分类号_____　　密级 _____
U D C _____　　编号 10741

硕 士 学 位 论 文

（专业学位）

论文题目　基于 Stacking 集成学习的地震人员死亡评估研究

研　究　生　姓　名：韩旭昊

指导教师姓名、职称：赵煜 教授　陈文凯 正高级工程师

学 科、专 业 名 称：统计学　应用统计硕士

研　究　方　向：大数据分析及应用

提　交　日　期：2024 年 6 月 3 日

独创性声明

本人声明所呈交的论文是我个人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

学位论文作者签名：韩旭昊　　签字日期：2024.6.3

导师签名：赵炬　　签字日期：2024.6.3

导师(校外)签名：陈文学　　签字日期：2024.6.3

关于论文使用授权的说明

本人完全了解学校关于保留、使用学位论文的各项规定，同意（选择“同意”/“不同意”）以下事项：

1.学校有权保留本论文的复印件和磁盘，允许论文被查阅和借阅，可以采用影印、缩印或扫描等复制手段保存、汇编学位论文；

2.学校有权将本人的学位论文提交至清华大学“中国学术期刊（光盘版）电子杂志社”用于出版和编入 CNKI《中国知识资源总库》或其他同类数据库，传播本学位论文的全部或部分内容。

学位论文作者签名：韩旭昊　　签字日期：2024.6.3

导师签名：赵炬　　签字日期：2024.6.3

导师(校外)签名：陈文学　　签字日期：2024.6.3

Research on Earthquake Casualties Assessment Based on Stacking Ensemble Learning

Candidate : Han Xuhao

Supervisor: Zhao Yu Chen Wenkai

摘 要

地震灾害人员死亡快速评估对地震应急响应至关重要。准确了解地震后人员死亡情况，并通过科学评估指导应急救援工作，对于降低经济损失和减少因救援不及时导致的人员死亡具有重要意义。中国大陆位于环太平洋地震带和欧亚地震带之间，地震灾害发生频繁，地震灾害人员死亡影响因素众多，如地震震级、地理环境、人口密度等。这些因素直接影响地震人员死亡结果，如：地震震级直接关系到地震灾害程度；地区人口密度越高，可能导致更多的人员死亡；地震次生灾害也会加剧人员死亡。鉴于以上考虑，本文利用 1950 年至 2022 年的中国大陆地震灾害损失评估资料，使用随机森林算法选择地震人员死亡影响因素，基于 Stacking 集成学习算法建立地震灾害人员死亡快速评估模型，为各级政府和应急管理部门应急指挥决策提供技术支持。该模型有助于在地震后及时部署救援资源，最大程度减少灾害损失。主要工作内容如下：

1. 震害数据的选取。收集中国大陆历史破坏性地震烈度图、人员死亡、灾区人口等基础数据，并根据需要进行整理，获得地震不同烈度区的震区面积和人口密度，同时甄别数据的准确性。

2. 基于随机森林、分类与回归树、梯度提升决策树、自适应提升算法选择地震人员死亡影响因素，选择算法时基于交叉验证评估每个算法的性能，根据特征重要性分析地震人员死亡影响因素，选择最相关的特征。

3. 选取 1950-2022 年间发生的破坏性地震，对地震死亡人数对数处理，使用生成对抗网络（GAN）将数据扩充到 2000 条，根据重要性分析选出的因素分别代入 Lasso、SVR、XGBoost、RF、LightGBM 模型进行训练，采用网格搜索算法确定参数值。用上述模型对地震人员死亡人数做出预测，分析各模型的 RMSE 值、MAE 值、MAPE 值，发现 LightGBM 预测效果最好，Lasso 多元线性回归模型最差。基于上述模型构造 Stacking 集成模型，将多个基础学习器的输出作为输入，通过另一个模型（元学习器）进行最终的预测，根据评价指标得出 LightGBM-Stacking 预测最准确，效果最好。

4. 为分析 LightGBM-Stacking 模型的评估效果，随机选取验证震例。充分考虑地震对不同地区造成的差异性影响，将中国大陆划分为西北、西南和东部三个区域，按地震烈度对样本进行分类，为增加样本的多样性，利用 GAN 扩充样

本，对模型结果进行分析，并与其他评估方法进行比较，验证模型的性能和准确性。这一方法的创新点在于考虑到地震在不同地区的特殊情况，通过细致的区域划分和样本分类，以及 GAN 网络的应用，提高模型的鲁棒性和泛化能力。

关键词：地震 人员死亡 Stacking 集成学习

Abstract

Rapid assessment of human casualties in earthquake disasters is crucial to earthquake emergency response. Accurately understanding the casualties after an earthquake and guiding the emergency rescue work through scientific assessment is of great significance in reducing economic losses and minimizing casualties caused by untimely rescue. Mainland China is located between the Pacific Rim Seismic Belt and the Eurasian Seismic Belt, where seismic disasters occur frequently, and there are many factors affecting the casualties of seismic disasters, such as earthquake magnitude, geographic environment, and population density. These factors directly affect the results of earthquake casualties, such as: the earthquake magnitude is directly related to the degree of seismic hazard; the higher population density of the region may lead to serious casualties; and secondary disasters of earthquakes may also aggravate casualties. In view of the above considerations, this paper makes use of the earthquake disaster loss assessment data from 1950 to 2022 in mainland China, uses the random forest algorithm to select the factors affecting earthquake casualties for modeling, and establishes a rapid assessment model of earthquake disaster casualties based on the Stacking integrated learning algorithm, which provides technical support for the earthquake emergency response command and decision-making of the governments at all levels and the emergency management departments. The model helps to deploy rescue resources in time after an earthquake and minimize disaster losses. The main work is as follows:

1. Selection of seismic data. Collect basic data such as historical destructive earthquake intensity maps, deaths, and population of the affected areas in mainland China, and organize them as needed to obtain the area of the seismic zone and population density in different intensity zones of the earthquakes, as well as to screen the accuracy of the data.

2. select the factors affecting earthquake casualties based on Random Forest, Classification and Regression Tree, Gradient Boosting Decision Tree, and Adaptive Boosting algorithms, evaluate the performance of each algorithm based on cross-validation when selecting

the algorithms, and identify the factors affecting earthquake casualties based on the analysis of the importance of the features, and select the most relevant features.

3. Destructive earthquakes occurring between 1950 and 2022 are selected, the number of earthquake fatalities is logarithmically processed, and the data are expanded to 2,000 using GAN, and the factors selected according to the importance analysis are substituted into the Lasso Multiple Linear Regression, SVR, XGBoost, RF, and LightGBM models for training, respectively, and the lattice search algorithm is used to Determine the parameter values. The above models were used to predict the number of earthquake deaths, and the RMSE, MAE, and MAPE values of each model were analyzed, and it was found that LightGBM had the best prediction effect, and Lasso multiple linear regression model was the worst. The Stacking integrated model is constructed based on the above model, and the outputs of multiple base learners are used as inputs to make the final prediction through another model (meta-learner), and according to the evaluation index, it is concluded that LightGBM-Stacking prediction is the most accurate and the best.

4. In order to analyze the evaluation effect of the LightGBM-Stacking model, randomly selected validation earthquakes were used. Taking into full consideration the differential impacts of earthquakes on different regions, the Chinese mainland is divided into three regions: northwest, southwest and east, and the samples are classified according to the seismic intensity. To increase the diversity of the samples, the samples are expanded by using the GAN, and the results of the model are analyzed and compared with other methods of assessing the casualties, so as to validate the performance and accuracy of the model. The innovation of this method is to consider the special situation of earthquakes in different regions, and to improve the robustness and generalization ability of the model through careful regional division and sample classification, as well as the application of GAN networks.

Keywords： Earthquake; Casualties; Stacking Ensemble Learning

目　录

1 引　言

1.1 研究背景和意义

1.1.1 研究背景

地震因其突发性和难预测性，是对人类威胁最大的自然灾害之一。中国大陆位于环太平洋地震带和欧亚地震带之间，是世界上地震活动最频繁的地区之一[1]。由于地震频繁发生，中国大陆一直面临地震所带来的人员伤亡和财产损失的威胁。20 世纪以来，中国发生 6.0 级地震 1000 多次，其中大陆地区 6.0 级以上地震发生约 470 次。地震不仅直接导致社会经济损失，还威胁着人们的生命和财产安全。地震发生时，除地震本身的能量大小，人口密度、房屋结构、发生时间、次生灾害等都是导致人员死亡的重要因素。中国大陆主要以浅源地震为主，浅源地震造成了大量的财产损失和人员死亡。如 1920 年宁夏海原 8.5 级地震中，约 23.4 万人死亡；2008 年四川汶川 8.0 级地震造成 69227 人死亡[2]。

准确评估地震灾后人员死亡情况对地震应急响应和地震灾害风险评估工作至关重要。基于机器学习的方法进行人员死亡评估，可以提高应急响应能力和地震灾害风险评估，地震人员死亡评估模型不仅在地震应急响应、地震灾害风险评估中发挥重要作用，在城市规划、建筑设计、应急管理等方面也得到广泛应用。尽管已有一些研究关注地震灾后人员死亡评估，但现有方法往往依赖于传统的经验统计模型或单一的机器学习算法。基于 Stacking 集成学习的方法可以结合多个模型的优势，提高预测准确性和稳定性，从而弥补现有研究的不足。

1.1.2 研究意义

理论意义：现有地震人员死亡评估模型通常需要综合考虑多个因素，包括地震参数（如震级、震源深度）、地震影响范围、建筑结构、人口密度等。这要求模型更全面地评估地震对人员的潜在影响，但地震人员死亡评估涉及众多不确定性因素，例如地震参数的准确性、建筑结构的强度和韧性、人员分布的变化等。这些不确定性会对评估结果产生影响，影响预测的精度和可靠性。传统经验统计

模型在考虑地震人员死亡因素时存在一定的局限性，无法完全适应各种地震场景和灾区特点。导致模型在某些情况下的适用性和精确性受到限制。

本文通过分析地震人员死亡影响因素，基于重要性权重评估，确立不同烈度下的地震人员死亡评估模型，鉴于地震人员死亡数据集的复杂性以及模型的准确性要求，使用 stacking 集成学习挖掘数据中的规律，提高地震人员死亡评估的精确性。

现实意义：地震人员死亡评估的研究可以帮助确定地震发生后的人员死亡情况，提供数据支持并确保准确性，有助于优化资源配置、增强救援效率，并提高应急响应能力；通过了解地震对人员伤害的主要因素，可以制定更有效的建筑规范、灾害预警系统和应急计划，以减少未来地震可能造成的人员死亡；地震人员死亡评估研究有助于深入理解地震的机理和影响因素，通过对历史地震事件和人员死亡数据的分析，可以提高地震预测和预警的准确性，为社会提供更有效的地震风险管理和应对措施。

综上所述，地震人员死亡评估的研究背景和重要性表现在提供准确的死亡评估、促进灾后救援行动、改进灾害管理策略以及促进地震研究和预测等方面。这些研究成果对于减少地震灾害带来的人员死亡和财产损失，具有重要的实际应用价值。

1.2 国内外研究现状

1.2.1 影响因素研究方面

地震人员死亡评估时，无论选择哪一种方法，地震人员死亡因素作为参数进行建模分析，选取合适的影响因素直接影响模型的预测效果。地震人员死亡的影响因素众多，并非每个因素都有决定性意义[3]。由于数据各因素存在一定相关性，选择过多的因素进行评估反而会增加冗余性，因此，对因素之间的相关性进行分析并筛选出适当的因素尤为重要。研究者会根据经验直接给出重要的影响因素[4]和选取的房屋类型[5]，或者根据线性模型得出地震人员死亡因素[6]。线性模型是常用的特征分析的方法之一，Chen 等基于经验模型给出了地震人员死亡与烈度、人口密度、震级等因素之间的关系[7]；Maqsood、Schwarz 提出在研究地震人员

死亡时需要考虑地震震动、强度和其他地震动参数[8]；Huang 等采用相关性分析与主成分分析选择地震人员死亡的影响因素，最后选定震级、震中烈度、人口密度、地震发生时刻和建筑物破坏面积分别作为变量和修正参数[9]；Aleskerov 等使用情景方法计算在不同地震烈度下参数关系[10]；Park、Shin、Cho 基于 HAZUS-MH 软件得知损坏建筑物中的室内人口比例和占用类型的空间分布是影响地震人员死亡的重要因素[11]；So 等基于剑桥地震影响数据库中人员死亡数据，得出建筑物和死亡人数分布与建筑物类型、建筑物破坏状况、烈度的关系，对比空间计量模型（SAR）、线性回归等模型对人员死亡影响因素的评估分析结果，结果表明人员死亡主要取决于建筑物的破坏程度[12]。

经验统计模型虽然能得到一些确切的结论，但由于传统数据的不确定性和模糊性，经验模型在处理人员死亡预测这类高度非线性问题时有所缺陷。因此，一些新的方法开始涌现。Hu 等使用神经网络模型探究汶川地震儿童死亡率的相关因素的贡献度，敏感性分析表示，地震坡度、地质因素、地震烈和人均收入对死亡率贡献较大[13]；Wang 等基于汶川地震使用灰色关联理论对影响因素进行分类，最终得到的人员死亡影响因素相关度系数[14]。一些研究人员将集成学习算法引入特征重要度评估问题中，以提高传统经验模型的准确性和泛化能力。Chen 等提出集成算法在预测能力和泛化能力上的优异性能比经验统计模型更好[15]。

1.2.2 评估方法研究方面

地震人员死亡评估研究方法主要分为三类：经验统计模型、机器学习模型和其他方法。

经验统计模型主要使用震级、人口密度、震区面积、发震时间等参数建立回归模型，肖光先等以房屋破坏间数为主要参数，同时考虑烈度和人口密度为辅助参数，给出了近似的经验公式[16]；马玉宏等详细综述了不考虑易损性的 5 种方法和以易损性为主要参数的 22 种方法[17]；Jaiswal 等建立双参数（烈度的均值与标准差）对数分布公式得到国家地震死亡率的经验模型[18]；刘金龙等建立以震中烈度为主要参数，震级和人口密度作为辅助参数进行修正的人员死亡预测模型[19]；尹之潜根据中国大地震中人员死亡数提出了计算人员死亡比的经验模型[20]；张莹

等采用多元非线性回归法建立了基于多因素的地震灾害人员死亡评估模型[21]；Huang 等提出了一种基于修正的局部高斯曲线的预测模型[22]。

机器学习模型可以通过学习历史数据预测未来事件的可能结果。在地震人员死亡模型中，机器学习模型使用历史地震和人员死亡数据训练模型，并将预测结果与实际发生的人员死亡情况进行比较验证，以评估模型的准确性和适用性。于山等以 20 次地震灾害的地震发生时刻、震级、震中烈度等 7 个评价指标建立了三层 BP 神经网络地震灾害人员死亡预测模型[23]；钱枫林等通过主成分分析提取了人员死亡主要影响因素，构建了 3 层 BP 神经网络预测模型[24]；周德红等利用遗传算法优化的 BP 神经网络预测模型进行训练并预测[25]；Oktarina 等使用人工神经网络（ANN）预测地震造成的受损建筑物和人员死亡数量[26]。Li 等提出了一种基于区域差异的空间划分方法和基于支持向量回归（SVR）的分区死亡预测方法[27]。

