NATURAL DISASTER MANAGEMENT USING MACHINE LEARNING

Thesis submitted in fulfilment of the requirements for the Degree of

DOCTOR OF PHILOSOPHY

By

ARUSH KAUSHAL

Department of Computer Science & Engineering and Information Technology

JAYPEE UNIVERSITY OF INFORMATION TECHNOLOGY
Waknaghat, Solan – 173234, Himachal Pradesh, INDIA

March, 2025

DECLARATION BY THE SCHOLAR

I hereby declare that the work reported in Ph.D. thesis entitled **"Natural Disaster Management using Machine Learning"** submitted at **"Jaypee University of Information Technology, Wakhnaghat, Solan (H.P), India"**, is an authentic record of my work carried out under the supervision of **"Prof. (Dr.) Vivek Kumar Sehgal and Prof. (Dr.) Ashok Kumar Gupta".** I have not submitted this work elsewhere for any other degree or diploma. I am fully responsible for the contents of my Ph.D. Theses.

Arush Kaushal
Enrolment No.: 216204
Department of Computer Science & Engineering
Jaypee University of Information Technology, Wakhnaghat,
Solan (H.P), India
Date:

SUPERVISOR'S CERTIFICATE

This is to certify that the work reported in the Ph.D. thesis entitled **"Natural Disaster Management using Machine Learning"** submitted by **Mr. Arush Kaushal**, Enrollment no. 216204 at **Jaypee University of Information Technology, Wakhnaghat, Solan (HP), India**, is a bonafide record of her original work carried out under my supervision. This work has not been submitted elsewhere for any other degree or diploma.

Prof. (Dr.) Vivek Kumar Sehgal (Professor)
Department of Computer Science & Engineering and Information Technology
Jaypee University of Information
Technology, Wakhnaghat,
Solan (H.P), India
Date:

Prof. (Dr.) Ashok Kumar Gupta (Dean A&R)
Department of Civil Engineering
Jaypee University of Information
Technology, Wakhnaghat,
Solan (H.P), India
Date:

ACKNOWLEDGEMENT

With the providential grace of "**Almighty God**", the expedition of my Ph.D. came to an end and my heart is overflowing with appreciation towards each and every person who has lend a hand in the form of support, believe and efforts to accomplish this journey.

It is an immense pleasure to express my profound gratitude towards my supervisors **Prof.** (**Dr.) Vivek Kumar Sehgal**, **Professor**, Department of Computer Science & Engineering and Information Technology, JUIT, Wakhnaghat, and **Prof. (Dr.) Ashok Kumar Gupta,** Department of Civil Engineering, JUIT, Wakhnaghat, who's graciously gave me the opportunity to work under their guidance. I am thankful for their patience, continuous support, optimistic approach, never-ending deliberations, time to time guidance and liberty throughout this course. I will always stay indebted to them for bearing my shortcomings with their immense sense of awareness, maturity, thorough knowledge of the specific field and consistency.

I would like to express my gratitude to our Honourable Vice Chancellor **Prof. (Dr.) Rajendra Kumar Sharma** and Dean (Academics and Research), **Prof. (Dr.) Ashok Kumar Gupta** to promote the research and facilitate resources in the institution. I would also like to pay my gratitude to the DPMC members **Dr. Ekta Gandotra, Dr. Hari Singh, and Prof. (Dr.) Tiratha Raj Singh** for their thought-provoking interactive assessments, queries and opinions. Their valuable motivation, help, suggestion, affirmative vision, magnificent supervision and enormous confidence in my abilities made me face tough circumstances during the progress of the research work.

I am deeply grateful to my father for finalizing this decision for me and to my mother for her unwavering support throughout my PhD journey. Their belief in me has kept my spirits and motivation high during this process. I am obliged for all the support received from all the faculty members and staff of Department of C.S.E & I.T for their scholarly support and guidance. I thank my fellow Ph.D. friends for their consistent help and valuable discussions.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AUC	Area Under the Curve
ADT	Augmented Dickey-Fuller Test
AdaBoost	Adaptive Boosting
ANN	Artificial Neural Network
BERT	Bidirectional Encoder Representations from Transformers
CatBoost	Categorical Boosting
CNN	Convolutional Neural Network
CNN-LSTM	Convolutional Neural Network - Long Short-Term Memory
DCN	Deep Cross Network
DEM	Digital Elevation Model
DNN	Deep Neural Network
DT	Decision Tree
ELM	Extreme Learning Machine
FFT	Fast Fourier Transform
FCN	Fully Convolutional Network
GBDT	Gradient Boosting Decision Tree
GBR	Gradient Boosting Regression
GIS	Geographic Information Systems
GPS	Global Positioning System
GRU	Gated Recurrent Units
IoT	Internet of Things
K-means	K-means Clustering
KNN	K-Nearest Neighbors
LDA	Linear Discriminant Analysis
LEWS	Low-Cost Landslide Early Warning System
LR	Logistic Regression

LSTM	Long Short-Term Memory
MAE	Mean Absolute Error
ML	Machine Learning
MLA	Machine Learning Algorithms
MLR	Multiple Linear Regression
MSE	Mean Squared Error
NIR	Near Infrared
NDVI	Normalized Difference Vegetation Index
OBIA	Object-Based Image Analysis
PCA	Principal Component Analysis
QDA	Quadratic Discriminant Analysis
RF	Random Forest
RFR	Random Forest Regression
ROC	Receiver Operating Characteristic
RNN	Recurrent Neural Network
SAR	Synthetic Aperture Radar
SARIMA	Seasonal Autoregressive Integrated Moving Average
SVR	Support Vector Regression
SVM	Support Vector Machine
SWIR	Shortwave Infrared
TCN-AR	Temporal Convolutional Networks - Autoregressive
TFF	Time to Failure
TS	Time Series
T-SNE	t-Distributed Stochastic Neighbor Embedding
UNet	U-Net Architecture
VAE	Variational Autoencoder
VHR	Very High Resolution
XAI	Explainable Artificial Intelligence
XGBoost	Extreme Gradient Boosting

ABSTRACT

Natural disasters like landslides and earthquakes are one of most common natural disaster in nature that have capability to cause threat to human life, infrastructure and economical damage. Both of these disasters are random which makes it difficult to predict and manage by providing early warnings. Traditional methods for disaster prediction are not that accurate and have limited ability to predict accurately and timely. So, this thesis addresses these issues or gap by providing advanced hybrid machine learning models which uses different technologies including computational and different type of data like real-time, remote sensing, and historical data to improve the prediction and forecasting of disasters. The aim is to improve disaster management system and make them more reliable by integrating these different ways and techniques, provide better early warning and enable more effective risk mitigations. The research primarily focuses on development of hybrid machine learning models designed for detecting and prediction landslide and earthquake. The main contribution is creating real-time landslide prediction model which collects data from real-time Wireless Sensor Networks (WSNs) in a laboratory setup, the sensors consistently monitor all useful factors such as soil moisture, vibration, temperature, humidity, angular acceleration, angular velocity other various other parameters. The data is processed using a predictive system which combines hybrid machine learning model such as Multiple Linear Regression (MLR) and Long Short-Term Memory (LSTM) to analyse this data in real time. Using this hybrid model for landslide prediction improves the accuracy by identifying patterns that provides insights that a landslide may occur and also offers early warning alerts for area prone to landslides.

As progressing forward, landslide detection using remote sensing data is introduced with advanced segmentation and feature extraction methods. A deep learning model UNet-Pyramid is used to capture minute details in the images and also analyzing high resolution images to grasp the change in landscape like angle displacement, shift in vegetation cover indicating a risk of landslide event. To deeper analysis of remote sensing Object-Based Image Analysis (OBIA) is used for feature extraction, which works by forming groups of small pixels together to identify larger objects such as displaced soil cover or vegetation area. Moreover, the Swin Transformer architecture is applied, that helps to capture features in images more effectively using window-based mechanism and provides more detailed segmentation. The main benefit by using combined feature extraction techniques like OBIA, and Swin Transformer helps to handle high-resolution satellite images better and detect complex surface changes. This

cumulative technique including segmentation, feature extraction and applying hybrid models all together provides a detailed, accurate and reliable solution in identifying risky areas and detecting landslides events efficiently.

Further, the complexity of seismic activities in earthquake prediction prompted the shift as earthquake being another frequent disaster in nature, so a hybrid model which combines the features of two different techniques such as SARIMA (Seasonal Autoregressive Integrated Moving Average) and XGBoost (Extreme Gradient Boosting) is applied. This hybrid model uses sequential data for earthquake prediction which contain both short-term and long-term seismic trends. Firstly, the SARIMA helps to capture seasonality trends and patterns which are commonly cyclic in nature in the time series data and then XGBoost supports to model complex, non-linear relationships between provided variables to improve the effectiveness of earthquake prediction. So, it starts with Exploratory Data Analysis (EDA), that is applied for understanding intrinsic patterns and insights of dataset. For feature engineering process, EDA helps in identifying the most useful features and data augmentation and feature engineering techniques are applied to further enhance the prediction accuracy, by integration of these techniques, provides more accurate earthquake forecasts by considering different kind of patterns in the data. As research progressed, it explores another hybrid model which combines CatBoost and Support Vector Regression (SVR) for earthquake prediction using LANL earthquake dataset. Here, CatBoost uses gradient boosting method to optimize and handle categorical data. The signal based LANL earthquake dataset comprises of acoustic data and Time to Failure (TTF) which uncover important patterns from acoustic data used to analyze significant features and patterns which contributes to accurate earthquake prediction. In CatBoost, multiple decision trees are built on top of other to improve the prediction accuracy by reducing error at each stage and Support Vector Regression (SVR) captures the residuals from CatBoost and further process them using its support vector-based mechanism to capture non-linear relationships in the data, that cannot be modeled by simple boosting approaches. Finally, a precise and reliable earthquake prediction system is developed using hybrid model that contributes in generating early warning systems and improves disaster preparedness methods. The integration of these hybrid machine learning models represents a substantial advancement in the field of natural disaster prediction. Finally, the prediction technique that combine real-time sensor data, remote sensing data, and time series analysis, provides with a comprehensive framework disaster prediction which is accurate, timely and lifesaving. Various models considerably improve the early warning systems for both disasters, by providing precise resource allocation, informed decision-making, and optimized disaster response strategies. The proposed model has deep and

far-reaching impact in disaster preparedness, as it could be applied to different natural disaster scenarios, minimizing the loss of life and infrastructure. So, these data driven solutions provide a promising pathway for safeguarding lifestock and infrastructural damages from disaster events.

CHAPTER 1

INTRODUCTION

1.1 *Introduction*

Disasters are sudden catastrophic events that result in fundamental disturbances, loss of property, human lives, and the environment. These events can be caused by man, such as chemical leaks, nuclear accidents, and industrial accidents, or can be natural, like storms, floods, and earthquakes. Disasters affect local communities or have global consequences, and their impact differs depending on factors such as geographical position, population density, and readiness of the affected areas. They often lead to extensive destruction, health crises, economic loss, and long-term environmental damage, which makes effective disaster management. The coordinated strategy is used to manage disasters to avoid, prepare for, respond to, and recover from disasters by relieving readiness, reactions, and recovery. These are its four primary phases. Through measures such as the creation of robust infrastructure or forcing laws to minimize environmental damage, mitigation is to reduce disaster probability or its effects. Planning, teaching, and community education about the risks of disasters and effective. While the recovery phase focuses on reconstruction and returning to normal after the incident, the reaction phase deals with urgent measures taken. during the disaster to maintain lives and provide assistance. The disaster management requires cooperation between governments, local authorities, humanitarian organizations, and communities. The importance of disaster management cannot be overestimated because it minimizes the negative effects of disasters and guarantees a rapid and organized reaction. Appropriate planning and disaster reactions can save lives, reduce injuries, and reduce infrastructure and environmental damage. It also helps to maintain social order and stability due to a disaster, which allows communities to recover faster and return to normal activities. In addition, catastrophe management promotes resistance by preparing companies to better address future challenges, minimize the vulnerability of endangered populations, and support sustainable development. Investing in disaster management is necessary for the protection of communities and building a safer and more resistant world. A landslide is the movement of a rock, soil, mud, or debris on a slope, usually caused by natural events such as severe precipitation, earthquakes, volcanic activity, or gradual weakening of the earth's materials. Soil landslides can occur in various forms, such as

rockfalls, debris flows, or landslides, and differ in size from small to large movements of soil—destructive events that destroy the whole community. These natural disasters are most common in mountain areas or in regions with steep slopes, where the stability of the country is easily endangered by external forces such as precipitation, seismic activity, or human activity such as deforestation and construction.

Landslides have devastating effects, resulting in death, damage to property, and interruption of vital infrastructure, including motorways, railways, and communication systems. Soil landslides often avoid rivers, which leads to a flood or construction of temporary dams that can use and cause more damage downstream. Landslide events deeply affect places where the local economics, tourism and agriculture depends on slope stability specifically in rural areas. The risk associated with these events needs an early detection method to save livestock, reduce injuries and infra structural losses associated with them. When identifying danger of extinction and understanding the triggers of landslides, authorities can perform specific strategies to reduce vulnerability and prevent disaster. In addition, the ML and DL models allow monitoring and appropriate real-time warning systems that immediately emphasize the populations that are in danger.

This helps to protect communities, maintain infrastructure and reduce the financial burden on recovery efforts. In addition, landslides detects before they happen, it contributes to the overall more efficient management of disasters, improves response time and ensures that suitable sources are available, when and where they are needed. Early detection and proactive planning, supported by machine learning and deep learning technologies, are necessary to protect lives, protection of property and support sustainable development in areas susceptible to landslides.

1.2 Motivation

Landslide prediction is motivated by the need to lessen the destructive effects of these natural disasters, which can result in a large loss of infrastructure, property, and human life. Geological conditions, precipitation, seismic activity, and human activity are just a few complicated factors that could affect the occurrence of landslides that are often not expected. Due to the destructive nature of landslides, especially in vulnerable regions such as mountain or coastal areas, early detection and predictions are essential for saving lives, preventing injury, and reducing economic losses. Due to the dynamic and complex nature of landslides, it historically depends

on field observation and expert interpretation, which has often proved to be inadequate. Because ML and DL can process and evaluate a huge amount of different data, they have become effective tools for the prediction of landslides. To identify trends and predict future events, ML algorithms can learn from past layout incidents and environmental factors (such as slope, collision, soil moisture, and seismic activity). DL uses neural networks and automatically finds a comprehensive association in large data sets, which allows even more advanced analysis. The abundance data present requires proper processing and analyzing so that these advanced models can do continuously learning and adapting to new environmental condition. Also, rapid changes in these factors are observed, so by working on data from various sources such as real time, remote sensing data, historical data can allow model to provide precise predictions and timely alerts to authorities for issuing early warning and also initiate preventive measures. The main goal revolves around reducing economic losses and building durable communities to prevent risk of landslide with limited information of historical trends of landslide. Since these technologies are still progressing, they offer promising opportunities to improve our understanding of landslide dynamics and create a proactive approach to disaster management. One of the greatest risks for the local and global economies, as well as for human settlements, is geological risks. The most common geological dangers include landslides that include the movement of rock, dirt, mud, or debris. Natural occurrences, such as earthquakes or intense rains, can often cause landslides, especially in areas with hydrological, geological, and geomorphological characteristics. However, the mechanics of landslides also depend strongly on other key elements such as weather, soil head, and in situ tension. Topography, forests, soil characteristics (such as consistency, structure, density, and temperature), and infrastructure, such as roads and agriculture, can be significantly affected by landslides in mountain areas. The size of landslides determines how serious these effects will be. Finding vulnerable areas and understanding mechanisms of landslides over the past 20 years has become more important in landslide research. This research has led to valuable knowledge about the analysis of geomorphological, tectonic, geological, climatic, and human-induced factors. Historical records show that it experienced the highest number of deaths in the land of landslides, with a total of 132 deaths. Risk assessment relies strongly on the location of landslides and their risk assessment. Research on landslides has been significantly advanced in recent years using new technologies and techniques, especially in crisis management for mountain regions or those that are vulnerable to such risks. Number of quantitative techniques are tested and evaluated to create accurate and reliable model for improved landslide prediction.

1.3 Contributions

The major contributions of this thesis can be summarized as follow:

- Designing a threshold-based real-time landslide prediction system utilizing IoT networks: This system efficiently monitors key environmental parameters such as, soil humidity, slope displacement, and rainfall, providing real-time data and triggering alerts if hazardous thresholds are reached to enhance early warning capabilities for landslide-prone areas.
- Developing an inexpensive landslide early warning system based on IoT for continuous landslide monitoring, especially in regions with constrained resources, our contribution offers a scalable and affordable approach that ensures real-time data collection and analysis for efficient risk reduction and disaster management.
- Creating a semantic segmentation system using a UNet-pyramid architecture for landslide prediction. This framework improves the accuracy of detecting landslide-prone areas by utilizing remote sensing data from the Landslide4Sense dataset, allowing for precise and reliable landslide hazard evaluations.
- Integrating SARIMA and XGBoost for spatial earthquake forecasting. This hybrid approach improves earthquake prediction by accounting for both spatial and temporal dependencies in seismic data. This offers more accurate forecasts and better risk management for subsequent earthquake areas.
- Improving landslide hazard mapping through deep learning-based semantic segmentation. The implementation of deep learning models for analyzing remote sensing data contributes to more reliable identification and classification of landslide hazards, , increasing the assessment of risks and land use in affected areas.
- Contributing to disaster preparedness and mitigation through advanced early warning systems: The thesis provides advance methodologies and smart tools that allows accurate real-time landslides and earthquakes prediction, reducing the potential impacts of natural disasters and improving resilience in vulnerable communities.

1.4 Thesis Outline

This thesis consists of seven chapters, each of which focuses on different aspects of landslide and earthquake prediction using advanced machine learning and deep learning techniques. Chapter 1 represents the research topic and outlines the significance of predicting natural

disasters and the need for innovative methods. Chapter 2 presents an extensive overview of existing literature on landslide and earthquake prediction, focusing on hybrid models that integrate deep learning and machine learning approaches. Chapter 3 discusses a threshold-based real-time landslide prediction system which is designed for hilly areas, along with the enhancement of a low-cost Landslide Early Warning System (LEWS) utilizing Internet of Things (IoT) networks for regions susceptible to landslides. Chapter 4 presents a novel semantic segmentation framework using UNet-pyramid for landslide prediction, using remote sensing data. Chapter 5 examines the synergy between SARIMA (Seasonal Autoregressive Integrated Moving Average) and XGBoost for spatiotemporal earthquake time series forecasting, highlighting the potential of combining statistical and machine learning models. Chapter 6 examines the LANL earthquake dataset and the hybrid CatBoost and SVR model for earthquake forecasting. By providing a summary of the main conclusions drawn from the simulation and experimental data, Chapter 7 sums up the thesis and proposes future study avenues in the areas of earthquake and landslide prediction.

CHAPTER 2

RELATED LITERATURE AND BACKGROUND

2.1 Introduction

Recently, the most promising advancement in the disaster management field is introducing ML and DL. This demonstrated exceptional possibilities for good data analysis, disaster prediction, and resource optimization. The permeation of these technology processes has resulted in a field of dynamic and transformation-based data to remove the influence of disasters on infrastructure and human life. These tools and methods allow disaster management systems to improve proactive and adaptive strategies in the direction of risk assessment, early warning systems, and accurate disaster forecasting systems. These approaches are integrated into the key aspects of reaction to disasters such as evacuation planning, damage assessment, and logistics optimization, which provide efficiency and accuracy. By reviewing existing research, this study highlighted the enhancement and advancement in disaster management and remaining major challenges that still lack to provide better management results in disaster management and prediction, providing insight into the future abilities of these technologies in the domain of disaster management and identifying important opportunities for further progress in this critical area.

2.2 Foundation of Landslide Prediction using Machine Learning

The combination of DL and hybrid ML models with wireless sensors and Internet of Thing (IoT) has led to a revolutionary approach for the prediction of landslides and early alert systems. Fundamentally, this new strategy aims to predict and reduce the catastrophic effects of landslides, which continue to be one of the most unexpected and devastating natural disaster, using the synergy of ML, DL and sensor network. With the possibility of more precise, reliable and real -time landslides prediction methods, this technological integration means a shift from conventional monitoring systems to more advance hybrid approaches.

By combining different methods and architectures, "hybrid models of ML and DL aim to improve the precision of prediction and detection by combining various sources and data models. The critical or important data for these models are provided by wireless sensors and internet devices of things that constantly monitor environmental parameters, including temperature, rain, soil moisture and soil movement. When combining these sensor networks with advanced computing methods we introduced a novel powerful framework for prediction of landslides.

The core foundation of this approach is found in several key areas. Wireless sensors and IoT devices play a key role in data collection in real time distant and often dangerous areas. These sensors monitor vital parameters of the environment such as rain, soil moisture level and seismic activity, which are critical indicators of landslides. To evaluate big database and detect complex patterns, ML models like Decision Trees, Support Vectors and Random Forests are combined with DL architecture like convolution neuron network (CNN) and recurring neural network (RNN). By combining the best characteristics of both paradigms, this hybrid technique increases the accuracy and durability of the landslide detection. Among the techniques of deep learning, a fully convolutional network (FCN) has appeared as a powerful tool for predicting landslides, especially when working with remote sensing geospatial data, and very good in tasks such as segmentation, which means classification of each image or map in different categories. They are made to handle predictions at the pixel level. For the analysis of satellite images, topographic maps and other sources of geospatial data that are frequently used in the evaluation of earth landslides risk, this capacity makes FCN especially useful. Regions with a high risk of landslides can be identified with precision by using FCN, which consider environmental characteristics such as the type of soil, the steepness of topography and historical landslides events. The flexibility of the FCN in the processing of space data provides an additional precision layer to the models of prediction of landslides, particularly in complex and large -scale geographical areas.

IoT sensors and machine learning models work together to process real -time data, allowing dynamic changes in early warning and risk assessment. This characteristic is crucial to give the appropriate authorities and communities so that they can take precautions before a crisis occurs. The integration of many data sources is one of the main obstacles in the prediction of landslides. Hybrid models can produce more complete risk profiles combining information from many

sensors, weather reports and geographic information systems (SIG). In addition, feature engineering methods help determine which factors are most important to the precise prediction.

The IoT technology combination with ML and DL offers scalable solutions that can be implemented in vast and difficult to achieve. This allows monitoring large geographical areas continuously and efficiently, improving the general resilience of the regions which are prone to landslides. This literature survey aims to explore these fundamental aspects of the prediction of landslides through hybrid machine learning models, focusing on the integration of wireless sensors, IoT networks and advanced techniques such as FCN. When reviewing the current state of the investigation, we will highlight the advances, challenges and future opportunities in this field. Through this exploration, we seek to provide an integral understanding of how these technologies can work together to improve disaster preparation and response, offering new possibilities to save lives and minimize destruction caused by landslides.

To achieve best output, the application of ML is used to analyze data through clear process using several phases. One phase begins with the collection and preparation of data as illustrated in Figure 2.1. Raw data is gathered from many sources like databases or sensors then, data cleaning is done to missing values or other outliers present in data. Further suitable model for training is taken into consideration.

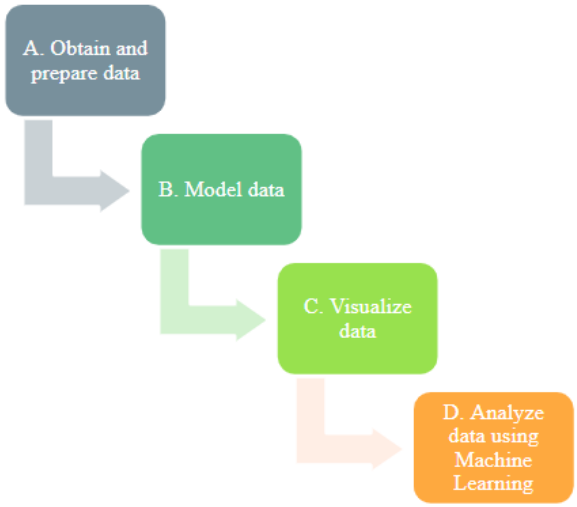


Figure 2.1 Overview of the key stages in the machine learning pipeline

After modelling, the next phase is visualization in which the data must be visualized using graphs or graphs to appear patterns or correlations. Then model evaluation is done using various metrics which determines the usefulness of model. Each phase of this process is

interconnected, and the decision made in one step affects the outcome biased in the following phases.

N Casagil et al. [1] designed a model for detecting and monitoring landslides, which emphasizes the importance of remote detection techniques (RST) in control of the grounds associated with landslides. Landslides are generalized phenomena that can cause considerable damage, especially if they occur near inhabited areas and infrastructure. Precise detection, continuous monitoring and reliable prediction are essential for risk alleviation. The first, such as satellite -based observations, laser scanning and earthly interferometry, are helpful in these processes. Extrusion of landslides can be detected and measured on various spatial and time schedules thanks to a thorough overview provided by satellite RST. On the other hand, although they are focused on smaller regions, ground sensors such as Lidar, Doppler radar and interferometric radar, offer excellent accuracy, frequent data collection and customizable settings. The use of these systems for early warnings of landslides and monitoring in real time is growing. Special needs, including the size of the affected region, the type of landslide and possible risks, determine which first is the best. Each first has a unique set of advantages and disadvantages. Integration of multiple technologies is therefore often the most effective approach for laying and risk control. Better communication with residents in areas susceptible to landslides, deployment of intelligent sources and the use of large data are necessary for more efficient control of landslides.

J Barman et al. [2] designed a model for creating landslide forecasting zonation in Lunglei Mizoram district using bivariate statistical techniques based on GIS. After a multicollinearity test for landslide susceptibility zonation, 17 factors were selected for the study. 234 occurrences of landslides, divided into 70% training and 30% of data sets were used to create a map of soil inventory. Nine main factors such as altitude, gradient, dimension, curvature, normalized difference vegetation index (NDVI), geomorphology, road length, distance from the line and river - it was found to have substantial weights for landslide susceptibility zonation using entropy index (IOE) model. It was found that other elements such as geology, collisions and soil and cover use were of very importance. Two models have been improved Scenario 1 and assessed nine factors, and scenario 2, which contained all 17 factors. The results indicated that 16% and 14% of the district area was classified as very highly susceptible to the screenplay in the 1 and screenplay 2. Accuracy of model present, with area under the Curve (AUC) Values of 0.947 for Scenario 1 and 0.922 for Scenario 2, Indicating Better Performance for Scenario

1. Mapping landslide susceptibility zonation from Screenplay 1 is considered to be the most suitable for management of policy creation in land risk management with regard to these findings.

MT Riaz et al. [3] developed a model to use the efficiency of different techniques of distribution of landslides sensitivity modeling (LSM) and was used in the Pakistani district of Muzaffarabad. The model was used 961 landslide samples which further split into training samples of 70%, 672 and testing samples of 30%, 289. The training samples were processed using the average method of the nearest neighborhood index (ANNI), revealing a sliding distribution pattern for landslides. Among the training samples of 79% showed the behavior of the cluster, while 21% showed random behavior. With 17 parameters of geoenvironmental parameters, five machine learning algorithms were used in clusters and conventional random training samples to evaluate the prediction force of clustering. Using AUC-ROC, sensitivity, specificity, accuracy and index Kappa was evaluated model performance. With AUC-ROC values, it varies from 0.96 to 0.86, Kappa index between 0.76 and 0.60 and accuracy between 0.90 and 0.83, the cluster distribution method showed greater predictive potential. On the contrary, the approach of random segmentation also did not work. AUC (0.962), accuracy (0.902) and Kappa Index (0.755) were highest for the Random Forest (RF) based on cluster training samples that overcome the XGBoost. The outcome demonstrate that the division of clusters enhances LSM precision, especially for complex Himalayan terrain, and emphasizes the potential advantages of using data sets based on cluster over the traditional random division in LSM.

