此识别各类地震观测数据中的未知特征或规律, 对地震的发生时间、可能震级等进行预测。本文对国内外机器学习在地震预测领域中的应用进行总结发现, 数据的稀缺性、数据收集的质量以及同一模型在不同位置的性能可变性等问题都会显著影响模型的性能。在这里, 我们将讨论地震预测研究中面临的困难, 并提出克服这些困难的一些方法。

(1) 机器学习能够发挥计算机处理大数据的优势, 可发现人工难以直观观察到的某些震前特征。机器学习技术的应用, 开拓了针对非线性复杂系统的地震孕育发生过程的新的研究途径, 对提升地震预测能力起到重要的促进作用, 其中以 ANN 为基础发展起来的神经网络是最常见的地震预测模型。

(2) 目前已有的研究大多采用基于区域地震活动特征的特征提取方法, 由于每个地震带都有其独特的地震活动特点, 建立精确的区域地震预测模型时, 需要考虑特征参数的区域性差异。这种“个性化”的方法在不同区域的实际应用中其结果差异尤其明显, 在一个区域表现得较好, 可能在另一个区域则表现不佳。由于机器学习地震预测模型主要基于数据驱动, 而地震发生又是小概率事件。对于一个特定的区域, 其模型训练和检验样本数据一般都匮乏, 这不仅需要未来在更多地区、更长时段、对更多震例的积累和研究, 还应创建基准地震数据集, 对地震模型进行测试, 以简化不同模型的比较过程 (Al Banna *et al*, 2020)。

(3) 地震预测包括时间、地点及震级 3 个维度。由于地震孕育过程的极端复杂性及三要素预测的极端困难, 目前大多数机器学习地震预测研究主要集中于指定时空域内地震活动的震级预测。随计算机处理速度的提升、AI 技术的发展和可用资料的进一步积累, 未来应思考如何能够同时提升地震时、空、强三要素的预测能力。

(4) 地震与区域构造环境、地震之间具有非常复杂的非线性相互作用。当前的机器学习并不能清晰地解释这种高度的非线性关系, 因此仍需

要进一步研究各个地震参数的相互关系,找出它们的影响因素和物理成因,并结合现有对地震及地震前兆成因、特点、震例的认识,考虑物理约束的半监督学习算法,采用物理模型约束下的数据驱动方式,可能是机器学习等 AI 技术在地震预测领域能够更好地发挥作用的重要途径之一。当前,基于机器学习对输入特征重要性的进一步认识,也可说明某些强相关特征与未来强震活动之间可能存在物理关联,这也为深化地震前兆物理机理的认识提供了一条可能的途径。

(5) 相对于真正的大数据,用于地震预测的数据和学习样本仍十分匮乏,加之同一模型、相同特征在不同区域的预测效能差异极大,因而机器学习地震预测研究仍有许多问题亟待解决,这些问题显著影响着机器学习地震预测模型的预测效能。尽管基于神经网络的预测方法更直接地结合了结构化地震目录完整的时空结构,还利用了诸如 DNN 等算法的自动特征提取能力,但这些研究大多仍处于实验阶段,侧重于比较不同深度学习或经典机器学习预测因子的相对性能。深度学习地震预测模型是否为能够为地震可预测性提供新的见解,以及它们是否比传统的“统计+经验”的方法表现更好,仍有待进一步的研究和对比(Mousavi, Gregory, 2022)。

当前机器学习地震预测算法主要基于地震目录等地震观测资料,基于形变、电磁、流体等地球物理观测资料的机器学习地震预测算法或模型尚不多见。我国长期的地震预测实践表明,形变、电磁、流体等地球物理观测资料对地震短期预测可能具有不可替代的作用。地球物理观测资料是以秒或分钟为单位的等间隔采样数据,因而时间连续,大量成熟的机器学习时间序列预测算法可尝试应用于基于地球物理观测资料的地震异常识别和地震预测;地球物理观测资料易受环境和人为因素影响,机器学习异常检测类算法可用于地球物理观测资料的异常识别以及异常机理或影响因素分析,这方面的探索近期已有成果。因而,基于地球物理观测资料的机器学习地震预测研究是一个值得探索的领域。

#### 参考文献:

蔡煜东,甘俊人,姚林声. 1993. 人工神经网络在地震综合预报中的应用[J]. 地震学报,15(2):257-260.  
陈一超,曾三友,张好春,等. 2008. 基于遗传神经网络的地震预测研