其他方法在地震人员死亡评估时各有优缺点，张文娟提出并设计了基于移动通信大数据的地震灾害人口死亡评估系统，通过地震发生前后的两次定位数据进行回归分析，该系统在时效性和稳定性方面均优于传统评估系统，但技术的实现环节发展仍不成熟[28]；曾婷婷等基于历史案例的灾情加权综合评估模型，引入地震断层矩量化空间相关程度，通过改进模型评估的历史震例的权重提高了评估精度，但只能得到评估区域的总体估计值，没有具体空间分布信息[29]；吴昊昱等采用幂律分布发现汶川地震死亡人数的增长呈现出分段规律，能够在地震发生几天后对死亡人数的规模趋势进行推断，但需要数据详细到县，对数据的精细程度要求较高[30]；Fang 等综合考虑掩埋情况和救援过程，基于建筑物损坏率构建地震死亡预测方法[31]；Cui 等综合多种因素提出了一种基于堆叠集成学习和改进的群智能算法的有效预测方法[32]；Xia 等基于建筑物抗致死水平，构建人员死亡快速评估矩阵[33]；Badal 等利用地理信息系统（GIS）环境中开发的应用程序来计算不同强度区域内的人员死亡数量[34]；Zhang 等将所有灾害损失评估模型集成到地理信息系统（GIS）中，系统评估的死亡人数平均准确度为 66.1%[35]。

1.2.3 文献述评

进行地震人员死亡评估时，影响地震人员死亡的影响因素有很多，一般都是根据经验直接给出重要的地震人员死亡影响因素[36]。经验统计模型虽然能得到一些确切的结论，但由于传统数据的不确定性和模糊性，经验模型在处理人员死亡预测这类高度非线性问题时有所缺陷。在方法选取方面，基于专家经验和历史震害数据的经验统计分析法，难以有效提取复杂的地震数据特征；机器学习模型能够较好的进行非线性分析，但对历史震害数据量要求较高，模型收敛到全局最优较难；基于 GIS 技术、高分辨率卫星图像等新兴技术需要依赖于更高分辨率的图像数据。

1.3 研究目的与研究内容

1.3.1 研究目的

第一，提高人员死亡评估准确性。引入基于 Stacking 集成学习的方法，通过网格搜索选择最优参数，提高地震灾后人员死亡评估的准确性。传统的评估方法往往依赖于单一模型或经验，无法充分利用不同模型的优势。通过集成学习模型的预测结果，期望得到更准确的估计，以提高地震应急响应能力。

第二，提高人员死亡评估的稳定性。除了准确性，稳定性也是评估模型的重要指标。在面对不同地震场景和数据不确定性的情况下，传统的评估方法可能不稳定。通过 Stacking 集成学习，利用多个模型集成，减少预测的波动性，从而提高人员死亡评估的稳定性。

第三，通过集成算法筛选影响地震人员死亡的核心因素。地震灾后人员死亡评估涉及众多复杂的影响因素，如地震震级、震源深度、建筑结构、人口分布等。传统的评估方法往往只能考虑其中一部分因素，难以全面综合考虑多种影响因素的作用。通过 Stacking 集成学习，利用多个模型，考虑不同的影响因素，从而更全面地进行人员死亡评估。

1.3.2 研究内容

整理 1950-2022 年中国大陆地震灾害数据、人口数据和基础地理数据，基于 GIS 空间分析和数学统计方法，从时间、震级、地区等不同的角度研究中国大陆地区地震灾害和人员死亡规律，探讨各影响因素对地震人员死亡的影响。通过集成学习算法筛选影响地震人员死亡的主要因素，对比不同的地震人员死亡评估模型，分析不同集成学习算法的优缺点，分析不同区域不同烈度下地震人员死亡情况，尝试建立不同烈度的分区 Stacking 模型，提高地震人员死亡评估模型的实用性。

第一章为引言，介绍了地震人员死亡评估模型的背景和意义，对近些年国内外的研究现状总结，最后阐述了本论文的研究目标以及本文各个章节的内容。

第二章是基本理论，详细介绍了本篇论文用到的各模型的原理和特点。

第三章是数据来源及处理，对数据的来源及属性进行解释说明，并分析影响地震人员死亡的因素，根据中国大陆幅员辽阔特点进行样本划分，增加预测的准确性。

第四章是单一机器学习模型的搭建，分析各基学习器的评估效果，引入三种常用的模型评价指标，分别是 RMSE、MAE、MAPE，根据各个模型的 RMSE、MAE、MAPE 对预测效果进行对比分析，根据 Stacking 集成算法确定第一层的基学习器和第二层的元模型。

第五章是 Stacking 集成学习的预测，将 Stacking 集成模型用于实际震例的预测，得出最后的预测效果，并与其他模型对比分析。

第六章是总结与展望，对本论文进行总结，并对不足之处进行阐述，提出展望。

1.4 研究创新点

第一，对地震影响因素和死亡人数进行相关性分析和重要性评估。采用相关性分析和随机森林等方法，对地震影响因素与死亡人数之间的关系进行深入研究，并评估各影响因素的重要性，以确定对地震人员死亡影响最为显著的因素。

第二，分区域分烈度建模。考虑到中国地域辽阔、地震活动频繁的特点，将中国大陆划分为西北、西南、东部地区。根据烈度特征将历史震例进行分类。通过对中国大陆进行区域分析和烈度划分，使地震人员死亡评估更加准确和可靠。

2 相关理论与方法

2.1 地震人员死亡评估理论

地震人员死亡评估理论是一种用于预测地震人员死亡情况的基本原理和方法。该理论基于地震的关键属性，旨在通过综合分析地震烈度、建筑物脆弱性、人口分布等因素，提供对人员死亡的评估和预测。建立不同参数的理论模型，综合考虑震级、地震烈度、人口密度等因素，从而预测可能发生的人员死亡情况[37]。该理论主要考虑以下关键要素：

地震属性：地震属性是地震数据内在特性的量化表达，涵盖了地震在几何、运动学、动力学和统计学等多个维度上的特征，地震属性是从地震测量数据中经过分析计算提取出的具体参数指标。其中，震级和烈度是地震能量属性的具体指标，与地震人员死亡评估密切相关。震级是地震能量的量度，反映了地震在震源处释放的能量大小。震级与多个因素有关，地震波振幅是用确定震级的直接变量，振幅越大，通常震级越高。不同类型的地震波（如 P 波、S 波和表面波）对震级的计算也有所不同。地震烈度是衡量地震强度和地震影响程度的指标。地震人员死亡评估理论基于地震烈度分析不同震级和震源距离下的影响范围，进而估计人员死亡情况。

建筑物脆弱性：建筑物脆弱性指地震造成的建筑物倒塌或部分破坏，是导致人员死亡的主要原因之一。关注建筑物的结构特征、材料性能和抗震能力等因素，分析建筑物在地震中的破坏程度可以有效评估因房屋倒塌造成的人员死亡。

人口分布：人口分布考虑地震事件发生时人口的分布情况，包括人口密度、建筑物类型和使用情况等因素。结合地震烈度分布，可以分析在不同地区和建筑物类型下的人员死亡情况。

综上所述，地震人员死亡评估理受到震级、地震烈度、建筑物脆弱性和人口分布等因素共同影响。地震人员死亡评估理论可以帮助预测地震事件中可能发生的人员死亡情况，为灾害应急相应和地震灾害风险评估提供决策依据。然而地震人员死亡评估理论仍面临着挑战，例如在地震预测准确性和数据可靠性等方面，需要进一步研究和改进。

2.2 数据增强方法

数据增强是一种通过对原始数据进行变换或扩展的技术手段，以增加模型的训练样本数量。更多的数据能够帮助机器学习模型更好地泛化未见过的数据。数据增强通常涉及到对图像、文本或其他类型的数据进行一系列的变换，例如旋转、翻转、裁剪、缩放等。通过这些变换可以生成新的训练样本，同时保持样本的标签不变，从而为模型提供更多的学习材料。尽管数据增强在提高模型性能和泛化能力方面具有许多优势，但存在一些潜在的缺陷。以下是一些常见的数据增强缺陷：

（1）引入不真实样本。一些数据增强技术可能会引入在实际场景中不存在的特征，使模型学到的信息在真实情况下不够准确。

（2）计算成本增加。在大规模数据集上训练时数据增强方法需要更多的计算资源和时间。

（3）过度变换可能损害模型性能。过于激进的数据增强可能导致模型对噪声过度敏感，从而影响其性能。

因此，在应用数据增强技术时，需要权衡其优势和局限性，以确保模型的性能和泛化能力。生成对抗网络（GAN）是一种深度学习模型，其独特的结构和训练方式使其在图像生成、风格迁移等领域取得了显著的成功。生成对抗网络由生成器和判别器两部分相互协作组成，形成一种博弈的结构，通过不断的竞争和学习，生成器逐渐提高生成样本的质量，而判别器则不断进化以更好地区分真实样本和生成样本。生成器负责接收随机噪声或输入数据，并生成类似于训练数据的新样本，由多个层组成，每一层负责学习特定级别的特征。判别器接收来自真实数据和生成器的样本，然后尝试区分哪些是真实的，哪些是生成的，判别器由多个层组成，通过学习提高对真伪样本的判别能力。GAN 在训练过程中反复迭代生成器和判别器的权重，在不断迭代过程中，生成器试图生成更逼真的样本，而判别器试图更好地区分真实样本和生成的样本，这种博弈过程使生成器不断提高生成样本的逼真度，判别器不断提高判别的难度。生成对抗网络最开始是在无监督学习的领域提出，但经证明对半监督学习、完全监督学习、强化学习也有效，所以同样适用于地震数据。

在具体博弈过程中，生成器试图最小化一个损失函数，而判别器试图最大化相同的损失函数。整个博弈过程可以看成一个极小极大值问题。生成器的目标是最小化生成样本与真实样本的差异，使判别器无法有效地区分两者。生成器的损失函数通常被定义为判别器错误的负值，即生成样本被错误地判别为真实样本的概率。生成器的目标可以表示为：

$$\min_{G}\max_{D}V(D,G)=E_{x\sim p_{data}(x)}[\mathrm{logD(x)}]+E_{z\sim p_z(z)}[\log\left(1-\mathrm{D(G(z))}\right)] \quad (2.1)$$

式中：G代表生成器，D代表判别器，x代表真实样本，z代表随机噪声，通常服从高斯分布，$p_{data}(x)$是真实样本的分布，$p_z(z)$是随机噪声的分布。

训练过程通过交替更新生成器和判别器来实现。首先，判别器通过最大化判别错误的概率来提高自己的性能。

$$\max_{D}V(D,G)=E_{x\sim p_{data}(x)}[\mathrm{logD(x)}]+E_{z\sim p_z(z)}[\log(1-\mathrm{D(G(z))})] \quad (2.2)$$

接着，生成器通过最小化生成样本被错误判别为真实样本的概率来提高自己的性能：

$$\min_{G}V(D,G)E_{z\sim p_z(z)}[\log(1-\mathrm{D(G(z))})] \quad (2.3)$$

在生成对抗网络的训练进程中，通过交替优化机制，生成器和判别器的能力会逐步增强，并最终趋向一种动态均衡。在这个均衡点上，生成器创造出高度真实的样本，从而判别器难以准确地区分真实样本与生成样本之间的差异。在此过程中，生成器致力于最小化其损失函数以提升伪造样本的真实性，而判别器最大化损失函数以强化自身的分辨能力。在训练过程中持续驱动着生成器和判别器共同提升和发展，构成了生成对抗网络特有的动态博弈特征。

2.3 评估方法

2.3.1 随机森林

随机森林是一种基于 Bagging（Bootstrap Aggregating）的集成学习方法，通过组合多个决策树进行预测和分类。每个决策树基于随机抽样和特征选择构建，通过投票或平均等方式集成各个决策树的结果[38]。

随机森林在特征选取方面主要包括以下步骤：

（1）采用 Bootstrap 重抽样技术从原始样本有放回地随机抽取数据构造多个样本集。

（2）从输入特征中随机抽取 m 个特征$X_1, X_2, ..., X_m$，在 m 个特征中选择最佳特征用于分割节点，构造决策树各个分支，直到这棵树能够准确表示分类或遍历所有属性。

（3）通过 m 棵决策树模型进行投票得到分类结果。

特征选择过程是对特征重要程度进行排序的过程，特征重要性评分用 VIM 表示，选择基尼指数作为衡量特征分割的效果，假设集合 T 中包含 N 个不同类别的样本：

$$Gini(T) = 1 - \sum_{i=1}^{N} P_i^2 \tag{2.4}$$

式中：P_i表示节点中第i类样本的概率，特征X_m在节点q的重要性就是节点前后$Gini$指数变化量：

$$VIM_{mq}^{Gini} = GI_q - GI_l - GI_r \tag{2.5}$$

式中：GI_l和GI_r分别表示分枝后两个新节点的$Gini$指数，

将所有特征的重要性评分归一化处理：

$$VIM_j^{Gini} = \frac{VIM_j^{Gini}}{\sum_{i=1}^{m} VIM_i^{Gini}} \tag{2.6}$$

随机森林处通过随机选择特征和样本训练,在特征选择过程中减少了维度的影响，在集成多个决策树的过程中，引入随机性和多样性，从而降低模型的过拟合风险，使随机森林在训练数据上表现出较好的泛化能力。

2.3.2 XGBoost

XGBoost（Extreme Gradient Boosting）是一种梯度提升树算法，基于决策树集成学习，通过迭代训练多个决策树逐步提升模型的预测能力。每一轮迭代中，新的决策树用来训练并纠正前一轮迭代中模型的残差。将所有决策树的预测结果进行加权融合，每棵树的权重由训练过程中的性能表现（损失函数减少的程度）决定，得到最终的模型预测结果[39]。

XGBoost 使用的损失函数由两部分组成：损失函数的期望值和正则化项。对于回归问题，损失函数可以定义为：

$$\text{Obj} = \sum_{i=1}^{n} l(y_i, \hat{y}_i) + \sum_{k=1}^{K} \Omega(f_k) \tag{2.7}$$

$$\Omega(f) = \lambda T + \frac{1}{2}\lambda \|w\|^2 \tag{2.8}$$

式中：n是样本数量，y_i是第 i 个样本的真实值，$\hat{y}_i$是模型的预测值，$l(\cdot)$是损失函数，用于衡量预测值与真实值之间的差异，K是树的数量，f_k是第k棵树，$\Omega(\cdot)$是正则化项，用于控制模型的复杂度。

$$\hat{y} = \sum_{k=1}^{K} f_k(x_i) \tag{2.9}$$

$$F = \{f(x) = w_q(x)\}(q: R^m \to T, w \in R^T) \tag{2.10}$$

式中：q代表 CART 树的结构，T是决策树的子节点的个数，$f(x)$为决策树结构，模型实质上是通过特征的提取，构建决策树，确定树的结构q和权重w。

2.3.3 LightGBM

LightGBM（Light Gradient Boosting Machine）是一个梯度提升框架，专为大规模数据和高效训练而设计。基于树模型通过迭代地训练一系列决策树提高模型性能[40]。由微软开发，LightGBM 在训练过程中引入了一些创新性的技术，提高了训练速度和模型性能。LightGBM 使用梯度提升算法，核心是最小化损失函数。在回归问题中，典型的损失函数是均方误差；而在分类问题中，可以使用交叉熵损失函数等。下面是主要的技术原理：

（1）典型的均方误差损失函数：

$$L(y, F(x)) = \frac{1}{2}\Sigma_{i=1}^{n}(y_i - F(x_i))^2 \tag{2.11}$$

LightGBM 使用基于树的弱学习器，通过递归地建立决策树来逼近损失函数。每一步，选择能够最小化损失函数的特征和分割点，决策树的建立过程中，损失函数的负梯度被用来更新叶子节点的预测值。

（2）直方图

使用直方图来近似和加速梯度的计算。将连续的特征值划分成一系列的范围，每个范围称为一个桶，这样可以将特征值分桶成离散的直方图，减少计算量。

（3）Leaf-wise 生长策略

采用叶子生长策略，每次选择能够最大化增益的叶子进行生长。有助于降低树的深度，减小过拟合的风险。

（4）特征捆绑

引入特征捆绑的概念，通过自动发现相关性强的特征将它们捆绑在一起，减少特征空间的维度，提高训练速度。

总体而言，LightGBM 因其高效的训练速度和在大规模数据上的出色性能而受到广泛欢迎，其快速的训练速度使得在实际应用中能够更快地部署和更新模型，同时在数据科学竞赛中提供了一个强大的工具，因此在工业界和各种数据科学竞赛中备受青睐。