P. C Huang et al. [4] developed a model which integrates a seepage flow model and slope stability model to get the spatial and temporal shifts in areas. The input of the model, the dynamic cumulative shifted area, is derived from the seepage flow and soil slope stability analysis, ensuring that the physical meaning of the ML based model is preserved. The approach effectively predicts when landslides are likely to occur in different areas.The results show that only specific regions, which experience significant changes in instability during rainfall, need to be closely monitored. The model’s prediction accuracy is high, with the mean relative error of the predicted landslide periods (Ps) and initial time (Ts) controlled within 5.19%, and an R2 consistently greater than 0.889. The model successfully predicted around 82.2 percent of the study area's recorded landslide incidents.

P Varangaonkar et al. [5] introduced an innovative framework that uses a long -distance survey to automatically detect landslides and localize regions. The system includes pre-processing, segmentation, feature extraction, and classification. In the preliminary processing phase, the image is denominated by 2D medium filtering, atmospheric and geometric corrections are performed and excess areas are eliminated. ROI is then extracted by dynamic image segment. For the automatic extraction of the elements, convolutional neural network (CNN) layers are used and use the ResNet50 to enhance accuracy and minimize computing expectations. Long -term short memory and artificial neuron network classifiers are used to predict land landslides. The potential placement of landslide is identified in the subsequent processing step if landslides are predicted. According to experimental data, the proposed CNN-LSTM model works better than current solutions in terms of accuracy, score F1, accuracy and download. It also reduces computational complexity by 35% and increases the overall accuracy of prediction by 2% compared to the latest techniques.

Y A Nanehkaran et al. [6] introduced a model which implemented artificial neural networks to evaluate the risk of landslides along riverbanks. This model is alligned with the Sustainable Development Goals of the UN, specifically Goal 11: Sustainable Cities and Communities. The article examines how ANNs are increasingly being used to map landslide vulnerability in riverbank regions, emphasizing how well they outperform conventional techniques. Better risk management and increased community resilience to geohazards are made possible by the incorporation of ANNs into landslide assessments, which promotes sustainable and disaster-resilient urban development. In order to promote sustainable development, better risk management techniques, the review focuses on the most widely used neural network algorithms for riverside landslide prediction. By expanding knowledge and creating safer, more resilient communities, the use of ANNs supports the SDGs.

C Zhou et al. [7] designed a new, cost-effective framework for landslide prediction that combines a MT-INSAR with machine learning techniques. MT-INSAR is used in terms of extracting time series of shifts from the Corernicus Sentinel-1A SAR images. Then the displacement series is divided by means of wave transformation into trend, periodic and noise components. The trend and periodic shifts are predicted by the machine learning model known as the gated recurring units (GRU). These predictions are summarized to estimate the overall shift. GRU overcomes other algorithms such as long short -term memory networks and an extreme kernel -based learning machine, with an Adam algorithm. The results of the prediction

show low errors (the RMSE of 3,817 and 5.145 for Shuping and Muyubao landslides). The proposed framework effectively integrates MT-InSAR and machine learning and offers cost-effective solutions for prediction of moving landslide on large areas.

S Alqadhi et al. [8] developed a thorough strategy to improve the prediction of landslides by a combination of explaining approaches to artificial intelligence (XAI) with deep neuron networks, 1D convolutional neural networks and DNN-CNN. XAI increases the interpretation of deep learning and makes it easier to decide. To evaluate how the variables affect the prediction of landslides, the DNN model uses the game theory. The study identifies the sensitivity zones to a high and very high landslide and shows that the DCN model overcomes CNN and DNN with AUC of 0.97, compared to 0.94 for CNN and 0.9 for DNN. XAI reveals significant remnants at the back of CNN despite its high AUC. The key parameters for accurate prediction include precipitation, inclination, soil texture and line density, while the game theory emphasizes the line density as the primary influential factor, observed by a topographic moisture index, curvature and inclination.

H Ishibashi et al. [9] proposed an approach for assessing the economic risk of structures impacted by collisions caused by landslides with a focus on increasing resistance to extreme rainfall events. The study used ML, specifically random forests and LightGBM, to develop models of landslides and include spatial division of conditioning and trigger factors. The precipitation index, which considers time differences in precipitation, was used to assess the intensity of precipitation and the risk curve was estimated by the generalized distribution of extreme value to represent the connection between the precipitation index and its annual probability of crossing. To evaluate the sensitivity to the landslide with landslides of landslides of landslides with landslides, the risk curve was created for economic loss of structural damage. The results showed that LightGBM exceeded the random forest in predicting a collision caused by a soil landslide.

L. Liu et al. [10] introduced a study to investigate the effiectiveness of classical models in predicting landslide failure-time using displacement monitoring data, with a focus on dynamic prediction. Since landslide monitoring continuously updates the data, predictions should be re-evaluated in real-time. The study examined the limitations of classical models, using data from four real landslides. To improve prediction accuracy, an ensemble model was developed, integrating classical models through a machine learning-based meta-model. A new indicator, the "discredit index (β)," was introduced, where higher β values indicate poorer prediction

quality. Results showed that Verhulst and Saito models had higher β values, while GM (1,1) had the highest MAE. In comparison, ensemble models, particularly the decision tree regression-based ensemble, performed better and provided more accurate predictions.

H. Harsa et al. [11] developed landslide prediction models in Indonesia using artificial intelligence algorithms and ML. These ML models were trained using precipitation data from global satellite observations and landslide occurrence data provided by the Indonesian National Board for Disaster Management. The model was trained with two distinct approaches, leading to the creation of 52 and 72 model candidates for each approach. The best-performing models from each method were selected, with the generalized linear model excelling in the first method and DL outperforming others in the second. The top models achieved AUC values of 0.828 and 0.836, with log-loss values of 0.156 and 0.154, respectively. The second method, which included data transformation, yielded superior results.

Z Chang et al. [12] introduced a study to forecast the LSP using slope units retrieved by the multi-scale segmentation (MSS). They dealt with the question of neglect of heterogeneity of conditional variables in slope units, which can lead to incomplete input variables in LSP modeling. The authors introduced a new approach that includes internal variations of conditional factors (diameter, standard deviation and range) into slope units. Using the Chongyi Country as a case study, the study has expanded 15 original conditional factors to 38 considering their internal variations. The authors compared models of machine learning, including random forests (RF) using slope units with and without internal variants, as well as conventional grid -based models. The finding revealed that the models representing the internal variations within the slope units overcame models based on the grid, which shows greater directional and practical usability. This approach emphasizes the importance of incorporating the heterogeneity of conditional factors into slope units for more accurate and thorough modeling of the sensitivity of landslides.

L. Nava et al. [13] introduced the evaluation of seven DL algorithms to predict landslide shif ts. This study compared 1D CNN and LSTM architectures CONGLSTM combining 1D CNN and LSTM architectures via multilayer shortterm memory, repeating units, 1D folding netwo rks, 2XLSTM, bidirectional LSTM, and slides from four countries with different geographic and geographical conditions. Two of these landslides were affected by artificial reservoirs, w hile the other landslides were driven by precipitation. The results showed that MLP, GRU, an d LSTM models were reliable in all scenarios. This allows the CONPLSTM model to be best

run on seasonal Baishuiheslides. MLP was excellent at predicting top shifts, while LSTM and GRU models were effective at lower shift peaks. These results recommened that these DL m ethods can significantly improve landslides.

Y. Shen et al. [14] proposed a new type of landslide model of machine learning, neuronal net works, and topography indexes to improve accuracy. A study conducted in Western Baijan, Ir an evaluated 16 factors related to geology, environment and geomorphology and analyzed 16 0 landslides. A 30:70 training ratio of the test data was used with four support vector algorith ms and an artificial neural network. The results showed that over 80% of landslide areas were extremely sensitive. Geological factors such as trends, increases and precipitation played an i mportant role with sensitivity of 100%, 75.7%, 68%, and 66.3%. This study assessed the perf ormance of the model using AUC, classification matrix and sensitivity, accuracy and specifici ty metrics, and found that the algorithm surpassed other methods for machine learning. The S VM and Kernel-Sigmoid algorithm achieved the highest accuracy with a performance value of 1.

Z. Chang et al. [15] introduced a new methodology to examine the uncertainty in the selection of sensitivity sample for LSP, which does not include landslides. In order to create LSP models and calculate various soil sensitivity indices, this framework uses machine -based machine - based models in which samples of non -domestic soils are randomly selected many times (n = 1, 10, 100, 500, 1000, 5000) from places outside the soil. The maximum probability analysis (MPA) is used to lower the unpredictability of identification of the ideal sensitivity level for each slope unit, while the statistical analysis is used to display landslide susceptibility indexes uncertainty based on different selections. A study conducted in China Chongyi County in China used LR models and SVM with 16 conditioning factors. The power of the model was evaluated by the accuracy. The results showed that landslide susceptibility indexes monitored normal distribution instead of constant value and was effectively represented by uncertainty when choosing a sample without land.

L. Achu et al. [16] developed a new methodological system for quantification of uncertainty in the prediction of the landslide sensitivity by means of files by eight machine learning techniques (MLT). This framework has been tested in the southern western Ghats area in Kerala, India, a area susceptible to frequent landslides. Fourteen factors of landslides have been identified and correlated with 671 historical landslides. Four models were used in the study: Committee diameter, weighted probability diameter, median probability and probability

average. Based on the operating functions of the receiver, real skills and areas under the curve, the weighted probability average was determined as the most successful model between them. A variation coefficient was used to analyze the uncertainty and a confident map was created to show the zone of the sensitivity of landslides with different scales of uncertainty. The results revealed that 74% of past landslides fell into high uncertain zones with low susceptibility. The study concluded that using such a micro -level zone could increase the efficiency of soil sensitivity maps and provide planners valuable tools for formulating adaptation strategies of landslides.

Y. Wie et al. [17] introduced an improved method to assessing the soil sensitivity by integrating ML, including random forest, a tree decision -making tree and logistics regression, with interferometric synthetic radar technology. This combined approach was compared with the original models and the results showed improved accuracy of prediction with reduced FN-false negative and FP- false positive mistakes. The LR-InSAR has shown the best performance, especially when identifying areas with high susceptibility, both in regional and smaller scales. The results of the modeling were verified by means of data from the unmanned aerial vehicle (UAV) flights.

L. Chen et al. [18] developed a better technique of landslide -based landslides that integrates the machine learning models into the spatiotemporal Knowledge graph. This method deals with the difficulty of integrating data from multiple sources of long -distance survey and creating a consistent prediction process, which is often the disadvantage of contemporary models. This technique chooses the best ML model to forecast landslides in places with a smaller figure when it takes into account environmental similarities between areas. Compared to conventional machine learning techniques, experimental results showed 93% improvement of processing efficiency and 29% increase in score F1. In addition, this approach has solved the problems of Subpar prediction, which caused a lack of data, especially for forecasts performed at the region level. Especially in regions with limited data, this strategy offers thorough information to create more effective techniques of landslide.

C. Chen et al. [19] examined the effect of selection of contributing factors on the precision of landing sensitivity forecasts using machine learning and deep learning models. The study has explored four methods of selection of factors: the ratio of information profit, recursive elimination of elements, optimizing particle swarm, least absolute shrinkage and operators' selection and optimization Harris Hawk, along with an auto -gap factor for deep learning. The

results showed that the selection of significant assisting factors enhanced the accuracy of models. But the result of the DL models has improved when the autoencoder architecture was used to select factors. The study concluded that the choice of factor selection method was more significant than specific factors providing to increasing the precision of the permission sensitivity.

N. Nocentini et al. [20] developed a dynamic approach to analyze the susceptibility of landslides by combining the static sensitivity index with dynamic variables using the random forest algorithm (RF). This methodology integrates the likelihood of spatial landslide (static) with dynamic factors such as seasonality and precipitation in different periods to increase land landslide forecasts. The RF model was applied in the metropolitan city of Florence in Italy, where the importance of variables and verification of the consistency of the model with observed trigger mechanisms used out-of-Bag errors and charts of partial dependence. The aim of the study was to fulfil dataset of training and test datasets with space -time data, identify relevant variable precipitation for timing and location of landslides, and test the dynamic RF application for forecasts. The results showed that the dynamic model precisely reflected the triggers of physical landslides, especially short and intense precipitation, and identified promising configurations for future regional applications in the assessment of the probability of landslides and early warning systems.

T. Xiao et al. [21] developed an innovative framework of ML to predict the landslides caused by rain in space and time that dealt with the challenge of incomplete landslides, especially the lack of accurate timing of landslides. This study systematically compared various methods based on data-based, statistical and machine learning-for soil landslides and introduced a probability model of landslide, which can be used even when timing data is missing. The integrated model provides a useful tool for timely warning systems and real -time decision -making accurately estimated the risk of soil and predicting the spatial development of landslides during rain storms. The model beat previous data -based approaches in terms of accuracy and predicational ability after being verified against 35 years of data on Hong Kong.

K. Doerksen et al. [22] designed a method using machine learning techniques (ML) and deep learning (DL) of artificial intelligence (AI) for predicting landslides in Nepal at the level of the district with 7, 10- and 14 days of time resolution. This approach uses an open source, space data, including calibrated precipitation and geomorphic data. The study showed predictive power of random forest and U-Net models to predict land landslides and provided scientific

knowledge through the analysis of significance. This method improves predictive abilities and offers valuable tools for disasters and solves the challenges of a complex causal chain of landslides in Nepal, where large earthquakes and intensive monsoon collisions are common triggers.

H. Hong et al. [23] designed five integration models that combine locally weighted learning (LWL) with various classifiers such as radial basis function classifier, decision tree, credal decision tree, quadratic discriminatory analysis, Fisher linear discriminatory and classifier with radial basis. The study conducted in China in the Yongxin district used 364 landslides and 15 environmental factors. The results showed that the LWL-RS-ADT has surpassed others in terms of reliability and stability. Among the environmental factors, NDVI, lithology and altitude were identified as the most important in predicting the sensitivity of landslide. The proposed integration models have been induced as effective tools for the prediction of soil landslide.

S. Aldiansyah et al. [24] designed a model of foresting sensitivity, which combines the techniques of resampling, including cross validation, bootstrap and random subsampling, with a series of machine learning models such as support vector machine, random forest, generalized linear model, maximum regression tree Discriminating discriminatory disgraceful analysis, flexible discriminatory analysis, flexible disgraceful analysis, flexible disgraceful analysis, maximum regression tree, and regression. Probability and maximum entropy. The methodology was used in Kendari, an area affected by destructive erosion. The predictive accuracy of the model was assessed using metrics like AUC, TSS, COR, NMI and CCRTHE, achieved impressive power metrics with AUC of 0.97, COR of 0.99, NMI of 0.50, TSS of 0.97 and CCR of 0.93. The study concluded that these integrated models provide promising results to predict the landslide sensitivity and could be successfully applied in other regions.

M. Dahim et al. [25] focused on predicting the sensitivity of landslides in the area of Saudi Arabia by means of machine learning and deep learning algorithms, along with the sensitivity and analysis of uncertainty. The study took advantage of a random forest as a model of machine learning and a deep neural network as a model of deep learning, both of which were enthusiastically tuned through the grid search. The operating characteristics of the receiver, score F1 and F2, the Gini value and the accuracy curve were used to verify the models. Analysis of sensitivity and uncertainty made using the DNN model revealed the impact and uncertainty of various parameters on the occurrence of landslides. The results indicated that the RF and

DNN models predicted 35.1-41.32 km² and 15.14–16.2 km² of high and very high soil sensitivity zones. The DNN model reached 0.96, while the RF model won 0.93. The sensitivity analysis emphasized that the most important factor is collision, followed by the topographic wetness index, curvature, inclination, soil texture and lines density.

N. Sharma et al. [26] developed a probability framework for mapping sensitivity to landslides that deals with limitation of existing maps, such as small-scale data, heuristic methods, low, small study areas and spatial resolution. The framework combines the techniques of handling the imbalance and techniques of machine learning by means of support of vector machine synthetic oversampling technique to solve class imbalances and generate smaller representative data for model training. The technique of mixing file is used for lower uncertainty, which includes support vector machines, random forests and hyperparameter tuned ANN. This methodology is provided by the probability and class of landslides. With a resolution of 0.001 ° (~ 100 m), the frame was used to create the first Indian landslide maps of landslides at the national level and was divided into five levels. The map achieved of sensitivity With a sensitivity of 97.08%, accuracy of 95.73% and correlation coefficient Matthews 0.915 on test data showed an excellent generalization, robustness and accuracy. The model found new high -risk regions, including the Eastern Ghats regions that were not previously reported. The Indian map of sensitivity to landslide will be assumed to help model prediction models and reduce the risks of disaster.

T. Zeng et al. [27] investigated the impact of grading factors in landslide prediction modeling, addressing the subjectivity and randomness typically linked to this method. Focusing on the Wanzhou section of the Three Gorges Reservoir area, the research evaluated the performance of various machine learning models under different grading strategies, including non-grading, equal intervals, and natural breaks. The results indicated that the optimal grading strategy varies depending on the model used. For instance, the SVM model performed best with level 8 grading using natural breaks, while decision tree models were more effective without any grading. Deep learning models, such as Multi-Layer Perceptron Neural Networks and Convolutional Neural Networks, showed better results with natural breaks grading beyond 8 levels. Gated Recurrent Unit and Deep Neural Networks performed more effectively with equidistant grading of over 12 levels, while Long Short-Term Memory Networks excelled with equidistant grading exceeding 16 levels.

P. Priyanka et at. [28] proposed a model for soil movement prediction in areas susceptible to landslide Himachal Pradesh in India, where climate change intensifies the risks of landslide. The study has used models like long short-term memory, a convolutional neural network long short-term memory, convolution LSTM, encoder-decoder LSTM and the new model of the ensemble, Multi-LSTM-SVM, which is critical for understanding for landslides. Research, which was carried out in the Kamand valley with extensive monitoring systems, found that the SoilSense Multi-LSTM-SVM has reached 88.1% accuracy, overcame other models such as the LSTM and CNN-LSTM, which achieved 82.26% accuracy. adaptation. The study suggests that the refining of the fine models can further improve predictions and eventually help to reduce the risks of landslides and investigate lives and property in vulnerable areas.

A. Saha et al. [29] developed a model which used susceptibility of landslides combining the statistical model AHP and ML model SVM to predict landslides in Darjeeling district, West Bengal, India. The study identified 114 placement of land shells and divided them into training sets (70%) and validation (30%). Ten training factors were assessed, including rain, soil texture, waiting and geomorphology for analysis. The AHP-SVM using linear, polynomial, radial bases (RBF) and sigmoid algorithms generated four maps of the sensitivity of ground landslides. Among them, the AHP-SVM. Sigmoid showed the highest prediction performance and achieved a prediction capacity of 86.2%. This study concluded that the AHP-SVM Sigmoid model is a promising technique for mapping sensitivity to landslides that offers valuable ideas for local planning and decision-making and can be used for other regions for similar studies.

Q. Ge et al. [30] investigated how different elements of element selection techniques affect the productivity of machine learning model in prediction of offspring soil deposits in a deposit area of China. The research focused on the landslides of shuping and Baishuihe as case studies and evaluated four automatic learning algorithms: backup neuronal network, support vectors, short -term memory and closed recurring unit. Three characteristic engineering approaches were used: unprocessed multi-variation timing, autocorrelation functions of the maximum information coefficient and partial coefficient and relational analysis of GRA-PACF. The results revealed that static automatic learning models have improved significantly with the selection of suitable characteristics, while dynamic models such as LSTM and GRU, which inherently represent temporary formulas, showed only a slight improvement with engineering other features.

Y. Wang et al. [31] introduced various machine learning models for accessing the sensitivity in the Wushan region using models like random forest, logistics regression to compare the results and best model. The mentioned dataset uses 19 conditional parameters and train & test split is done is 80:20 ratio for further analysis. Different performance metrics were implemented to calculate the accuracy of model. The results provide insights that random forest outperformed all other models with 0.848 accuracy, 0.904 for area under curve, F1 score as 0.740. The research summarized as the random forest provided most efficient results and that proved to be one of most useful approach to assess the sensitivity and other parameters of landslides.

H. Wu et al. [32] the study included factors such as landslides, triggering factors and dams, and developed six predictive models using logistics regression, K-nearest neighbours, Support Vector Machine, Naive Bayes, decision tree and random forest. These models considered five factors, including the parameters of the geometry and attribute properties, and were compared with the dimensionless blockage index (DBI). The results showed that while the machine learning models corresponded to the accuracy of DBI, they provided benefits in situations where DBI cannot be used. Among the models achieved random forest with the highest performance, with 89% accuracy, 7% error rate, 15% false alarm rate and without uncertainty.

H. Shahabi et al. [33] evaluated the efficiency of three machine learning algorithms, such as decision tree (DT), random forest (RF) and support vector machine (SVM) to map the sensitivity of landslides, focusing on Kamyaran -Sarvabad Road in Iran and Kurdistan, and an area that was often influenced with Landslides. Fourteen factors of landslides, including inclination, aspect, height, river density, disorders and topographic indexes, were used as inputs for the MLA. The study identified 64 landslide seats using 70% for model training and the remaining 30% for verification. The model of the decision -making tree reached the highest area under the operational characteristic curve of the receiver 0.94, exceeded random forest (0.82) and support vector machines (0.75).

G. Tang et al. [34] developed a AutoML-based framework for global landslide sensitivity prediction (LSP) in two spatial resolutions (90 m - 1000 m), reaching the area under the operational characteristics of the receiver (AUC) over 0.96. Global prediction results were validated using regional landslides from three countries, three provinces and two prefecture data files. In addition, global LSP results at 90 m were used to increase regional predictions by incorporating areas with low and very low susceptible as samples of non -domestic soils. The

model has shown improved performance compared to the original global predictions. This study emphasizes the potential of intelligent learning methods for reliable global LSP applications.

S. Meng et al. [35] developed a deep learning framework that integrates the LSCDBN-WOA with Laplace function sparse regularized continuous deep belief network. This model addresses issues such as feature homogenization of continuous input variables, limited samples of landslide, and local optima during the training phase. Using a comprehensive database of 18 landslide conditioning factors, the study demonstrated that the LSCDBN-WOA model AUC = 0.964, RMSE = 0.174 outperformed the LSCDBN-GWO model AUC = 0.952, RMSE = 0.182 and the standalone LSCDBN model attain AUC = 0.913, RMSE = 0.291. The proposed LSCDBN-WOA framework also surpassed traditional machine learning models SVM, BP, RF, and LR and deep learning models RNN, CNN. The outcome highlights the effectiveness of the LSCDBN-WOA framework for landslide susceptibility assessment.

C. Chen et al. [36] proposed a deep learning model, Deep-Attention-LSF, designed for mapping of landslide susceptibility. This model assigns relevance scores to input contributing factors at local levels, improving the understanding of the factors influencing landslide events. DeepLIFT was implemented as an attribution branching network to interpret the relationship between the factors and landslide events. The model, which combines convolutional neural networks and long short-term memory networks, was tested on the Three Gorges Reservoir Area, using 18 landslide-related factors. The Deep-Attention-LSF model achieved high performance with accuracy - 0.9645, precision - 0.9676, recall - 0.9583and F1-score - 0.9522, outperforming other models such as self-attention LSM, random forest, and gradient boosting decision tree.

C. Yang et al. [37] developed the Bayesian optimization technique to maximize the sample ratio of landslide to a non-landslide assessment to assess the soil sensitivity based ML. The study focused on the edge of Anhua in Hunan province in China, which is the area susceptible to landslides. Three ML models, such as random forest, support vector machine and gradient increase, were used to assess sensitivity to landslides. The use of Bayesian optimization algorithm identified the optimum sample ratio, which improved the performance of the model. The finding has shown that higher power was the result of an optimized P/N ratio, with the RF gaining maximum or AUC 0.840, followed by GBDT 0.831 and SVM 0.775. In the LSA Study

models, the Bayes optimization technique works well to maximize the P/N sample ratio, while RF and GBDT are more suitable for solving imbalance problems.

M. A. Hussain et al. [38] updated the inventory of land landslides along the Karakoram highway (KKH), a critical route connecting South Asia, Central Asia and China, which is highly sensitive to landslides due to extreme geological conditions. The study was used by SBAS-InSAR and PS-InSAR technology and processed Sentinel-1 data from June 2021 to June 2023 to identify and measure slope deformation (Vslope). Among the 571 landslides that were found were 24 new landslides and some of the pre -defines the existing existing ones. The soil -sensitivity model was developed using an updated inventory that combined land landslides to causing factors. To evaluate deep learning models such as deep learning models such as CNN 2D, RNN, RF and XGBoost, 70% of training and 30% test part were used. The mapping was considered a total of fifteen elements causing landslide. The CNN 2D made the best and a map of landslide susceptibility that has been produced offers a useful risk control and risk prevention tool and helps to assess and alleviate risks.

Y. Liu et al. [39] developed a method of assessing the sensitivity of landslides, which combines information models with machine learning (ML) for more accurate forecasts and solves the problem of sample selection outside Landslide. The study focuses on the selection of samples without the land of the first screening of influential factors using a correlation analysis and then using a model of the value of the information value (IV) to define low and relatively low sensitivity. IV-ML models, such as IV-Logistic regression, IV-Random forest, IV-Support vector machine and IV-artificial neural network, were used to assess the sensitivity of landslides in the province of the province of Dabie in Anhui province. Compared to traditional ML models such as LR, RF, SVM and Ann, IV-ML models have shown significantly better performance in terms of accuracy, with improved ACC, AUC and Kappa values. This emphasized the increased efficiency of the proposed method for evaluation of soil landslide.

D. Sun et al. [40] focused on mapping of landslide sensitivity using interpretable machine learning, specifically exploring topographic differentiation. The study area included two different regions in Chongqing: zone I (corrosion layered high and middle mountain areas) and zone II (middle mountain area with strong regional feet). Bayes optimization was used to increase the parameters of the LightGBM and XGBoost models, with the most accurate model selected for soil sensitivity mapping. The SHAP (Shapey additive explanation) was applied to examine the molding mechanisms in both regions. The results showed that LightGBM

overcame XGBoost, with AUC values 0.8525 and 0.8859 for zones I and II. Common dominant factors for the occurrence of soil in both zones included altitude, soil use, section depth, distance from roads and annual collision.

This literature survey examines the application of ML and DL hybrid models to predict landslide using remote sensing images and data in real time collected from several sensors. The integration of satellite images, LiDAR, radar data and other technologies remote detection with sensor networks plays an essential role in the identification and predictions of areas susceptible to landslides. An overview emphasizes how hybrid models combining traditional algorithms, including support vector machines (SVM), random anticipation and K-nearest neighbors (KNN), along with advanced deep learning techniques such as deep beliefs networks (DBN), fully convention networks (FCN), offer improved prediction capabilities for landslide prediction and detection and risk assessment. In addition to remote exploration data, the overview focuses on data collection in real time of various sensors located in regions susceptible to landslides. These sensors capture key environmental factors such as soil moisture, rainy intensity, soil relocation, seismic activity and atmospheric conditions. Real time data is collected by ground sensors, meteorological stations and, among other things, unmanned aerial vehicles (UAV) and then feed on hybrid models of machine learning. Incorporating real -time sensor data improves the ability to monitor dynamic environmental changes and detect possible landslides and provide valuable information for early warning systems. Further advances are essential for improving early warning systems and efforts to prepare disasters. Table 2.1 provides a detailed summary of review of literature on landslide prediction using hybrid models of automatic learning and deep learning with sensor data in real time and remote sensor images.