究[J]. 计算机应用与软件,(4):135-137.  
韩晓飞,潘存英,罗词建. 2012. 基于遗传算法的广义回归神经网络在地震预测中的应用[J]. 华北地震科学,30(1):48-53.  
李荣峰. 2000. 福建及其周边地区地震活动人工神经网络模型的构建[J]. 台湾海峡,(1):107-112.  
马士振,刘宏志,牟磊. 2020. 在地脉动数据上应用分类算法的地震预测实验[J]. 地震,40(1):159-171.  
石绍先,范杨. 1990. 多级判别在地震综合预报中的应用[J]. 地震研究,13(4):8-14.  
石绍先,沈斌. 1986. 聚类分析在地震综合预报中的应用[J]. 地震研究,9(1):57-64.  
史翔宇. 2021. 基于机器学习回归算法的地震预测研究及其在中国地震科学实验场的应用[D]. 北京:中国地震局地震预测研究所.  
王炜,林命邈,马钦忠等. 2006a. 支持向量机及其在地震预报中的应用前景[J]. 地震工程学报. 28(1):78-84.  
王炜,刘悦,李国正等. 2006b. 我国大陆强震预测的支持向量机方法[J]. 地震学报,28(1):29-36.  
王炜,宋先月,王琤琤. 2000. 人工神经网络在地震中期预报中的应用[J]. 地震,22(1):10-16.  
王炜,吴耿锋. 2000. 神经网络在地震学方法综合预报中的应用[J]. 地震学报,22(2):189-193.  
中国电子技术标准化研究院. 2018. 人工智能标准化白皮书(2018版)[EB/OL]. (2018-01-24)[2022-09-20]. <http://www.cesi.cn/201801/3545.html>.  
Adeli H, Panakkat A. 2009. A probabilistic neural network for earthquake magnitude prediction[J]. Neural networks, 22(7):1018-1024.  
Al Banna M H, Taher K A, Kaiser M S, et al. 2020. Application of artificial intelligence in predicting earthquakes: state-of-the-art and future challenges[J]. IEEE Access, 8:192880-192923.  
Alarifi A S N, Alarifi N S N, Al-Humidan S. 2012. Earthquakes magnitude prediction using artificial neural network in northern Red Sea area[J]. Journal of King Saud University - Science, 24(4):301-313.  
Alves E I. 2006. Earthquake forecasting using neural networks; results and future work[J]. Nonlinear Dynamics, 44(1):341-349.  
Asencio - Cortés G, Martínez - Álvarez F, Morales - Esteban A, et al. 2016. A sensitivity study of seismicity indicators in supervised learning to improve earthquake prediction[J]. Knowledge - Based Systems, 101:15-30.  
Asencio - Cortés G, Morales - Esteban A, Shang X, et al. 2018. Earthquake prediction in California using regression algorithms and cloud-based big data infrastructure[J]. Computers & Geosciences, 115:198-210.  
Asim K M, Awais M, Martínez - Álvarez F, et al. 2017a. Seismic activity prediction using computational intelligence techniques in northern Pakistan[J]. Acta Geophysica, 65(5):919-30.  
Asim K M, Idris A, Iqbal T, et al. 2018a. Earthquake prediction model using support vector regressor and hybrid neural networks[J]. PLoS one, 13(7):e0199004.  
Asim K M, Idris A, Iqbal T, et al. 2018b. Seismic indicators based earthquake predictor system using Genetic Programming and AdaBoost