2.3.4 Lasso 回归

LASSO（Least Absolute Shrinkage and Selection Operator）是一种基于线性回归的正则化方法，用于进行特征选择和模型简化。通过在损失函数中添加 L1 正则化项，将模型系数的绝对值作为惩罚项，只保留对目标变量具有显著影响的特征，从而实现特征选择和稀疏性[41]。LASSO 的目标是找到一个最小化损失函数和正则化项的解，即最优的模型系数。可以通过求解以下优化问题来实现：

$$\min_{\beta_0,\beta_1,\ldots,\beta_n}\{\frac{1}{2m}\sum_{i=1}^{m}(y_i-\beta_0-\sum_{j=1}^{n}\beta_j x_{ij})^2+\alpha\sum_{j=1}^{n}|\beta_j|\} \tag{2.12}$$

式中：m是样本数量，y_i是第 i 个样本的真实值，n是特征数量，β_j是第j个特征的系数，x_{ij}是第 i 个样本的第j个特征值，α是正则化参数，用于平衡损失函数和正则化项的重要性。

通过使用优化算法（如坐标下降法或梯度下降法）求解上述优化问题，可以得到最优模型系数，非零系数对应对目标变量具有重要影响的特征，零系数对应被 LASSO 特征选择剔除的特征。LASSO 通过优化问题求解最优模型系数的同时平衡损失函数和正则化项，在特征选择、模型简化和处理高维数据等方面具有重要的应用价值。

2.3.5 支持向量回归

支持向量回归（SVR）是一种基于支持向量机（SVM）的回归方法。与传统的回归方法相比，SVR 通过最大化边界内的数据点数量来寻找最优回归函数，从而更好地处理非线性关系和异常值[42]。支持向量机适用于分类和回归问题，通过寻找一个最优超平面将数据点分隔开或拟合数据，在回归问题中，SVR 通过寻找一个最优回归函数拟合数据。与 SVM 类似，SVR 使用核函数将数据从原始特征空间映射到一个高维特征空间，从而更好地处理非线性关系。常用的核函数包括线性核、多项式核、高斯径向基核（RBF 核）等。

建立优化问题：SVR 的目标是找到一个最小化模型复杂度和误差的回归函数。通过求解一个凸优化问题寻找该回归函数，解决该问题的目标是最小化模型参数和松弛变量之间的平衡。

$$\min_{w,b,\xi,\xi^*} \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{N}(\xi_i + \xi_i^*) \tag{2.13}$$

$$subject\ to \begin{cases} y_i - w\cdot\phi(x_i) - b \le \varepsilon + \xi_i \\ w\cdot\phi(x_i) + b - y_i \le \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \ge 0, i = 1, \ldots, N \end{cases} \tag{2.14}$$

式中：w是回归函数的权重向量，$\phi(\cdot)$是核函数将数据映射到高维空间，b是偏置项，ξ_i和ξ_i^*是松弛变量，y_i是样本x_i的真实标签，C是正则化参数，ε是边界宽度参数。

SVR 通过最大化边界内的数据点数量寻找最优回归函数。通过引入一个松弛变量允许一些数据点落在边界内，用来处理噪声和异常值。最大化边界内的数据点数量寻找最优回归函数，利用核函数处理非线性关系，通过解决优化问题求解最优回归函数。SVR 在处理非线性关系和异常值时表现出较好的性能。

2.3.6 集成算法

集成算法是一种机器学习方法，通过组合多个基础模型的预测结果来提高整体的预测性能，集成算法通常可以分为 Bagging 算法、Boosting 算法和 Stacking 算法三种主要类型。

Boosting 称为提升算法，通过迭代训练多个弱学习器，加权组合预测结果，以构建一个性能更好的集成模型。Boosting 算法通过反复调整训练样本的权重和

弱学习器的权重，逐步提升模型的预测能力。在每一轮迭代中，Boosting 关注并纠正之前迭代中错误分类的样本，使下一轮迭代中模型更关注于分类错误的样本，从而逐步减小模型的误差。由于学习器每一步的迭代都与上一个学习器有很大关联，所以该算法各学习器之间不是并行训练[43]。因为对错误预测的样本更加重视，所以 Boosting 的效果要比 Bagging 更好，对误差较大的学习器分配更低的权重。Bagging 是一种集成学习方法，通过对原始训练集进行有放回的随机采样，训练多个独立的弱学习器，通过简单的投票或平均等方式组合预测结果。在并行训练多个独立的弱学习器过程中，每个弱学习器使用自助采样的训练数据集进行训练，然后对弱学习器的预测结果进行组合，得到最终的集成模型。通过采样和组合的方式，Bagging 可以降低模型的方差，提高预测的稳定性和泛化能力。使用不同的方式组合弱学习器的预测结果，如投票、平均等。对于分类问题，通过多数投票的方式确定最终的类别标签；对于回归问题，通过取平均值的方式得到最终的预测结果。

在理论上，Stacking 相对于 Bagging 和 Boosting 有一些优势。Bagging 和 Boosting 算法是通过组合多个基础模型的预测结果提高性能，但它们都是直接使用基础模型的预测结果进行加权平均或投票，相比之下，Stacking 选择构建一个新的模型，再次训练多个学习器的预测结果，这使 Stacking 能够更灵活地组合各个基础模型的预测能力。Stacking 算法通过两层模型结构来进行预测。在第一层，多个基学习器被训练使用，每个基学习器针对原始训练集进行独立训练。在第二层，元学习器使用基学习器的预测结果作为输入特征，进行最终的预测。

Stacking 算法的一般步骤如下：

（1）将原始训练集分为训练集和测试集。训练集用来训练整体的 Stacking 集成学习模型，测试集用来测试模型。

（2）每个学习器基于训练集进行独立的学习和预测，将基学习器在训练集上的输出作为新的输入特征，形成一个新的训练集并输入到第二层。使用测试集对基学习器进行预测，并将这些预测结果作为新测试集，通过元学习器进行最终的预测。

在第二层中，使用基学习器的预测结果，训练一个元学习器。元学习器可以是任何机器学习模型，如逻辑回归、支持向量机等。训练元学习器的公式和方法

与单个模型的训练类似，用交叉验证选择最优超参数。为防止过拟合情况，第一层基学习器在训练数据时，采用K折交叉验证法筛选数据,将数据集随机地划分为K等份，每次取 $1/K$作为测试集，余下的作为训练集进行训练，重复K次，得到K组不同的模型参数(每次训练样本和测试样本都在发生变化)，将K组参数值不同的模型用于同一验证集进行预测，得到最终的预测结果。

具体的 Stacking 算法实现过程如图所示，对已经划分好训练集和测试集的样本数据，利用K折交叉验证法将训练集划分K份，K一般取 5，然后轮流将其中 4 份作为训练数据，剩下的 1 份作为验证集训练模型。产生 5 组不同的训练集和验证集，同时得到同一模型的 5 种不同的参数配置，将 5 中不同参数配置的同一模型分别对各自对应的验证集进行预测，将 5 组验证集的预测结果拼接在一起组成新的训练集，对同一测试集，5 种不同参数配置的同一模型又会得到 5 组预测值，将 5 组预测值取平均值，得到针对该模型在测试集上的预测值。假设第一层有三个基学习器，仍然使用 5 折交叉验证法，在上述交叉验证的基础上需要重复两次，总共得到 3 组新的训练集，3 组测试集，然后分别将这 3 组训练集和 3 组测试集合并，得到最终的训练集和测试集，数据的规模相当于原来的 3 倍。将最后的训练集和测试集分别用于第二层元模型的训练，即可得到第二层模型。两层模型的参数均确定后，预测时直接将数据代入即可。

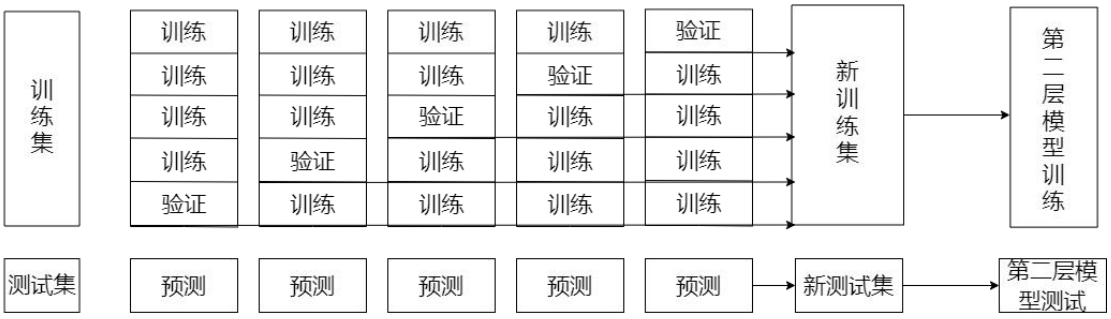

图 2.1 Stacking 集成学习流程图

2.4 模型评价指标

在实际地震人员死亡的预测和评估中，通常将预测值和实际值放在同一数量级上比较，数量级在地震学领域广泛使用，普遍采用数量级概念来衡量地震人员死亡的评估效果，是业内人士通用且广泛认可的指标。小规模的地震和大规模的地震破坏程度差异巨大，将预测值和实际值放在同一数量级上，可以更好地进行

比较，观察模型对地震强度的预测是否处于合理的范围，有助于提高模型结果的适用范围和可信度。地震的人员死亡通常呈现不均匀的分布，将预测和实际值放在同一数量级上，有助于缩小地震人员死亡数据的范围，使得模型的预测结果更加可控和准确。

基于统计学分析指标，选取了三个用于衡量模型性能和预测能力的指标：均方根误差（RMSE）、平均绝对误差（MAE）和平均绝对百分比误差（MAPE），这些指标常被用来评估模型的预测效果。

表 2.1 评价指标

	指标	计算方法
均方根误差	RMSE	$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\hat{y}_i)^2}$
平均绝对误差	MAE	$\frac{1}{n}\sum_{i=1}^{n}\lvert y_i-\hat{y}_i\rvert$
平均绝对百分比误差	MAPE	$\frac{1}{n}\sum_{i=1}^{n}\left\lvert\frac{y_i-\hat{y}_i}{y_i}\right\rvert$

其中n是样本数量，y_i是实际观测值，$\hat{y}_i$是模型预测值，$\bar{y}$是观测值的平均值。RMSE 是观测值与预测值之间差异的平方的平均值的平方根，RMSE 越小，模型预测性能越好。MAE 是观测值与预测值之间差异的绝对值的平均，MAE 衡量预测误差的平均大小，对异常值不敏感，MAPE 的值越小，表示预测模型的准确性越高。

3 中国大陆地震灾害数据来源及处理

3.1 数据来源

本文整理各地震局发布的烈度图，梳理李懿龙等人编制的1950-2018年中国大陆破坏性地震综合目录（Mainland China Composite Damaging Earthquake Catalog，MCCDE-CAT）[44]，并整理中国大陆地震灾害损失汇编[45-49]和历史震例，补充震级、震中烈度、发震时刻、震中位置、震源深度、各烈度区面积和死亡人数等信息，建立矢量化的地震人员死亡数据库。因地震人员死亡主要发生在震中烈度为Ⅷ度及以上的地震中，本文选取1950-2022年117次震中烈度Ⅷ度及其以上的地震进行建模。人口数据包括MCCDE-CAT的地震受灾人口数据和人口栅格数据[50]，人口栅格数据来自美国橡树岭国家实验室（https://landscan.ornl.gov），空间分辨率为1 km。

3.2 数据处理

已有的地震数据存在一些冲突的地方，如：地名变换引起的错误，地名在不同的语言中存在不同的表达方式，会导致地名的混淆或错误，应使用标准化的地理信息系统或地名数据库，确保在数据处理中使用一致的地名标准；在地震数据的采集、录入或处理阶段，人为错误是一个常见的问题，为减少人为错误，可以通过使用自动化工具来规范和验证数据，确保数据的一致性和正确性；技术层面的错误，在数据采集和处理的技术过程中，会发生各种技术层面的错误，传感器故障、数据传输问题等，定期的系统检查、校准更新是确保数据质量的关键。坐标系统不一致问题，地震数据可能涉及到不同的坐标系统，而坐标系统的不一致会导致定位错误，在整合数据时，确保使用相同的坐标系统，或者进行必要的坐标转换，以避免误差[51]。

3.2.1 缺失值处理

缺失或不完整的数据，会导致分析的不准确性。在处理数据时，需要考虑并处理缺失数据的情况，可以采用插值或其他方法来填补缺失值。导致地震数据缺失的原因多种多样，在不同时间、地点和设备上采集数据的过程中会出现数据不完整的情况；数据录入或传输过程中可能发生错误；技术限制导致无法获得完整的地震数据等。处理步骤如下：

（1）识别缺失值，地震数据主要是死亡人数、受灾人口、受灾面积等字段存在缺失值，通过 excel 来检查每个特征的缺失情况。

（2）对于缺失的数据，查阅官方发布的烈度图并进行矢量化，以获取补充数据。对于年代较远的地震，没有烈度图无法矢量化时，在保证数据对整体分析影响较小时可以选择直接剔除。

3.2.2 异常值处理

地震数据的异常值主要发生在数据处理过程，在数据清洗、单位转换或分析阶段，操作不当或单位不统一均会导致异常值的产生。处理步骤如下：

（1）以中国地震台网中心数据为基础，对存在异常的数据检查并修正。

（2）对于地震台网中心未提供的数据，对该地震的烈度图进行矢量化，通过矢量化修正异常值。

3.2.3 数据矢量化

数据矢量化是利用 Arcgis 内置的矢量化工具将栅格数据（烈度图）转换为矢量要素的过程，用矢量化的数据解决数据缺失和数据异常的问题。烈度图包含死亡人数、受灾面积、烈度、震级等信息，考虑到数据字段长度的限制，将矢量后的新字段设置为双精度型数据，并将修正的信息输入到新字段中。通过地理配准将不同的烈度图转换到相同地理坐标系统，将数据集的位置、比例尺和方向与标准地图对齐。为准确定位地震烈度图在地理空间中的位置，选择 7 对控制点进行地理配准，配准后的烈度图可与其他地理数据进行叠加分析。当烈度图上没有明显的经纬度标记时，通过便民查询网（https://jingweidu.bmcx.com）搜索烈度

图上的明显地物（比如村庄、城市、道路等）并确定它们的地理坐标，使用这些地物作为控制点进行地理配准。面分割工具适用于对面状要素进行分割，基于几何或属性条件将面状要素分割成更小的面状要素，为获取不同烈度区的受灾人口和受灾面积，使用面分割工具将烈度图划分，矢量后将数据导出便于进一步分析。

地理信息系统（GIS）处理的是地球上的空间信息，正确理解和使用坐标系统直接影响数据的精度。坐标系统主要划分为地理坐标系统和投影坐标系统两大分支，在地理坐标系统中，其定义依赖于两个关键元素：大地基准面及其相关的地图投影参数。大地基准面的选择涉及特定的地球椭球体模型以及与之配套的转换参数集，这些参数共同描述地球形状和尺寸的近似数学模型。鉴于地球表面并非完美的数学曲面，而是一个复杂的三维地形，为便于精确的空间定位、测量和地图制作，需要将这个不规则曲面通过某种数学变换映射到可以精确描述和表达的平面上，通常情况下，使用椭球体来近似地球的形状，椭球体的几何特性主要由赤道半径、极半径和扁率三个核心参数描述，赤道半径代表椭球体沿赤道方向的最大直径；极半径是地心到两极点的距离，通常小于赤道半径；扁率是衡量椭球体扁平程度的关键指标，描述了椭球体相对于完美的球体在形状上的偏离程度。这三要素共同决定近似地球形状的数学椭球体模型的具体形态。常见的地图投影包括等角投影、等积投影和等距投影等。