Table 2.1 Summarization of literature review for Landslide Prediction

Author	Technique	Problem Statement	Performance Analysis	Limitation
N. Casagil et al. [1]	Remote Sensing Techniques (RSTs)	Landslide detection and monitoring using satellite and ground-based sensors to manage landslide risks	Multiple RSTs offer effective monitoring with high spatial and temporal flexibility. Ground-based systems offer accuracy for small areas	Integration of multiple RSTs is complex and resource-intensive
J. Barman et al. [2]	GIS-based Bivariate Statistical Approach, Index of Entropy (IOE)	Landslide susceptibility zonation (LSZ) in Lunglei, Mizoram	Scenario 1 (9 factors) outperformed Scenario 2 (17	Limited by the geographical scope of the study,

		to predict landslide-prone areas	factors), with AUC of 0.947 (Scenario 1) and 0.922 (Scenario 2)	overlooking other factors
M.T. Riaz et al. [3]	Machine Learning Algorithms, Clustering Partitioning	Evaluate alternative partitioning techniques for landslide susceptibility modeling in Muzaffarabad	Cluster-based partitioning method improved predictive accuracy with AUC-ROC values up to 0.962 for Random Forest model	Specific geographical features of Muzaffarabad may affect generalizability
P.C. Huang et al. [4]	Neural Network Algorithm, Clustering Method	Predict shallow landslides using geomorphological features and clustering methods	High prediction accuracy (mean relative error: 5.19%), with $R^2 >$ 0.889 and 82.2% of observed landslide events predicted	Relies heavily on geomorphological data, limiting generalization to other regions
P. Varangaonkar et al. [5]	Remote Sensing, CNN, LSTM, ANN, SVM	Automatic landslide detection and region localization using remote sensing images	CNN-LSTM model outperformed traditional methods, improving accuracy by 2% and reducing complexity by 35%	Requires computationally intensive resources for real-time applications
Y.A. Nanehkaran et al. [6]	Artificial Neural Networks (ANNs)	Assess riverside landslide susceptibility for better urban planning and disaster resilience	ANNs improved risk management strategies for riverside areas, supporting sustainable urban development	Limited to riverside areas, unsuitable for non-riverside regions
C. Zhou et al. [7]	Multi-Temporal InSAR, Gated Recurrent Units (GRU)	Cost-effective displacement prediction for landslides using InSAR and machine learning	GRU model provided high accuracy in displacement prediction with RMSE values of 3.817 and 5.145	Dependent on satellite imagery, which may be inaccessible or expensive in some regions
S. Alqadhi et al. [8]	Deep Neural Networks (DNN), 1D CNN, DCN, XAI	Enhance landslide prediction by integrating deep learning and explainable AI	DCN model achieved AUC of 0.97, outperforming DNN (AUC: 0.9) and CNN (AUC: 0.94)	XAI increases complexity and computational demand
H. Ishibashi et al. [9]	Machine Learning (Random Forest, LightGBM), Rainfall Hazard Curve	Assess economic risk from rainfall-induced landslides and improve resilience of structures	LightGBM outperformed Random Forest with higher accuracy for rainfall-induced landslide susceptibility	Does not account for non-rainfall-induced landslides in areas with different triggering factors
L.L. Liu et al. [10]	Classical Models (Verhulst, GM (1,1), Saito), Ensemble Models	Predict landslide failure time using displacement monitoring data	Ensemble models, especially decision tree regression-based, outperformed classical models in prediction accuracy	Classical models had high prediction errors (e.g., GM (1,1) with high mean absolute error)
H. Harsa et al. [11]	Machine Learning and AI, Precipitation Data	Landslide event prediction in Indonesia using	Best models: Generalized Linear Model (AUC:	Performance varies with different satellite data or

		satellite precipitation data	0.828) and Deep Learning (AUC: 0.836)	geographic conditions
Z. Chang et al. [12]	Multi-Scale Segmentation (MSS), Machine Learning	Landslide susceptibility prediction addressing the heterogeneity of conditioning factors	Models incorporating internal variations within slope units performed better than grid-based models	High complexity due to multi-factor modeling and data requirements
L. Nava et al. [13]	Deep Learning (MLP, LSTM, GRU, Conv-LSTM)	Forecast landslide displacement and improve early warning systems	Conv-LSTM performed best in predicting displacements, particularly in seasonal landslides	Does not generalize well to all types of landslides
Y. Shen et al. [14]	Machine Learning, Neural Networks, Geomorphological Indices	Enhance landslide susceptibility mapping with machine learning and geomorphological data	SVM and Kernel Sigmoid algorithms achieved the highest accuracy	Does not generalize well outside the studied geographical area
Z. Chang et al. [15]	Slope Unit-based Machine Learning Models	Study uncertainty in selecting non-landslide samples for landslide susceptibility prediction	Maximum probability analysis (MPA) reduced uncertainty effectively	Dependent on large-scale data collection, not suitable for small areas
L. Achu et al. [16]	Ensemble Machine Learning Techniques	Quantify uncertainty in landslide susceptibility prediction	Weighted mean of probabilities model was the most effective for uncertainty quantification	High computational demands for large-area predictions
Y. Wie et al. [17]	Machine Learning (Random Forest, Logistic Regression, GBDT), InSAR	Improved landslide susceptibility assessment using machine learning integrated with InSAR	LR-InSAR model showed superior performance in identifying high-susceptibility areas	Limited by data availability and unsuitable for regions with minimal InSAR data
L. Chen et al. [18]	Spatio-Temporal Knowledge Graph, Machine Learning	Address challenges in organizing multi-source remote sensing data for landslide prediction	29% increase in F1 score and 93% improvement in processing efficiency	Struggles in areas with poor data availability or inaccessible regions
C. Chen et al. [19]	Machine Learning and Deep Learning Models, Factor Selection Methods	Investigate the impact of contributing factor selection on landslide prediction accuracy	Autoencoder-based factor selection improved DL model accuracy	Limited to factors studied, unsuitable for other regions or datasets
N. Nocentini et al. [20]	Random Forest (RF), Dynamic Variables	Combine static and dynamic factors for landslide susceptibility assessment	RF model with dynamic variables accurately forecasted landslide occurrence based on rainfall data	Does not perform well in regions with limited rainfall data or varied climatic conditions
T. Xiao et al. [21]	Machine Learning	Incomplete landslide	Probabilistic landslide model	Model performance is dependent on the

		inventories, particularly missing landslide timing data.	outperformed other data-driven methods in predicting the spatio-temporal evolution of landslides during rainstorms and assessing landslide risk.	availability of landslide data for validation.
K. Doerksen et al. [22]	Machine Learning, Deep Learning, AI	Complex causal chain of landslides in Nepal due to large earthquakes and intense monsoon rainfall.	Random Forest and U-Net models demonstrated strong predictive power, with feature importance analysis providing insights into the causal factors.	Limited by the availability and resolution of open-source space-based data.
H. Hong et al. [23]	Locally Weighted Learning (LWL) with various classifiers (RBF, FLDA, QDA, CDT, ADT, RS)	Need for reliable landslide susceptibility models in areas with limited data.	LWL-RS-ADT model showed superior reliability and stability, with NDVI, lithology, and altitude as key predictive factors.	May be sensitive to the choice of classifiers and model parameters.
S. Aldiansyah et al. [24]	Machine Learning models (GLM, SVM, RF, BRT, CRT, MARS, MDA, FDA, MaxEnt, MaxLike) integrated with resampling techniques	Landslide susceptibility prediction for regions with destructive erosion.	Resampling algorithms enhanced the performance of models; Bt-RF model showed highest statistical performance (AUC=0.97).	Limited to the performance of the resampling algorithm used in model integration.
M. Dahim et al. [25]	Hyper-tuned Machine Learning (RF) and Deep Learning (DNN)	Landslide susceptibility prediction in Saudi Arabia with sensitivity and uncertainty analysis.	RF and DNN models achieved high prediction accuracy (AUC: RF=0.93, DNN=0.96).	Sensitivity and uncertainty analysis highlight limitations in parameter influence.
N. Sharma et al. [26]	Ensemble machine learning (SVM-SMOTE, ANN, RF, SVM)	Issues with limited data, low spatial resolution, and small study areas in landslide mapping.	Developed a national-scale landslide susceptibility map with high accuracy (95.73%) and sensitivity (97.08%).	Model may require refinement for broader application in different regions.
T. Zeng et al. [27]	Machine Learning (SVM, DT, DNN, GRU, LSTM) with grading conditioning factors	Subjectivity and randomness in grading strategies for landslide susceptibility modeling.	Grading strategy improved performance of deep learning models; model performance varied with different models.	Performance varies based on the choice of grading strategy and model used.
P. Priyanka et al. [28]	Machine Learning (LSTM, CNN-LSTM, Conv-LSTM, Multi-LSTM-SVM)	Escalating landslide risks due to climate change in Himachal Pradesh, India.	SoilSense Multi-LSTM-SVM model outperformed other models (88.1% accuracy).	Performance may be limited by the availability of accurate antecedent rainfall data.

A. Saha et al. [29]	Ensemble models (AHP + SVM)	Landslide susceptibility mapping in Darjeeling, India.	AHP-SVM model showed 86.2% prediction accuracy.	Model effectiveness may depend on the choice of algorithms and input variables.
Q. Ge et al. [30]	Machine Learning (BPNN, SVM, LSTM, GRU) with feature selection techniques	Impact of feature selection techniques on landslide displacement prediction.	Static ML models benefited significantly from feature selection; dynamic models showed marginal gains.	Optimal feature selection varies by model and specific landslide characteristics.
Y. Wang et al. [31]	Machine Learning (RF, Logistic Regression, Extreme Gradient Boosting)	Landslide susceptibility assessment in Wushan County.	Random Forest model outperformed others with higher AUC, F1 score, and accuracy (0.848).	Dependent on the quality of soil thickness and other conditioning factors.
H. Wu et al. [32]	Machine Learning models (Logistic Regression, KNN, SVM, Naïve Bayes, DT, RF)	Landslide dam life span prediction.	Random Forest showed highest performance (89% accuracy) for predicting dam life spans.	Models may not address all types of landslide dam scenarios.
H. Shahabi et al. [33]	Machine Learning models (RF, DT, SVM)	Landslide susceptibility mapping for the Kamyaran–Sarvabad road in Iran.	Decision Tree model performed best (AUC=0.94).	Limited by available input variables and spatial scale of study.
G. Tang et al. [34]	AutoML-based framework	Need for global landslide susceptibility prediction across different resolutions.	The model achieved an AUC > 0.96 and improved regional predictions.	Results are dependent on the availability of regional landslide inventories for validation.
S. Meng et al. [35]	Deep Learning (LSCDBN, GWO, WOA)	Challenges such as feature homogenization and local optima in landslide susceptibility modeling.	LSCDBN-WOA model outperformed others with AUC = 0.964 and RMSE = 0.174.	Limited by the quality of landslide conditioning factors used in model training.
C. Chen et al. [36]	Deep Learning (Deep-Attention-LSF, CNN, LSTM)	Need for interpretable models in landslide susceptibility mapping.	Deep-Attention-LSF model achieved high accuracy (0.9645) and was more interpretable than other models.	Interpretability is limited to the quality of input factors.
C. Yang et al. [37]	Machine Learning (SVM, RF, GBDT) with Bayesian optimization	Optimizing sample ratio for machine learning models to address landslide susceptibility.	Bayesian optimization improved performance, with RF achieving the highest AUC of 0.840.	Performance dependent on optimizing the P/N sample ratio for each case.
M. A. Hussain et al. [38]	ML and DL (CNN, RNN, RF, XGBoost)	Updating landslide inventory along the Karakoram Highway (KKH).	CNN 2D model demonstrated best performance in landslide	Limited by the quality and availability of geospatial data.

			susceptibility mapping.	
Y. Liu et al. [39]	Machine Learning (IV-Logistic Regression, IV-Random Forest, IV-SVM, IV-ANN)	Non-landslide sample selection in landslide susceptibility evaluation.	IV-ML models significantly outperformed traditional ML models in terms of accuracy and other metrics.	Model's effectiveness may vary across different regions and datasets.
D. Sun et al. [40]	Interpretable Machine Learning (LightGBM, XGBoost, SHAP)	Topographic differentiation in landslide susceptibility mapping.	LightGBM outperformed XGBoost, with AUC values of 0.8525 and 0.8859 for Zones I and II, respectively.	The model is limited to the selected zones and topographic factors.

2.3 Fundamental of Earthquake Prediction using Machine Learning

The phrase earthquake prediction refers to the process of predicting seismic occurrences using real-time data and sophisticated computing methods. Therefore, the goal of earthquake prediction is to increase the accuracy of earthquake predictions by utilizing data from several sensors, including accelerometers, seismometers, and GPS devices. So, it begins with data collection from different sources like real-time sensors, historic data and acoustic signal captured over time. The signals and values obtained from them provides the early alert signals for evacuation and safety measures. Most commonly used prediction technique specifically for time series forecasting involves the use of SARIMA (Seasonal Autoregressive Integrated Moving Average), this model helps in capturing temporal patters formed and variations which are seasonal in nature. Further, to capture the non-linear relationships XGBoost model is used to improve prediction by analyzing large and complex dataset to find the patterns. Finally, after completion of training model is ready for prediction of probability of earthquake and lastly provide necessary strategy for warning and alerts. The main objective of this literature survey is to assess various aspects related to earthquake prediction. In analysis phase we tried to present the essential principals of earthquake prediction. Our goal is to provide with contribution that surrounds its potential revolution in disaster preparedness and strategical responses.

V. Macchiarulo et al. [41] developed approach for evaluation of regional-scale post-earthquake damage using post-event very high-resolution, synthetic aperture radar imagery and

machine learning. The method utilizes supervised learning on specific datasets and it was tested on random study area to assess the ML model adaptability. The model outperformed with achieving 72% accuracy by classification of collapsed building in that region. The framework provided potential for improving disaster preparedness and management techniques.

F. H. Chen et al. [42], developed earthquake recognition and warning systems that integrate Arduino, sensors and transmission technologies to improve the security of earthquake attacks in Taiwan. The system uses a capacitive 3-axis acceleration meter to measure vibrations and early earthquake warnings. Additionally, it includes an IR flame sensor to identify fires and an MQ series air quality sensor to monitor harmful gas concentrations. When it recognizes an earthquake, the system warns individuals of evacuation. When the harmful gas exceeds critical levels, the system activates a warning light and exhaust gas to extract the toxic gas. If a flame occurs, an alarm is triggered to arrange for a quick evacuation. The proposed system is affordable and easy to offer, providing immediate notifications. This provides a valuable tool for improving disaster response and security in earthquake zones.

E. M. A Alcantara et al. [43] proposed an approach to forecast the damage state of buildings in reinforced concrete (RC) resistance moment frames using machine learning technology (machine learning). This study includes the design of structural members of RC buildings, with X and Y story counters and direction numbers using virtual working methods. The purpose of this study is to split earthquakes, construct data records, and split tests of data records to predict damage conditions in new buildings and to reduce bias, multiple random selections were made, and prediction accuracy was measured using the mean and standard deviation. The study utilized 27 Intensity Measures based on ground and roof sensor data, which included acceleration, velocity, and displacement to analyze building behavior. The input data for the machine learning methods consisted of IMs, the instance of stories, and spans, while the outcome data was the maximum inter-story drift ratio. Seven machine learning methods were tested to identify the optimal combination of training buildings, IMs, and methods for the highest prediction accuracy.

B. Tian et al. [44] reviewed the development and application of movement-detection sensors, emphasizing their importance in understanding surface movement and tectonic activities. Modern sensors have significantly contributed to various aspects of earthquake management, including monitoring, prediction, early warning, emergency response, search and rescue, and life detection. The study classified sensors based on earthquake timelines, their physical or

chemical mechanisms, and sensor platform locations. The analysis focused on sensor platforms that have become widely used in recent years, particularly satellites and UAVs. The study's findings offer valuable insights for improving future earthquake response, relief efforts, and research aimed at reducing earthquake-related risks.

M. E. Tusun et al. [45] developed a sensor system utilizing strain gauge technology, optimized for detecting earthquake waves and structural vibrations, to address the limitations of traditional damage detection methods in buildings. These methods, such as cross-sectioning of columns, are time-consuming, costly, and lack continuous monitoring capabilities. Accelerometers, although useful, are inadequate for detecting low-frequency earthquake waves. The proposed system collects vibration data, applies Fourier transform to obtain the frequency response of the structure, and detects shifts in this response to classify the structure's damage condition as intact, slightly damaged, or very damaged. This classification is performed using a deep learning model running on a low-power microcontroller. The results demonstrate that the developed sensor is more effective than accelerometers in assessing structural health and enables real-time damage evaluation using Fourier transform and machine learning techniques.

M. Bhatia et al. [46] developed a cooperative monitoring and prediction system focused on collaborative IoT-Edge-centred, which combines marginal and cloud computing to provide warnings of early earthquakes for high-risk areas. Real -time sensor data collects the system using Internet of Things and is sent to a marginal layer for categorizing functions using the unique Bayes belief method. To predict magnitude earthquakes, the cloud layer also uses an adaptive neuro-fuzzy inference system (ANFIS). Through achievement of good classification performance with an accuracy of 92.52%, sensitivity 91.72%and specificity 91.01%showed experimental simulation efficiency of the frame. The system also showed a significant reduction in computing delay (23.06s) through the edge calculation. In addition, the model showed high reliability (95.26%) and stability (92.16%), ensuring increased accuracy and permeability for the prediction of the earthquake.

P. Govindarajan et al. [47] developed a new way of predicting real -time earthquake in Chile a combination of machine learning techniques with intelligent technologies. Due to the serious threats that the earthquake represents Chilean people and infrastructure, the study uses AI and ML to overcome the deficiencies of conventional prediction techniques and increase the speed and accuracy of prediction. The proposed technique combines improved analysis of neuron

network time series, LSTM-IC (long short-term memory inverse correlations), with modified cluster strategy, LMSCAN (local maximum based spatio-cluster). This strategy uses sensor network, sophisticated predictive algorithms and previous seismic data, provides early warnings, improves response to disasters and resistance. The model achieved a remarkable accuracy of 95%, showing its exceptional learning and adaptability, distinguishing it from other methods of forecast and offering significant progress in the prediction of the earthquake.

MS. Abdalzaher et al. [48] proposed the integration of the EEWS Early Warning System into intelligent cities to improve disaster management and preservation of human lives. The system uses the Internet of Things to collect data from various EEWS entities and machine learning technology to analyze this data for effective decision making. The article examines the key EEWS components, starting with the IoT role in monitoring the earthquake parameters. He then classifies ML models to linear or non -linear categories and discusses the assessment of metrics focusing on seismology. The study represents a taxonomy that emphasizes the emerging efforts of ML and IoT for EEWS, followed by the design of the EEWS generic architecture based on these technologies. Finally, the article examines how ML can increase the observations of the earthquake parameters, which eventually leads to a more efficient EEWS.

W. Huang et al. [49] introduced the use of data from granular fault tests for the creation of a ML method for the prediction of earthquake. They gathered data dynamics, such as shifting and speed, on 2203 sensors, and used the combined method of finite discrete method (FDEM) to model two -dimensional cut granular failure system. The LightGBM algorithm has been trained using this data to predict the Gouge-Plate coefficient, which represents the behavior of the wand and the friction state of the error. This study optimized the data by assessing the importance of input elements and selected the most important for prediction. The model reached a value of R^2 0.94, showing high accuracy. Additionally, values were calculated to evaluate the contribution of each input function of the prediction. The results show that LightGBM, along with form values, can accurately predict the frictional state of laboratory faults and identify the most important input functions for predicting earthquakes. This study provides potential knowledge about natural earthquake prediction and the use of artificial intelligence to study earthquake predecessors.

P. Lara et al. [50] prepared a model using P wave data capturing by a single station in less than 3 seconds, the Earthquake early warning system (E3WS) is designed for identifying,

locating and estimating the amount of the earthquake. The ensemble ML algorithms, which consist of the E3W, have been trained using data on the time series of ground acceleration from the global data set, Peru, Chile and Japan, analysis of temporary, spectral and Cepstral attributes. The three steps of the system operation are detection, selection of P-phase and the characterization of the source that includes azimuth, depth, size and estimate of epicenter distance. Without false positives and several false negatives (only for the earthquake $M \leq 4.3$), the E3SS has an amazing 99.9% success in distinguishing between earthquakes and noise. For collecting P-phase, the average system's absolute error is 0.14 s sufficient for early warning of the earthquake. Compared to the current one with one station, the E3Ws offer objective and extremely accurate estimates to characterize the source, especially for the size and somewhat improved accuracy of the earthquake location. By updating estimates, E3Ws offer E3Ws estimates -dependent -dependent -dependent and provides faster predictions than current multiple stations systems, which provides more time for protective actions.

M. S. Abdalzahar et al. [51] proposed a new method for EEWS (ML) early warning systems that use machine learning techniques (ML) to analyze seismic activity in two seconds after P-wave begins to quickly assess the intensity of the earthquake. The 2S1C1S evaluates the force of the earthquake using data from one component and one station. After being trained on a large data set known as an "instance", which contains information from hundreds of stations in the Italian National Seismic Network (INSN), the model examines 50,000 occurrences or 150,000 seismic windows after two seconds. With a stunning rate of accuracy of 99.05%, the algorithm predicts the severity of the earthquake by identifying important elements from the tracks of the wave shape. The centralized IoT system, which includes the 2S1C1S paradigm, enables rapid transfer of alarm to the authorities for early response. Compared to traditional manual techniques, the 2S1C1S with extreme gradient boosting (XGB) works better in estimates than a number of comparative machine learning values, which shows its usefulness to EEWS applications.

A. Joshi et al. [52] designed a cross-region prediction model called SeisEML (Seismological Ensemble Machine Learning) to forecast peak ground acceleration (PGA) at a specific location during an earthquake. The hybridization models, tree regression algorithms, kernel-based algorithms, and other regression techniques are all combined in the SeisEML model. A study on model ablation was carried out to assess the effectiveness and choice of meta-machine learning models in SeisEML. There are 20,852 and 6,256 accelerograms from the Kyoshin

Network in Japan that make up the training and testing datasets. SeisEML reduces both the mean absolute error and root mean square error by around half when compared to traditional attenuation relations, according to a comparison of the model's performance using these metrics. The iso acceleration contour map of Japan was created using the model for three earthquakes of magnitudes of 7.4, 6.6, and 6.1. A qualitative comparison of the iso acceleration contours from actual and predicted PGA showed that SeisEML can reliably predict PGA distributions. Additionally, the model was tested for Iranian earthquakes, outperforming regional attenuation relations in terms of MAE and RMSE.

W. Zhu et al. [53] developed a chain machine learning models (ML) to predict multiple seismic reactions to the center of the distribution device during strong earthquakes by means of intensity (IMS) measures. The models are designed to predict answers to multiple vulnerable positions like porcelain insulators and connection flanges, by connecting several individual models in the sequence. One model is a simple chain where the output of one model becomes the input for the next, while the other combines the intensity measures with the previous output as the input for the next model. The training of these ML chain models is optimized using bio-inspired multi-objective techniques for selecting hyperparameters. A case study involving a 1100 kV transformer bushing is used to establish ANN-gradient boosting regression and ANN-kernel ridge regression models for predicting peak stresses at the most susceptible positions. The results, including evaluation indicators and shaking table tests, demonstrate that both ML chain models provide accurate predictions. These models are effective for detecting initial equipment damage and can be used to support post-earthquake rapid judgment and relief efforts.

K. C. Sajan et al. [54] designed an approach based on artificial intelligence to predict the intervention of damage and rehabilitation after the earthquake, especially after the earthquake Gorkha in Nepal in 2015. The study analyzes comprehensive information on the building of 549 251 impacts on buildings and intensity of ground shocks on the use of ML methods (ml) to predict the scope of damage. The models were created, and their performance was evaluated by four popular machine learning algorithms: logistics regression, XGBoost, Random Forest and decision tree. The finding has shown that if the building collapse was predicted and the need to strengthen, XGBoost led better than other algorithms. In addition, 19 of the 20 best features were found to predict the degree of injury and rehabilitation therapy using an analysis of important importance from the XGBoost. Compared to typical fragility functions, which are

often ambiguous and difficult to use in specific locations, our method provides a more accurate forecast.

C. E. Yavas et al. [55] analyzed machine learning and neural network to develop a new method for earthquake prediction of Los Angeles earthquake in California. They connected previous work with new information for improving accuracy of model. They achieved high results by forming a matrix that predicts the estimation of highest size of earthquake. This study emphasizes how machine learning and neuron networks can revolutionize the accuracy of the prediction of the earthquake, which significantly increases seismic risk management and readiness.

K. A. Yusof et al. [56] explored the potential of geomagnetic anomalies as precursors to the earthquake (EQs) and focused on the development of practical models of earthquake prediction using AutoML. The work has used geomagnetic field data from 131 global magnetometer observatory over 50 years. To extract functions that were then fed into models that were optimized by asynchronous consecutive algorithm (ASHA) and automatic methods of methods and enhancement of hyperparameters and automatic methods and automatic methods of methods and hyperparameter improvements Selection of methods and automatic methods of method and automatic selection of methods and wavelet selecting transform (WST) and after optimized asynchronous consecutive half algorithm. With an accuracy of 83.29%, the model of the neural network (NN) exceeded the five other classification methods tested. The results show that even for complicated systems such as lithospheric and seismo-induced geomagnetic processes, the automobile can facilitate useful models of earthquake prediction.

K. Qaedi et al. [57] examined how to improve the accuracy of the Earthquake prediction (EQ) by reducing the complexity of the global data of the geomagnetic field by analyzing the principal component analysis (PCA). Further for prediction of EQ intensity multiple classes (ML) were built also SMOTE analysis was executed to solve the imbalanced data. The final results show promising values as accuracy of 77.50%, F1-score of 76%. So, the principal component-based ML model is used for prediction of earthquake with optimal accuracy.

S. OMMI et al. [58] designed a model for predicting large earthquakes by studying changes in seismicity and the potential occurrence of significant seismic events in the seismic zone. This research is not only necessary for seismological studies, but also for informed decisions on crisis management. To analyze it, they tested several machine learning techniques (ML),

including artificial neural networks (Ann), Random Forest (RF) and supporting vector machines (SVM). The study focused on the seismic catalog of northern Zagros, seismically active areas with large cities. Nine seismic parameters were used to predict the likelihood of the large earthquake that occurs within a month. The accuracy of the models was evaluated by four statistical measures: evocation, accuracy, accuracy and score F1. The results revealed that the Ann method overcame others, especially for predicting larger earthquakes.

A. Berhich et al. [59] designed a model of earthquake prediction dependent on the spot using recurring neural network algorithms. This approach includes clustering of seismic data based on geographical parameters (length and latitude) using the K-Means algorithm. Each cluster is further divided into two subgroups: one for events between 2 and 5 and the other for those who have more than 5 sizes. This cluster allows models to focus on specific regions and more precisely capture regional trends. In addition, large earthquakes that have less events are trained independently to prevent interference from smaller, more common. LSTM, GRU and Hybrid models LSTM-GRU tested data from Morocco, Japan and Turkey. Their performance is evaluated using metrics such as an average absolute error, an average error for the second and the root diameter error. The results show that models overcome other existing methods, especially when predicting large earthquakes.

M. H. Al Banna et al. [60] proposed systematic research of AI based techniques to predict the earthquake. The study reviews 84 scientific papers from various academic databases that report the use of artificial intelligence methods in predicting earthquake. These techniques include methods based on rules, shallow machine learning and deep learning algorithms. The post provides an overview of these methodologies and offers a comparative analysis of their performance with regard to the data and evaluation metrics used. The aim is to help select the most suitable techniques to predict an earthquake -based comparison. In addition, the contribution deals with continuing challenges and potential future directions of future research in this area.

S. Mujherjee et al. [61] proposed a novel Ensemble Earthquake Prediction Method (EEPM) aimed at improving the accuracy, reducing variance, and minimizing errors in earthquake prediction. The method uses a combination of continuous data parameters collected from India and Nepal, along with categorical surveyor's data (precursors) gathered from India, Nepal, and Kenya over five years. The data is preprocessed by merging both types of information. EEPM focuses on detecting early signs of earthquakes and calculating the probability of occurrence

in specific regions. The results show that EEPM outperforms individual machine learning models, achieving a higher R2 value, lower variance, and less error, with an accuracy rate of 87.8%. This prediction model not only helps alert society but also aids organizations in understanding the potential magnitude and dynamics of an earthquake's occurrence.

R. Yuan et al. [62] designed a seismic prediction model that uses clustering of global earthquake data. In order to deal with the limitation of traditional clustering K-Simple-for example, the need to define the number of clusters, any selection of initial centers and the lack of parameter. This study introduced an improved algorithm K-significant. This improvement takes into account the maximum minimum distance of STM and the distance space distance when selecting the initial cluster centers. The number of evaluation criteria, such as the sum of square errors, calculates the number of Davies - Bouldin clusters, the Calinski - Harabasz and the Silhouette coefficient. In addition, the model uses an artificial neural network to predict the earthquake in conjunction with the findings of clustering. When the improved technique was applied to global seismic data USGS from 1900 to 2019, the accuracy of clustering over K-Means conventional approach. In addition, this method worked well for the analysis of local seismic risks and showed a promise to predict future earthquakes.