- classification [J]. *Soil Dynamics and Earthquake Engineering*, 111: 1 – 7.
- Asim K M, Idris A, Martínez – Álvarez F, *et al.* 2016. Short term earthquake prediction in Hindukush region using tree based ensemble learning [C]//International conference on frontiers of information technology (FIT). IEEE, 365 – 370.
- Asim K M, Martínez – Álvarez F, Basit A, *et al.* 2017b. Earthquake magnitude prediction in Hindukush region using machine learning techniques [J]. *Natural Hazards*, 85(1): 471 – 486.
- Asim K M, Moustafa S SR, Niaz I A, *et al.* 2020. Seismicity analysis and machine learning models for short – term low magnitude seismic activity predictions in Cyprus [J]. *Soil Dynamics and Earthquake Engineering*, 130: 105932.
- Aslam B, Zafar A, Khalil U, *et al.* 2021. Seismic activity prediction of the northern part of Pakistan from novel machine learning technique [J]. *Journal of Seismology*, 25(2): 639 – 652.
- Corbi F, Sandri L, Bedford J, *et al.* 2019. Machine learning can predict the timing and size of analog earthquakes [J]. *Geophysical Research Letters*, 46(3): 1303 – 1311.
- DeVries P M R, Viegas F, Wattenberg M, *et al.* 2018. Deep learning of aftershock patterns following large earthquakes [J]. *Nature*, 560(7720): 632 – 634.
- Florido E, Martínez – Álvarez F, Morales – Esteban A, *et al.* 2015. Detecting precursory patterns to enhance earthquake prediction in Chile [J]. *Computers & Geosciences*, 76: 112 – 120.
- Gulia L, Wiemer S. 2019. Real – time discrimination of earthquake foreshocks and aftershocks [J]. *Nature*, 74(7777): 193 – 199.
- Hinton G E, Salakhutdinov R R. 2006. Reducing the dimensionality of data with neural networks [J]. *Science*, 313(5786): 504 – 507.
- Hu W S, Nie H L, Wang H. 2012. Applied research of BP neural network in earthquake prediction [J]. *Applied Mechanics and Materials*, 204: 2449 – 2454.
- Huang J, Wang X, Zhao Y, *et al.* 2018. Large earthquake magnitude prediction in Taiwan based on deep learning neural network [J]. *Neural Network World*, 28(2): 149 – 160.
- Hulbert C, Rouet – Leduc B, Johnson P A, *et al.* 2019. Similarity of fast and slow earthquakes illuminated by machine learning [J]. *Nature Geoscience*, 12(1): 69 – 74.
- Kail R, Burnaev E, Zaytsev A. 2021. Recurrent convolutional neural networks help to predict location of earthquakes [J]. *IEEE Geoscience and Remote Sensing Letters*, 19: 1 – 5.
- Kanarachos S, Christopoulos S – R. G, Chronos A, *et al.* 2017. Detecting anomalies in time series data via a deep learning algorithm combining wavelets, neural networks and Hilbert transform [J]. *Expert Systems with Applications*, 85: 292 – 304.
- Kennedy D, Norman C. 2005. What don't we know? [J]. *Science*, 309(5731): 75.
- Külahcı F, Inceöz M, Doğru M, *et al.* 2009. Artificial neural network model for earthquake prediction with radon monitoring [J]. *Applied Radiation and Isotopes*, 67(1): 212 – 219.
- Li R, Lu X, Li S, *et al.* 2020. DLEP: A deep learning model for earthquake prediction [C]//International Joint Conference on Neural Networks (IJCNN). IEEE, 1 – 8.
- Madahizadeh R, Allamehzadeh M. 2009. Prediction of aftershocks distribution using artificial neural networks and its application on the May 12, 2008 Sichuan earthquake [J]. *Journal of Seismology and Earthquake Engineering*, 11(3): 112 – 120.
- Martínez – Álvarez F, Reyes J, Morales – Esteban A, *et al.* 2013. Determining the best set of seismicity indicators to predict earthquakes. Two case studies: Chile and the Iberian Peninsula [J]. *Knowledge – Based Systems*, 50: 198 – 210.
- Mignan A, Broccardo M. 2019. One neuron versus deep learning in aftershock prediction [J]. *Nature*, 574(7776): E1 – E3.
- Mignan A, Broccardo M. 2020. Neural network applications in earthquake prediction (1994 – 2019): Meta – analytic and statistical insights on their limitations [J]. *Seismological Research Letters*, 91(4): 2330 – 2342.
- Morales – Esteban A, Martínez – Álvarez F, Reyes J. 2013. Earthquake prediction in seismogenic areas of the Iberian Peninsula based on computational intelligence [J]. *Tectonophysics*, 593: 121 – 34.
- Mousavi S M, Beroza G C. 2020. A machine – learning approach for earthquake magnitude estimation [J]. *Geophysical Research Letters*, 47(1): e2019GL085976.
- Mousavi S M, Gregory C B. 2022. Deep – learning seismology [J]. *Science*, 377(6607): eabm4470.
- Moustra M, Avraamides M, Christodoulou C. 2011. Artificial neural networks for earthquake prediction using time series magnitude data or seismic electric signals [J]. *Expert systems with applications*, 38(12): 15032 – 15039.
- Nicolis O, Plaza F, Salas R. 2021. Prediction of intensity and location of seismic events using deep learning [J]. *Spatial Statistics*, 42: 100442.
- Niksarlioglu S, Kulahci F. 2013. An artificial neural network model for earthquake prediction and relations between environmental parameters and earthquakes [J]. *International Journal of Geological and Environmental Engineering*, 7(2): 87 – 90.
- Panakkat A, Adeli H. 2007. Neural network models for earthquake magnitude prediction using multiple seismicity indicators [J]. *International journal of neural systems*, 17(1): 13 – 33.
- Pandit A, Panda S. 2021. Prediction of earthquake magnitude using soft computing techniques: ANN and ANFIS [J]. *Arabian Journal of Geosciences*, 14(13): 1 – 10.
- Perol T, Gharbi M, Denolle M. 2018. Convolutional neural network for earthquake detection and location [J]. *Science Advances*, 4(2): e1700578.
- Reyes J, Morales – Esteban A, Martínez – Álvarez F. 2013. Neural networks to predict earthquakes in Chile [J]. *Applied Soft Computing*, 13(2): 1314 – 1328.
- Rouet – Leduc B, Hulbert C, Bolton D C, *et al.* 2018. Estimating fault friction from seismic signals in the laboratory [J]. *Geophysical Research Letters*, 45(3): 1321 – 1329.
- Rouet – Leduc B, Hulbert C, Lubbers N, *et al.* 2017. Machine learning predicts laboratory earthquakes [J]. *Geophysical Research Letters*, 44(18): 9276 – 9282.