在地理坐标系统中，地面点位通常借助经度和纬度标识，这些坐标既可以采用十进制度数表达，也可以转换为更传统的度分秒格式。然而在实际的测量和制图过程中，必须将球面坐标点转换为平面坐标。投影坐标系统在这个过程中起到桥梁作用，通过数学转换使用直角坐标或极坐标表示平面上点的位置信息。在 GIS 中，地理坐标系统和投影坐标系统相辅相成，对空间分析和地图制图至关重要。在 ArcGIS 桌面产品中，用户可以根据实际需求选择合适的坐标系统，并进行相应的转换和投影，以确保数据的准确性和可视化效果。墨卡托投影和阿尔伯斯投影是两种常用的地图投影方法，在映射地球表面到平面上的方式和特性上各有所长。墨卡托投影是一种等角投影，可以在地图投影上保持任意两点的角度不变[52]。阿尔伯斯投影是一种等积投影，在投影过程中保持地图上不同区域的面积比例[53]。墨卡托投影在面积上有较大的变形，而阿尔伯斯投影在形状和方向上可能有一些变形，但主要目的是保持面积比例。

本文先后使用墨卡托和阿尔伯斯投影计算各烈度区面积，发现两者相差较大，墨卡托投影得到的面积偏差较大，而阿尔伯斯投影作为等面积投影，在横跨较大纬度范围的地理区域投影时，能够确保投影至地图上每个区域的面积与其在地球表面的实际面积保持一致，该投影在标准纬线附近形状和方向保持较好，但随着远离标准纬线，形状和方向的变形会增加。在 Arcgis 中，默认的阿尔伯斯投影有亚洲北半球阿尔伯斯等面积投影，然而该投影中央经线为 95 度，双标准纬线分别为 15 度、65 度，而中国地图的中央经线位于东经 105 度，两条标准纬线分别为北纬 25 度和北纬 47 度，故根据中央经线和标准纬线新建一个符合中国区域的阿尔伯斯等面积投影，具体参数如下：

投影坐标系： Asia_North_Albers_WGS_China
投影： Albers
False_Easting: 0.0
False_Northing: 0.0
Central_Meridian: 105.0
Standard_Parallel_1: 25.0
Standard_Parallel_2: 47.0
Latitude_Of_Origin: 0.0
线性单位： Meter

由于地理坐标系统在高纬度处的纬度线汇聚，未经投影的地图在高纬度地区会出现形状和面积的扭曲，从而导致图上的对象形状和大小出现变形。选择合适的投影方法取决于地图的使用目的以及研究区域的特点。

为验证使用合适的投影坐标系，本文以 2010 年 4 月 14 日青海玉树 7.1 级地震为例，计算该地震各烈度区的面积，世界墨卡托投影（World Mercator）是一种圆柱投影，通常用于显示全球范围的地图。通用横向墨卡托投影（UTM）是一种锥形投影，用于分带投影地球表面。每个带都有自己的投影。

表 3.1 不同投影坐标系计算的各烈度面积

烈度	VI	VII	VIII	IX
WGS_1984_World_Mercator/km^2	28990.57	6598.04	2039.69	243.12
WGS_1984_UTM_Zone_47N/km^2	20338.31	4618.08	1427.23	170.37
Asia_North_Albers_WGS_China km^2	20337.61	4617.92	1427.18	170.36

2010年4月14日青海玉树7.1级地震震中位于北纬33.1度、东经96.6度，WGS_1984_UTM_Zone_47N坐标系的精度范围（东经）为96度-102度，中央经线经度为99度，相对契合玉树地震，由表3.1可知世界墨卡托投影得到的面积与通用横向墨卡托投影和阿尔伯斯投影相差较大，不适用于计算震区面积，而通用横向墨卡托投影在不同经纬度地区需要用不同分带投影，相对繁琐，中国幅员辽阔涉及多个分带投影。为方便统一计算，本文使用阿尔伯斯投影来计算震区面积。

最终矢量化得到的中国大陆地震烈度图显示了地震在不同地区的强度分布，从图 3.1 中可以看出，地震在地理空间上的分布有明显的地域特征。对于估计可能造成的死亡人数和灾害管理的规划至关重要。

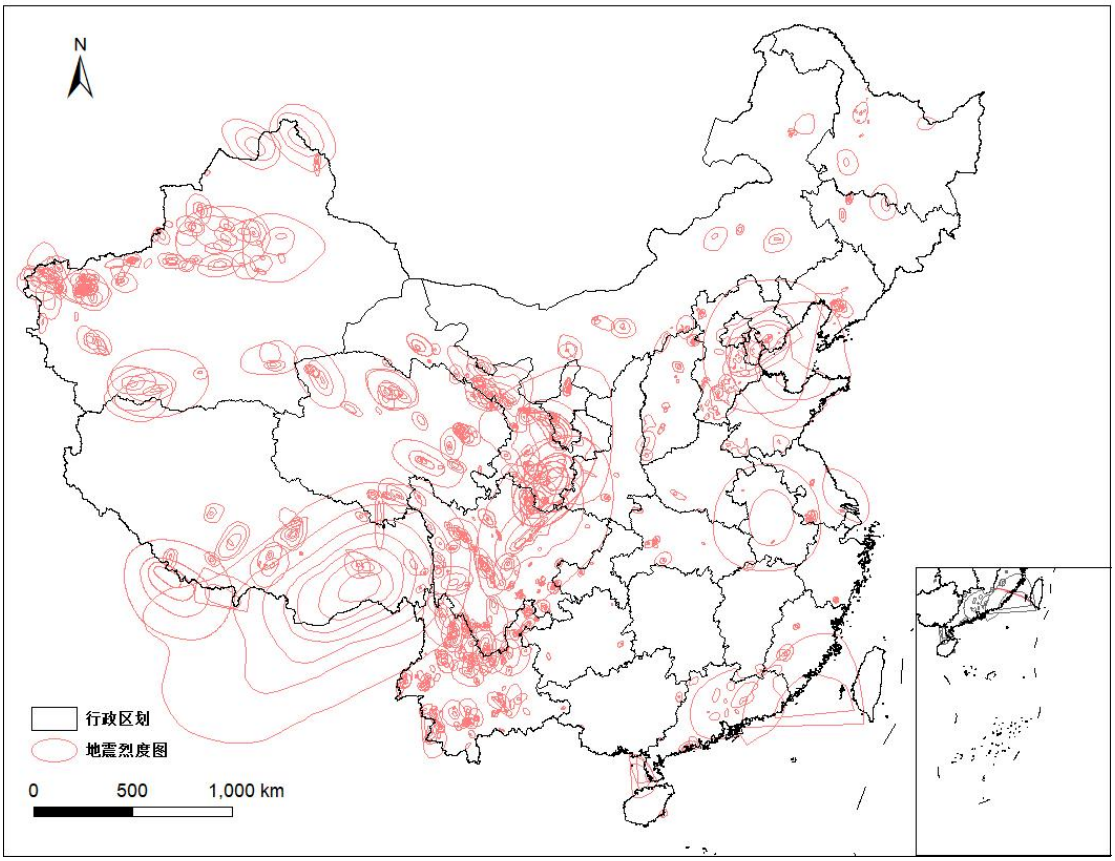

图 3.1　1950-2022 中国大陆地震烈度图

注：该图基于全国地理信息资源目录服务系统（https://www.webmap.cn）的矢量地图数据制作，底图无修改

4 地震人员死亡影响因素

4.1 地震人员死亡因素概述

地震是由地球内部的地壳运动引起的自然现象，地震发生的地方称为震源，通常是地壳中断层面的滑动或地壳板块之间的相对运动，地震数据的基本特征包括震级、震源深度、震中烈度、人口密度和发生时间等。这些特征在地震灾害研究中扮演着关键的角色，对于了解地震的影响和制定相应的灾害管理策略至关重要[54]。

地震震级是用来衡量地震活动强度的度量标准，它揭示了地震发生的剧烈程度。震级数值越大，意味着地震释放出的能量越庞大，带来的破坏程度也愈发显著增强。地震能量的释放对于地表构造和建筑物的影响至关重要，正是这种能量的大小直接影响到地表结构遭受破坏的程度以及建筑物在地震中的受损状况，从而对灾区居民造成不同程度的影响和伤害。

震源深度指地震发生的深度，分为浅源地震、中源地震和深源地震，震源的深度不同，地震的性质和影响也会有所不同。浅源地震（0 到 60 公里深度）通常对地面建筑物造成较大破坏，而深源地震的影响相对较小，深入了解震源深度有助于预测地震的破坏程度和可能的影响范围。

烈度表示地震在不同地点产生的感知强度，通常使用烈度图表示，震中烈度表示地震中的最高烈度，反映震中区域的破坏程度。烈度通常分为十二级，从较轻的破坏（V 级）到毁灭性的破坏（XI 级）。震中烈度越高，地震对建筑物和人口的影响越严重。烈度图反映地震在地表上的影响程度，通过分析烈度图，可以了解地震的空间分布、强度分布以及受影响区域的特征，而烈度图的制作通常需要地理配准和矢量化。

人口密度表示单位面积上的居住人口数量。在地震研究中统计受灾人口时，人口密度是一个关键的考虑因素，人口密集区的地震意味着更多的人员死亡和财产损失，因此在分析地震影响时需要考虑到这一因素。

地震发生时间通常分为白天和黑夜，白天或黑夜与死亡人数之间关系密切，白天人口密度较高，因为白天人们通常在城市中工作、学习和进行其他活动。地震发生在白天，会导致更多的人受到影响，白天人们通常在公共建筑内，而夜晚更多在家中，不同建筑结构和质量对地震造成的人员死亡产生不同影响。白天时紧急救援和应急服务通常更容易组织和展开，受灾人员更容易获得及时医疗救援，从而减少死亡人数。白天时，人们更容易察觉到地震的迹象，更容易采取逃

生行动。相比之下，夜间人们处于睡眠状态，在地震发生时反应较慢，增加了死亡风险。同时不同地区对白天和夜晚的划分，因纬度、季节、文化习惯不同而存在差异，这种差异性导致在一些地区，白天和夜晚的时间划分相对较为固定，而在另一些地区则随着季节的变化而有所不同。

在进行地震数据的分析时，更全面地了解地震的性质，有助于制定相应的防灾和救援策略。地震数据的分析涉及多个学科领域，包括地震学、地质学、地理信息系统、地震工程等，综合分析有助于深入了解地震的性质、准确评估地震风险，并制定相应的防灾措施。

在地震数据中使用 seaborn 的多变量联合分布图（pairplot）可帮助直观地理解不同变量之间的关系。seaborn 是一个基于 matplotlib 的 Python 数据可视化模块，pairplot 通常用于绘制多个变量之间的散点图和直方图，以及显示变量之间的相关性。引入虚拟变量区分地震发生的时段，将发生时间分为白天（06:00-20:59）和黑夜（21:00-05:59）两个时间段，0 表示白天，1 表示夜晚，通过 hue 参数对数据进行颜色编码，根据发生时间的不同值来区分不同的数据点。

如图 4.1 所示，对角线显示各变量的分布，散点图显示两个变量之间的相关程度。通过联合分布图观察变量之间的线性或非线性关系，变量之间相相关性越高，两个变量之间的散点图越呈线性分布。震级和震中烈度存在明显的线性关系，当两个相关性强的变量同时参与建模时，容易引发多重共线性，降低最终预测结果。同时由于地震死亡人数从几人到上万人离散程度较大，个别大地震在图中表现为离群点，如图中明显看到死亡人数为 24 万多人的唐山大地震，离群点影响对整个图形趋势的判断。总体来说，pairplot 提供了一个全面的视图，帮助理解地震数据中不同变量之间的关系，发现模式和异常情况。通过观察图中的趋势，更好地了解地震数据的特征和相互关系。

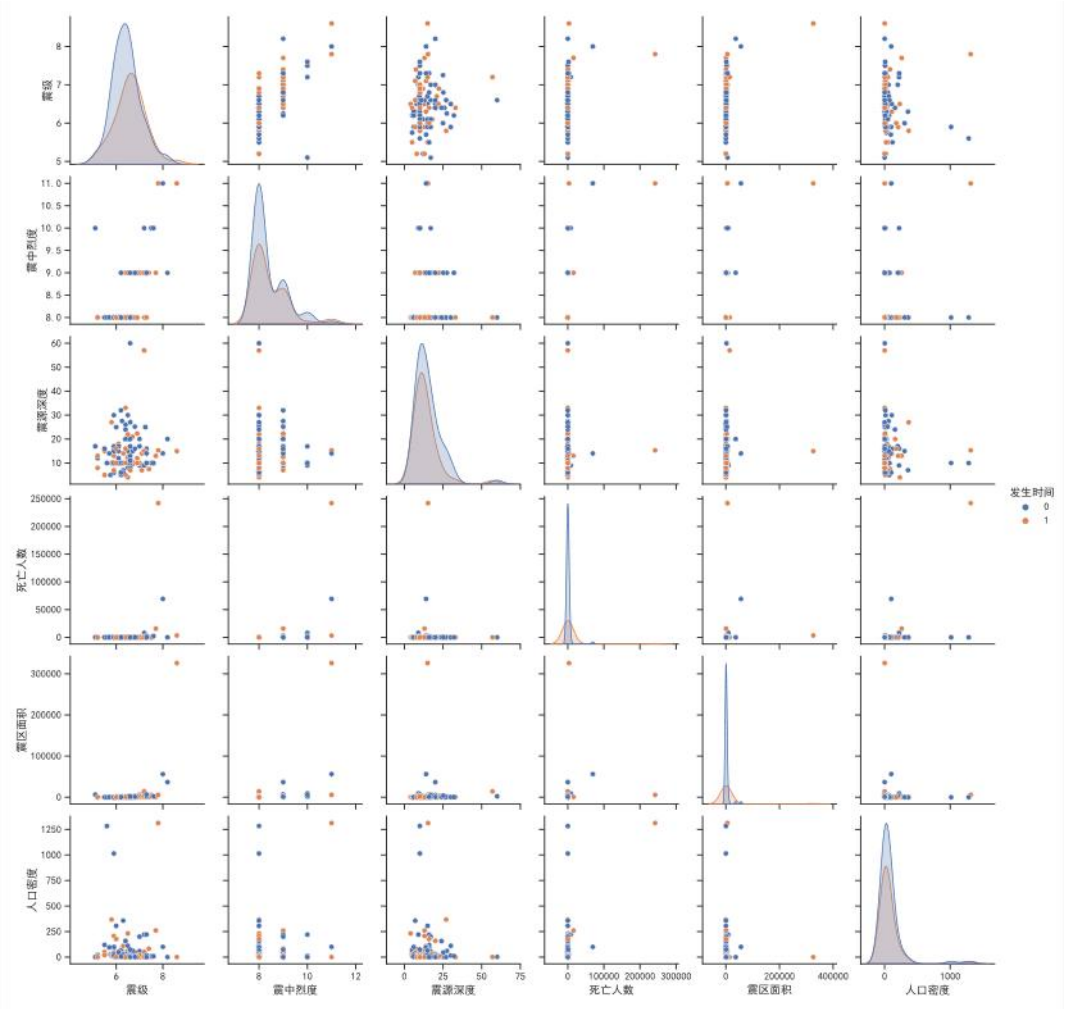

图 4.1 多变量散点图

针对离群点问题，对死亡人数进行对数处理，降低数据的尺度，使变量间的关系更容易可视化和理解，使用对数变换后的数据创建散点图，对数变换有助于拉伸较小数值，同时压缩较大数值的差异，从而更好地展示数据的变化。如图 4.4 与图 4.5 所示，经过对数变换后，地震死亡人数与震级和震中烈度的线性关系更明显，但这种关系受到较大的离散性影响，存在一定的不确定性。与之相反，震区面积和人口密度与死亡人数之间并没有清晰的对应关系，难以通过简单的函数描述二者之间的关联。随着震中烈度和震级的增大，死亡人数上升趋势明显，这表明地震的死亡人数与震级和震中烈度之间的关系更为密切，成为人员死亡评估的重要参数。