A. Berhich et al. [63] examined a long short -term memory network (LSTM), which is based on the attention for predicting the location, size and timing of large earthquakes. The predication of the earthquake features is difficult due to their complexity and lack of different formulas. While the attention mechanism focuses on the extraction of significant patterns and information from input characteristics, LSTM is used to record time correlations. Because the region is experiencing a lot of seismic activities and large earthquakes, the Japanese data file for earthquakes, which lasts 1900 - October 2021, was used. Metrics including MSE, RMSE, MAE, R -squared and accuracy were used to assess the performance of the model. As MSE increases by almost 60% of the date, the proposed model works noticeably better than alternative empirical methods and the chosen baseline.

A. A. Mir et al. [64] focused on predicting anomalies in the concentration of soil radon gas caused by seismic activities using various methods of files and individual machine learning. The study used file methods such as a strengthened tree, a poaching trolley and a strengthened linear model, along with individual methods such as a SVM with linear and radial cores and KNN. The methods were tested on the radon time series collected from the failure line in Muzaffarabad between March 1, 2017 and 11 May 2018, which contained data from nine

earthquakes. To minimize noise in performance estimates, the models were evaluated using a ten -fold cross validation process, which was repeated ten times. Metrics such as RMSE, RMSLE, MAPE, PB and MSE were used to evaluate performance. Setting 1 was best for SVM using a radial core that produced the lowest RMSE 1381.023. SVM worked best on setting 3, where RMSE varied from 1262,864 to 1409,616. The model of the strengthened tree had the lowest map (0.056) and RMSE (1573,174) in settings 4. The study found that the method of strengthened tree is particularly accurate for automatic predictions from environmental parameters and SVM core and strengthened the activities.

C. Wang et al. [65] designed a model for monitoring and collecting signals by precursor anomaly before the earthquake for seismic prediction, was created and acoustic and electromagnetics for artificial intelligence in China. To find the enhanced architecture for the prediction of the earthquake, this study evaluates a number of traditional models of time series and non-time series. The neuron network of long short -term memory (LSTM) was selected to predict the real -time earthquake during the 16 -week period, as it brought the best results based on AETA from precursor anomalies of signals.

B. Zhang et al. [66] developed an EPT, a deep learning data model to predict the earthquake that overcomes the disadvantages of earlier methods that only local seismic data used. To improve the prediction of the main Mainshock in specific areas, the model uses closed blocks of elements to extract basic patterns in the movements of the plate and by crustal movement from global historical seismic catalog data. Using this method, up to 50% more predictions are performed. The model also overcomes the difficulties of LSTM network, which it encounters in the processing of long -term data using multi -headed self -confidence to identify long -term dependencies in regional time series. The EPT was verified on five provincial data sets and in all cases, it achieved more than 90% accuracy.

Q. Wang et al. [67] dealt with an important problem for the prediction of the earthquake by means of long short -term MEMORY (LSTM) network to use spatial correlations between earthquakes in different places. Traditional methods of prediction, such as mathematical analysis, decision -making trees and support vector machines, often fight for the dynamic and unpredictable nature of the earthquake. However, the authors acknowledge that the earthquakes are influenced by the movement of the crust, and suggest that predictions not only consider local data but also historical data from a larger area. The results of their simulation show that

the LSTM network using a two -dimensional input effectively captures these space -time relationships, leading to an improvement in the accuracy of prediction.

Z. Zhang et al. [68] designed a new method for the prediction of the earthquake that combines a ConvLSTM with a sequence. This network gains knowledge of global time and spatial correlations of seismic data. This method overcomes the limitation of existing approaches, which are often limited to local areas and fully remove spatial correlations and resolution. The proposed approach includes the creation of a Spatiotemporal series from global seismic maps with high resolution, spatial distortion by turning maps, and using a weighted MSE-MAE to focus on the area of the earthquake. It also integrates a 4-D data file that includes earthquakes and depth. The output demonstrate that the method exceeds existing models and achieves an average download of 51.83% and an accuracy of 64.54% per test kit, with a pixel resolution of 72.92 × 67.71 km. These findings emphasize the ability of the model to predict an earthquake with higher resolution and accuracy and provide valuable knowledge about global seismic samples of activity.

M. Akhondzadeh et al. [69] reviewed the progress and challenges in the prediction of an earthquake using satellite data, emphasizing the potential for creating earthquake warning systems. Due to the limits of data on in-Situ, including the quantity, location and expenditure of ground stations, the precise forecast of the earthquake has not yet been carried out despite extensive research. However, with the development of satellites with a long -distance survey, statistical research of the earthquake precursors has dramatically increased and focused on unusual changes in physical and chemical parameters that occur one to thirty days before significant earthquakes. The report emphasizes recent developments, such as an increase in satellites devoted to earthquake research, the availability of different earthquake precursors and creating more methods of identification and prediction. In addition, progress in cloud data storage and processing services (such as Google Earth Engine and Giovanni), together with the creation of intelligent integration systems for integrating and analyzing multiple precursors in the near future increased optimism with low uncertainty.

Z. Ye et al. [70] presented a long-term short-term memory model (LSTM) to predict seismic size along with elite genetic selection (EGA) with genetic algorithm function (EGA-LSTM). The time series structure of seismic data and dual properties provides the difficulties of overcoming this method. To find serious correlations, the technology pre-combine electromagnetic and acoustic data from Roulette-based EGA and AETA systems. LSTM uses

selected features to estimate the size. Entire procedure includes fitness components such as RMSE and the ratio of selected properties. Using data from four different locations in China, the models were evaluated taking into account different periods and weights of fitness functions. The results show that EGA-LSTM exceeds several metrics including EV, Mae, MSE, RMSE and R2. Non-parametric testing confirms that EGA-LSTM significantly exceeds the standard LSTM model.

The literature overview deepens the diverse applications of the ML and DL hybrid models in the prediction of the earthquake and emphasizes the integration of historical seismic data records, geophysical data and sensor networks. Also, explore how historical earthquake data is critical to the training of predictive models that provide information about samples, trends and relationships in space-time, which may not be recognized in real time. This overview emphasizes the development and improvement of hybrid models combining several algorithms, such as decision -making trees, neural networks, SVM, KNN and random forests, to determine the decision to improve the precision and provision of real -time predictions. Advanced architecture for deep learning, such as CNN, RNN, long short time network (LSTM) and generative contradictory networks (GAN), also use to extract functions, analysis of time series and detect anomaly in seismic data. The installation of historical seismic data records, often covering decades or centuries, plays a key role in enhancing the predictive power of ML and DL models. These data sets contribute to the capture of long -term trends and seismic formulas, so the model can identify the correlation between predicted events and the occurrence of earthquakes. This overview also analyzes the importance of preliminary data processing techniques for noise processing, missing data and irregularities in historical data sets of earthquakes. In addition, to improve robustness and modeling, the use of methods, including pulses and closure techniques, can be examined.

The survey describes the use of hybrid models, including a combination of SVM and neural networks and integration of DBN with RNN for relationships with uncomplete lines in seismic data. The integration of strengthening learning (RL) and transmission learning to predictive earthquake models is also examined because they promise to adapt to a new seismic environment with minimal data. Literature also describes the value of the inclusion of specific domain knowledge, such as tectonic plates and geological properties, along with historical data sets to enhance the accuracy and interpretability of the model. This includes the difficulty of receiving high quality data and the challenge of converting historical formulas into processable

predictions. Despite these challenges, hybrid models discussed in this literature have an important potential for progress in techniques of earthquake predictions supported by historical earthquake data. These models provide promising solutions for early alarm systems and disasters readiness. Table 2.2 includes a detailed summary of an overview of literature for predicting earthquakes by means of ML and hybrid models of DL.

Table 2.2 Summarization of literature review for Earthquake Forecasting

Author	Technique	Problem Statement	Performance Analysis	Limitation
V. Macchiarulo et al. [41]	Machine Learning (ML) & Very High-Resolution (VHR) SAR imagery	Post-earthquake damage assessment using SAR imagery and ML for regions not previously studied.	Achieved 72% accuracy to classify standing and collapsed buildings in new regions.	Limited to specific earthquake cases (2021 Nippes, 2023 Kahramanmaraş).
F. H. Chen et al. [42]	Arduino, Sensors (Accelerometers, IR Flame, Gas sensors)	Earthquake detection and safety system for Taiwan, using a combination of sensors for early warning, fire detection, and gas monitoring.	System issues alerts for evacuation and safety, providing effective real-time notifications.	Limited to Taiwan; sensors face operational limitations in certain environments.
E. M. A Alcantara et al. [43]	Machine Learning (ML) for RC building damage prediction	Predicting damage condition of Reinforced Concrete buildings using ML and time-history analysis.	Determined the best combination of ML methods and input data for accurate predictions.	Results are highly dependent on the quality of earthquake records and may not apply to all building types or locations.
B. Tian et al. [44]	Movement Detection Sensors	Importance of movement-detection sensors in earthquake management, including monitoring and early warning systems.	Contributed significantly to earthquake management areas like early warning, monitoring, and life detection.	Focuses on sensor technology and does not address all earthquake scenarios.
M. E. Tusun et al. [45]	Strain Gauge Technology & Machine Learning (Deep Learning)	Detecting earthquake waves and structural damage through advanced vibration sensors and ML-based classification.	Demonstrated superior effectiveness in detecting low-frequency earthquake waves and real-time damage classification.	Requires specialized equipment and is costly for large-scale implementation.
M. Bhatia et al. [46]	IoT–Edge Computing & Bayesian Belief Model	Smart earthquake monitoring using IoT and edge computing for early warning at vulnerable locations.	Achieved high classification accuracy (92.52% precision), reduced computational delay (23.06s), and	Relies on edge computing, which may face limitations in some environments or regions.

			high reliability (95.26%).	
P. Govindarajan et al. [47]	Machine Learning & Modified Clustering Approach	Real-time earthquake forecasting for Chile using AI and ML techniques.	Achieved 95% accuracy in forecasting earthquakes, improving prediction speed and accuracy.	Limited to Chile; may not generalize to other regions with different seismic behavior.
M. S. Abdalzaher et al. [48]	IoT & Machine Learning for EEWS	Integration of IoT and ML in earthquake early warning systems for smart cities.	Enhanced decision-making and earthquake parameter observation using ML and IoT.	Requires advanced infrastructure and may not be feasible in all cities or regions.
W. Huang et al. [49]	Machine Learning & Finite-Discrete Element Method (FDEM)	Predicting fault friction states for earthquake prediction using sensor data and FDEM simulations.	Attained high prediction accuracy with an R^2 value of 0.94, using LightGBM with SHAP values for feature importance analysis.	Limited to laboratory-based fault models and may not fully replicate natural earthquake conditions.
P. Lara et al. [50]	Ensemble Machine Learning Algorithms	Earthquake detection and source characterization using single station data and P-wave arrival times.	Achieved 99.9% success in distinguishing earthquakes from noise with minimal false positives and an accurate source characterization.	Works best with high-quality P-wave data and may not be effective in low-seismicity areas.
M. S. Abdalzahar et al. [51]	Machine Learning (2S1C1S Model)	Earthquake intensity estimation within 2 seconds after P-wave onset using ML.	Achieved 99.05% accuracy for earthquake intensity forecasting with the 2S1C1S model, outperforming conventional methods.	Requires high-quality seismic data and may not work well in regions with sparse networks.
A. Joshi et al. [52]	Machine Learning (SeisEML Model)	Cross-region prediction of peak ground acceleration (PGA) for earthquake hazard mapping using hybrid ML models.	Reduced MAE and RMSE by approximately half compared to conventional attenuation relations; demonstrated good predictive performance for multiple regions.	May not be suitable for all tectonic settings; heavily dependent on the availability of quality regional data.
W. Zhu et al. [53]	Machine Learning Chain Models	Predicting seismic responses of substation equipment during earthquakes using intensity measures (IMs).	Provided accurate predictions of peak stresses for substation equipment, supporting post-earthquake rapid judgment.	Focuses on specific substation equipment; not applicable to other infrastructure types.

K. C. Sajan et al. [54]	Machine Learning Algorithms (XGBoost, Decision Tree, etc.)	Predicting damage grade and rehabilitation interventions for buildings after earthquakes using ML.	XGBoost outperformed other algorithms in predicting damage and rehabilitation needs. Identified top features for prediction.	Limited to a specific earthquake event (2015 Gorkha) and building types in Nepal.
C. E. Yavas et al. [55]	ML & Neural Networks	Earthquake detection using advanced ML models and neural networks, focusing on Los Angeles' seismic risk.	Successfully predicted the maximum potential earthquake magnitude with Random Forest algorithm.	Limited to Los Angeles and may not be applicable to other regions with different seismic behaviors.
K. A. Yusof et al. [56]	Geomagnetic Anomalies & AutoML	Investigating geomagnetic anomalies as earthquake precursors and developing models using AutoML.	Achieved 83.29% accuracy using a neural network model, with effective feature extraction through wavelet scattering transform and optimization via ASHA.	Requires over 50 years of geomagnetic field data; limited to geomagnetic anomalies as precursors.
K. Qaedi et al. [57]	Principal Component Analysis (PCA) & Multi-class ML	Enhancing earthquake prediction accuracy by applying PCA to geomagnetic data and utilizing ensemble and SVM models for multi-class classification of earthquake intensity.	Achieved 77.50% accuracy with ensemble models, surpassing SVM models in all evaluation metrics.	Requires PCA and SMOTE for data balancing; may not apply universally to all seismic regions.
S. Ommi et al. [58]	Machine Learning (ANN, RF, SVM)	Predicting large earthquakes by analyzing seismicity changes in the Zagros seismic zone.	ANN outperformed RF and SVM models, showing the highest accuracy in predicting large earthquakes.	Focused on a specific seismic region, which may not generalize to other earthquake-prone areas.
A. Berhich et al. [59]	Recurrent Neural Networks (RNN), K-Means Clustering	Location-dependent earthquake prediction by clustering seismic data based on geographical parameters using RNN models.	Hybrid LSTM-GRU model outperformed other methods, improving prediction accuracy, particularly for larger earthquakes.	Limited to specific earthquake regions (Morocco, Japan, Turkey); clustering may not capture all seismic patterns.
M. H. Al Banna et al. [60]	AI-based Techniques	Systematic review of AI methods used for earthquake prediction, comparing various algorithms.	Provides a comprehensive comparative analysis to aid in selecting the most effective AI techniques for	Limited to a review; does not present new empirical results or direct application models.

			earthquake prediction.	
S. Mujherjee et al. [61]	Ensemble Earthquake Prediction Method (EEPM)	Novel approach combining continuous and categorical data to improve earthquake prediction accuracy.	EEPM outperforms individual models with 87.8% accuracy, higher R^2, and lower error variance.	Relies on data from India, Nepal, and Kenya; may not generalize well to other regions.
R. Yuan et al. [62]	Improved K-Means Clustering & Neural Networks	Clustering global earthquake data for better seismic prediction, addressing the limitations of traditional K-means clustering.	Achieved better clustering accuracy compared to traditional methods, combined with neural network predictions for seismic risk analysis.	Limited to global earthquake data; may not work well in regions with sparse data or different seismic characteristics.
A. Berhich et al. [63]	Attention-based LSTM Networks	Predicting the time, magnitude, and location of large earthquakes using attention-based LSTM models.	LSTM with attention significantly outperformed empirical methods, improving MSE by 60%.	Limited to Japan's earthquake dataset; may not generalize to other regions or earthquake types.
A. A. Mir et al. [64]	Ensemble & Individual Machine Learning (SVM, K-NN)	Predicting anomalies in radon gas concentration caused by seismic activity, using ML methods to detect earthquake precursors.	Boosted tree and SVM with a radial kernel were the most effective in predicting radon anomalies, with the lowest RMSE and MAPE.	Limited to radon data from Muzaffarabad; may not apply to other precursor types or regions.
C. Wang et al. [65]	Acoustic and Electromagnetic Data & LSTM	Real-time earthquake prediction using a self-designed AETA system, combining acoustic and electromagnetic data with LSTM networks.	LSTM network outperformed other prediction models, effectively processing precursor anomaly signals for real-time prediction.	Focused on AETA data from China; may not apply to regions with different seismic behaviors.
B. Zhang et al. [66]	Deep Learning (EPT Model)	Data-driven earthquake prediction using a deep learning model to analyze global seismic catalog data for plate movements and crustal motion.	Achieved over 90% accuracy in predicting earthquake magnitude and location, outperforming traditional methods.	Limited to Chinese seismic datasets; may not generalize to other global regions with different seismic activity.
Q. Wang et al. [67]	LSTM Networks	Predicting earthquakes by modeling spatio-temporal correlations among seismic events across different locations using LSTM networks.	LSTM network improved prediction accuracy by effectively incorporating spatio-temporal relationships.	May not be applicable to regions with insufficient seismic data or where spatio-temporal correlations are less defined.

Z. Zhang et al. [68]	ConvLSTM Networks & Sequence-to-Sequence Framework	Earthquake prediction using a ConvLSTM network for spatiotemporal analysis of seismic data on a global scale.	Achieved high prediction accuracy with better recall and precision compared to existing models, improving resolution and capturing global seismic patterns.	Limited to global seismic maps; may not perform well in regions with sparse or less detailed seismic data.
M. Akhoondzadeh et al. [69]	Satellite Data & Multi-method Algorithms	Review of earthquake prediction progress using satellite data, focusing on remote sensing and anomaly detection algorithms.	Emphasized advancements in satellite technology and multi-method anomaly detection for improving earthquake prediction systems.	No concrete prediction model; focuses on the challenges and advancements of using satellite data.
Z. Ye et al. [70]	Elite Genetic Algorithm (EGA) & LSTM	Earthquake magnitude prediction using a feature selection method combined with LSTM, addressing the challenges of redundant features in seismic data.	EGA-LSTM outperformed other models across multiple evaluation metrics, providing highly accurate earthquake magnitude predictions.	Limited to AETA data from China; may not generalize to other regions or different seismic conditions.

CHAPTER 3

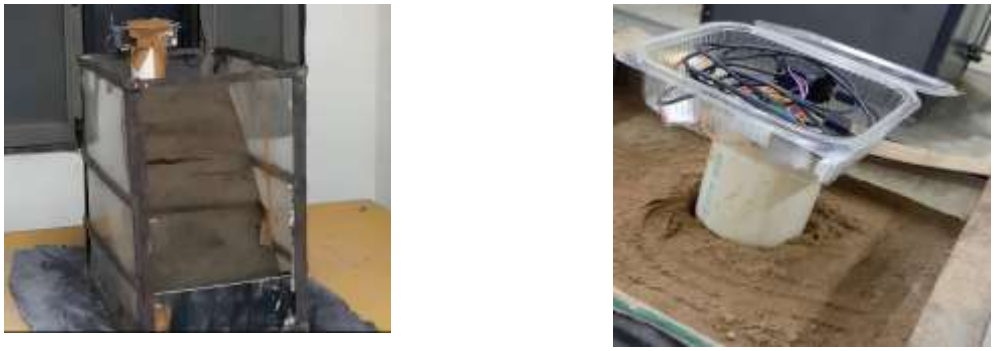
LOW-COST IOT-BASED THRESHOLD-DRIVEN LANDSLIDE PREDICTION AND EARLY WARNING SYSTEM FOR HILLY AREAS

3.1 Introduction

Landslide is one of most commonly occuring natural disaster in nature, which leads to significant loss to life and property damage. It is essential to mitigate their harmful effects due to the destructive consequences of landslides. Early warnings allow authorities to take proactive measures to evacuate people and infrastructure protections in time. However, regional forecasting is difficult due to the numerous parameters that affect its occurrence. The key factors such as local geology, physical telephones, precipitation patterns, and hill trend structure take part in the sensitivity of landslides on Earth. Therefore, these factors need to be measured and understood to predict when and where landslides will occur.

Different technologies are being used to improve landslide prediction to monitor and recognize landslides. Mainly it comprises of remote recognition techniques that use satellites or aerial photographs to observe changes in the landscape using GPS (Global Positioning System) and Geographic Information Systems (GIS) that map and analyze land elements. It monitors fiberglass, radar, wireless sensor technologies, changes in floor movement and actual environment changes, laser and acoustic technologies to measure and recognize early signs of landslides [71]. However, there are also limitations such as high costs, environmental limitations, or technical issues that can affect reliability. Due to the limitations of individual technologies, scientists focused on developing an integrated online surveillance system from landslides. This system combines various geosensors and wireless sensor networks to measure factors that influence slides [72]-[75] to continuously collect data on critical parameters. Additionally, multivariate statistical analysis allows for the processing of complex data records and the identification of formulas that can predict future landslides. Simultaneous analysis of several parameters such as precipitation, soil moisture, and climbing stability allows the system to more accurately predict landslide events.

The aim is to provide early warnings to local residents and authorities that will allow preventive measures such as strong infrastructural strength and evacuation routes before landslides. This research is a comprehensive system for Real-time landslide that integrates several components, including wireless sensors, microcontrollers, cloud servers, and hybrid algorithms. The system collects sensor data in the environment and is then transferred to a cloud server for storage and analysis. An important part of the system is the use of visualization software such as Tableau, which contributes to improving data representation, making it easier to interpret and understand patterns related to landslide. Algorithmic process alongwith the data obtained from sensors are responsible for generating warnings for possible landslides which are already recognized. The objective of this multi-component system is to give impactful and automated forms of prediction and warning for future landslide prone areas. The study also includes laboratory experiments that can simulate real conditions that could lead to landslides.



(a) (b)

Figure 3.1 Laboratory setup of landslide prediction system (a) Front view of landslide laboratory setup (b) Sensor's placement.

The laboratory setup is intended to mimic environmental factors that contribute to landslide events, as shown in Figure 3.1, several sensors are used to collect the critical environmental data needed to predict these events. For example, soil moisture sensor (S1) measures soil moisture content, as changes in soil moisture can have a significant impact on gradient stability and can contribute to landslides. Ultrasonic sensors (S2) are used to measure distance and recognize all physical changes in the monitoring area. Similarly, the temperature sensor (S3) measures humidity and temperature and checks the ambient temperature. Other sensors in the system are vibration sensors (S4) that demonstrate vibrations in the environment and measure changes in tension in response to vibration. Also, accelerometer and a gyroscope sensor (S5) are used to understand the movement and direction of the slope and these sensors also help to monitor changes in acceleration or angular velocity, which are key indicators of possible

landslides [76]. The ESP-32 microcontroller collects data for all these sensors, which act as a central processing unit for sensor inputs [77]. The microcontroller processes and forwards it to a cloud server (Thingspeak), a cloud platform that acts as a memory and analysis center [78]. Thingspeak adds time marks to the incoming data, that enables the system to analyze and monitor trends over time, which allows to detect any pattern that can indicate a greater probability of a landslide.

3.2 Proposed work

In the context of landslide prediction and monitoring, each sensor in the system plays a vital role in detecting key environmental parameters that can signal the potential for a landslide. These sensors are methodically used to collect data related to the soil, terrain movement, climate, and other factors that influence slope stability as shown in Figure 3.2.

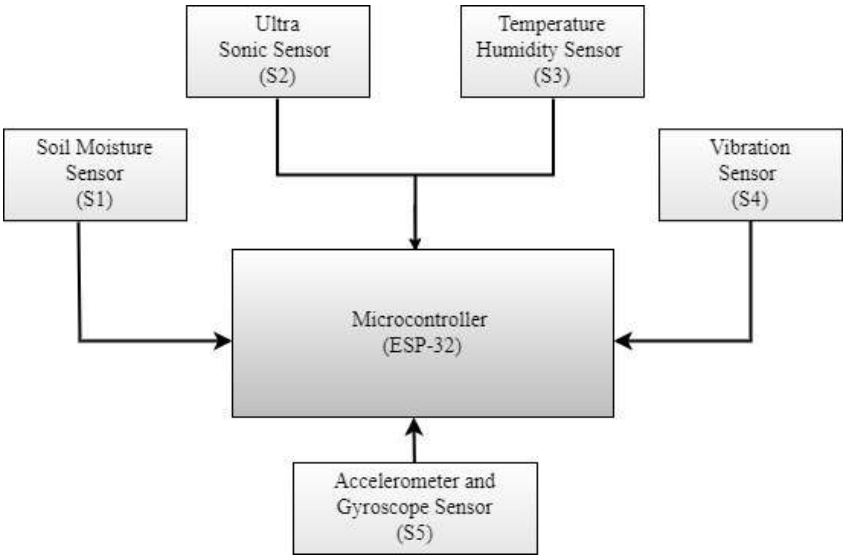


Figure 3.2 Microcontroller Integration for Multi-Sensor Data Collection.

Below is a detailed explanation of how each sensor functions in the landslide prediction system:

i. **Soil Moisture Sensor (S1):** The soil humidity sensor is crucial to detect changes in soil moisture content. Excessive humidity of rain or other sources can significantly weaken the soil, reducing its resistance to cutting and increasing the risk of landslide. The monitoring of soil humidity levels allows early identification of unstable slopes that can be prone to failure. The sensor uses electrodes placed in contact with the ground. As the soil absorbs water, the electrical resistance between the electrodes changes. The more

water present, the less the resistance the sensor can measure. These data are transmitted to the microcontroller, which processes the information and sends it to the cloud server. An increase in soil moisture can be a strong indicator that a slope is at risk of sliding.

ii. Ultra Sonic Sensor (S2): The ultrasonic sensor measures any change in the physical displacement of the soil, the rock or other materials on the slope. This sensor can detect small displacements or changes in the ground, providing early alert signals. The sensor emits ultrasonic waves and measures the time it takes for the waves to recover after hitting a surface. The distance between the sensor and the surface is calculated depending on the moment. If the slope changes or moves, the distance will change, that the sensor can detect in real-time.

iii. Temperature and Humidity Sensor (S3): The temperature and humidity sensor detects the environmental conditions by measurement by thermistor and humidity with a capacity or resistance element. If there is a significant change in these factors, especially in areas susceptible to rain or temperature fluctuations, the sensor can alert the system to the possible risks of landslides. For example, high humidity can indicate saturated soil, a critical factor of landslide instability.

iv. Vibration Sensor (S4): The vibration sensor detects any vibration in the soil, which may be caused by seismic activity, movement on the ground or external force. These vibrations are often related to landslide because it shows instability in the field. The sensor uses a piezoelectric element or accelerometer to detect mechanical vibrations in the environment. As the vibration is observed in the sensor, there comes an electrical signal, which is then sent to the microcontroller. By monitoring the frequency and vibration intensity, the system can identify abnormal formulas that may indicate that the inclination begins to fail, providing early warnings.

v. Accelerometer and gyroscope Sensor (S5): These sensors are designed to detect changes in movement and terrain orientation. By measuring acceleration and angular velocity, they provide essential data on any change or change of slope that may indicate an immediate land sliding. The accelerometer measures the speed of acceleration or slowing down the ground along one or more axes (usually in three dimensions: x, y, z). It works detecting capacity changes as the soil moves. The gyroscope measures angular speed or how quickly the slope revolves around the axis, which helps identify changes

in the slope angle. The combined data of these sensors provide detailed ideas about the movement of earth, allowing the system to detect whether the inclination moves in a way that suggests that it could collapse.

vi. Microcontroller (ESP-32): The microcontroller is known as the brain of the system. It collects data from all sensors, processes them and transmits information to the cloud server for a subsequent analysis. The microcontroller ensures that data is accurately recorded and can trigger immediate responses if necessary. The ESP-32 microcontroller is responsible for receiving data from each sensor in real time. It processes this data, analyzes trends, and checks if any thresholds indicative of a landslide have been surpassed (e.g., a significant increase in soil moisture or movement). It then transmits this data to the cloud server for storage and further analysis. The microcontroller also communicates with the cloud platform (such as ThingSpeak) for alert generation. The ESP-32, Bluetooth Ultra-Low Power Consumption, Dual Core + 38Pin Development Board WiFi formally known as ESP-32, the entire solution uses the least amount of PCB space to the integration of the ESP32 with RF baluns, antenna switches, low-noise amplifiers, power amplifiers, filters, and management modules. Using TSMC's ultra-low power consumption 40nm technology, the 2.4 GHz Wi-Fi plus Bluetooth dual-mode chip offers the best power dissipation and RF performance, and is safe and dependable, and is simple to adapt to a wide range of applications. It helps send data to other servers through wireless means.

vii. Cloud Server (Thingspeak): The cloud server serves as the central data repository and analysis platform. It stores all sensor data, timestamps it, and allows machine learning algorithms to process the data and generate landslide predictions as shown in Figure 3.3. The server is also responsible for issuing alerts when potential landslide conditions are detected. ThingSpeak is a cloud-based platform that collects data from the sensors via the microcontroller. Once the data reaches the cloud server, it is timestamped, allowing the system to track changes over time. The server then runs machine learning models that analyze the data for patterns or anomalies that indicate an increased risk of a landslide. If the system detects such risks, it can issue early warnings to local authorities and residents, potentially saving lives and minimizing property damage.