- Rouet – Leduc B, Hulbert C, McBrearty I W, *et al.* 2020. Probing slow earthquakes with deep learning [J]. *Geophysical Research Letters*, 47(4) :e2019GL085870.
- Shodiq M N, Barakbah A R, Harsono T. 2015. Spatial analysis of earthquake distribution with automatic clustering for prediction of earthquake seismicity in Indonesia [C]//The Fourth Indonesian – Japanese Conference on Knowledge Creation and Intelligent Computing (KCIC), 24 – 26.
- Shodiq M N, Kusuma D H, Rifqi M G, *et al.* 2017. Spatial analysis of magnitude distribution for earthquake prediction using neural network based on automatic clustering in Indonesia [C]//International Electronics Symposium on Knowledge Creation and Intelligent Computing (IESKCIC). IEEE, 246 – 251.
- Shodiq M N, Kusuma D H, Rifqi M G, *et al.* 2018. Neural network for earthquake prediction based on automatic clustering in Indonesia [J]. *JOIV: International Journal on Informatics Visualization*, 2(1) : 37 – 43.
- Suratgar A A, Setoudeh F, Salemi A H, *et al.* 2008. Magnitude of earthquake prediction using neural network [C]//Fourth International Conference on Natural Computation. IEEE, 2:448 – 452.
- Wang K, Johnson C W, Bennett K C, *et al.* 2021. Predicting fault slip via transfer learning [J]. *Nature communications*, 12(1) :1 – 11.
- Wang Q, Guo Y, Yu L, *et al.* 2017. Earthquake prediction based on spatio-temporal data mining: an LSTM network approach [J]. *IEEE Transactions on Emerging Topics in Computing*, 8(1) :148 – 158.
- Wang Y, Li X, Wang Z, *et al.* 2022. Deep learning for magnitude prediction in earthquake early warning [J]. *Gondwana Research*, doi:org/10.1016/j.gr.2022.06.009.
- Zhang L, Si L, Yang H, *et al.* 2019. Precursory Pattern Based Feature Extraction Techniques for Earthquake Prediction [J]. *IEEE Access*, 7: 30991 – 31001.
- Zhang S, Shi Z, Wang G, *et al.* 2022. Application of the extreme gradient boosting method to quantitatively analyze the mechanism of radon anomalous change in Banglazzhang hot spring before the Lijiang  $M_w$  7.0 earthquake [J]. *Journal of Hydrology*, 612:28249.

## Research Progress in Field of Earthquake Prediction by Machine Learning Based on Seismic Data

WANG Jinhong<sup>1</sup>, JIANG Haikun<sup>2</sup>

(1. Institute of Earthquake Forecasting, China Earthquake Administration, Beijing 100036, China)

(2. China Earthquake Networks Center, Beijing 100045, China)

### Abstract

Machine learning's prominent implicit feature extraction and complex task processing capabilities are driving the science of earthquake prediction. In order to systematically understand the development status of machine learning technology in the field of earthquake prediction, this paper focuses on the application of machine learning in the field of earthquake prediction at home and abroad in recent years from the three aspects of earthquake magnitude prediction, earthquake location and earthquake occurrence time estimation in a specified time – space window, among which AI is the most widely used in earthquake magnitude prediction. In addition, this paper summarizes the main characteristic parameters, models and evaluation related issues of machine learning earthquake prediction, and explores the influence of seismicity parameters on earthquake prediction results from various evaluation mechanisms. Finally, a preliminary discussion and outlook on the problems existing in the field of earthquake prediction will be conducted. In the foreseeable future, the introduction of AI technology and the expansion of application fields are likely to lead the continuous development of the field of earthquake prediction.

**Keywords:** earthquake prediction; machine learning; feature extraction; model evaluation