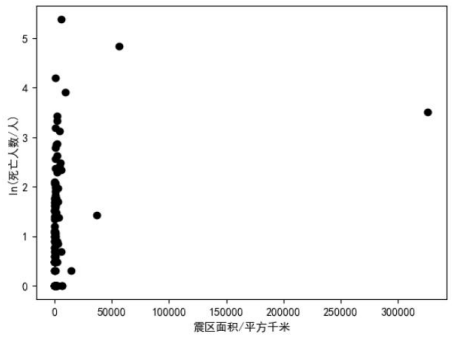

图 4.2 震区面积与死亡人数散点图

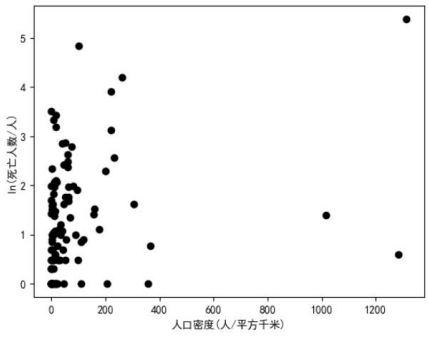

图 4.3 人口密度与死亡人数散点图

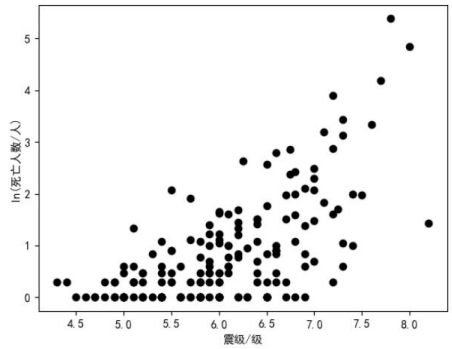

图 4.4 震级与死亡人数散点图

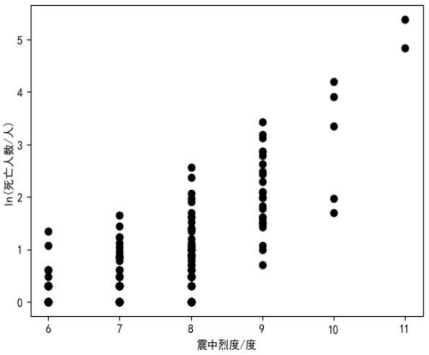

图 4.5 震中烈度与死亡人数散点图

另一方面，死亡人数与发震时刻之间的关系，呈现出模糊的定性关系。这种情况可能是因为选择的地震数据是Ⅷ度区及其以上的震例，都是相对较大规模的地震，因此地震的发生时间可能对死亡人数的影响不太明显，这种情况下，地震的死亡人数可能更多地受到其他因素的影响，需要进一步考虑其他影响因素，如紧急救援响应、人口密度等因素。鉴于上述分析，深入研究地震人员死亡的评估方法，同时在下一节采用严谨的特征选取理论综合考虑地震人员死亡因素，以建立更为准确和全面的人员死亡评估模型，有望提高地震风险评估的精度和可靠性。

4.2 要素选取

根据区域灾害系统理论，灾害是致灾因素、承灾体和孕灾环境之间的相互作用的结果[55]。地震引发的人员死亡受多种因素共同影响，但所有因素都纳入考虑可能导致信息冗余和增加误差，进而影响模型的准确性。

在地震人员死亡评估研究中，输入变量的选择直接影响评估效果好坏，由于地震数据各变量之间的关系复杂，单一的线性关系往往难以拟合出良好的评估模

型，过去研究通常依赖于历史震害经验来确定关键的人员死亡影响因素，但这种方法缺少严谨的数据支撑，并且难以处理复杂的非线性关系，此外，各变量之间还存在一定程度的相关性，因此如何选择变量成为评估研究的重难点。

机器学习在处理地震数据时具备从数据中学习复杂模式的能力。地震数据通常具有大量的非线性特征和复杂结构，传统的线性方法往往难以捕捉这些复杂规律，相比之下机器学习模型能够自动学习和识别数据中的非线性模式，使其更适应地震数据的特征。机器学习算法具有自动从数据中学习最相关的特征能力，而无需手动定义特征。由于地震数据可能包含大量的维度和特征，其中一些可能对地震预测或模拟更为重要。地震数据的性质可能随时间和地理位置的变化而变化，机器学习模型能够适应这些变化，无需手动调整模型的参数。这种适应性使模型更具泛化能力，可以用于不同地区和时间范围的地震数据。然而机器学习模型的性能受到数据质量、模型选择、特征工程等因素的影响，在应用机器学习地震数据分析之前，必须深入了解数据、优化数据质量、选择合适的特征，并谨慎选择和优化模型以解决特定的地震科学问题。鉴于机器学习在处理非线性问题上具有灵活性和适应性，并能够识别复杂模式，本节采用机器学习中的随机森林算法对影响因素的重要性进行评估，从而提高对地震死亡评估的预测能力。

特征的重要性分析是选择特征的一种常见方法[56]。通过评估每个特征对模型性能的贡献，可以识别和选择对任务最有帮助的特征，提高模型的泛化能力并减少过拟合的风险。在对历史地震灾害数据进行系统梳理和严谨筛选的基础上，首先分析影响因素的来源，并对潜在的影响因素进行初步探讨与分析，根据理论指导实践经验，使用集成学习算法对这些影响因素进行重要性评估，下面是使用集成算法分析特征重要性并进行特征选择的一般步骤概述：

（1）训练模型。使用机器学习模型（例如随机森林、梯度提升决策树等）训练数据集。

（2）获取特征重要性。模型训练完成后，通过模型提供的特征重要性分数来评估每个特征的重要程度，分数基于特征在模型训练中的贡献程度来确定。

（3）选择重要特征。根据特征重要性进行排序，选取排名靠前的特征作为最终使用的特征集。

（4）验证选择的特征。使用交叉验证来验证所选特征集是否改善了模型性能，尝试不同的特征集组合以找到最佳组合。

特征重要性的计算方法和可视化工具因模型而异。在随机森林中，特征重要性通常是通过观察树的分裂点上特征的不纯度减少来计算的。

对数据集进行数据预处理，使用 Anaconda Navigator 软件的 jupyter 模块分别建立 RF、CART、GBDT 和 AdaBoost 模型，在 Jupyter Notebook 中进行 Z-score 标准化需要使用 Python 的 scikit-learn 库，该库提供了 StandardScaler 预处理方法来进行标准化处理。

$$x_{normalization} = \frac{x-\mu}{\sigma} \tag{4.1}$$

式中：x表示特征参数，μ为特征参数的平均值，σ表示特征参数的标准差，$x_{normalization}$表示归一化后的特征参数。

特征选择和参数调整需要根据实际问题和数据集的情况进行，机器学习算法 scikit-learn 库中大多有默认参数，默认参数在许多问题上能够提供良好的性能，针对模型性能对模型参数调整。在进行参数调优时，先尝试使用默认参数，然后根据模型在验证集上的表现进行调整，使用交叉验证等技术评估不同参数组合的性能，交叉验证可以用 GridSearchCV 或 RandomizedSearchCV 来实现。对于 AdaBoost 和 GBDT 算法主要有两方面需要调参：框架参数和基础学习器（弱分类器）参数。n_estimator 为决策树的数量，learning_rate 为学习率，本文决策树的数量设置为 50，学习率设置为 0.5，基础学习器的参数指每轮迭代中弱学习器的参数，弱学习器的参数与框架参数相互关联，在 scikit-learn 库中，DecisionTreeClassifier 是常用的基础学习器，每个决策树的最大深度默认为 None，Subsample 表示每个树训练样本的子采样比例，设置为 0.8 可以避免过拟合，alpha 参数只有 GradientBoostingRegressor 有，将参数设为 0.9 控制损失函数。随机森林决策树数量设置为 100，n_jobs 设置为-1 表示使用所有可用的 CPU 核心，max_features 设置为 None 表示考虑所有特征，criterion 表示节点的划分标准，选用 gini 系数。

表 4.1 四种集成学习算法的参数值

Random Forest	CART	GBDT	AdaBoost
n_estimators = 100	max_depth = 10	n_estimators = 100	estimator= trees classifier
n_jobs = -1	max_features = None	subsample = 0.8	algorithm=SAMME.R
max_features = None	n_estimators = 100	learning rate = 0.5	learning rate = 0.5
criterion = gini	criterion = gini	alpha = 0.9	number of estimators = 50

确定好模型参数后，使用交叉验证训练优化模型，采用简单交叉验证法将数据 80%分为训练集，20%为测试集，模型在划分成不同训练集和测试集的数据上进行评估。AdaBoost 和 GBDT 都是基于加法模型和前向分布算法，在原理上有一些相似之处，但也有一些关键的区别。AdaBoost 通过迭代训练弱学习器，每一轮都调整样本权重，使前一轮错误分类的样本在下一轮得到更多关注。最终将各个弱学习器的加权求和，权重由各个弱学习器的表现决定。GBDT 也通过迭代训练弱学习器，但每一轮目标是拟合前一轮的残差，由于地震中离群点的数据较多，离群点被错误分类，AdaBoost 受其影响在迭代中获得更高的权重，最终预测准确性略低一些。

CART 是一种决策树算法，随机森林是多棵决策树的集成，GBDT、AdaBoost 通过组合多个弱学习器来构建一个强学习器，虽然四种算法都是基于决策树的方法，但工作原理和训练方式有一些关键的区别，CART 作为一种单独的学习器，受限于其单一模型的表达能力，随机森林引入了随机性，通过在每个节点处随机选择特征子集来构建树，这增加了树之间的多样性，有助于提高模型的泛化性能。GBDT、AdaBoost 通过迭代学习的方式使集成学习算法更加灵活，能够更好地适应数据，随机森林、GBDT 和 AdaBoost 有许多可调参数，例如树的数量、深度等，经过仔细调整参数后，集成模型能够在性能上超过单一的 CART 模型。

总体而言，决策树是一种简单而直观的模型，适用于某些类型的数据。但在地震数据集上，相较于随机森林，其平均精度略低，决策树容易过拟合，对于复杂问题的泛化能力可能有限。GBDT 有更大的模型复杂度，每一轮都在拟合前一轮的残差，使它有更强的拟合能力，在训练集上的表现更好，但会增加过拟合的风险。AdaBoost 对异常值更敏感，AdaBoost 每一轮通过调整样本权重来关注错误分类的样本，如果某个样本异常，它可能会在后续轮次中得到更多的关注，导

致模型过度关注异常值。如表 4.2 所示，随机森林的平均精度为 0.825，比 CART 算法的 0.784、GBDT 的 0.803 和 AdaBoost 算法的 0.798 要高，表明在处理地震数据上，随机森林表现更好，随机森林是一种强大的集成学习方法，通过组合多个决策树来提高模型的性能。随机森林表现良好，具有较高的平均精度，说明随机森林对于复杂数据集和特征的处理能力较强，所以在特征选择过程中使用随机森林算法选取特征。

表 4.2 四种集成学习算法的准确度

Algorithm	Random Forest	CART	GBDT	AdaBoost
Mean accuracy	0.825	0.784	0.803	0.798

随机森林的特征重要性权重可以帮助确定哪些特征对模型性能贡献最大。特征重要性通过随机森林中各个决策树的节点分裂贡献累积计算得出，衡量每个特征对模型预测能力的影响程度。本文使用基尼系数计算特征重要性，一个特征的重要性越高，说明在模型中使用该特征进行划分时，对最终预测结果有更大的影响。通过识别重要性较低的特征，可以帮助简化模型，减少过拟合的可能性。特征重要性的解释会受到一些限制，如存在共线性的特征会导致两者之间重要性分配不准确。因此，在使用特征重要性进行分析时，综合考虑基尼系数，确定哪些特征对模型性能的贡献最大，根据特征重要性的排序选择最重要的特征，从而降低模型的维度，提高模型的解释性，并改善模型的泛化性能。

传统线性模型在处理非线性、复杂、高维的数据时，无法准确反映变量之间的关系，随机森林确定特征重要性的过程中，会考虑特征之间的相互作用和非线性关系，因此能够更好地处理非线性、复杂和高维度的数据。这是随机森林的优势之一，可在评估特征重要性时更全面地考虑数据的结构。但当特征之间存在相

关性时，重要性会相互抵消，影响对特征的理解，散点图已经观察到震级和震中烈度存在明显的线性特征，两者的相关性削弱了他们的重要性权重，所以在图中表现为人口密度重要性权重最大，震区面积次之，震级和震中烈度排第三和第四。

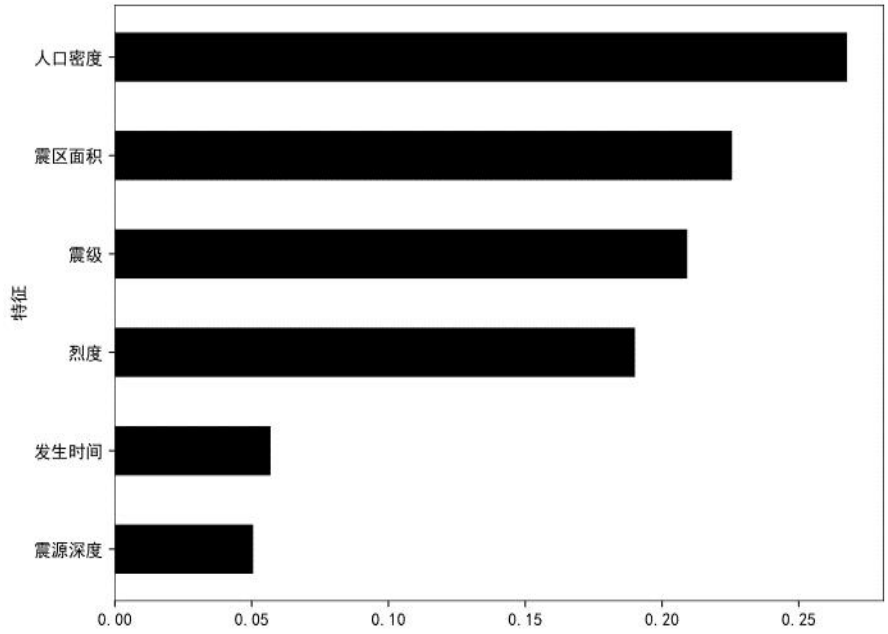

图 4.6 特征的重要性权重

考虑到地震人员死亡主要发生在Ⅷ度区及以上，震区面积选择所有达到和超过Ⅷ度烈度区的总面积之和，人口密度通过计算Ⅷ度及以上烈度区内的人口数量除以相应区域总面积获得，为探究时间因素对地震死亡人数的影响，引入虚拟变量区分地震发生的时段，将发生时间分为白天（06:00-20:59）和黑夜（21:00-05:59）两个时间段, 赋予白天时段的虚拟变量值为 0，黑夜时段的虚拟变量值为 1。通过设置虚拟变量后，利用随机森林算法量化发生时间的重要性，从而更准确地评估不同因素对地震人员死亡的影响程度。从图 4.6 的结果可以看出，震级、震中烈度、人口密度和震区面积是主要的影响因素。在特征选择过程中，特征之间的相关性是一个关键考虑因素，当特征之间存在相关性时，重要性会相互抵消。震级和震中烈度之间存在较强相关性，为避免信息冗余，选择具有更高权重的震级作为模型的输入参数。高人口密度地区地震会导致更多的人员死亡和财产损失，因此人口密度也是一个重要参数。震区面积往往容易被忽略，而较大的震区面积意味着更广泛的地区受到震后潜在影响。根据重要性评价结果，遵循快速评价和避免信息冗余的原则，最终选择震级、人口密度和震区面积作为模型的输入参数。