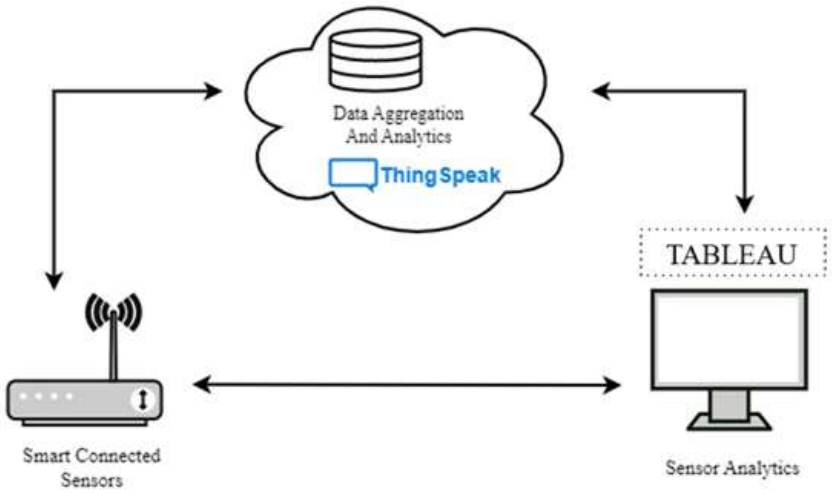


Figure 3.3 Smart Sensor Network with Data Aggregation, Analytics, and Sensor Data Insights.

viii. Real-time data: In real-time landslides prediction systems, several sensors are used to collect critical environmental data that can help identify possible landslides events. These sensors include soil moisture sensors, ultrasonic sensors, temperature and humidity sensors, vibration sensors, accelerometers and gyroscopes [79]-[83]. Before collecting data, important environmental factors are considered like slope angle, soil type, precipitation and the use of the land to improve the accuracy of the data collected. Steep slopes, certain types of soil, heavy rains and human activities can affect the likelihood of a landslide. The data in Table 3.1 are collected before and throughout terrestrial landslides and include several environmental and geophysical determinants crucial for landslide understanding and prediction. These sensors monitor various physical parameters like temperature, displacement, acceleration in three axes (X, Y, and Z) and angular rates, with readings taken at regular time intervals. They are time - marked and give an accurate record of the behaviour of the system over time.

The data being collected are critical in order to assess the status of the system so that immediate interventions or corrective actions can be taken. It is also fine-tuned to ensure the quality and accuracy remains consistent. This preprocessing can involve data cleaning and standardization of features and engineering to prepare the data so that it is reliable and relevant for the analysis. After processing, the final data record serves as the basis for advanced analysis, trend detection, anomaly identification, and future actions of the system. Ultimately, data records can help enable well-discovered decisions, optimizing system performance or support more research into system

dynamics. This data leads to fixed lines and is extremely important for analysis of factors monitoring current events to improve predictive modeling. Once the data is collected, a microcontroller transmits data, where it is updated and stored continuously in real-time.

Table 3.1 Real time landslide data

Time stamp	**S1 (%)**	**S2 (cm)**	**S3 (°C)**	**S4 (%)**	**S5 X (ms^{-2})**	**S5 Y (ms^{-2})**	**S5 Z (ms^{-2})**	**S5 X (°/s)**	**S5 Y (°/s)**	**S5 Z (°/s)**
09:40:59	24.5	2.75	22.5	45	0.02	0.05	0.03	9.78	0.02	0.03
09:41:18	24.6	2.80	22.6	46	0.03	0.06	0.04	9.77	0.03	0.02
...	...	...	...	...	...	...	...	...	...	...
10:10:34	25.0	3.00	23.0	50	0.07	0.10	0.08	9.73	0.07	0.07
10:10:53	25.1	3.05	23.1	51	0.08	0.11	0.09	9.72	0.08	0.08

The process of compilation and use of real -time sensor data for machine learning models (ML) or deep learning (DL) in the prediction systems of landslides implies several steps to verify that the data is captured, pre -processed, transmitted and used effectively for prediction. Initially, several sensors, like soil moisture sensors, ultrasonic sensors, humidity and temperature sensors, vibration sensors, accelerometers and gyroscopes, are placed in strategic locations to collect real-time environmental data that indicates a possible landslide activity. Then, ESP-32, acts as an interface between the sensors and the cloud server. The microcontroller analysis the unprocessed signals received from the sensors and prepares the data for the transmission. Further, the microcontroller transmits the data to the cloud server for further analysis. The next crucial step is data preprocessing, which is important to clean the unprocessed data in a usable format for automatic learning models or deep learning. During preprocessing, the data goes through several steps, like sometimes it may be necessary to scale or standardize the sensor data to take into account variations in the measurement unit so that the model can operate with uniform inputs. This is especially important as the raw data collected by various sensors can change and the model must operate at comparable scale using input values to avoid distortion. The characteristics such as moisture levels, temperature, slope angle and soil type are extracted as individual unprocessed attributes or characteristics. The preprocessing data set can be feed on ML models, such as random forests, support vectors (SVM) or deep learning models, such as

convolutionary neuronal networks (CNN), which can detect complex patterns and correlations in data.

After training, the model can make predictions on real-time sensor data, giving early warnings to residents and local authorities when the system detects conditions that lead to landslide. When continuously updating the sensor data and feeding them to the model, the system guarantees timely and precise predictions, which can be used for alerts generation, allowing the appropriate efforts for evacuation or mitigation before a landslide occurs.

3.3 Data Preprocessing

Once the real-time data from the sensors is transmitted to the cloud server or data storage system, it undergoes a crucial stage of data preprocessing to make it suitable for ML or DL model analysis [84]. Data pre-processing requires a series of steps to clean, reshape, and organize the unprocessed sensor data to ensure it is accurate, consistent, and ready for model training or prediction. The following detailed steps explain the preprocessing process:

3.3.1 Data Cleaning:

The initial step in preprocessing is data cleaning, whose objective is to eliminate any noise or inconsistency in unprocessed data that may affect model's performance. Unprocessed sensor data may include missing values, duplicate inputs or atypical values that do not reflect real world conditions. For example, a sensor might malfunction and provide a reading that is far beyond the expected range (for example, an ultrasonic sensor detects an extremely high distance due to a faulty signal). To address this, the following techniques are commonly used:

i. **Handling missing data:** If sensor data is missing in certain time intervals, this can be due to communication failures or sensor malfunction. Missing data can be handled in different ways, such as imputing (filling) missing values using techniques such as average imputation (filled with the average of nearby values) or interpolation (estimating missing values based on the tendency of surrounding data points). In some cases, missing values can be ruled out if they are too frequent, but this depends on the context and the proportion of missing data.

ii. **Outlier detection:** Outlier are data points that fall far from the expected range, often due to sensor errors or unusual environmental conditions. Statistical methods such as the Z score or the IQR- interquartile range method is used to detect and eliminate or adjust these atypical values. If atypical values represent unusual genuine events (for example, a seismic event), they can be maintained in the data set, since they could indicate a possible discarding of landslides.

iii. **Noise extraction:** Some sensors, such as accelerometers or vibration sensors, can produce noise due to environmental factors (for example, wind, vehicles that pass). Smoothing techniques can be applied as moving averages or Gaussian filters to eliminate short -term fluctuations and highlight the general trend in the data.

3.3.2. Normalization and Standardization:

Different sensors may output data in various units and scales. For instance, soil moisture could be measured in percentage, while vibration might be in voltage or acceleration units. To make these different data types comparable, it is important to normalize or standardize the data: Normalization: This involves rescaling the sensor readings to a fixed range, typically between 0 and 1. The normalization is represented as:

$$Y_{normal} = \frac{Y - Y_{minimum}}{Y_{maximum} - Y_{minimum}}$$

where Y is the value of the original sensor, and $Y_{minimum}$ and $Y_{maximum}$ are the minimum and maximum values of the data set. The standardization ensures that all data characteristics have the same scale, which is particularly important for distance-based models such as K-Nearing (KNN) neighbours and neural networks.

3.3.3 Feature Engineering:

The feature engineering is crucial because it determines the model's ability to detect patterns in the data:

i. **Extraction of Temporary Features:** Real-time landslide prediction generally implies monitoring changes over time, so temporary characteristics such as day time, averages per hour/daily or rolling windows (for example, moving average of 30 minutes) can be useful for capturing trends and cyclic patterns. For example, the accumulation of rain

or soil moisture in recent hours may be more indicative of an imminent landslide than a unique isolated reading.

ii. **Feature Interaction:** Some parameters might interact with each other, such as soil moisture and temperature. Creating new features that combine these parameters, such as the humidity temperature ratio, could reveal non-obvious patterns of individual characteristics. Deriving additional features: For sensors such as accelerometers and vibration sensors, calculate the speed of exchange or frequency domain features (for example, FFT - Fast Fourier Transform) can provide deeper information about motion or vibrations on the slope, which could indicate instability.

3.3.4. Data Transformation and Reshaping:

In some cases, the raw sensor data might need to be reshaped to fit the input format required by certain machine learning models. For instance, Time-series data: Since landslide prediction often depends on trends over time, organizing the sensor readings as time-series data is essential. This may involve transforming the data into sequences where each instance consists of sensor readings at a specific time step (e.g., 5-minute intervals). The input features would consist of previous time steps, helping the model to learn temporal dependencies. Categorical Data Encoding: Some features, like soil type or land use, might be categorical. These need to be encoded into numerical values using techniques like one-hot encoding or label encoding before being fed into the model.

3.3.5. Data Split and Model Preparation:

After the data has been cleaned, normalized and transformed, it is usually split into test, train, and validate datasets. This step ensures that the ML or DL model are able to learn from a data set (training), tune hyperparameters based on another set (validation) and evaluate its performance in a completely invisible data set (test).

3.3.6. Final Dataset Creation:

Once the preprocessing is completed, the final data set consists of characteristics that represent the conditions that lead to landslides, such as soil moisture, temperature, vibrations and other

environmental factors. This structured data set is now ready to be admitted to models , which can detect patterns in the data and predict the future trends on the risks of future landslides.

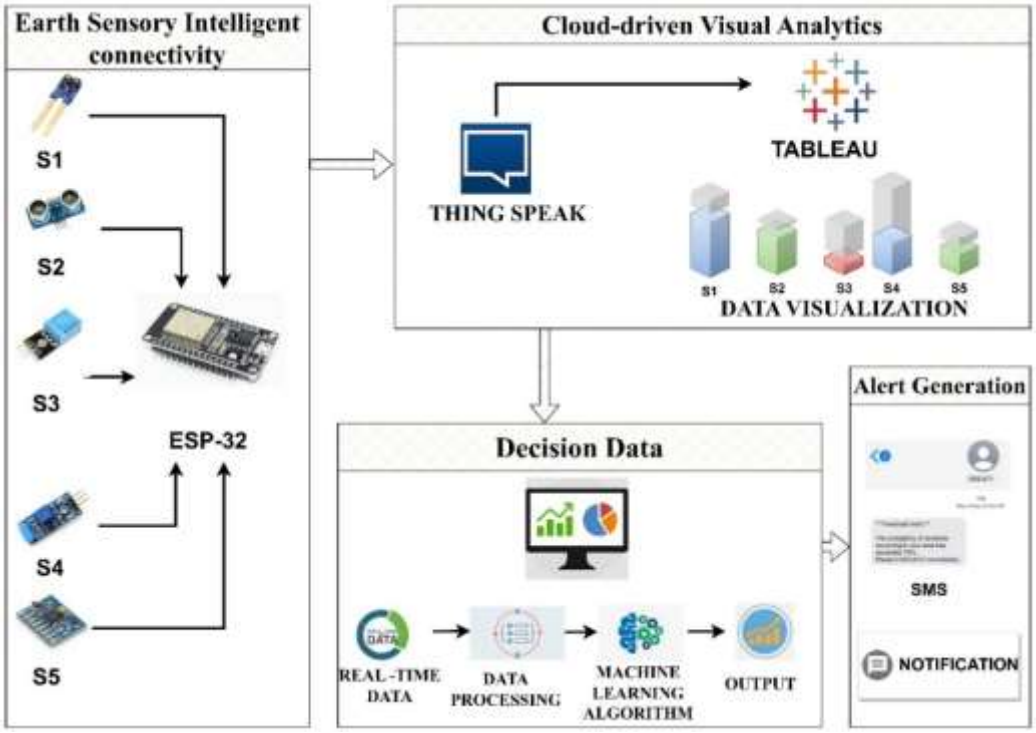


Figure 3.4 Proposed workflow of real-time landslide prediction.

Through this data cleaning process, standardization, characteristics and transformation engineering, unprocessed sensor data is prepared and optimized for use in predictive modeling, which allows precise and reliable land landslides predictions and alert generation as shown in Figure 3.4.

3.4 Machine Learning Algorithm

In this section, attention focuses on the application of several algorithms to predict landslides using real -time sensor data. These algorithms RFR, MLR, GBR, LSTM and XGBOOST are used to evaluate their predictive performance in real world scenarios. The aim is to compare the performance of these algorithms for accurate prediction of landslides to provide the ability to process data in real time and predict it quickly.

3.4.1 Multiple Linear Regression (MLR):

MLR is one of the simplest and most widely used machine learning algorithms to predict a persistent result derived from multiple input features [85]. In the context of the prediction of landslides, MLR can be used to model the relationship between the sensor data and the probability of a landslide occurring. The model presumes that the correlation between the input features and the outcome is linear. It computes a set of weights (coefficients) for each feature, with the goal of minimizing the difference between predicted values and actual observations. While MLR is easy to implement and computationally efficient, it may struggle with capturing complex, non-linear relationships in the data, which is common in environmental systems like landslides.

Equation for model fitting:

$$X = a_0 + a_1Y_1 + a_2Y_2 + \cdots + a_mY_m + \in \quad (3.1)$$

Here a_1, a_2, a_3….a_m represents change in X that is linked with one unit of increase in value of corresponding independent value. Also, X is DV and Y_1,Y_2,…,Y_n are IV, a_0 is constant and a_1, a_2, a_3 are coefficients of IV and $\in$ represents error rate. Least square method is mostly approached by linear regression [34].

Estimation of coefficients: $$\gamma = (Y^TY)^{-1}Y^TX \quad (3.2)$$

Here estimation of $\gamma_o, \gamma_1, \gamma_2, .., \gamma_p$ can be done using given equation. γ corresponds to p+1 coefficient vector, Y represents n-dimensional vector corresponding to dependent variables, $(Y^TY)^{-1}$ denotes inverse of n*n matrix of Y^TY and Y^T represents transpose of Y.

Evaluation of model: For evaluation of model coefficient of determination (R^2) and adjustment of R^2 is done. $$R^2 = 1 - (S\ residual - S\ total) \quad (3.3)$$

Here S residual corresponds to squared sum of residual i.e. difference of predicted and actual value and S total corresponds to total squared sum of difference between actual and mean values.

$$Adj.\ R^2 = 1-[(1-R^2) * (n-1)/(n-p-1)]. \quad (3.4)$$

No. of independent variables (Adj. R^2) can be represented as shown in equation 3.4, here p denotes number of independent variables and n represents total no. of observations.

Prediction values: Now model is ready to make prediction after being evaluated properly.

$$\hat{X} = a_0 + a_1Y_1 + a_2Y_2 + \cdots + a_mY_m \quad (3.5)$$

Where $\hat{X}$ is predicted value of dependent variable for set of independent variables.

One hot encoding (OHE): OHE is technique of translating categorical data into format that may be input into machine learning algorithms to boost prediction accuracy. Each category is converted into a binary vector of zeros and ones during this procedure, with 1 denoting the presence of a category and 0 denoting its absence. In OHE, we set single entry to one and all other corresponding must be zero. The one-hot vector a is binary vector of length n.

$$a \in \{0,1\}^n \quad \sum_{i=1}^{n} a_i = 1 \tag{3.6}$$

3.4.2. Random Forest Regression (RFR):

RFR is a method that builds multiple decision trees during training and fuses its results to produce a more precise and stable prediction [86]. Each decision tree is trained in a random subset of the data, and the final prediction is carried out averaging the outputs of all trees. RFR is particularly effective to handle complex and high-dimension data such as sensor readings for the prediction of landslides, since it can capture non-linear correlations and interactions between the characteristics. In the prediction of landslides, RFR can take into account numerous factors, such as soil properties, slope angles, humidity levels and rain patterns to predict the probability of a earth slide event. The advantage of RFR lies in his robustness towards its ability to handle a combination of numerical and categorical data.

We take average prediction of decision tree (DT), the prediction for data point y, taking decision Tree X, can be written as:

$$f(y, X) = \Sigma\, q_j * I(Z \in B_j) \tag{3.7}$$

Here q_j is the predicted value for j^{th} leaf node of decision tree and B_j is region under j^{th} leaf node. The function I() return 1 if y is in B_j, and 0 and is termed as indicator function.

Using random forest overall prediction for data point y can be given as:

$$f(y) = \left(\frac{1}{m}\right) * \Sigma f\left(y_{L_j}\right) \tag{3.8}$$

Where m is number of DT present in RF, and L is the jth DT in forest. In this equation average of predictions of all DT is used to make final predictions.

Criteria for splitting in DT is Gini index (G), which can be represented as:

$$G = \Sigma K_j * (1 - K_j) \tag{3.9}$$

Here K_j is proportion of data points in j^{th} node of DT. Lower the value of G represents a pure node.

Prediction space is divided into non overlapping or distinct regions $P_1 \ldots . P_x$. Predicted mean of all observations is P_j. Root of sum of squares (RSS) is:

$$\sum_{j=1}^{R}.\sum_{j\in P_j}(x_i - x_{Pj})^2 \quad (3.10)$$

Here within j^{th} region, P_j is mean response of observation.

3.4.3. Gradient Boosting Regression (GBR):

Gradient Boosting Regression (GBR) is another powerful machine learning algorithm that builds an ensemble of decision trees [87]. Unlike RFR, GBR constructs trees sequentially, where each tree tries to correct the errors made by the previous one. This process results in a model that is more accurate over time. GBR focuses on minimizing the residual errors in predictions by optimizing a loss function using gradient descent. For landslide prediction, GBR can be particularly useful as it handles non-linear relationships effectively and can model complex interactions between variables like soil moisture, rainfall, and slope angle. The model is highly flexible and can provide better accuracy than simpler models, particularly when fine-tuned with hyperparameters. However, GBR can be computationally expensive and prone to overfitting if not properly regularized.

$$\text{Initialize } Lo(a) = \arg\min\rho \sum_{i-1}^{m} F(b_{i,}\rho) \quad (3.11)$$

For n=1 to N do:
Step 1: Computing the negative gradient:

$$\hat{b_i} = [\frac{\delta F\left(b_{i,,}L(a_i)\right)}{\delta L_{a_i}}] \quad (3.12)$$

Step 2: Fit the model:

$$\beta_n = \arg\min_{\beta,\gamma} \sum_{i}^{M}[b - \gamma h(a_i;\beta_n)]^2 \quad (3.13)$$

Step 3: Gradient descent step size selection:

$$\varrho_n = \arg\min_{\varrho} \sum_{i-1}^{M} F(b_{i,}L_n - 1(a_i) + \rho h(a_i;\beta)) \quad (3.14)$$

Step 4: Updating the estimation of L(a):

$$L_n(a) = L_{n-1}(a) + \rho_n h(a;\beta_n) \quad (3.15)$$

3.4.4. XGBoost

XGBoost (Extreme Gradient Boosting) is an optimized and tuned version of the Gradient Boosting algorithm. It was developed to improve computational efficiency and prediction accuracy, especially for large data records [89][90]. XGBoost accomplishes this with advanced techniques such as parallelization to accelerate model training and reduce over-fitting at the same time. When predicting landslides, XGBoost is very effective when processing sensor

data. This is because nonlinear relationships and complex interactions between multiple properties can be modeled. XGBoost is extremely popular due to its robustness, flexibility and efficiency in the machine learning competition.

3.4.5. Long Short-Term Memory (LSTM) Networks

LSTM is a type of recurrent neuronal network (RNN) designed for the processing of continuous data. LSTM is particularly useful for time series predictions where previous observations (such as sensor values) affect the outcome [88]. In relation to landslide prediction, LSTM models can analyze temporary data. LSTM networks can learn long-term dependencies by training with real-time order data, the LSTM model can learn the patterns and time-dependent features given by future landslides, this ability to capture long-term relations, especially when actual monitoring of time is important, makes LSTM a powerful tool for predicting landslides.

3.4.6. MLR-LSTM

The MLR-LSTM hybrid model is a very powerful algorithm for landslide prediction, as it combines the characteristics of both models to make final predictions. First, MLR provides a linear relationship between environmental factors. The LSTM layer improves these predictions by including the temporal nature of the data. In contrast to traditional models that rely solely on historical data, LSTMs can be continuously trained from real-time sensor data flows. Dynamically updates new data when it arrives so that the model can quickly adapt to changing conditions. This makes LSTM particularly suitable for temporary serial data. Also, the sensor values, develops patterns over time and LSTM models there too captures the nonlinear relationships and complex patterns in the data making it more effective in situations where complicated relationship is there between variables.

The LSTM layer comprises of several components that helps to manage temporary units effectively. These include input, output and forget gate which together control the flow of information from a time step to the next, ensuring that the relevant data of the previous steps are retained while irrelevant data is discarded. This allows the model to remember long-term patterns and make precise predictions even if the data is incomplete. During the training and prediction phases, MLR-LSTM hybrid model optimizes parameters, adjusts the weight, distortion and states of LSTM cells to improve their ability to learn both linear relations identified by MLR and complex time patterns captured by LSTM. By integrating the strengths

of both models, the hybrid MLR-LSTM offers a comprehensive approach to the prediction of landslides. It provides immediate insight into risk factors through MLR and at the same time it represents developing risks with LSTM over time. This combination results in a more accurate, more reliable and sensitive tool for prediction of landslides, which is essential for generating early warnings and strengthening the efforts to alleviate disasters. Real-time data collected from multiple sensors ensures that the model remains upto date and is able to provide important information to help prevent and manage land landslides.

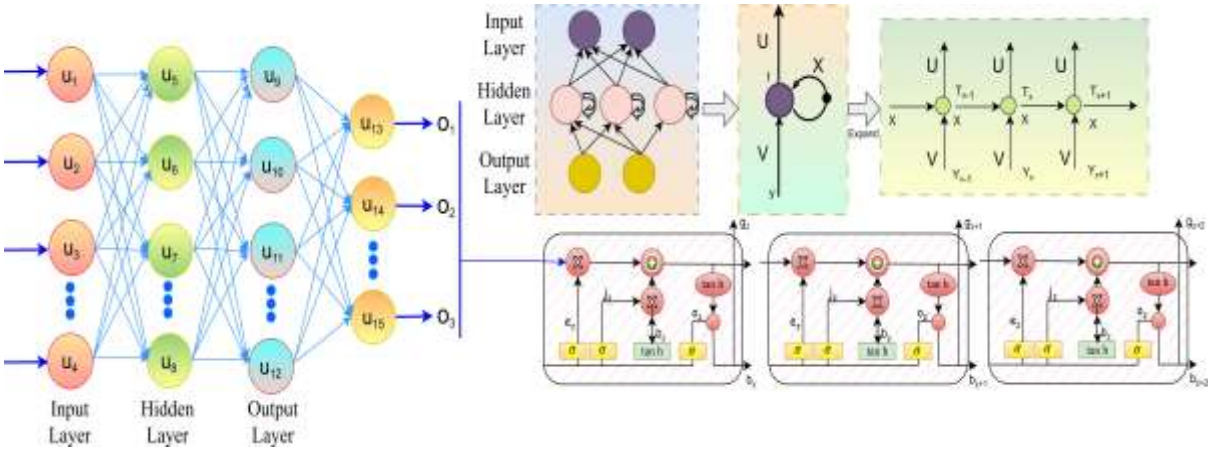


Figure 3.5 Architecture of MLR-LSTM model.

3.5 Results and Analysis

In the testing phase of the landslide prediction system, the process of calculating and optimizing values for each of the key sensor parameters (S1, S2, S3, S4 and S5) began. These thresholds act as critical decision-making points that help determine when the system must induce warnings in order to alert people with immediate risk of landslide. The thresholds for these five fields obtained were 0.77, 0.55, 0.60, 0.40 and 0.50. These threshold values were calculated on the basis of historical data, deducting sensors in real time and observing how the parameters correlate with the occurrence of landslides. The system uses these thresholds to classify sensor data in different risk areas depending on the severity of the values. The first step in the categorization process means comparing the sensor data in real time with predefined thresholds. The values are assigned in specific ravages that correspond to variable levels of landslides. For example, reading a sensor between 0 and 50 is considered a "safe zone", which means that the environment is stable and that the probability of landslides is minimal. Reading between 50 and 60 is classified as a "yellow zone", suggesting that the system has detected

certain differences in parameters and residents must remain careful. If the reading drops between 60 and 80, the system enters the "red zone", a critical phase that indicates that the conditions are rapidly dangerous and more likely to be landslide. Finally, any value greater than 80 will start the most urgent warning, marked as an "evacuated zone" where people are recommended to leave the area immediately to seek safety. Once the thresholds are defined for each zone, the system continuously monitors real-time sensor data and compares them with these thresholds. If one of the sensor parameters exceeds the threshold values defined for a particular zone, the system generates a warning to informing population and local authorities. For wider areas, the system uses digital communication methods such as SMS messages and notifications sent through mobile applications or other digital platforms. These reports contain details of the risk of landslides, affected areas and instructions for evacuation, which ensures that people who may be further from the immediate danger zone are still informed and can take measures if necessary. The combination of real time data monitoring, predefined thresholds and multiple communication channels ensure an integral and efficient system of early and efficient warning system for prediction and prevention of landslides.

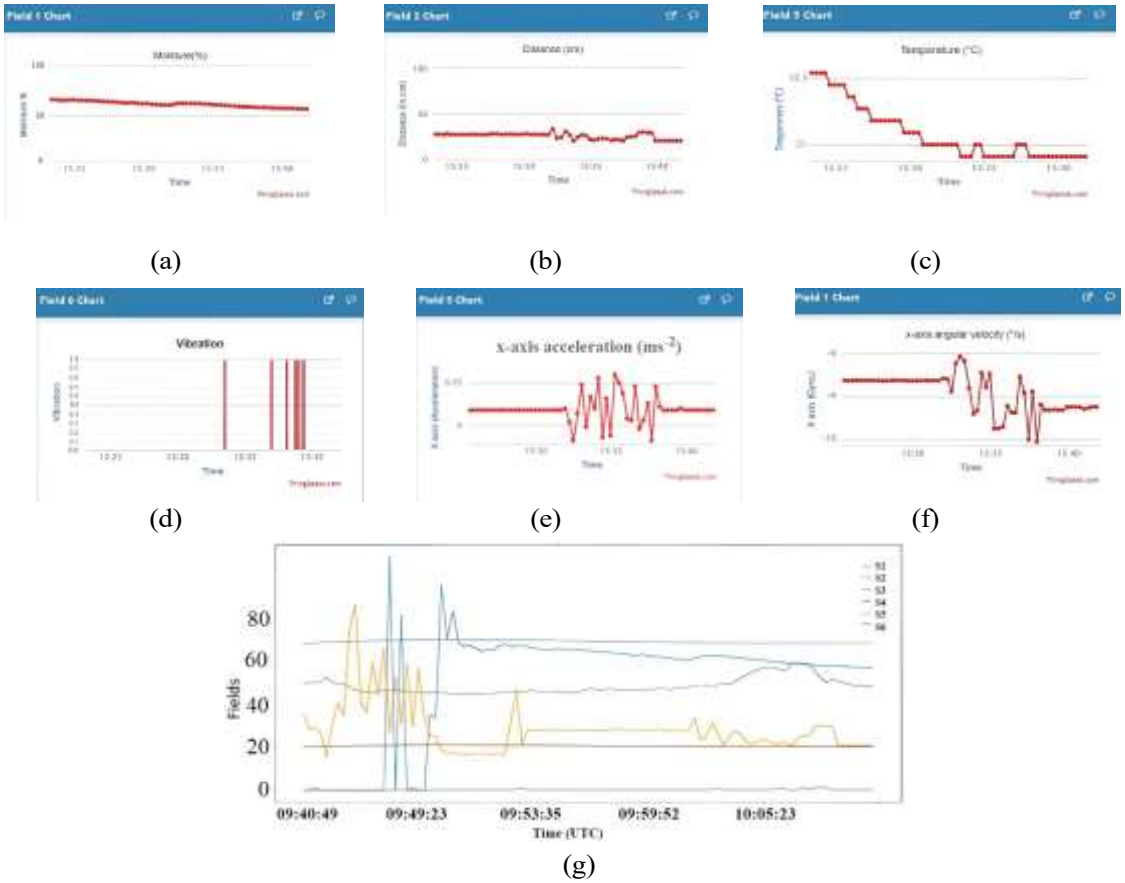


Figure 3.6 Cumulative displacements obtained through (a) Moisture, (b) Distance, (c) Temperature, (d) Vibration, (e) Gyroscope, (f) Accelerometer, (g) Fluctuations observed through sensors.

The experimental evaluations of the sensor system, as illustrated in Figure 3.6, highlight the fluctuations observed in the sensor readings through continuous monitoring of various environmental factors. These fluctuations are normalized to make the data comparable across different sensors. The data from each sensor is collected in real-time and continuously assessed to track changes in environmental conditions. The sensors in this study monitor a variety of variables, including soil moisture, temperature, vibration, and other indicators associated with landslide prediction. To scale the data for these sensors to an equivalent scale, the values are normalized and the sensor measurements are directly compared and made available for use in predictive models. The model is trained with five sensors, all contribute to the final prediction. The training process involves the weight, distortion and adaptation of the MLR and LSTM components to optimize prediction results. For a part of the MLR model, each sensor is assigned an individual weight that reflects the relative significance of predicting the probability of landslide. The sensor weights are set to 0.70, 0.55, 0.60, 0.40 and 0.50 and the sensor determines the amount of the output. These weights record the strength and orientation of the contribution of each sensor to the final prediction.

In the LSTM model, parameters are modified by the flow of information through neural networks. These parameters include input, forget, output gateway weights, and recurrence matrices that determine the relationships between different LSTM network layers. For example, the entrance gate (W_{in}), Forgotten (W_{Forget}), and the start gate (W_{Out}) are assigned weights: 0.75, 0.55, 0.7, and 0.6. These weights adjust the flow of information each time and determine the number of sensor data from previous steps to the current prediction. Preloads connected to the gateway, such as B_{in}, forget gate, output gate etc., are set to 0.1, -0.2, or 0.2 to control information flow, improve the model's capabilities, and capture complex time addictions.

Finally, the combined layer of the hybrid model integrates the MLR and LSTM component outputs. In this combined layer, additional weights (W_{comb}) and distortion are used to fine -tune the final output and ensure that the model creates accurate prediction in real time. The MLR and LSTM integration allows the model to use both linear relationships identified by the MLR and non-linear time dependencies captured by LSTM, which provides robust access to landslide prediction. The ability of this hybrid model to modify its real-time parameters based on incoming sensor data makes it a powerful tool for predicting landslides and release early alerts to relieve the risk.

3.5.1 Evaluation and Performance parameters

i. **Mean Absolute Error (MAE):** It is used for evaluating the efficiency of the regression model. In addition to this, it is used to calculate the average of absolute differences between the predicted value and the actual value simply by taking the average of absolute difference between the predicted and actual values for each data point.

$$\text{MSE}= \left(\frac{1}{m}\right) * \sum(xi - xi\hat{}) \hat{} 2 \quad (3.16)$$

Here m represents number of observations. For i^{th} observation, x_i is actual value of dependent variable and $xi\hat{}$ is the predicted value of dependent variable.

ii. **Mean Squared Error (MSE):** It is obtained by averaging the squared differences between the actual value and the predicted value for each data point. It helps measure the average square differences between predicted and actual values.

$$\text{MAE} = \left(\frac{1}{\text{m}}\right) * \sum|\text{i} = 1 \text{ to m}||\text{x}_\text{i} - \text{x}_\text{i}\hat{}| \quad (3.17)$$

Here m denoted total number of samples. For i^{th} sample x_i is actual value of target variable and $xi\hat{}$ is the predicted value of target variable.

iii. **Root Mean Squared Error (RMSE):** RMSE is basically used when errors are expected to be distributed normally and it considers square root of average squared difference.

$$\text{RMSE} = \text{sqrt}(\text{sum}((\text{predicted values} - \text{actual values})^2)/\text{m} \quad (3.18)$$

Here m represents number of data points, predicted values obtained from regression model denoted predicted values and actual value are those that are directly obtained from dataset.

In the context of landslide prediction, data was gathered from five key variables that represent distinct environmental factors, all of which change dynamically when a landslide is at its peak. Each field corresponds to a specific type of sensor reading that was critical for understanding the behavior of the land at the time of a landslide event. Field 1 represents soil moisture, which plays a pivotal role in landslides, as soil with higher moisture content tends to lose its cohesion

and stability. When rainfall or other water sources saturate the soil, it becomes more prone to slipping, especially on steep slopes. Soil humidity sensors capture real-time data for monitoring these changes that are vital to detect early symptoms of potential landslide. Field 2 represents a distance measured by an ultrasonic sensor that is usually used to monitor physical changes or shifts in the environment. As the landslide progresses, the ultrasonic sensor can be detected by physical shifts such as cracks, shifts or off-road movements. This can help assess how far the earth moves and identify the scope of landslides in real time. Field 3 monitors the level of moisture in the air, which can contribute to the accumulation of water in the soil. The temperature in the degrees of Fahrenheit is measured by Field 4, which can affect the behavior of the soil and water content. Sudden drops or temperature increases can lead to changes in water retention in the soil, which could contribute to instability. Field 5 monitors ground vibration using a vibration sensor that detects oscillation on the surface. These vibrations are often early indicators of landslides because the movement of the surface begins before the real large slide occurs.

As soon as the data from these five areas were collected, the comparison of the behavior of these parameters at the time of landslide and when no landslide was presented. The aim was to identify data formulas that could indicate when landslide was immediate. For example, increased soil humidity, combined with a sudden increase in humidity and temperature, Similarly, increased vibrations and displacement of the ultrasonic sensor indicates that the landslide event was about to occur. By analyzing fluctuations in the sensor values, the key characteristics that defined the platform phase of landslide and detect specific data formulas that signal when the land is probably landslide were isolated and to further increase the accuracy of the prediction, sensors of accelerometer and gyroscope were used for more detailed monitoring of ground movements.

Table 3.2 represents five different sensors used to monitor the key parameters of the environment that could indicate the onset of landslides. Each sensor provided vital data points, which allows to detect changes that could signal threatening landslides. The sensors included soil moisture sensor, ultrasonic sensor, temperature sensor, vibrating sensor and other movement sensors in the form of accelerometer and gyroscope. These sensors continuously capture data during normal and critical conditions and help scientists to observe the fluctuation of the environment that occur before, during and after landslide.

The soil humidity sensor was used to measure the water content in the soil, a critical factor of stability of landslides. In front of soil landslide, soil moisture values usually ranged from 10% to 20%, reflecting normal soil conditions where the moisture content was not too low or too high. The ultrasonic sensor was used to monitor the distance between the sensor and the object, the detection of any changes was observed by the sensor. After the landslide, these shifts became more pronounced because the surface moved significantly, with the sensor detecting more distance changes that indicated the substantial displacement of the terrain. The temperature sensor was used to measure ambient temperature fluctuations, which can affect the retention of humidity in the soil. However, the temperature remained relatively stable, usually in the range of 60 ° F to 70 ° F. After the landslide, the temperature may remain stable or fluctuate on the basis of weather conditions.

The vibration sensor played a decisive role in detecting the movements of the land. In front of the landslide, the vibrations were generally small, with values in the range of 0.01 to 0.1 m/s², reflecting the minimum shifts of the Earth. As the soil became increasingly unstable, these vibrations grew by frequency and intensity. During the landslide, the vibrations increased significantly and achieved 0.2 to 1.5 m/s² or higher, indicating greater and significant disturbance of the Earth. These sharp fluctuations in vibration data were key indicators that landslide began or actively occurred.

The accelerometer and gyroscope sensors were used to monitor the movement of the surface in more detail. The accelerator measured the acceleration or moved in motion along the axis X, Y and Z and even captured minor movements. At time of the landslide, these sensors recorded small shifts, with readings such as 0.01 m/s² (axis X), 0.09 m/s² (axis Y) and -0.03 m/s² (axis Z). These subtle changes testified of slight movement in the ground when destabilized. After the landslide, the accelerometer detects much greater movements, with significantly higher accelerations, especially along the axes. Similarly, the gyroscope captured the rotary movements of the Earth, which reflected any tilt or rotation when the terrain moved. At time of landslide, the gyroscope values were minimal, with values such as 9.26 ° (X axis), 2.87 ° (Y axes) and -1.32 ° (Z axis), indicating slight tilting or rotation on the surface. These rotations become more pronounced during and after the landslide because the surface has experienced greater shifts and tilting. The data collected by these five sensors together with the accelerometer and the gyroscope provided a comprehensive overview of the environmental conditions before and after the landslide.

Table 3.2: Sensor Readings Before and After Landslide

Sensor	**Before Landslide**	**After Landslide**
Soil Moisture	10-20%	50% or higher
Ultrasonic (Distance)	5.0 meters (stable)	4.8 meters or lower (displacement)
Temperature	60°F to 70°F	Fluctuating, depending on conditions
Vibration	0.01-0.1 m/s² (small shifts)	0.2 to 1.5 m/s² (large shifts)
Accelerometer (x-axis)	0.01 m/s²	Increased (larger shifts)
Accelerometer (y-axis)	0.09 m/s²	Increased (larger shifts)
Accelerometer (z-axis)	-0.03 m/s²	Increased (larger shifts)
Gyroscope (x-axis)	9.26°	Increased (larger tilts)
Gyroscope (y-axis)	2.87°	Increased (larger tilts)
Gyroscope (z-axis)	-1.32°	Increased (larger tilts)

Through the continuous monitoring, these sensors made it possible to develop a brighter understanding of changes in the environment that preceded landslides. The study by monitoring soil moisture, distance, humidity, temperature, vibration and additional data from accelerometer and gyroscope is to increase early detection systems and provide more accurate warnings to alleviate the effects of landslides.

Table 3.3 Parameters resulting in landslide.

S. No.	**Fields**	**Values**
1.	Moisture in Soil	56
2.	Distance (by Ultra-sonic)	20
3.	Humidity	45
4.	Temperature (degree F)	68
5.	Vibration	1
6.	Accelerometer (X axis)	0.01
7.	Accelerometer (Y axis)	0.09
8.	Accelerometer (Z axis)	-0.03
9.	Gyroscope (X axis)	9.26
10.	Gyroscope (Y axis)	2.87
11.	Gyroscope (Z axis)	-1.32

Table 3.2 shows the relationship of different parameters contributing to the events of landslides, with the corresponding values for each parameter. This table contains measurements such as soil moisture, distance (measured by ultrasonic sensors), humidity, temperature, vibration,

accelerometers, and gyroscope data for various fields. These parameters are important to understand factors that influence landslide occurrence and to understand the monitoring conditions that may indicate immediate events. Thingspeak allows real-time monitoring and allows to collect, store and analyze sensor data in one centralized system. The values of the recorded gyroscope and accelerometer sensors, along with other environmental factors, were constantly updated and provided an overview of the situation in real time. By observing these fluctuations and comparing values with historical data, the system could predict whether the conditions are correct. Once the specific thresholds were determined from the data collected during previous events, these thresholds were determined as a reference point for future forecasts. For example, if moisture or vibration levels approach the values observed during past landslides, the system may issue warnings to local authorities or inhabitants, provide early warnings and enable timely evacuation or safety measures.

(a) (b)

Figure 3.7 An illustration of a system sending an alert via, (a) Notification and (b) SMS services.

A reading between 50 and 60 falls into the "yellow zone" signalling a cautionary state that requires close monitoring for potential risks. A reading between 60 and 80 is designated as the "red zone" where the situation is more critical, and immediate attention is necessary. When sensor reading exceeds 80, it enters the "evacuation zone", where immediate evacuation and relocation to a safe place are required. Once these thresholds are defined, they generate warnings by comparing the sensor values in real time with predefined ranges. As shown in Figure 3.7, in local areas, these warnings are communicated through visual indicators such as turn signals and sound alarms such as Hooters, which ensures immediate awareness of the situation. SMS messages and pressure notifications are sent for a wider area to inform the larger population of level changes, which keeps all of the affected areas updated in real time. These thresholds are based on changes observed from historical data and are essential for categorizing risk levels from safe conditions to the need for evacuation. The system ensures that people are immediately informed about changing conditions through various warning channels, which provides a structured and reliable method for readiness and response to disasters. By defining

these risk zones and monitoring sensor data in real time, authorities can take timely steps to alleviate potential threats and protection of public security.

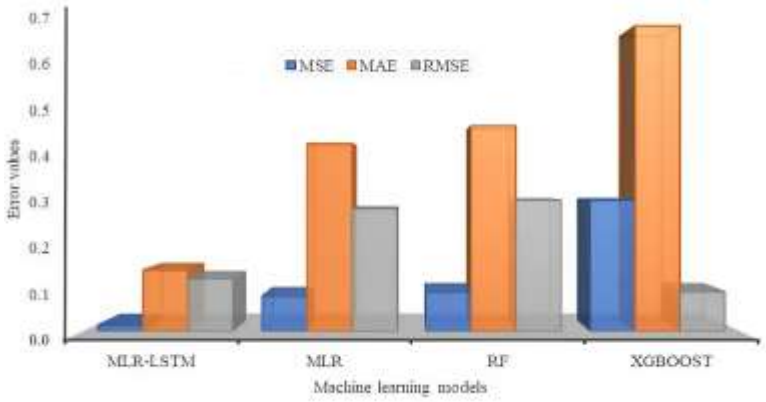


Figure 3.8 Comparison of performance across various machine learning techniques.

Data-based models, such as our hybrid MLR-LSTM model, analyze real-time data from several environmental sensors to discover underlying patterns, which significantly improves the accuracy of landslide predictions. These models use predefined thresholds to establish reference criteria that help classify data at different levels of risk. The thresholds define safe, warned and high-risk areas, forming the basis for the initial evaluations of the conditions. While these thresholds provide a clear structure, the hybrid MLR-LSTM model goes beyond considering the dynamic and real-time factors, improving early alert systems and improving disaster mitigation strategies, particularly in areas prone to landslides.

Our proposed hybrid model demonstrates a clear advantage over traditional models, including individual MLR models, as well as more complex models such as random forest (RF) and XGBoost, in the prediction of landslides, as shown in Figure 3.8. This higher performance is attributed to the combination of two powerful modeling techniques: MLR and LSTM. The MLR component captures linear relationships in the data, offering simple but effective ideas about factors such as soil and rain moisture. On the other hand, the LSTM component stands out in the modeling of complex and non-linear interactions and temporal dependencies present in the data, such as the changing dynamics of environmental conditions over time. Individual models such as MLR, RF and XGBOOST often do not accurately model the complex and dynamic nature of landslides due to numerous interdependence and correlations between different environmental factors. However, the MLR-LSTM hybrid model is able to capture both linear patterns with MLR and linear formulas more complex through LSTM. This results in a more holistic and accurate prediction system that provides a more reliable early warnings

for landslides and improved the general disaster management process. By integrating linear and non-linear modeling techniques, the MLR-LSTM hybrid model offers an integral approach to the prediction of landslides that exceed each model, ensuring better preparation and alleviating in high-risk areas.

Table 3.4 Performance comparison of landslide prediction.

Author	Algorithm	MSE	MAE	RMSE
Kumar et al. [91]	BS-LSTM	–	0.160	0.270
D. Zhang et al. [92]	TCN-AR	–	0.280	0.250
Kshirsagar et al. [93]	LR	0.045	–	0.126
Proposed Model	MLR-LSTM	0.014	0.140	0.120

This table provides insights into the work already been carried out by comparing the values with the previous research on basis of MAE, MSE and RMSE values. MSE is an average square difference between expected and actual values, with less value indicating better accuracy. MAE measures the average absolute difference between the anticipated and actual values, which gives a direct indication of the accuracy of the model prediction without considering the direction of errors. RMSE is the square root of MSE and provides a scale to understand the size of a prediction error, with a lower value indicating a better performance.

Kumar et al. [91] used the BS-LSTM model (two-way short-term memory), the model showed MAE 0.160 and RMSE 0.270. Zhang et al. [92] used the TCN-AR algorithm (a time convention network with the author), also without the value of MSE. The performance of the model resulted in MAE 0.280 and RMSE 0.250. Kshirsagar et al. [93] implemented the model of linear regression (LR), which showed MSE 0.045 and RMSE 0.126. The MSE and RMSE values indicate that it worked better than the BS-LSTM and TCN-AR models, but still had higher errors than the proposed model. The proposed model is a combination of multiple linear regression (MLR) and long short -term memory (LSTM), which outperforms all other models at MSE 0.014, MAE 0.140 and RMSE 0.120. This suggests that the MLR-LSTM hybrid approach provides the most accurate forecast of landslides compared to other methods tested in this study.

Comparison of different models of landslide prediction highlights the efficiency of the proposed MLR-LSTM hybrid model. Performance metrics including MSE, MAE, and RMSE indicate that the MLR-LSTM model overcomes other models such as BS-LSTM, TCN-AR, and linear regression. In particular, for MSE (0.014), MAE (0.140), and RMSE (0.120), a significant reduction in error values suggests that this model provided more accurate and

reliable predictions for landslides. This accuracy is essential for early warning systems and helps minimize and mitigate the impact of landslides in areas at risk. Integrating several sensors in real time, including soil moisture, ultrasound, temperature, air humidity, vibration, and accelerometer sensors will improve soil capacity. By constantly monitoring environmental factors and including them in machine learning models, the system recognizes the changes that indicate future landslides. Generating warnings related to actual data processing systems and decisions using threshold-based limits for early warning communities in the early stages to save lives and prevent significant infrastructure damage. The results check the feasibility and efficiency of MLR-LSTM access and demonstrate the possibilities of real-time applications.

The real-time IoT-based landslide monitoring system has been fully implemented and tested within a laboratory environment, proving its ability to effectively collect, analyze, and transmit environmental data for early warning purposes. Transitioning this system to deployment in real-world landslide-prone regions involves addressing several critical operational challenges to ensure consistent performance and reliability. Field deployment requires systematic installation of sensor nodes across rugged and often inaccessible terrain. This involves detailed site surveys to identify optimal sensor locations that comprehensively cover vulnerable slopes and critical points. Installation teams must navigate difficult access routes and unstable ground conditions, using specialized equipment for secure sensor placement. Regular maintenance schedules are essential to inspect and replace sensors affected by environmental wear and damage. The sensors and associated hardware must be ruggedized and enclosed within protective casings to withstand harsh environmental conditions, including heavy rainfall, extreme temperatures, humidity, soil movement, and vegetation growth. These protective measures prevent physical damage and reduce sensor drift caused by environmental exposure, ensuring accurate and consistent measurements over extended periods. Power supply is a fundamental challenge in remote areas lacking grid infrastructure. The system employs solar panels combined with rechargeable battery storage to provide continuous power. This setup ensures uninterrupted sensor operation day and night and during adverse weather conditions. Battery capacity is sized to sustain operation for extended periods without sunlight, minimizing maintenance frequency. Communication between sensors and cloud servers relies on robust wireless networking adapted to remote and challenging environments. The deployment utilizes long-range communication technologies such as LoRaWAN or mesh networks to ensure reliable data transmission across complex terrain with limited cellular coverage. Local data buffering in sensor nodes stores measurements during connectivity interruptions, preventing

data loss and enabling synchronization when the network is restored. Sensor calibration and health monitoring protocols are integral to maintaining data integrity. Automated self-diagnostic routines detect sensor anomalies or failures, triggering maintenance alerts. Calibration is conducted periodically in the field to adjust for sensor drift caused by environmental factors. Finally, integration with local disaster management agencies and community engagement are critical for effective deployment. The system's alert dissemination framework delivers timely warnings through SMS, mobile applications, and sirens. Training programs for local authorities and residents ensure that alerts translate into prompt evacuation and mitigation actions, maximizing community safety. This deployment framework addresses all key challenges associated with real-world operation of the IoT-based landslide monitoring system. The laboratory-validated prototype evolves into a resilient, autonomous, and maintainable network capable of providing reliable early warnings in landslide-prone regions. This implementation advances practical disaster risk reduction and sets a foundation for large-scale field applications.

3.6 Summary

The integration of various sensors and advanced algorithms had a major impact in real time on the accuracy and efficiency of landslide prediction. The study focuses on solving problems related to prediction of such disasters so by collecting data from various sensors provides with examined study of possibilities of different machine learning models for predicting landslides. These models, such as multiple linear regression (MLR), XGBoost (XGB), and Random Forest (RF), have been tested and evaluated against each other to identify the most effective methods for accurate predictions. Of the models tested, the MLR-LSTM hybrid was listed as the most promising and this hybrid model combines MLR intensity when dealing with linear relationships between time dependencies and LSTM networks, when analyzing non-linear forms. The LSTM function of the sequential storage function allows models to be analyzed and learned from time-dependent data such as soil moisture, temperature, and other environmental factors that represent key indicators to threaten landslides. This approach made it possible to identify and predict the events of landslides more efficiently, especially when considering the dynamic and fluctuating nature of data collected in real time. Moreover, the successful implementation of the hybrid model of generating real-time warning has shown its potential to help the authorities and communities early measures to minimize the impact of landslides. The results of this study emphasize the importance of using the advancements of machine learning

and sensor data to improve the strategies of prediction and management of disasters, and eventually save lives and reduce the economic and social costs associated with landslides. The thesis effectively utilizes real-time sensor data to simulate and monitor environmental conditions that can lead to landslides, offering a practical framework for early detection. However, despite the advantages of real-time data, this method of data collection might still have a few limitations. One key issue is sensor coverage—in a controlled laboratory environment, sensors are positioned optimally, but in real-world settings, full coverage of the affected terrain may not be feasible due to geographical constraints, installation challenges, or cost limitations. As a result, certain critical changes in environmental conditions may go undetected, affecting the comprehensiveness of the data. Additionally, sensor reliability poses another concern. Sensors are prone to malfunctions, calibration drift, or temporary failures due to harsh weather conditions, physical damage, or interference from surrounding elements. For example, an ultrasonic sensor might give false readings if obstructed by debris, or a moisture sensor might fail to respond accurately if embedded in compacted soil. Moreover, data transmission delays or losses—especially when using wireless communication and cloud platforms like ThingSpeak can lead to time lags or gaps in the data stream, which in turn affect the real-time responsiveness of the system. These limitations, if not accounted for, can reduce the accuracy and reliability of the landslide prediction model, particularly in dynamic and unpredictable environments. Therefore, while the system is robust in a lab context, acknowledging these potential real-world limitations is important for future improvements and deployment.

CHAPTER 4

A SEMANTIC SEGMENTATION FRAMEWORK WITH U-NET-PYRAMID FOR LANDSLIDE PREDICTION USING REMOTE SENSING DATA

4.1 INTRODUCTION

Landslides are frequent natural events triggered by several factors, including earthquakes, heavy rains, river erosion, the cutting of slopes for the construction and activity of groundwater induced by natural and human water [94][95]. Due to climate change, the increase in urbanization and increased seismic activity, the frequency of landslides has constantly increased, which leads to long-term impacts, such as the destruction of property, infrastructure and loss of life [96][97]. A landslide occurs when rocks, soil and debris move downhill due to gravitational force. These movements can be different in speed and size, usually resulting in serious damage to buildings and to human security. Traditionally, landslides were detected through field work on the site, where experts physically examined areas prone to landslides, paying attention for signs like unstable slopes, cracks on the ground or past landslides remains [98]. Although this method was useful, it was intensive in labour, slow and limited, especially in areas difficult to access or remotely. With the development of geospatial technologies, the detection of landslides has become much more advanced. Techniques such as aerial photogrammetry use high resolution aerial images to create precise landscape models, allowing an in-depth analysis of the characteristics of the earth and the possible risks of landslides [99]. In addition, satellite remote sensing allows continuous supervision of large regions and identifies displacement in landscape that might indicate the probability of immediate landslide. Compared to the traditional methods, these modern technologies offer more efficient and accurate way of identifying areas susceptible to landslides, which improves both speed and accuracy of detection.