5 基于 Stacking 集成学习的地震人员死亡评估模型

5.1 数据增强

数据增强旨在提高模型性能和增强其对不同样本的泛化能力，数据增强是通过对原始数据集进行多样性的变换来生成更多样本的技术，被广泛应用于计算机视觉和自然语言处理领域。选择使用数据增强的主要原因是在数据有限的情况下提高模型的鲁棒性。

本文使用 GAN 网络进行对有人员死亡的 117 组数据进行数据增强，具体参数设置如下：

生成器的架构：

输入层：4 个神经元，对应输入噪声的维度；

隐藏层 1：64 个神经元，使用 Tanh 激活函数；

隐藏层 2：128 个神经元，使用 Tanh 激活函数；

输出层：6 个神经元对应生成的数据的维度，使用 Tanh 激活函数。

判别器的架构：

输入层：6 个神经元对应生成器输出和真实数据的维度；

隐藏层 1：128 个神经元，使用 LeakyReLU 激活函数；

隐藏层 2：64 个神经元，使用 LeakyReLU 激活函数；

输出层：1 个神经元，使用 Sigmoid 激活函数。

使用二元交叉熵作为判别器和生成器的损失函数，采用 Adam 优化器，学习率为 0.0001，训练周期设置为 10000。

表 5.1 原始与增强数据集各项统计指标对比

数据集	指标	震级	烈度	人口密度	震区面积	死亡人数
原始数据集	均值	6.5157	8.3684	77.3807	4769.3070	3084
	变异系数	0.6121	0.7753	2.0092	3.0896	23.4353
增强数据集	均值	6.5561	8.3760	50.5863	4206.8237	4618
	变异系数	0.7110	0.5411	2.7652	2.3257	20.4667

由表 5.1 可知，原始数据集与增强数据集各项指标较为接近，而死亡人数均值相差较大主要是因为死亡人数离散程度较大，因而考虑在建模过程中对死亡人数进行对数处理，而增强数据集的变异系数比原始数据集更小，证明数据增强后增加了数据的稳定性，更有利于机器学习拟合地震数据中的非线性规律特征。

5.2 模型选取

（1）Lasso 多元线性回归

Lasso 多元线性回归的参数通常通过交叉验证和网格搜索进行调整。Lasso 回归的主要参数是 alpha，表示正则化项的权重。正则化项有两种形式，L1 正则和 L2 正则，而 Lasso 回归采用的是 L1 正则。以下是确定 Lasso 多元线性回归参数的一般步骤：

①设定候选参数列表。选择一系列α值作为候选参数，通常从一个较小的范围开始，候选列表设为[0.0001, 0.001, 0.01, 0.1, 1, 10]。

②使用交叉验证。采用 10 折交叉验证，将训练数据划分为 10 个子集。然后对每个α值，使用 9 个子集进行训练，剩下的一个子集用于验证模型。这个过程重复 10 次，每次选取一个不同的验证集。

③计算性能指标。对于每个α值，通过交叉验证计算均方误差。

④选择最优参数。通过比较交叉验证的结果选择使性能指标最小化的α值。

最终通过反复的交叉验证和参数搜索，得到 Lasso 多元线性回归的正则化参数最优值为 0.01。

（2）SVR 模型

支持向量回归模型的参数调整涉及核函数选择及核函数内参数调整。在 SVR 中，常用核函数包括线性核、Sigmoid 核、径向基函数核和多项式核。对于支持向量回归模型，调参是一个实验性过程，需要不断尝试和评估。

以下是支持向量回归参数调整步骤：

①核函数的选择。线性核适用于线性可分情况，Sigmoid 核适用于生成神经网络，径向基函数核适用于线性不可分的情况，多项式核适用于处理非线性关系。对于地震数据复杂的非线性规律，本文选择径向基函数核。

②参数调整。针对径向基函数，调整的参数是 gamma 和 C，gamma 参数控制径向基函数核宽度，较小的 gamma 值意味着较宽的径向基函数核和相对平滑的决策边界，gamma 值较大导致较窄的径向基函数核，决策边界更复杂。C 是正则化参数，控制对错误的容忍度。C 值较小导致对误差的容忍度较高，导致模型欠拟合，C 值较大表示对误差的容忍度较低，容易过拟合。

③网格搜索。使用网格搜索在指定的参数范围进行组合搜索，找到使模型性能最优的参数组合。

④交叉验证。使用 10 折交叉验证，以评估每组参数的性能，防止过度拟合，并提供对模型泛化性能的更好估计。

候选值列表是在调参过程中指定一组可能的参数取值。使用候选值列表有助于更有效地搜索参数空间，最终得到的参数最优值如下表所示：

表 5.2 SVR 模型参数

参数	候选值列表	最优值
gamma	[0.1, 0.2, 0.3, 0.4, 0.5]	0.2
C	[1, 5, 10, 100]	10

（3）XGBoost 模型

XGBoost 参数分为三种类型：通用参数、booster 参数以及学习目标参数，通用参数和 booster 参数控制在提升过程中具体的提升方法，学习目标参数控制学习的场景，booster 多被用来指定弱学习器的类型，既可以使用基于树的模型还可用线性模型作为弱学习器。

进行调参时，为每个参数选择一组可能的取值，在模型超参数空间中进行有针对性的搜索，以找到最优的超参数组合，从而提高模型性能，同样使用交叉验证评估每个参数的性能，通过评估每个参数性能，既可以减少方差又可以减少过拟合。

表 5.3 XGBoost 模型参数

参数	说明	候选值列表	最优值
n_estimators	树的个数	[100, 200, 300, 400, 500]	200
max_depth	树的最大深度	[4, 5, 6, 7, 8]	6
min_child_weight	子节点的最小权重	[1, 2, 3, 4, 5]	2
gamma	节点分割最小阈值	[0, 0.1, 0.1, 0.3, 0.4]	0
subsample	下采样比例	[0.5, 0.6, 0.7, 0.8, 0.9, 1]	1
colsample_bytree	树的特征采样比例	[0.5, 0.6, 0.7, 0.8, 0.9, 1]	0.8
reg_alpha	L1 正则化系数	[0, 0.2, 0.4, 0.6, 0.8]	0.2
reg_lambda	L2 正则化系数	[0.2, 0.4, 0.6, 0.8, 1]	1
learning_rate	学习率	[0.1, 0.3, 0.4, 0.5]	0.3

（4）RF 模型

在使用随机森林进行集成学习预测时，有一些重要的参数需要进行优化。主要是对决策树的数量（n_estimators）和每棵树的最大深度（max_depth）进行参数调整。增加树的数量可以提高模型性能，但容易过拟合，限制树的深度有助于防止过拟合。通过交叉验证来选择一个合适的值，可以从较小的值开始尝试，逐步增加深度，观察模型性能。参数设置如表 5.4 所示。

表 5.4 RF 模型参数

参数	说明	候选值列表	最优值
n_estimators	树的个数	[100, 200, 300, 400, 500]	200
max_depth	树的最大深度	[5, 10, 15, 20]	15

（5）LightGBM 模型

LightGBM 中一些关键的调优参数可帮助提高模型的性能、泛化能力和速度。以下是对每个参数的简要解释和调优建议：

进行参数调优时，先选择一个较广泛的范围，通过网格搜索逐渐缩小范围，找到最佳的参数组合，最终的参数选择需要多次迭代，直到找到最优的参数组合。

表 5.5 LightGBM 模型参数

参数	说明	候选值列表	最优值
num_leaves	叶子的个数	[10, 20, 30, 40, 50]	50
max_depth	树的最大深度	[1, 3, 5, 7, 9]	5
min_child_samples	子节点的最小样本	[10, 20, 30, 40, 50]	10
subsample	下采样比例	[0.5, 0.6, 0.7, 0.8, 0.9, 1]	0.8
colsample_bytree	树的特征采样比例	[0.5, 0.6, 0.7, 0.8, 0.9, 1]	0.6
reg_alpha	L1 正则化系数	[0, 0.2, 0.4, 0.6, 0.8]	0.8
reg_lambda	L2 正则化系数	[0.2, 0.4, 0.6, 0.8, 1]	0.4
learning_rate	学习率	[0.1, 0.3, 0.4, 0.5]	0.1

（6）基准模型对比

利用 GAN 网络生成的 2000 组数据进行预测，对死亡人数进行对数处理，输入变量：震级、人口密度、震区，输出变量为死亡人数的对数，80%的数据用于训练，20%用于测试。使用基于最优参数的五个模型进行预测，每个模型预测结果的评价指标对比如表 5.6 所示。

表 5.6 模型效果

模型	RMSE	MAE	MAPE
Lasso	0.305	0.242	34.912
SVR	0.214	0.165	17.389
RF	0.209	0.148	15.378
XGBoost	0.175	0.135	13.934
LightGBM	0.169	0.121	13.335

由表 5.6 可知，通过比较模型的评价指标，对于每一个模型而言，Lasso 回归预测模型是最差的，RMSE 值为 0.305，说明对于该数据集 Lasso 回归模型不太适合；支持向量回归模型 RMSE 值为 0.214，说明模型的泛化能力还有提升的空间；RF、XGBoost 和 LightGBM 集成算法 RMSE 最低说明预测能力良好。综合而言，LightGBM、XGBoost 和 RF 这三个模型的预测能力最为接近，也是泛化能力最强的三个。

5.3 Stacking 集成学习基本思路

在大数据集上进行建模时，使用单一学习器可能导致泛化性能差，计算陷入局部极小。Stacking 模型是一种有效的综合学习框架，通过整合多个模型提高了模型性能，Stacking 模型综合了不同算法的优势，提高了整体预测效果，同时集成学习避免了过拟合风险，不需要过多的调参步骤和特征选择，因此，采用 Stacking 集成学习建模。

（1）对原始数据进行独热编码和归一化等预处理操作。

（2）模型选择和参数调优。第一层基分类器选择 RF、LightBGM、XGBoost、SVR。第二层通常只选用一个学习器作为元模型，通过网格搜索和交叉验证选择最优超参数。

（3）Stacking 集成过程。将训练数据分成 80%用于训练融合模型，20%用于评估模型性能。对每个基分类器使用五折交叉验证方法进行训练，将单个分类器的预测结果拼接在一起，拼接得到预测结果作为第二层分类器的新训练样本。使用新训练样本训练 Stacking 集成框架中的第二层元学习器。

（4）将测试集输入到模型中得到 Stacking 的预测结果。

实验环境和参数调整：

使用 Python 3.8.4 和 Tensorflow-gpu 2.0 版本。

对不同模型的参数进行调优，使用网格搜索等方法选择最优参数。

5.4 分区分烈度的 Stacking 集成模型

除了对单一模型的优化外，另一种提升性能的方法是采用模型融合，Stacking 是一种集成学习方法，基于异质集成将不同类型的个体学习器进行融合。Stacking 算法基本思想是先利用多个基模型对数据集进行预测，然后将基模型的输出数据合并，输入到元模型中，最终使用元模型输出预测结果。

Stacking 算法的流程分为两个层次。第一层由多个基模型组成，每个基模型使用的数据来自原始数据集。第二层元模型的训练数据来源于第一层中每个基模型对原始训练数据进行K折交叉验证得到的预测数据，对于测试数据，元模型的输入是每个基模型对原始测试集进行预测得到的结果，然后取均值并拼接而成的

数据集。这种多层次的组织结构使 Stacking 算法能更充分地利用各个基模型的优势，从而提高整体预测性能。

为进一步提升模型性能和泛化能力，本小节采用了经过处理的数据集和参数优化的模型。通过使用 Stacking 算法对单一模型进行双层融合，上一节评估了 LightGBM、XGBoost、RF、SVR 和 Lasso 这五个模型的预测效果。根据评估指标的比较，这五个模型的预测效果递减，Lasso 的预测效果最差，与其他模型相差较大，因此在融合中不使用该模型。对于 Stacking 算法，通常只选用一个学习器作为元模型，因此本文依次选用 LightGBM、XGBoost、RF 和 SVR 这四个模型中的一个作为元模型剩余的三个作为基模型，分别构建 LightGBM-Stacking、XGBoost- Stacking、RF-Stacking 和 SVR-Stacking 这四个双层模型融合的预测模型。

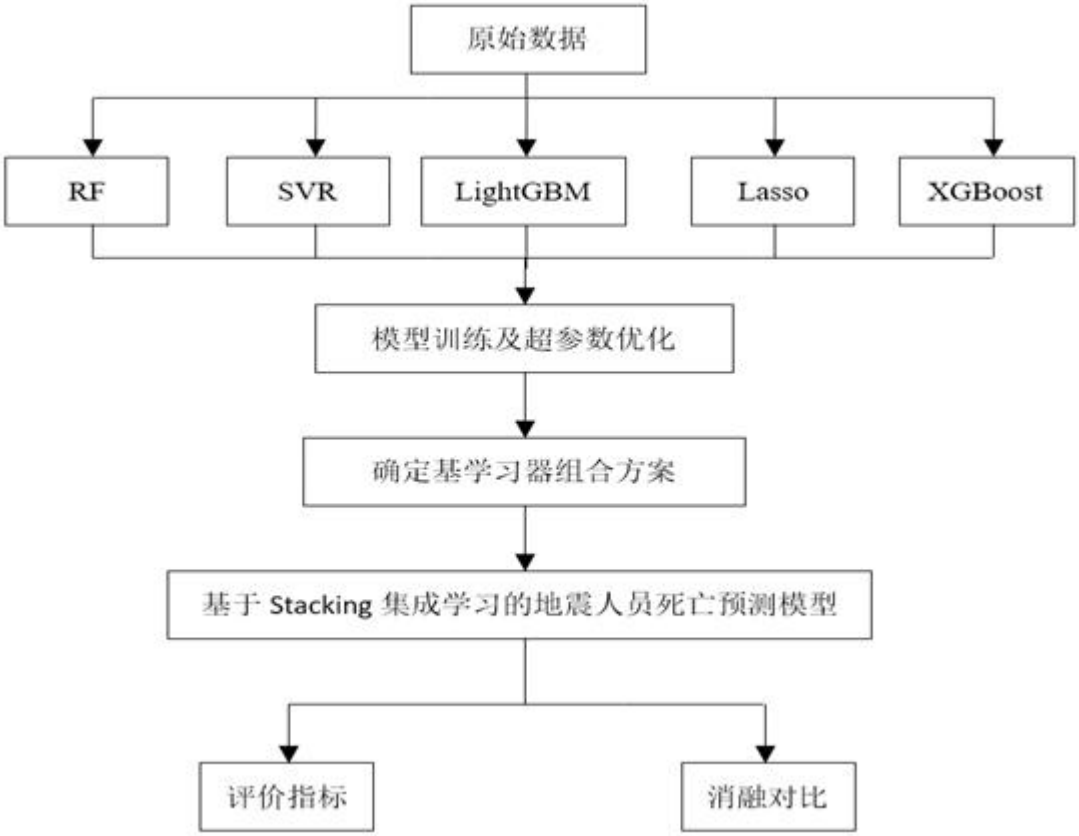

图 5.1 Stacking 模型流程图

5.4.1 地区划分

中国大陆地域的多样性和地震活动频率差异使地震死亡评估和应急准备变得非常重要。地震人员死亡评估需要考虑地理、地质、气候、人口分布、地震活动频率等因素，在地震死亡评估的研究中，将地区划分为不同的区域有助于更细致地分析地方特定的地震风险因素和应对策略。