Continuous landslide monitoring and detection is essential to reduce the risk and consequences of landslide on infrastructure, municipalities and the nature. Traditional landslide methods often rely on subjective interpretations of satellites or aerial photography. The ML algorithm

is trained on large remote sensing image data records, allowing to automatically identify samples assigned to landslide. These ML algorithms can detect important features like topographical changes, increase, vegetation patterns, and surface morphology. All of these can indicate the probability of a landslide. ML not only accelerates landslide prediction it also increases the reliability and accuracy of EWS [100]. This technology is necessary to reduce the damage caused by landslides, as it allows for proactive measures like evacuation plans, strengthening infrastructure and effective management of land use in high-risk areas. Due to the complexity of traditional methods, there is a growing demand for aim and more techniques which are automated to predict landslides. Many landslides occur on slopes with exposed soil or rock, which often have vegetation or other surface features that create complex patterns in optical images. Advances in remote sensing technology, in particular the enhancement of spectral and spatial resolution of satellite images from the platform such as the Sentinel series significantly improved the ability to monitor large areas susceptible to landslides in detail. Landslide prediction or detection techniques and algorithms can generally be classified into three main approaches: local field survey, machine learning algorithms and deep learning techniques. Local field survey means on site surveys and manual data collection, while ML analyzes significant quantities of geospatial images to detect trends and patterns. Deep learning proposes neural networks to analyze even more complex data and create very accurate predictions [101]. These developing technologies are the key to the detection of landslides and alleviate the potential risks or destruction they can pose. The field survey has been recognized efficiency in the prediction of landslides but faces significant challenges. These challenges include the risk of damage, time constraints deposited by huge areas that must be covered, high costs associated with sending equipment and helping teams to the affected sites, and potential inaccuracies when trying to attempt manually evaluate large and complex terrains. While manual exploration was once the main method of identifying landslides, the entering of geospatial technologies such as remote sensing RS and aerial images, provided new and more efficient forms for detecting areas susceptible to landslides. One of the key progresses is the use of high-resolution satellite and aerial images that have significantly improved accuracy of landslide prediction or detection. Satellite images are particularly valuable because they provide detailed spatial data that can detect essential indicators of potential activity of landslide. A remarkable technique used in this context is an Object-Based Image Analysis (OBIA) [102]. Unlike traditional pixels -based analysis, which focuses on individual pixels, OBIA groups pixels in significant objects based on their spectral properties (such as color and intensity) and spatial relations inside the image. This makes it possible to analyze not only

spectral characteristics, but also the context in which these pixels appear, such as their spatial disposition, form to other characteristics of the landscape. Combining this contextual information allows OBIA to identify consistent properties of land that can have geological properties and have vegetation density, topographic morphology, or instability. On the other hand, pixel-based analysis examines every single pixel based on its color or spectral properties, without considering the broader context or spatial relationships between adjacent pixels. This method is easier and works well in some scenarios, but can be problematic with more complex landscapes. This landscape neglects fine texture changes, spectral signatures, or transitions between different types of land coverings and geological layers. OBIA can counter these challenges with regard to the spatial composition of groups of pixels. This is especially effective when identifying and mapping landslides. The efficiency of OBIA in landslide detection has been further improved based on the availability of high-resolution satellite images and the building of highly developed tools and these tools can perform semi-automatic or fully autonomous analysis. This increases the need, speed and accuracy for manual interpretation of landslides. As a result, OBIA is an important part of the detection and classification. This means it will be effective in a comprehensive approach to monitoring and predicting landslides. Pixel-based models were often used to classify images and disaster surveillance, including landslide detection [103][104]. However, these approaches have limitations, especially when satellite images are processed at very high resolution. These images often create issues such as the "salt and pepper" effect. This effect leads to large fragmented pixels of detail [105]. This problem makes it difficult to accurately classify images and predict landslides, as individual pixels can be misclassified due to data complexity. OBIA has become a more effective alternative, especially when manipulating high-resolution remote-acquired images, to recognize the boundaries of pixel-based methods, particularly in image classification. OBIA changes the analysis approach of individual pixels to evaluation of shaping objects by summarizing pixels based on common characteristics such as texture, color, and spatial proximity. Segmentation plays an important role in OBIA by converting individual pixels into coherent image objects that represent important features of the landscape. These objects are identified and grouped after a combination of spectral, structural, morphological, and topographical characteristics. By segmenting images with critical units, OBIA improves the accuracy of predicting landslides and reduces false alarms. The segmentation process refers to the definition of factors that define the size and format of objects in an image. This can be a difficult task. Since the characteristics of the land features, including landslides, may vary significantly, often requires an iterative process where various segmentation techniques are

tested and refined through visual inspection to ensure that objects accurately represent characteristics of the real world. Once the segmentation process is completed the classification starts, where each object is classified according to specific criteria. To detect landslide, this include the manual thresholds setting for properties such as field changes, vegetation coverage or soil composition, which are generally associated with areas susceptible to landslides. While OBIA offers clear advantages over traditional pixels-based methods, it is not exempt from their own challenges, particularly when they are combined with machine learning models for more detailed applications. The characteristics of landslides can vary in the size of space and spatial context in different areas, making it more difficult to apply one single segmentation strategy. The challenge is defining the correct scale parameter to ensure that segmented objects accurately reflect the wide range of geological and environmental features present in the images. In addition, this variability means that OBIA methodologies should refine to consider the different analysis scales, which requires a continuous adaptation of segmentation techniques. The integration of OBIA with ML algorithms can help address these problems improving the ability to automatically classify and map landslides in different terrains. However, the success of this integration is largely based on improving the dependency of OBIA, as well as further advancement in ML models to better manage complex and multi-scale data that is characteristic of high-resolution satellite images. Despite these challenges, the combination of OBIA and ML offers a promising approach for landslide prediction. As advancement in OBIA methodologies and in ML algorithms, the potential for reliable and large-scale landslide detection will expand while providing better tools to mitigate the risk posed by landslides to infrastructure, communities and the environment.

4.2 Dataset Description

Deep Learning (DL) requires large number of labelled data to effectively understand multiple parameters with variations. According to research, when small, labelled training data set is used, it can degrade the performance of the classification, while a large training dataset is used DL models cannot cover all conceivable cases. To solve this problem, we used a benchmark dataset called Landslide4Sense, which contain study of sites that are affected by landslides from different regions, as shown in Figure 4.1.

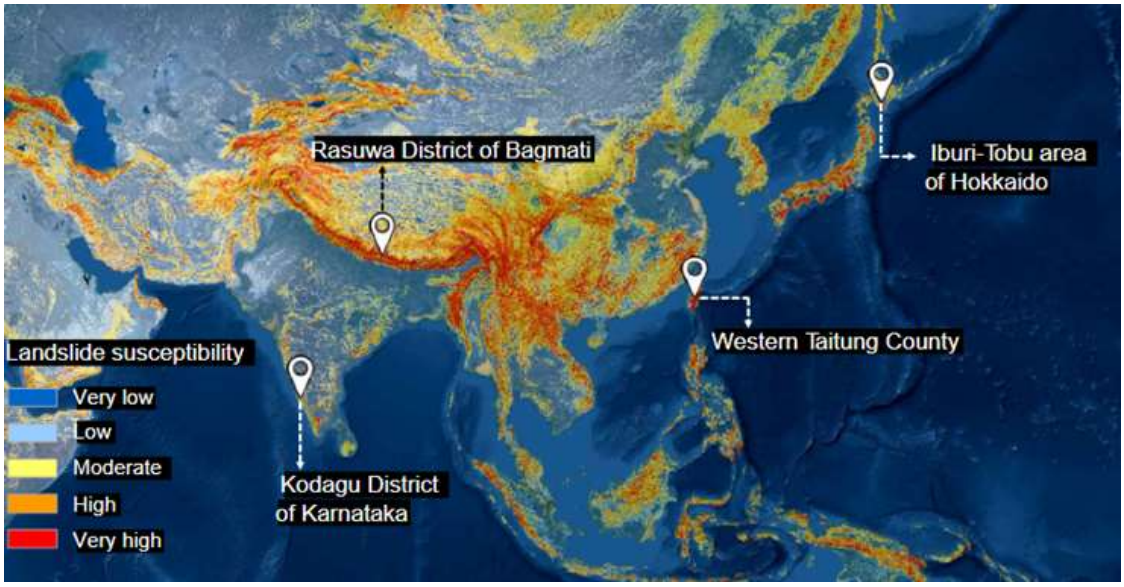


Figure 4.1 Geographical Locations for Landslide Susceptibility Dataset Collection.

Landslides4Sense dataset is a specialized benchmark collection created to increase the analysis of the landslide susceptibility by providing a comprehensive combination of terrain, slope and multispectral satellite data. The terrain height is a critical factor in the prediction of landslides, as height changes indicate areas where gravitational forces are likely to cause movement. In particular, steep slopes are more vulnerable to landslides, because the gravitational force acts more strongly on steeper tendencies, especially if the stability of the earth surface is endangered by external factors such as precipitation, seismic activity or human modification. The landslide data file includes ALOS PALSAR slope data, which provides detailed information of steepness of the terrain, essential for assessing the risk of landslides. This data is obtained from images based on radar satellite that offer high accuracy when measuring the surface slope, which increases understanding of areas susceptible to landslides. In addition to slope data, the dataset incorporates Sentinel-2 multispectral data, which includes band 1 to 12 and by combining information about the terrain height and slope with multispectral data, the Landslide4Sense dataset allows a more detailed and more accurate assessment of areas susceptible to landslides. The dataset is particularly valuable because it has been thoroughly marked for classification of landslides and non-landslide, each of which has changed to a resolution of about 10 meters per pixel. This high level of detail and accuracy allows efficient models and analysis of machine learning and helps to identify fine changes in the field that may indicate an increased risk of landslide. The combination of detailed height, slope and spectral data makes the Landslide4Sense dataset an important resource to improve the landslide detection or prediction and increase the reliability of early warning systems.

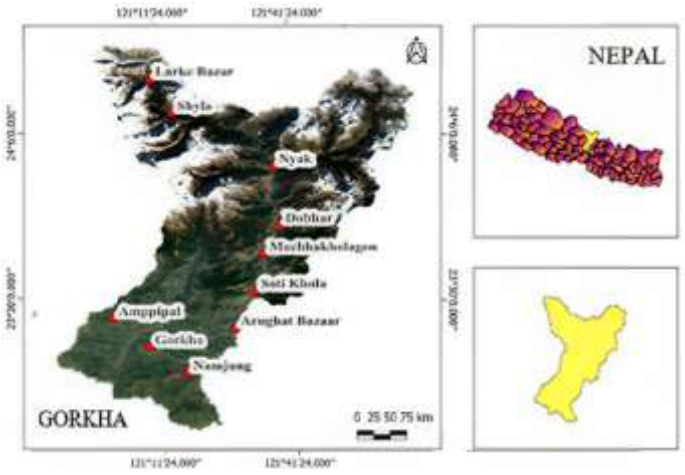


Figure 4.2 Gorkha District - Nepal.

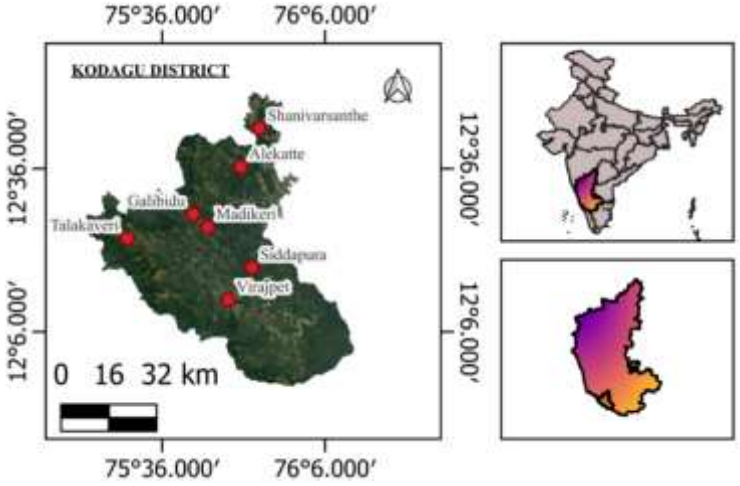


Figure 4.3 Kodagu - District of Karnataka

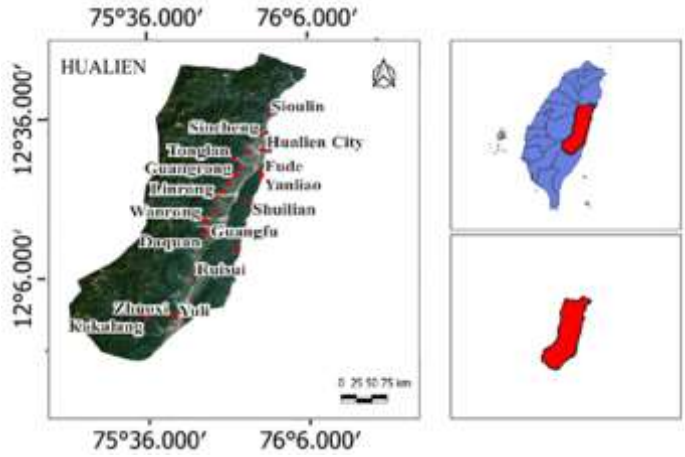


Figure 4.4 Hualien - Taiwan.

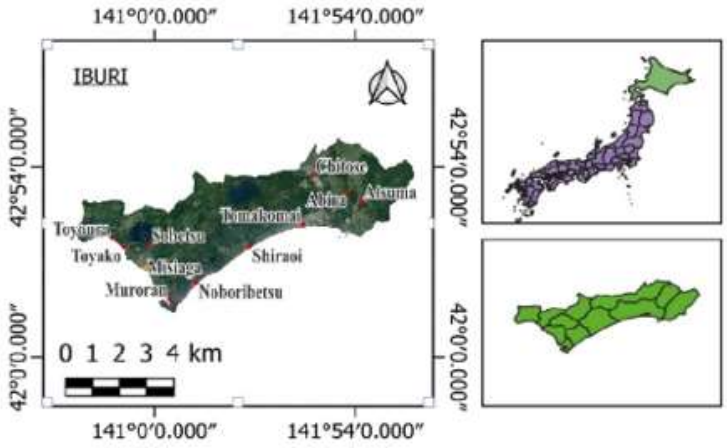


Figure 4.5 Iburi-Tobu

This dataset includes patch images containing landslide from different global locations. This dataset is split into three separated sets: training, testing and validation. This structured division is designed to strengthen the training of DL models and provide the potential to handle large range of new and unseen situations, especially those that differ from the data on which the model was originally trained. By exposing a model to diverse geographical regions and conditions, the dataset helps to enhance the efficiency of the model in different terrains, allowing it to be better generalized when applied to new data. The training set contain the data from four different regions around the world: Iburi-Tobu in Hokkaido (Japan), Kodagu in Karnataka (India), Rasuwa in Bagmati (Nepal) and Western Taitung (Taiwan), as shown in Figure 4.2-Figure 4.5. These regions were selected to represent different terrains, climate and landslide condition. Data from these areas are used to create patches approx. 3799 of 128x128 pixels. These image patches serve as training data for the model. In addition, this dataset contains verification and test set consisting 245 and 800 image patches of the same size (128x128 pixels). Sentinel-2 provides detailed images across different wavelength bands, from

SWIR, allowing analysis of multiple surface features. Landslide4Sense uses specific bands such as B2 is blue, B3 is green, B4 is red and B8 is almost infrared, which have a resolution detail of 10 meters per pixel. Other bands such as B5, B6, and B7 are vegetation red edge, B11, B12 are SWIR, as well as B1 is coastal aerosol, B9 is water vapor and B10 is Cirrus, have a different spatial distinction of 20 meters and 60 meters. These different resolutions allow for detailed analysis with multiple features such as vegetation, water and soil formations, all of which are important in the study of landslides.

In addition, Alaska satellite facility provides a high-resolution digital elevation model that is derived from the ALOS PALSAR system. DEM offers detailed topographic data, which is essential for understanding the terrain elevation. From this DEM, a layer of slope is formed, which represents the steepness of the terrain. Since the DEM and the slope layer both are important for the landslide prediction, they are modified into a spatial resolution of 10 meters, which ensures consistency with other data layers. These datasets are combined in 14 different layers in the Landslide4Sense dataset, which are then used for training and testing of DL models aimed at detecting landslides and analysis of sensitivity as shown in Figure 4.6. The combination of high-resolution images, and information on landslide gradients using Landslide4Sense provides a rich multidimensional dataset. By including data from different regions and different spatial resolutions, this data record creates a more robust and reliable model for predicting or detecting landslides that can be used at a global level.

i. Band 1 Sentinel-2: Blue spectral band data.
ii. Band 2 Sentinel-2: Green spectral band data.
iii. Band 3 Sentinel-2: Red spectral band data.
iv. Band 4 Sentinel-2: Near Infrared (NIR) spectral band data.
v. Band 5 Sentinel-2: Shortwave Infrared (SWIR) spectral band data.
vi. Band 6 Sentinel-2: Shortwave Infrared (SWIR) spectral band data.
vii. Band 7 Sentinel-2: Shortwave Infrared (SWIR) spectral band data.
viii. Band 8 Sentinel-2: NIR spectral band data.
ix. Band 9 Sentinel-2: Water Vapour (WV) spectral band data.
x. Band 10 Sentinel-2: Cirrus (CI) spectral band data.
xi. Band 11 Sentinel-2: SWIR spectral band data.
xii. Band 12 Sentinel-2: SWIR spectral band data.
xiii. DEM - Digital Elevation Model: Elevation information data.

xiv. Slope: Slope information.

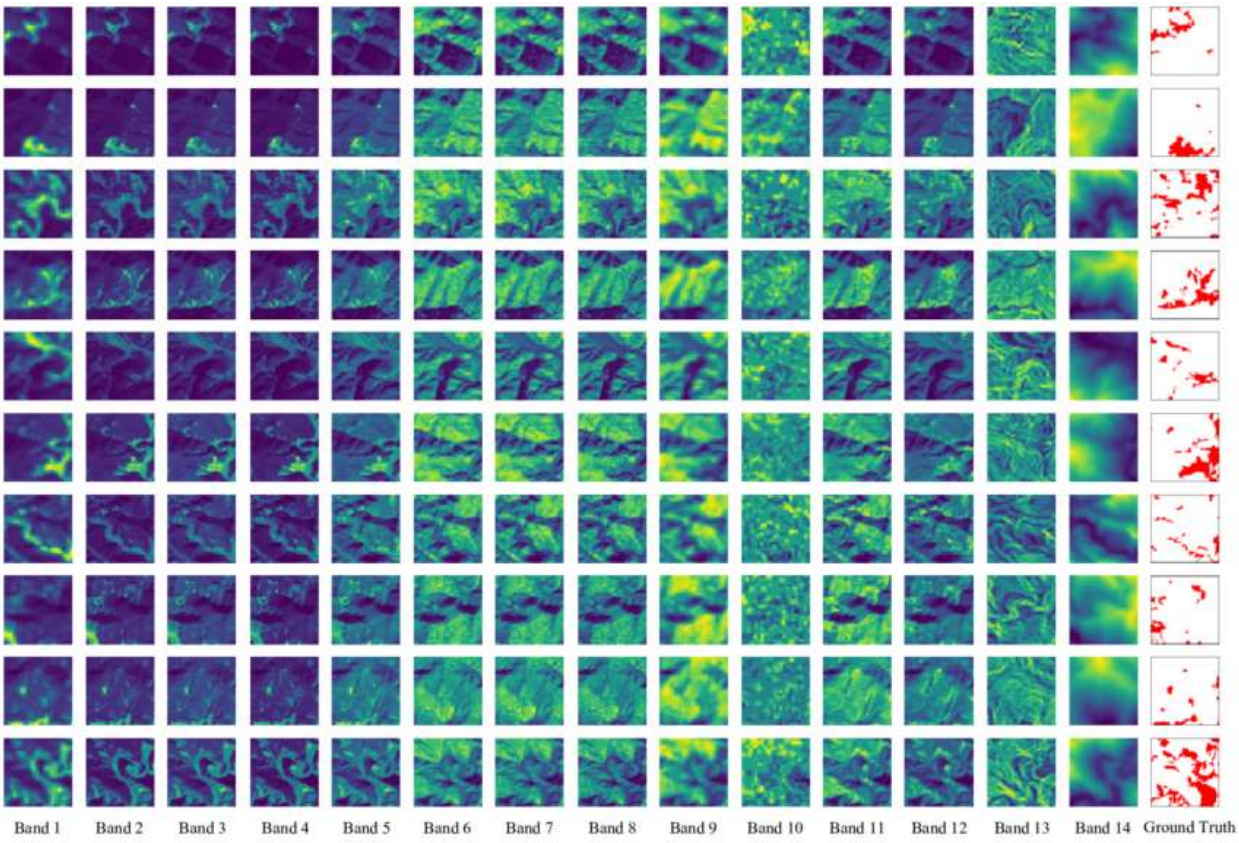


Figure 4.6 Visualize every unique layer inside the 128x128 window-size patches of the generated landslide dataset. The first 12 bands shows multi-spectral data from Sentinel-2, bands 13 and 14 shows DEM data and slope from ALOS PALSAR.

This dataset contains test, train and validation subsets, the subset which contain train data is collected from four different landslide-susceptible region. Table 1 provides an extensive description of this dataset consisting complete attribute description about data.

TABLE 4.1: Comprehensive Description of the Landslide4Sense Dataset

Sr. No.	**Attributes**	**Description of Attributes**
1.	Name of Dataset	Landslide4Sense
2.	Total Samples	Training sample-3799, Testing sample- 800, Validation sample- 245
3.	Response Variable	No landslide - 0, Landslide - 1
4.	Source of Data	Landslide detection using multi-sensors
5.	Geographic Regions	Rasuwa district, Kodagu district, Iburi-Tobu area, and western Taitung Country
6.	Pre-processing	Normalization, Empty values Removal

7.	Types of Features	Topographic, Meteorological, Geological, Geotechnical
8.	Data Format	In CSV format
9.	Feature resolution	Temporal and spatial measurements at specific locations

The dataset used in this study is well-structured and diverse, comprising image patches of landslides from four geographically distinct regions—Japan, India, Nepal, and Taiwan—which strengthens the model's capacity to generalize across different terrains and climatic conditions. However, the data collection process presents several limitations that affect the model's overall robustness and applicability. The geographical scope of the dataset, although varied, excludes many global terrains with unique geological and environmental characteristics. This lack of coverage limits the model's exposure to critical landslide types found in arid regions, mountainous rockslide zones, or coastal slopes. Expanding the dataset to include satellite imagery from additional regions addresses this limitation by improving terrain diversity and increasing the model's generalization capabilities.

Another limitation involves inconsistencies in spatial resolution among the Sentinel-2 image bands. With bands captured at 10 m, 20 m, and 60 m per pixel, the variation introduces spatial imbalance, making it difficult to uniformly detect smaller or more subtle landslide features. Standardizing all bands through image resampling to a common resolution, such as 10 m, ensures uniformity in data input. Additionally, implementing multiscale feature extraction techniques in deep learning models allows the integration of spatial details from different resolutions without compromising accuracy. Temporal and seasonal bias also affects dataset reliability. Images concentrated within a specific season or climate condition restrict the model's adaptability to changes in vegetation, lighting, and weather. Including multi-seasonal and multi-temporal satellite data enhances variability in the training process, ensuring the model performs consistently across different time frames and environmental settings. The patch generation process further introduces the risk of class imbalance, with uneven representation between landslide and non-landslide samples. This imbalance skews the model's learning process and reduces prediction accuracy. Applying data augmentation techniques such as flipping, rotating, scaling, and contrast adjustments expands the sample set and balances class representation. Incorporating synthetic data generation and implementing stratified sampling also ensures equitable distribution of classes during training. Addressing these limitations strengthens the model's reliability, increases prediction accuracy, and supports more effective application in real-world landslide detection scenarios.

4.3 Methodology

In this work, we developed a hybrid deep learning model to improve the accuracy of landslide forecasting models.

Landslides have serious threat for infrastructure and human lives highlighting the need for reliable and accurate models for prediction. Traditional approaches usually found it difficult to extract detailed space information from the satellite images, which motivates to explore more advanced techniques. Our approach takes advantage of the UNet model, a well-established model known for its accuracy in semantic segmentation performance. By integrating a pyramid grouping layer, our goal is to improve the capacity of the model to capture multiple scale characteristics, improving its performance in variables spatial resolutions. This approach aims to address the challenges related to landslide detection which offer a stronger solution for evaluating landslide susceptibility. The UNet design is appropriate for tasks that demand accurate segmentation of spatial characteristics like landslide detection. However, to further improve the model, we integrate the layers of pyramid groups. This layer adds several scales to the model, allowing the model to process information about various spatial resolutions. This way can get a sense of the fine grains and the broader landscape features that are essential for accurate landslide detection. The ability to analyze several scales ensures that the model is able to recognize equations that indicate the risk of landslides. This treatment at multiple scales improves the general performance of the model and helps to adapt to a variety of topographical conditions and soil cover types. This technique enables the model to better understand the correlation between different image segments. This is necessary to recognize complex patterns associated with landslide sensitivity.

OBIA improves the ability of model to identify features like variation in terrain height, vegetation density and soil texture, which are the key risk indicators of landslide. By combining deep learning with OBIA, we strive to increase the accuracy and reliability of landslides detection, especially in areas where traditional methods like pixel-based method struggle to capture simple on small variations in the landscape.

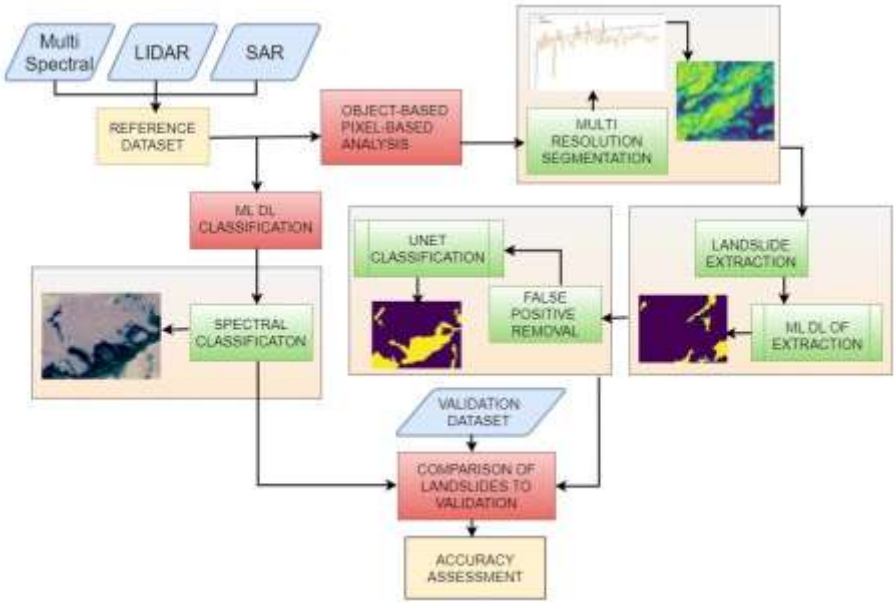


Figure 4.7 Approaches for mapping Geographical Feature with Rule-Based and Data-Driven models.

Landslide4Sense dataset is used to train and evaluate this hybrid models. Additionally, preprocessing procedures like normalization and noise reduction are used to ensure that the data in model training is consistently optimized. Additionally, data augmentation methods are used to increase the diversity of dataset. This will help the model to better generalize landslides in areas that were not seen during training. The purpose of this study is to improve landslide capabilities and convey valuable knowledge that can be useful for active disaster management and reduction strategies in landslide-sensitive areas.

4.3.1 Fully Convolutional Networks (FCNs)

FCN is a DL architecture that deals with the training closure challenges that usually arise when an additional layer of folding is added to increase model complexity [106]. This breakdown prevents functional networking functions of various shapes and sizes. FCN overcomes this restriction by replacing fully connected layers traditionally found in neural networks with convolutional layers and upsampling layers. This design causes FCN to be particularly suitable for tasks that require image mapping on image, such as landslides, where the output is also an image rather than a single label or classification. FCN is able to process input images of different sizes, allowing them to process image patches of any size and are therefore customized by different image resolution. The key feature of FCN is its ability to extract global and local contextual features through the skip mechanism between lower sampling and

upsampling layers [107]. Skipping the connection helps to maintain important information from the middle layers and pass it directly to the upsampling layers, allowing the network to create more accurate and detailed segmentation [108]. This is particularly important for complex tasks, such as detection of landslides, where there is need to keep fine field details for precise prediction. It also ensures that the model can maintain semantic features of high levels and low-level spatial features, which are essential for understanding complex landscape details that may indicate the risks of landslide. Among the different FCN models, UNet has shown that it is particularly effective for image segmentation tasks, especially if the training data is limited. This characteristic makes the UNet an ideal model for prediction of landslides, as it can often be demanding. UNet architecture with a combination of contractual and expanding routes allows the model to learn and produce accurate segmentation maps, although only a small number of training patches are provided. This is an important advantage in the prediction of landslide, where the marked data sets are obtaining sufficient training data which can be time-consuming and costly. UNet ability to generate accurate predictions with minimal data makes it a powerful tool for detecting landslides and other applications where there is limitation of training data. By using the strengths of convolution and upsampling layers, FCN can effectively process and segment complex landscapes and identify the features of terrain indicating the landslide risks [109]. The flexibility for handling various image resolution and image sizes along with efficient use of skip connection to maintain global and local features, makes FCN an ideal choice for this prediction task.

4.3.2 Swin Transformer

Swin Transformer is a specialized type of transformer vision that increases the efficiency and effectiveness of image processing in computer vision [110][111]. Unlike traditional transformers, which consider the image to be a sequence of non-overlapping patches and apply self-attention in all of them, the transformer receives a more structured approach. It divides the image into smaller, non-overlapping local region known as Self -attention window mechanism is applied only in each individual window. This localized approach significantly reduces computing complexity compared to the traditional method that would require computing attention throughout the image at once. By focusing on smaller areas, the swin transformer minimizes the computation number while still capturing all important local features. As a result, for the processing of large images it is highly efficient and implementing complex image recognition tasks. The factors that really distinguishes the swin transformer from traditional

models is the use of a shifted window where the windows move between different network layers [112]. This shift allows the model to collect both global and local patterns, local patterns in individual windows and global interactions between different regions of image. By moving the windows in each layer, the model gets a wider view of the image structure, allowing it to understand how different parts of the image are related. As the network deepens, the size of the windows increases, allowing the model to look at most of the image and capture the abstract features of a higher level. This hierarchical approach with the gradually growing window size allows the swin transformer to balance computing efficiency with the ability to capture complex features throughout the image.