影响地震人员死亡评估的因素并不是绝对的，每个影响因素互相关联，影响地震死亡的因素是复杂而多样的，包括内在地震因素和外在人为因素。例如有的地区拥有更为完善的紧急响应体系。包括警报系统、紧急疏散计划、急救措施等，有望减少地震伤亡。根据地震人员伤亡影响因素复杂的特点，依据地震数据表现出来的特征进行样本划分。现在主流的经验统计方法是一种基于历史地震事件的统计数据和进行人员死亡评估的方法，该方法依赖过去发生的地震事件，通过分析建立地震影响与人员死亡之间的关系模型，基于历史数据，通过统计学方法建立数学模型，了解地震参数与人员死亡之间的关系。例如回归分析、因子分析或其他统计建模方法。过去的地震历史和地震频率是划分地区的考虑因素之一，而频繁发生地震的地区面临更高的风险，因此划分地区可以使研究更有针对性。

根据人口密度、地理环境、建筑结构等情况，将中国大陆划分为西北(新疆、青海、甘肃、陕西、宁夏)、西南(四川、云南、西藏、重庆、贵州)和大陆东部(其余省份) 三个地区[57]。西北地区地震活动频繁，位于地理边界和板块交汇处，地形复杂多变，救援难度较大，因此具有更高的地震风险。西南地区同样是地震频繁的区域，其中四川盆地和川滇地区地震活动显著，西南地区的地理环境导致地震后容易发生次生灾害事件，比如山体滑坡、崩塌和泥石流。

5.4.2 烈度分档

在地震人员死亡评估模型的构建中，由于震级和震中烈度具有一定的相关性，共同参与建模会引起信息冗余反而导致预测精度下降，因此在建模时进选择震级、震区面积和人口密度作为参数，通过观察不同烈度区的死亡特征来进行烈度分档。

表 5.7 1950-2022 年中国大陆不同烈度死亡特征

烈度	V	VI	VII	VIII	IX	X	XI
有伤亡次数/次	14	153	127	82	27	5	3
有死亡次数/次	3	35	48	62	25	5	3
最大死亡/人	13	22	45	371	2698	15621	242000
最小死亡/人	1	1	1	1	1	94	3300
平均死亡数/人	1	1	2	19	356	4366	104842
死亡总数/人	16	94	253	1536	9618	26197	314527

根据震中烈度、死亡人数得到 1950 年至 2022 年的 411 次地震的人员死亡烈度分布特征（表 5.7）。如表 5.7 所示地震人员死亡主要发生在VIII度区及以上，VIII度区以下平均死亡人数不超过 2 人，震中烈度为VIII度的震例共 82 例，而震中烈度为IX、X、XI的震例共 35 例，在分区基础上将震例划分为VIII度和大于VIII度两类。

5.5 消融模型对比

在上一节基准模型对比中，利用 Stacking 集成学习算法对在测试集上预测效果排名前四的单一模型进行了双层融合共构建了四个融合模型，分别为 LightGBM-Stacking(基模型为 XGBoost、RF 和 SVR，元模型为 LightGBM)、XGBoost-Stacking(基模型为 LightGBM、RF 和 SVR，元模型为 XGBoost)、RF-Stacking(基模型为 LightGBM、XGBoost 和 SVR，元模型为 RF)以及 SVR-Stacking(基模型为 LightGBM、XGBoost 和 RF，元模型为 SVR)。本小节同样使用 RMSE、MAE 和 MAPE 这三个评价指标验证模型在测试集上的预测性能，接下来将对这四个融合模型以及用于构建融合模型的四个单一模型之间进行比较，选择出预测性能最好的预测模型，此外为更好验证本文构建的最优模型还将进行消融分析。

表 5.8 SVR-Stacking 消融模型评价指标

评价指标	RMSE	MAE	MAPE
SVR	0.214	0.165	17.389
RF-SVR	0.211	0.144	16.115
XGBoost-SVR	0.209	0.138	16.012
LightGBM-SVR	0.204	0.133	15.224
RF- XGBoost-LightGBM-SVR	0.197	0.128	14.595

表 5.9 RF-Stacking 消融模型评价指标

评价指标	RMSE	MAE	MAPE
RF	0.209	0.148	15.378
SVR-RF	0.203	0.139	14.773
XGBoost-RF	0.196	0.141	14.784
LightGBM-RF	0.186	0.135	14.654
SVR- XGBoost-LightGBM-RF	0.180	0.104	13.061

表 5.10 XGBoost-Stacking 消融模型评价指标

评价指标	RMSE	MAE	MAPE
XGBoost	0.175	0.135	13.934
RF- XGBoost	0.174	0.133	13.908
SVR-XGBoost	0.174	0.128	13.871
LightGBM- XGBoost	0.172	0.117	13.584
RF- LightGBM-SVR-XGBoost	0.167	0.112	12.125

表 5.11 LightGBM-Stacking 消融模型评价指标

模型	RMSE	MAE	MAPE
LightGBM	0.169	0.121	13.335
RF-LightGBM	0.153	0.119	12.861
SVR-LightGBM	0.150	0.118	12.335
XGBoost-LightGBM	0.149	0.112	12.098
RF-SVR-XGBoost- LightGBM	0.144	0.109	11.916

对比上面四张表可知融合模型预测性能普遍高于单一模型的预测性能，但是也有例外，如 LightGBM 单一模型的预测性能高于 SVR-Stacking 融合模型的性能，原因可能是 SVR 对异常值较为敏感，更适用于小数据预测，由此可知，对模型进行融合是一种提升模型预测性能的一种方法，但得到的结果不一定是最优的。本文构建的其它三个融合模型 RF-Stacking、XGBoost-Stacking 和 LightGBM -Stacking 的预测性能相比单一模型来说得到了提升，这四个模型中的预测性能排序为 LightGBM-Stacking、XGBoost-Stacking、RF -Stacking、SVR-Stacking，预测性能最优的是 LightGBM-Stacking 融合模型，其 RMSE 值为 0.144，MAE 值为 0.109，MAPE 值为 11.916。

5.6 评估效果分析

将构建的 LightGBM-Stacking 模型应用于地震人员死亡评估，模型输入为震级、震区面积、人口密度，输出为死亡人数。针对不同地区和不同烈度的样本数据，建立分区分烈度子模型，根据先前划分的烈度等级，按 8：1：1 划分训练集、测试集和验证集，该验证集不同于机器学习定义的验证集，而是将震例代入模型计算，验证评估模型的预测效果。使用 GAN 网络将训练集扩充到 2000 条，验证集数据及其对应的计算结果汇总于表 5.12 中，通过将模型评估结果与实际情况进行细致对照分析，可以对模型的性能及预测准确性作出分析。

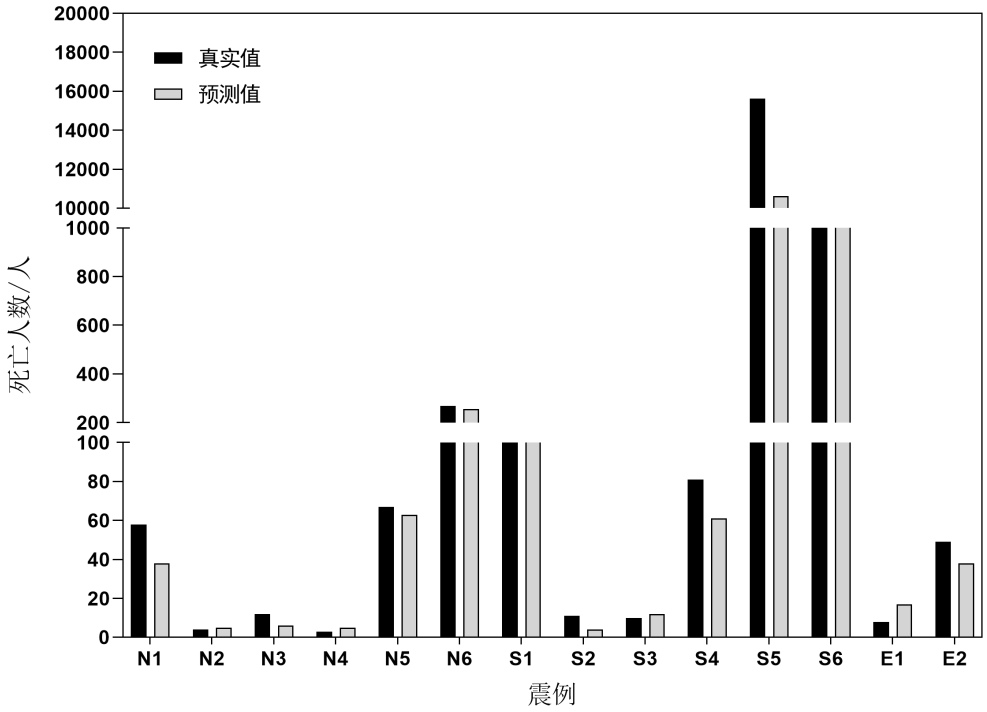

图5.2 Stacking 模型预测结果与真实值对比图

本文采用 Stacking 集成学习，针对不同区域调整相应的最优参数。在西北、西南和东部地区分别随机抽取了 6 个（N1-N8）、6 个（S1-S12）和 2 个（E1-E2）样本用于人员死亡预测，并与真实值进行比较，结果如表 5.5 和图 5.2 所示。由图 5.2 可知，N1、S4、S5 预测值与真实值相差较大。N1 震中为山西原平，地震发生时间为夜晚 10 点 24 分，而在模型特征选择部分没有把发震时间加入模型，可能是模型的预测值小于真实值的原因。S4 震中为云南彝良，S8 震中为云南通海，这两个地方都位于云南省，考虑到云南是一个地理环境变化较大，地质灾害多发，城乡经济差距大，地震造成的山崩滑坡和震后洪水灾害对地震灾害人员死亡影响较大，应进一步研究该地区地震人员死亡评估的方法。

表 5.12 Stacking 模型结果

地区	震中烈度	历史震例	震级	震区面积 (km^2)	人口密度（人/km^2）	死亡人数	预测值	数量级
西北	I = 8	1952-10-08 山西原平	5.5	112.95	83.85	58	38	Y
		1990-10-20 甘肃天祝	6.1	31.56	17.78	4	5	Y
		1995-07-22 甘肃永登	5.8	64.08	20.01	12	5	N
		2015-07-03 新疆皮山	6.5	1091.98	13.35	3	5	Y
	I > 8	1985-08-23 新疆乌恰	7.1	556.52	13.01	67	63	Y
		2003-02-24 新疆巴楚	6.8	2342.19	77.76	268	237	Y
西南	I = 8	1981-01-24 四川道孚	6.9	6.60	24.68	126	108	Y
		1995-07-12 云南中甸	7.3	227.22	10.14	11	4	N
		2008-10-06 西藏当雄	6.6	309.93	4.91	10	12	Y
		2012-09-07 云南彝良	5.7	270.74	121.59	81	61	Y
	I > 8	1970-01-05 云南通海	7.7	628.86	313.58	15621	10612	Y
		1973-02-06 四川炉霍	7.6	1301.43	12.95	2199	2584	Y
东部	/	1974-04-22 江苏溧阳	5.5	6.16	165.09	8	12	N
		1998-01-10 河北张北	6.2	114.88	110.75	49	38	Y

表 5.12 中，Y 代表死亡人数与预测值在同一数量级，N 表示二者不在同一数量级，因东部地区震例较少，“/”代表未对东部地区震例进行震中烈度分级，在西北地区，多数地震的实际死亡人数与评估结果在同一数量级，而 1995 年甘肃永登 5.8 级地震预测值为 5 人，实际死亡人数 12 人，主要是位于Ⅷ度区的圪瘩沟村震后山体滑坡造成 5 人死亡，导致死亡加剧。

在预测模型中多使用相对误差作为评判标准，根据相对误差公式计算 2015 年新疆皮山 6.5 级地震，预测值为 5 人，实际死亡人数为 3 人，计算的相对误差为 66.67%，但在实际震后快速评估中，这种误差并不影响应急指挥调度和救援物资调配，而该震例的相对误差会影响平均相对误差的计算，因此予以剔除，为方便计算，删除 0-9 这个数量级的震例。

$$\delta = \frac{|\bar{y} - y|}{y} \times 100\% \tag{5.1}$$

式中：δ表示相对误差，$\bar{y}$表示预测值，y表示真实值。

根据相对误差公式计算可得西北地区平均相对误差为 23%，西南地区平均相对误差为 28%，东部地区平均相对误差为 36%，在西南地区评估结果中，模型

预测的死亡人数与实际死亡人数之间的误差略微偏高。主要原因在于西南地区地质构造复杂且山区众多，地震发生时往往伴随着大规模的次生灾害，如泥石流和滑坡等，这些灾害连锁反应加剧了人员伤亡的程度，使模型在估算地震死亡人数时未能充分考虑到这些额外风险因素的影响。相反，在大陆东部地区，虽然地震活动相对较少且强度一般较低，但由于样本数据中造成人员死亡的震例数量有限，导致模型训练时代表性不足，无法准确反映该地区在极端情况下可能发生的灾害后果，因此对东部地区的评估误差也较高。

表 5.13 模型误差率

模型	西北地区误差率	西南地区误差率	东部地区误差率
GB/T 30352-2013	71	87	89
刘金龙模型	55	61	65
BP 神经网络	38	43	45
PCA-ELM	31	35	42
LightGBM-Stacking	23	28	36

考虑到不同地区地震频次不同，本文对五种模型进行分析，结果如表 5.13 所示。其中，地震灾情应急评估（GB/T 30352-2013）由中国地震局提出，适用于重大和特别重大地震灾害的灾情应急评估；刘金龙模型通过函数拟合与经典回归分析，应用简单，适于在震后利用震级和盲估烈度进行人员死亡的快速评估；BP 神经网络和 PCA-ELM 作为人工神经网络模型，可以用来解决复杂关系，是估计地震人员死亡的重要方法。五种模型中 LightGBM-Stacking 平均相对误差最小为 29%，表明 LightGBM-Stacking 集成学习模型在处理非线性数据时更具优势。

从预测结果可以看出，在预测地震人员死亡中，LightGBM-Stacking 在处理非线性复杂问题上相较于经验统计法和神经网络表现更优，表现出更小的误差率，相较于回归分析的有限规律，更能挖掘数据背后的非线性关系。以刘金龙模型为代表的经验统计法误差率较大，因为经验统计法很大程度上依赖于先前观察到的数据和规律，这些规律可能在未来情境中发生变化，导致模型的预测能力下降。经验统计法挖掘能力有限，不够灵活，难以适应复杂的非线性关系和数据背后的深层次模式。BP 神经网络在初级算法中相对较优，但在收敛和梯度等方面仍存在一定缺陷，因此结果表现略逊于极限学习机。

Stacking 集成模型误差率较小是因为 Stacking 作为一种集成学习方法，通过组合多个模型预测结果提高整体性能，能够综合各个模型的优势来降低误差率，集成学习的优势在于处理非线性复杂问题，可组合多个模型，每个模型负责捕捉问题的不同方面，挖掘数据中的模式和关系，综合了多个模型的信息。同时与经验统计法相比，集成学习具有更好的实时更新性，一旦建立好模型，只需即时更新数据，而无需重新建立模型。

在某些情况下，即使误差率较高，但在实际应用中仍是可接受的。这强调了误差率并不是衡量算法优劣的唯一标准，在不同应用场景中，对误差率的容忍度有所不同。不同的算法在某些情境下可能表现出色，而在另一些情境下较差。在集成学习选择学习器时，需要综合考虑模型的解释性、计算效率、可解释性、鲁棒性等因素。不同震例的选择对模型评估具有重要影响，导致模型在误差率上表现不同。因此，评估模型性能时应使用多个不同数据集，涵盖不同的情境和条件，在地震人员死亡预测中，考虑到死亡人数的巨大差异，需要确保模型在各种震例下都能够表现出鲁棒性。在地震人员死亡评估中真实值与预测值在一个数量级上的接近也很重要，特别是对于死亡人数巨大的事件，用相对误差来衡量并不是最优解，能够提供一个合理数量级的估计就能代表模型的有效性。

处理地震人员死亡这种非线性复杂问题，机器学习算法的应用更显优势，传统的经验统计模型难以捕捉地震人员死亡背后的复杂关系，根据人员死亡影响因素与伤亡数量之间的线性关系得到的规律相对有限。

6 结论与讨论

本文旨在通过总结地震人员死亡的影响因素、评估方法以及集成学习相关算法，从理论角度深入探讨。第二章介绍相关理论与方法。第三章，选择随机森林、CART、GBDT 和 AdaBoost 评估人员死亡的影响因素，最终确定随机森林算法作为选取地震人员死亡影响因素的最佳选择。随机森林算法在理论上有坚实的基础，并对人员死亡影响因素的重要度进行评估，为后续集成学习模型提供了有力的输入参数。第四章使用 GAN 进行数据增强，并进行单一模型预测，对比基准模型的预测效果。第五章确定基学习器组合方案，分别构建 LightGBM-Stacking、XGBoost- Stacking、RF-Stacking 和 SVR-Stacking 这四个双层模型融合的预测模型，根据评价指标预测性能最优的是 LightGBM-Stacking 融合模型，随机挑选震例验证该模型的预测效果。

6.1 结论

建模策略：考虑人口密度、地理环境和建筑结构等多个因素，采用分区域建模的方法。有死亡的地震往往发生在Ⅷ度区及以上，在分区域基础上，将地震数据分为Ⅷ度区和Ⅷ度区以上，使其能够更精准地反映不同烈度下的情况。采用 Stacking 集成学习模型，建立一种用于快速评估震后人员死亡的模型。该模型在考虑不同地震烈度的同时，兼顾多个影响因素，为灾后人员死亡的快速评估提供了一种可行且有效的方法。

由于地震成因复杂，包括地质、人口和环境等因素的影响，地震死亡预测充满不确定性。为应对这一挑战，从数据质量入手，以中国地震台网中心和各地震局发布的烈度图数据为准，采用分区域建模，运用集成学习算法挖掘地震数据中的非线性关系，建立 Stacking 集成学习模型。通过特征提取和参数优化，更准确地评估地震引发的人员死亡。实验证明，相较于其他模型，该模型在地震人员死亡评估中具有更高的精度和准确性。因此，该模型对地震造成的人员死亡快速评估具有一定参考意义，为未来的研究提供了新的思路和方法。

6.2 讨论

本文主要从数据入手，以地震台网中心数据和各地震局发布的烈度图为主建立数据库，提高数据质量，使用 Stacking 集成学习模型进行地震人员死亡评估，虽然在一定程度上提升了预测模型的性能和准确性，但在数据选取策略、参数优化调整等方面依然存在改进的空间。为进一步提升预测效果，后续研究分析应着重关注并解决以下几个方面的问题：

（1）模型的持续优化和适应性提升

随着每年不断出现的新地震事件，为确保该预测模型能够保持较高的适用性，必须对新增的震例数据进行及时补充，并据此进行模型的迭代更新与优化设计。同时可以考虑新影响因素，如手机热力图数据和地质条件因素，提高模型的准确度。

（2）克服集成学习的局限性

尽管集成学习在地震人员死亡评估中取得了成功，其局限性也需要克服。未来的研究可以通过调整学习模型参数、引入新优化算法、更新超参数和初始参数、以及适度调节学习框架，来提高模型的性能。此外，对集成学习的局限性进行更深入的研究，寻找更为灵活和适应性强的方法。

（3）次生灾害修正的进一步研究

次生灾害是一个复杂的问题，需要进一步深入研究。在每次地震评估时详细记录次生灾害导致的人员死亡情况，考虑地理环境等因素造成的次生灾害，并研究当地次生灾害的防治效果。而地震人员死亡评估实际需要综合考虑多种因素，因此在下一步研究中，将融入地理环境、建筑结构等因素，尝试多种模型算法，不断提高地震灾害人员死亡预测的准确性和可靠性。

参考文献

[1] 周汉杨,杜建军,张舒婷.非线性混合模型的优越性及其在中国典型地震强活动区的应用[J].地质力学学报,2023,29(2):264-275.

[2] 亓凤娇,苏鹤军,陈文凯,王紫荆,苏浩然.基于地震参数的人员死亡评估模型对比研究[J].地震工程学报,2021,43(1):123-130.

[3] 贾晗曦.基于机器学习算法的地震人员伤亡评估研究[D].中国地震局工程力学研究所,2020.

[4] Karimzadeh S, Miyajima M, Hassanzadeh R, et al. A GIS-based seismic hazard, building vulnerability and human loss assessment for the scenario in Tabriz[J]. Soil Dynamics and Earthquake Engineering, 2014, 66: 263-280.

[5] Wilson B, Paradise T. Assessing the impact of Syrian refugees on earthquake fatality estimations in southeast Turkey[J]. Natural Hazards and Earth System Sciences, 2018, 18(1): 257-269.

[6] Maqsood S T, Schwarz J. Estimation of Human casualties from earthquakes in Pakistan—an engineering approach[J]. Seismological Research Letters, 2011, 82(1): 32-41.

[7] Chen W, Sun Z, Han J. Landslide susceptibility modeling using integrated ensemble weights of evidence with logistic regression and random forest models[J]. Applied sciences, 2019, 9(1): 171.

[8] Maqsood S T, Schwarz J. Estimation of Human casualties from earthquakes in Pakistan—an engineering approach[J]. Seismological Research Letters, 2011, 82(1): 32-41.

[9] Xing H, Junyi S, Jin H. The casualty prediction of earthquake disaster based on Extreme Learning Machine method[J]. Natural Hazards, 2020, 102: 873-886.

[10] Aleskerov F, Say A I, Toker A, et al. A cluster‐based decision support system for estimating earthquake damage and casualties[J]. Disasters, 2005, 29(3): 255-276.

[11] Park J H, Shin M, Cho G H. A dynamic estimation of casualties from an earthquake based on a time-use survey: applying HAZUS-MH software to Ulsan, Korea[J]. Natural Hazards, 2016, 81: 289-306.

[12] So E, Spence R. Estimating shaking-induced casualties and building damage for global earthquake events: a proposed modelling approach[J]. Bulletin of Earthquake Engineering, 2013, 11: 347-363.

[13] Hu Y, Wang J, Li X, et al. Exploring geological and socio-demographic factors associated with under-five mortality in the Wenchuan earthquake using neural network model[J]. International journal of environmental health research, 2012, 22(2): 184-196.

[14] Wang Y, Dai J, Feng X. A grey-neural networks prediction model of death toll in “5.12” Wenchuan Earthquake[C]//2010 Sixth International Conference on Natural Computation. IEEE, 2010, 7: 3687-3691.

[15] Chen W, Sun Z, Han J. Landslide susceptibility modeling using integrated ensemble weights of evidence with logistic regression and random forest models[J]. Applied sciences, 2019, 9(1): 171.

[16] 肖光先.地震损失的预测方法[J].地震学刊,1987(1):1-8+81.

[17] 马玉宏,谢礼立.地震人员伤亡估算方法研究[J].地震工程与工程振动,2000(4):140-147.

[18] Jaiswal K, Wald D. An empirical model for global earthquake fatality estimation[J]. Earthquake Spectra, 2010, 26(4): 1017-1037.

[19] 刘金龙,林均岐.基于震中烈度的地震人员伤亡评估方法研究[J].自然灾害学报,2012,21(5):113-119.

[20] 尹之潜.地震灾害损失预测研究[J].地震工程与工程振动,1991(4):87-96.

[21] 张莹,郭红梅,尹文刚,等.基于多因素的地震灾害人员伤亡评估模型研究[J].震灾防御技术,2017,12(4):870-881.

[22] Huang X, Jin H. An earthquake casualty prediction model based on modified partial Gaussian curve[J]. Natural Hazards, 2018, 94: 999-1021.

[23] 于山,王海霞,马亚杰.三层 BP 神经网络地震灾害人员伤亡预测模型[J].地震工程与工程振动,2005(6):113-117.

[24] 钱枫林,崔健.BP 神经网络模型在应急需求预测中的应用——以地震伤亡人数预测为例[J].中国安全科学学报,2013,23(4):20-25.

[25] 周德红,冯豪,程乐棋,等.遗传算法优化的 BP 神经网络在地震死亡人数评估中的应用[J].安全与环境学报,2017,17(6):2267-2272.

[26] Oktarina R, Bahagia N, Diawati L, et al. Artificial neural network for predicting earthquake casualties and damages in Indonesia[C]//IOP Conference Series: Earth and Environmental Science. IOP Publishing, 2020, 426(1): 012156.

[27] Li B, Gong A, Zeng T, et al. A zoning earthquake casualty prediction model based on machine learning[J]. Remote Sensing, 2021, 14(1): 30.

[28] 张文娟.基于移动通信大数据的地震灾区人口伤亡获取与评估系统设计[J].地震工程学报,2019,41(4):1066-1071+1097.

[29] 曾婷婷,宫阿都,陈艳玲等.基于历史相似案例空间推演的地震伤亡人口评估方法研究[J].地球信息科学学报,2020,22(11):2166-2176.

[30] 吴昊昱,顾建华.汶川地震死亡人数的 Zipf 分布[J].国际地震动态,2008(11):168.

[31] Fang Z, Huang J, Huang Z, et al. An earthquake casualty prediction method considering burial and rescue[J]. Safety science, 2020, 126: 104670.

[32] Cui S, Yin Y, Wang D, et al. A stacking-based ensemble learning method for earthquake casualty prediction[J]. Applied Soft Computing, 2021, 101: 107038.

[33] Xia C, Nie G, Fan X, et al. Research on the rapid assessment of earthquake casualties based on the anti-lethal levels of buildings[J]. Geomatics, Natural Hazards and Risk, 2020, 11(1): 377-398.

[34] Badal J, Vázquez-Prada M, González Á. Preliminary quantitative assessment of earthquake casualties and damages[J]. Natural Hazards, 2005, 34: 353-374.

[35] Zhang S, Yang K, Cao Y. GIS-based rapid disaster loss assessment for earthquakes[J]. IEEE Access, 2018, 7: 6129-6139.

[36] 于晓虹,叶晶,洪赢政,等.地震死亡人数评估的投影寻踪回归建模研究[J].华北地震科学,2022,40(4):19-27.

[37] 马秀丹,郑宁宁,崔满丰等.2023 年 1—3 月全球地震活动述评[J].中国地震,2023,39(2):438-447.

[38] 段中满,贾亮亮,蒋明光等.基于不同特征选择方法和随机森林法的滑坡易发性评价——以湖南中西部地区为例[J].华南地震,2023,43(2):115-124.

[39] 李山有,陈欣,卢建旗等.基于 XGBoost 的现地地震烈度阈值实时判别模型[J].地球科学,2024,49(2):379-390.

[40] 谢军飞,张海清,李代伟等.基于 Lightgbm 和 XGBoost 的优化深度森林算法[J].南京大学学报(自然科学),2023,59(5):833-840.

[41] 胡记磊,吴文良,王璟等.基于自适应 LASSO 的逻辑回归砂土液化判别模型[J].三峡大学学报(自然科学版),2023,45(2):67-72.

[42] 刘赫奕. 基于机器学习的工程地震预警参数估计研究[D].中国地震局工程力学研究所,2023.

[43] 潘耀,邓浩宇,陈厦等.地震灾害应对决策的模型集成规则挖掘[J].自然灾害学报,2020,29(6):70-84.

[44] Yilong Li, Zhenguo Zhang, Danhua Xin; A Composite Catalog of Damaging Earthquakes for Mainland China. Seismological Research Letters 2021; 92 (6): 3767–3777.

[45] 中国地震局震灾应急救援司.1966-1989 年中国地震灾害损失资料汇编[M].北京:地震出版社,2015.

[46] 国家地震局,国家统计局.中国大陆地震灾害损失评估汇编:1990-1995[M].北京:地震出版社,1996.

[47] 中国地震局监测预报司.中国大陆地震灾害损失评估汇编:1996-2000[M].北京:地震出版社,2001.

[48] 中国地震局震灾应急救援司.2001-2005 年中国大陆地震灾害损失评估汇编[M].北京:地震出版社,2010.

[49] 中国地震局震灾应急救援司.2006-2010 年中国大陆地震灾害损失评估汇编[M].北京:地震出版社,2015.

[50] 段艳慧,郭伟,赵学胜,等.基于 Landsat 影像和统计数据的北京市人口密度制图[J].北京测绘,2022,36(8):1096-1101.

[51] 王冬年. 基于凸集投影算法的地震数据重建和噪声压制研究[D].东华理工大学,2020.

[52] 班寰宇.Python 在墨卡托投影与兰伯特投影坐标转换中的应用[J].电子技术与软件工程,2022(20):67-70.

[53] 李厚朴,李海波,唐庆辉.椭球情形下等角和等面积正圆柱投影间的直接变换[J].海洋技术学报,2019,38(5):15-20.

[54] 胡进军,石昊,谭景阳.西南地区水平向地震动设计谱参数特征分析[J].地震工程与工程振动,2023,43(6):1-13.

[55] 陈运泰.地震预测:回顾与展望[J].中国科学(D 辑:地球科学),2009,39(12):1633-1658.

[56] 杨杰英,李永强,刘丽芳等.地震三要素对地震伤亡人数的影响分析[J].地震研究,2007(2):182-187+205.

[57] 亓凤娇,李雯,苏鹤军等.中国大陆地震灾害分区人员死亡评估模型研究[J].地震,2022,42(1):70-84.

致 谢

岁月如诗，三年光阴荏苒，仿佛昨日刚踏进校园，时光如白驹过隙，这段时光的学习即将画上句号，站在研究生生涯的终点，感慨万分，满腔感激之情难以言表。

在这充满探索和奋斗的旅程中，首先我要衷心感谢我的导师，赵煜老师是我学术道路上的明灯。感谢您对我的耐心教导，您的言传身教将是我终生受益的财富。同时特别感谢陈文凯老师，作为我的校外导师，您让我接触到地震领域的学术研究，并提供了宝贵的学术指导。两位导师的引领让我看到了学术研究的广阔天地，这对我产生了深远的影响。

同时感谢甘肃地震局陈文凯老师团队的其他老师，在思绪交汇的岁月里，感谢孙艳萍老师，与您在同一环境下办公学习，从生活到学习都受益良多；感谢史一彤老师，在地震局实习期间您的照顾无微不至，在日常的交流中迸发很多灵感。感谢宁夏地震局的余思汗老师，您手把手教会我使用 Arcgis，在论文写作中给出建设性意见并在学习的间隙给出未来发展的意见。让我体验到团队的力量，每一个合作的瞬间都是心灵的碰撞。

其次，感谢在硕士期间的所有同门与同窗，在追求学术的路上并肩前行。一起奋斗，一起欢笑，你们是我科研生涯最美好的风景。感谢地震局的赵怀群师兄，探讨发现了很多我在学术研究中的不足。感谢同门刘迪、杨盛文，你们的陪伴和合作使得这段时光充满了欢笑和收获，在学术问题上的深入讨论和共同进步，你们的支持和友谊是我研究生生涯中的珍贵财富。

最后，感谢亲爱的家人和朋友们，是你们在风雨飘摇中给予我的温暖和鼓励。感谢百忙之中评阅论文以及毕业答辩的老师，感谢兰州财经大学。

在这个告别的季节，愿我们都能怀揣着理想，迎接新的征程。感激相遇，珍惜别离。未来的路漫漫，让我们携手前行，书写属于自己的故事。