4.3.3 Object Based Image Analysis

This approach focuses on grouping adjacent pixels in regions prior to classification, providing a more structured method for analyzing high-resolution satellite images. Particularly useful for remote sensing, this technology enables automated image analysis by describing image content based on specific object functions. One of the most important strengths of this method is its ability to combine spectral, structural, and spatial/context-related properties that allow pixels to group pixels with uniform and meaningful objects. In contrast to pixel-based methods that handle each pixel individually, this object-based approach enables objects that can be linked to real entities. The use of object-based image analysis (OBIA) for landslides was well established in previous studies. Previous research has shown how OBIA can be repeatedly applied to satellite images to create historical landslides and recognize landslides in various regions. OBIA was also used to modify landslides. One of the main advantages of OBIA compared to traditional pixel-based approaches is its ability to classify complex geospatial objects with large differences in size, shape and spectral properties. Landslides usually have other natural features, such as spectral characteristics similar to those that have been altered by people and agricultural areas etc. OBIA's ability to tackle this complexity and distinguish similarly visible features is particularly effective when OBIA is aware of landslides. In comparison, pixel-based approaches focus on individual pixels without considering the broader context that makes it difficult to classify such complex objects. In OBIA, two important principles control the analysis: segmentation and classification. Segmentation divides an image into smaller, meaningful objects based on spectrum, structure, and spatial properties. As soon as these objects are segmented, they are categorized based on specific features derived from them to allow for more detailed detection of landslides. For example, OBIA is effective at a spatial

resolution of 10 meters, with a minimum object size for identification being approximately 100 pixels. This provides a good approach to data analysis with medium resolution and substantial landslides. OBIA works on predicting landslides by dividing remote sensing imagery into small, meaningful objects, each one being a distinct feature of the landscape. For example, areas with steep areas, less coverage of vegetation, or the recent obstacles in that location can be identified as zones with high-risk zones for potential landslides. This is important because it grouped into more uniform areas, reducing noise and improving the model's ability to concentrate on wise characteristics related to landslides. OBIA's segmentation process typically uses an algorithm that groups adjacent pixels based on spectrum and spatial similarity. This step follows a classification in which segmented objects are assigned a specific name, such as "landslide" or "safety". Various criteria can be used to further improve classification. These classifications are often improved through the integration of additional remote sensing data, such as radar-based and LIDAR data, and can provide complementary information on surface deformations and hidden geological features that are invisible in optical images. The image processing is often performed in areas where there are subtle differences in fields that are difficult to record using pixel-based methods. For example, landslide boundaries can be irregular and the spectral signature of the landslide can overlap with other natural or human signatures. Through analysis of the entire image object rather than individual pixels, OBIA can take into account a wide range of contexts. Furthermore, OBIA can effectively handle extremely high-resolution images and recognize small landslides and their pioneers. Furthermore, OBIA could be improved by integrating ML techniques that allow for more sophisticated and adaptive classification. ML algorithms can be trained on large datasets with marked remote sensing images to automatically learn landslide distinction capabilities. These models are capable to apply learning patterns to new invisible images, improving landslide accuracy and robustness over time. This approach is particularly useful for monitoring large, or inaccessible areas where manual field inspections are time consuming and expensive. Further, recognizing limitations on pixel-based approaches, particularly for image classification, has gained the increased importance of OBIA (object-based image analysis). This is especially clear when considering high spatial resolution and very high resolution (VHR) remote recording data, as it can overcome the limitations of analysis per pixel. For satellite image processing, OBIA provides clear image analysis, for example, where focusing on the analysis of segmented image objects is instead of individual pixel values. This method allows for structured inspection of characteristics and events. One of Obia's core elements is segmentation that takes into account the spectral, structural, morphological, and topographical

properties of individual pixels and transforms them into objects [113]. This change reduces the frequency of false positive aspects and increases the accuracy of landslides. Segmentation and classification are usually two main steps in the OBIA method for landslide recognition. Determining criteria that indicate the size and shape of elements in one image makes segmentation a particularly difficult operation. To ensure that the final object is properly displayed by important landscape elements, this level requires an iterative process in which various segmentation techniques are evaluated and improved based on visual assessments. After segmentation, segmented objects are classified using criteria specified in the classification stage [114]. This classification is based on determining the different criteria's for landslide detection based on geological properties, changes in system coverage, field differences, or instability [115]. OBIA offers a variety of advantages over traditional pixel-based methods, integration into machine learning models (ML) poses the challenges of more complex applications. To find ideal scale parameters for examining factors related to geographical characteristics such as landslides is a major challenge. Satellite image landslides are very different and it is difficult to use a consistent segmentation approach to various criteria. This variability underscores the need for continuous improvement of the OBIA method and integration into ML algorithms to improve the accuracy of landslide detection and mapping in various environmental contexts.

4.3.4 Image Segmentation

Image segmentation is a critical component of OBIA. It is intended to define the basic unit or object of an image and is later analyzed in classification and interpretation [116][117]. OBIA's effectiveness in landslide prediction depends heavily on the quality of the segmentation process, as it directly affects how landscape features are identified. In segmentation, the goal is to group adjacent pixels into coherent segments based on similarity in spectral, spatial and context properties. These segments exceed individual pixel values by including additional statistics such as the mean, median, standard deviation, and range of values for each image [118]. This additional information layer makes segmentation more meaningful compared to pixel-based analysis, as it captures wider patterns within the image and identifies larger relevant features that show landslides.

Segmentation technology developed in the 1980s as part of a wider field of image processing and computer vision. Several algorithms have been developed for the processing of remote

sensing data [119]. These techniques often focus on adding spatial context information to traditional segmentation methods and improving the ability to segment characteristics based on their relationship to surrounding pixels. Methods frequently used for segmentation include regional algorithms, Markov models, surface catchments, hierarchical algorithms, and clustering techniques such as K-means. Each of these approaches has strengths in different contexts, but they all aim to group pixels into meaningful objects. This can be analyzed with specific features related to landslide detection, such as changes in topography, vegetation, or soil condition. One of the biggest challenges in segmentation is to deal with objects of different sizes in the same image, especially when applied to remote sensing data for landslide prediction.

Traditional segmentation methods sometimes struggle to determine subtle changes in smaller areas of landslides or wide areas affected by landslides. To overcome this challenge, a segmentation approach using several resolutions, such as regional merge technology, has been developed. This method adapts to different object sizes by merging smaller segments into larger ones, so that the segmentation process can recognize both fine and larger properties. Additional trigger segmentation ensures that objects representing landslides are accurately separated regardless of size. This is particularly important for landslide detection, and both small landslides need to be identified. When predicting landslides, effective segmentation is extremely important for extracting meaningful properties from remote sensing images that can display potential landslides. Furthermore, by segmentation [120], it allows for the integration of various data types, such as digital height models (DEMs) and multi-level images. By focusing on coherent image objects rather than individual pixels, segmentation allows for a more holistic view of the landscape, allowing for a more accurate assessment of landslide sensitivity. Therefore, segmentation plays an important role in improving the reliability and accuracy of landslide prediction or detection models using remotely acquired images.

4.3.5 UNet

UNet was originally developed for biomedical image segmentation, but later spread to different type of image segmentations, and landslide prediction or detection. Its effectiveness is based on a distinctive U-shaped architecture consisting of two units: an encoder and a decoder [121]. This architecture, shown in this distinctive "u" shaped, allows UNet to perform segmentation tasks efficiently. Additionally, the identification and classification of areas within the image is

ideally shown, as illustrated in Figure 4.8. The encoder and decoder structure is combined by skip connections to improve the important role of the transmission of key information between the two components and the model's capabilities and provide accurate predictions [122]. The context information is extracted from the input image using the UNet encoder component. This section consists of foldable layers, followed by a maximum pooling layer that reduces the spatial dimensions of the functional card and extracts from the image. By stopping data, maximum pooling aids the network with its most important properties, but the folding layer allows the model to identify different patterns and structures within the image. This helps to collect more comprehensive context-related data using the models needed to understand complex spatial connections that define properties such as landslides. The encoder downsampling process compresses and promotes the model to recognize and understand large properties and patterns [123].

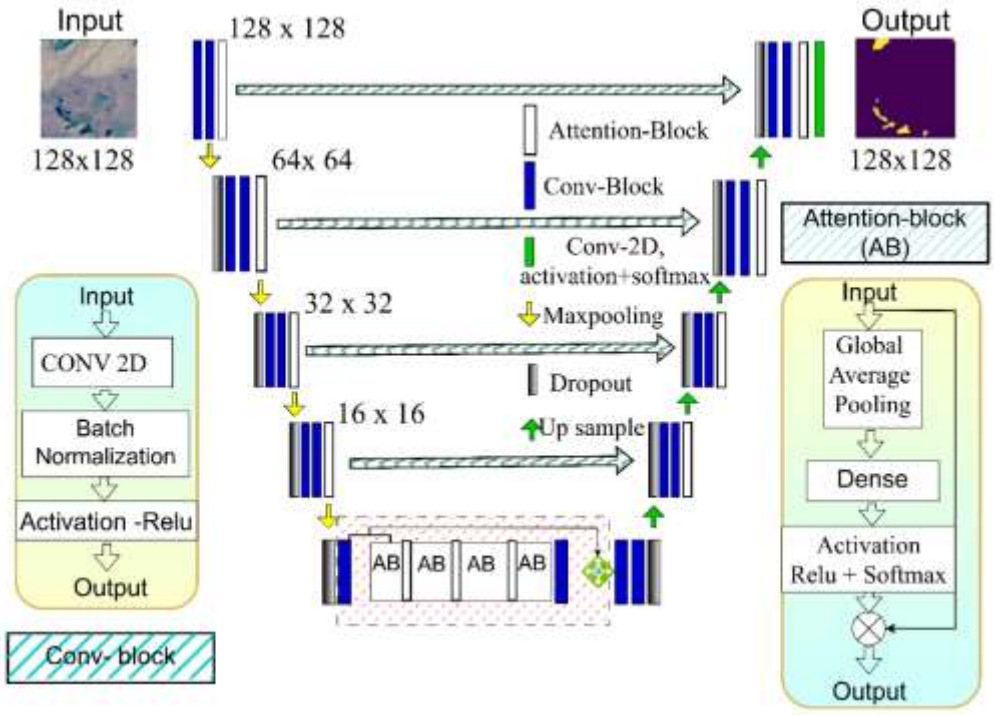


Figure 4.8 The Implemented U-Net: A Deep Learning Image Segmentation Model.

Skip connections are a key innovation in the UNet architecture. these connections link the corresponding levels of the encoder and decoder components, allowing important spatial information to be handed over directly between the two components. In this way, the decoder can effectively reconstruct the detailed segmentation by incorporating high resolution information from the encoder. The decoder uses upsampling values to increase the size of the

feature, while simultaneously reducing the number of channels, effectively reconstructing the image at a higher resolution. The folding layer in the decoder improves the segmentation and improves the accuracy of the final problem. This allows UNet to generate accurate predictions at the pixel level, allowing the gradual sampling and improvement process to clearly distinguish between landslides and unaffected areas. During the output phase, UNet creates a segmentation that assigns a probability value to each pixel. This shows the possibility of a particular class, how it belongs to a landslide. This allows for detailed analysis of the entire image and allows for the identification of specific areas of risk. It is a powerful tool in the areas of landslides and other applications for environmental monitoring, providing high accuracy and valuable insights for disaster risk management [124].

Image segmentation is the method of segmenting an image $N: \alpha \to Q$ into multiple regions $\alpha_{j_i}^{Q} = 1$, meeting requirements that don't overlap $\alpha_i \cap \alpha_j = \varphi, j \neq i$ and encompassing the whole image domain $K_{j=1}^{Q}\alpha_{j=}\alpha$, here α is image domain represented by the limited and open subset of Q2.

In image processing, the first step is to optimize a probabilistic posterior distribution in order to extract features from the image μ. By attempting to generate feature representations for a particular envision k, this process enables a more in-depth analysis and understanding of the content of image.

$$\begin{aligned} &\arg\max_{\mu} p(\mu \mid g;\delta) = \arg\max_{\mu} \log\, p(\mu \mid g;\delta) = \\ &\arg\max_{\mu} \log \frac{q(f|\mu;\delta)q(\mu;\delta)}{p(f)} = \\ &\arg\max_{\mu} \log\, q(g \mid \mu;\delta)q(\mu;\delta) \end{aligned} \tag{4.1}$$

The environmental factor is identified as being in terms of traditional unconscious inference. The prior probability p(μ;δ) may be well-modeled by the probability q(g ∣ μ;δ) and normal distributions. In particular, we have

$$\begin{aligned} &q(g \mid \mu;\delta) \propto e^{-\frac{1}{2\gamma^2}\int_{\alpha}(k\mu-g)^2} b\alpha = e^{-\rho\int_{\alpha}(k\mu-)^2} b\alpha, q(\mu;\delta) \propto \\ &e^{-\rho\int_{\alpha}\varphi(\nabla\mu)b\alpha} \end{aligned} \tag{4.2}$$

Therefore, the first step is to find a smooth approximation μ and decrease the multiphase generalizability. The following is a reformed version of this optimization problem:

$$\min_{\mu \in s^{1,2}(\alpha)} \int_{\alpha} (f - D\mu)^2 dy + \int_{\alpha} \varphi(\nabla\mu) dy \tag{4.3}$$

here μ is associated with function space $(g - B\mu)^2$ where B stands for blur operator, f for the provided picture, and φ for a geometric expression that uses the gradient μ. The trade-off between approximation smoothness to the original picture is controlled by the parameter μ. Finding the ideal μ that strikes a balance between these variables is the goal in order to produce a reliable and accurate solution for the picture segmentation problem.

In this case, $B: G^C \to G$ is a blur operator, and $(\varphi(\nabla\mu) = u|\nabla\mu|^2 + |\nabla\mu|$ represents the geometric prior of μ. Additionally, we have $\mu = \lambda\gamma$. Consequently, this method produces the nonlinear system that is given by:

$$F(\mu; \delta) = D^T D\mu - \tau\nabla(\varphi(\nabla\mu)) = b \tag{4.4}$$

where $a = D^T f$ and $\delta = (D, \nabla, \tau, u)$ all included in the parameter δ. In image segmentation tasks, this nonlinear system is essential, and its successful solution yields precise and significant predictions. Our suggested design is divided into two primary components from the variational segmentation model: the feature integration module $U_T(\mu: \emptyset_2)$ and the solution module $L_T(g: \emptyset_1)$ In the multi-stage example, extracting feature is handled by the solution module $L_T(g: \emptyset_1)$), whereas stage fusion is handled by the learnable feature fusion module $U_T(\mu: \emptyset_2)$. In this study, we provide UNet, a novel framework for explainable landslide prediction on images, using a nonlinear multigrid approach. The two modules function in this way:

$$\mu = L_T(g: \emptyset_1) \tag{4.5}$$

$$q = U_T(g: \emptyset_2) \tag{4.6}$$

Here, f stands for the input landslide picture, q for the truth partition prediction, and f for the feature map. These modules work together to provide the total approximation function:

$$q = P(L_T(g; \emptyset_1); \emptyset_2) \tag{4.7}$$

The parameters ∅1 and ∅2 in our suggested explainable UNet architecture must be learnt during the training phase to better understand the capabilities of the UNet-generated modules $L_T(g: \emptyset_1)$ and $p_T(\mu: \emptyset_2)$.

The UNet architecture and folding training techniques of neural networks have been thoroughly explained in various studies [125][128]. In this study, a fully convolutional neuronal network was constructed on an object basis for each pixel probability. In contrast to [125][126], they used traditional CNNs to classify landslides [129]. Based on previous research, we also propose a pyramid pooling layer, which can be seen in Figure 4.9. Empirically, this layer is the priority of the appropriate global context as being an important factor that influences the amount of contextual information used in deep neuronal networks is the coverage area. By collecting data at several scales, the pyramid pooling layer included in the architecture of system strategically enhance the benefits of context. A more comprehensive understanding of input data improves the system's ability to identify and predict complex patterns of landslides by this stepwise aggregation of context-related factors.

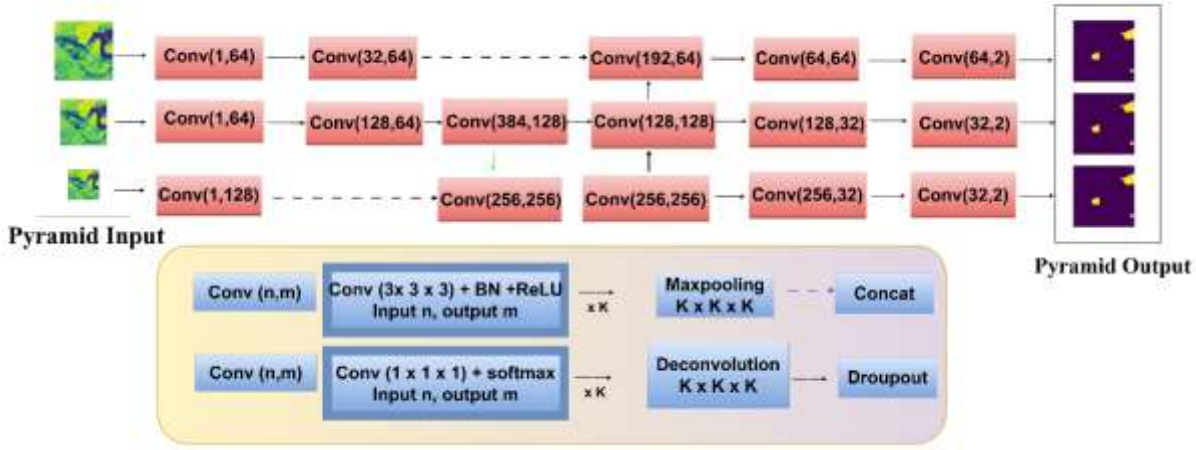


Figure 4.9 Visualizing UNet-Pyramid Layer Model for Multi-Scale Feature Extraction.

The UNet-Pyramid layer model is an extended and advanced variation of the traditional UNet architecture and was developed to improve the distinctive extraction and multi multi-scale information from the input image, making it more effective for tasks such as detection of landslides in remote sensing images. The most important innovation in this model is the addition of a pyramid pooling layer integrated into the contractual way of the network. These pyramid pooling layers allow the model to record information on several scales. This is very important for identifying objects or areas of different sizes and structures, such as landslides.

The encoder portion of the model focuses on low-level characteristics from images like edges, textures, simple patterns, and decoders are commissioned to record representations at a higher level, such as contract path of the UNet-Pyramid model includes several folding blocks, each consisting of 3x3 convolutions, increasingly extracting properties from the image. These blocks follow the SoftMax activation function that helps normalize the results, play 2x2 filters and feature cards, and use layers of maximum pooling layers in two steps that reduce spatial dimensions [130][131]. Pyramid pooling allows the model to capture multi-scale context-related information which improves the ability of the model to recognize objects of different sizes in the image. Each folding block in the contract path doubles the number of features and enhances the model's ability to understand abstract and high-level characteristics. In this way, networks not only capture fine details at a lower level, but also collect a higher level of broader contextual information, essential for accurate segmentation of complex objects such as landslides. Images are processed through pyramid pooling layers, so the network can extract features from different spatial resolutions, which better distinguish between landslides and other regions with similar textures or structures. This distinctive feature extraction is particularly useful for landslide detection. This can significantly change the characteristics of local sites, and models must accurately record both topographical features and large-scale details in order to make accurate predictions.

Given the feature map N, the element in n is represented by the notation $Y_{a,b,c}$, where a, b, and c stand for the indexed channel, row, and column, respectively. Whereas, the convolution process is as follows: In the convolution kernel K for the element $K_{a,b,c}$, a and b are the channels of the last slice (a) and the current slice (b), respectively, and c is the offset between two items.

$$N'_{a,b,c} = \left\{ N_{a,b,c} + n\left(\sum_m \sum_P N'_{a,b-1,c+P-1\times L'_{m,a,P}} \right) \right. \tag{4.8}$$

N is an activation function Softmax. Thus, through the convolution layers, the significance data from the feature map is transferred to the bottom segment of the feature map. An inverted version of the contracting path is the expanding path in the UNet-Pyramid layer model. To up sample and integrate the features with the matching output of the encoder block at the same level, it uses convolutional-transpose layers. Each decoder block results in a halving of the feature map's number.

$$\begin{aligned}&\forall n\varepsilon\left[1,2,\dots,p_Q^{[l]}\right]\\&\mathrm{Conv}\left(c^{[l-1]},L^{[n]}\right)_{x,y}\\&=\Psi^{[l]}\sum_{i=1}^{n_p^{[1-1]}}\sum_{j=1}^{[l-1]}\sum_{k=1}^{n_Q^{[-1]}}K_{i,j,k}^{[n]}a_{x+i-1,y+j-1,k}^{[l-1]}+b_n^{[l]}\end{aligned}\tag{4.9}$$

$$\mathrm{Dim}\left(\mathrm{conv}\left(a^{[l-1]},K^{[p]}\right)\right)=\left(n_T^{[l]},n_R^{[l]}\right)$$

$$\begin{aligned}a^{[l]}=\Big[&\Psi^{[l]}\left(\mathrm{Conv}\left(c^{[l-1]},L^{[1]}\right)\right),\Psi^{[l]}\left(c^{[l-1]},L^{[2]}\right),\dots.,\Psi^{[l]}\\&\mathrm{Conv}\left(c^{[l-1]},L\left[l_4^{[1]}\right]\right)\Big)\Big)\\&\mathrm{Dim}\left(a[1]\right)=\left(n_T^{[l]},n_R^{[l]},n_Q^{[l]}\right)\end{aligned}\tag{4.10}$$

Where,

$$\begin{aligned}&n_{\frac{T}{R}}^{[l]}=\left[\frac{n_T^{[l-1]}+2p^{[l]}-f^{[l]}}{s^{l]}}+1\right];s>0=\\&\frac{n_{\frac{T}{R}}^{[l-1]}+2pp^{[l]}-f^{[l]}}{1};s=0,n_Q^{[l]}=\text{Filters count.}\end{aligned}\tag{4.11}$$

Where $= c^{[l-1]}$ is input $\left(n_T^{[l-1]},n_R^{[l-1]},n_Q^{[l-1]}\right)$ is size of the input and filters $n_Q^{[l]},p^{[l]}$ and $s^{[l]}$ are the stride and padding values, $n_Q^{[i]}$ is the filter, where K(n) has dimension $\left(f^{[l]},f^{[l]},f_Q^{[l-1]}\right)$, $b_n^{[l]}$ is the n^{th} convolutions bias, $\Psi^{[l]}$ the activation function and finally a[l] is the output of this layer with size $\left(n_T^{[l]},n_R^{[l]},n_Q^{[l]}\right)$.

The 26 convolutional layers of the UNet-Pyramid layer model include 22 convolutional layers, 4 convolutional transpose layers, and additional pyramid pooling layers for gathering multi-scale data. Accurate and thorough segmentation results are produced by this architecture's ability to gather contextual information at various sizes effectively and efficiently [132].

$$\begin{aligned}&c_{i,i,k}^{[]}=\mathrm{pool}\left(c^{[l-1]}\right)_{i,j,k}\\&=\phi^{[l]}\left(\left(c_{i+x-1,j+y-1,k)}^{[l-1]}\right)\right)_{(x,y)\varepsilon\left[1,2,\dots,\dots,g^{[l]}\right]^2}\end{aligned}\tag{4.12}$$

$$p_{\frac{T}{R}}^{[l]} = \left[\frac{p_T^{[l-1]} + 2q^{[l]} - g^{[l]}}{t[l]} + 1\right]; t > 0$$
$$= p_{\frac{T}{R}}^{[l-1]} + 2p^{[l]} - g^{[l]}; t > 0$$
$$p_k^{[l]} = p_k^{[l-1]}$$

Where, $c^{[l-1]}$ is the input with size = $\left(p_T^{[l-1]}, p_R^{[l-1]}, p\right)$, $q^{[l]}$ and $t^{[l]}$ are the padding and stride value, $\phi^{[1]}$ is the pooling function which $g^{[l]}$ filter size. The $a^{[l]}$ with $\left(p_T^{[l]}, p_R^{[l]}, p_0^{[l]}\right)$ gives the output of pyramid pooling layer. We introduced the Binary cross-entropy loss function as the model's output contain multiple neurons. Loss = -1/N * Summation of [Ground truth label * log (Predicted probability) + (1 - Ground truth label) * log (1 - Predicted probability)]

ALGORITHM 4.1: ALGORITHM FOR FEATURE EXTRACTION USING PYRAMID POOLING LAYER

```
function pyramid_pooling_layer(input_feature_map, pool_sizes)
    Input_feature_map: a 3D tensor of shape (height, width, channels)
    pool_sizes: a list of integers specifying the pool sizes to use pooled_features = []
        for pool_sizeinpool_sizes do
            height = ⌈ input_feature map.height / pool_size ⌉
            width = ⌈ input_feature map.width / pool_size ⌉
            pooled_feature_mapmax_pooling
        (input_feature_map, pool_size, pool_size) resized_pooled_feature_map=
        resize (pooled_feature_map, height, width) pooled_feature_vector=
        flatten(resized_pooled_feature_map) pooled_features.append(pooled_feature_vector)


        concatenate (pooled_features, axis=channel_axis)
        return feature vector
end function
```

$$L = -\frac{1}{p}\sum_{i-1}^{p}\left\{\sum_{j-1}^{L+1}\left[x_{ij}\log\left(\hat{x_{ij}}\right) + (1 - x_{ij})\log\left(1 - \hat{x_{ij}^{N}}\right)\right] + \beta * Q_i\right\} \quad (4.13)$$

$$Q_i = \frac{1}{L}\sum_{k-1}^{L+1} x_{ik}\left[\hat{x_{ik}^{k}} + \mu - \sum_{j-1}^{L+1}\left(x_{ij}\hat{x_{ij}^{j}}\right)\right] \quad (4.14)$$

p is the batch size, while μ is a number between 0 and 1. The label for the i^{th} pair of pixel is $x_i^{\wedge}$. $Q_i's$ is a regularization term to ensure classification accuracy, and $x_i^{\wedge}$ is the expected value of the i^{th} pair of pixels. The weight coefficient is denoted by β. Deconvolutional layer, or convolutional-transpose layer:

Convolutional-transpose operation:

x = sigma (U * y + c) here,

x is output of feature map; sigma is an activation function, U is convolutional-transpose filter which is learnable, y is the input of feature map and c is bias term.

By using convolutional layers of 3x3 and as sliding windows here Softmax activation function is used, the UNet-Pyramid layer model scans the input picture and reduces spatial dimensions by half. In the UNet-Pyramid model, like in the traditional UNet model, the decoder route uses up-convolution layers and concatenates feature map from the appropriate locations in the encoder step to extract the spatial position.

$$z_p^{[i]} = \sum_{l=1}^{n_{(i-1)}} w_{p,l}^{[i]} a_l^{[i-1]} + b_p^{[i]} \tag{4.15}$$

$$a_p^{[i]} = \Psi^{[i]}\left(z_p^{[i]}\right) \tag{4.16}$$

a[i-1] input is convolution layer and pooling layer result with the dimension $\left(p_T^{[i-1]}, p_R^{[i-1]}, p_Q^{[i-1]}\right)$. In order to plug it into the fully connected layer:

$$p_{i-1} = p_T^{[i-1]} \times p_R^{[i-1]} \times p_Q^{[i-1]} \tag{4.17}$$

To effectively include the comprehensive information from the grid and account for encoded characteristics, the UNet is utilized for landslide prediction. This is accomplished by interpolating the coarse grid changes back to the fine grid:

$$v^{\wedge} \leftarrow v^i + k_{i+1}^i v^{i+1} - k_i^{i+1} v^{\wedge} \tag{4.18}$$

where the interpolation function L_{i+1}^i is approximated by the learnable upsampling operation k_{i+1}^i For this, we specifically employ a transposed convolution with p filters and a stride of 2.

Here, we have L-grid cycles with i = 1,..., i-1. We compensate for the information in the feature maps V^, by updating the fine grid approximation V^, using this transposed convolution. In order to recover features with more precise information, the transposed convolution develops an adaptive mapping.

These feature extraction parameters, which are expressed as follows, are learnt to approximate the feature solution:

$$\emptyset_1 = k_{i+1}^{i}, k_i^{i+1}, \left(k_{q,i}\right), \left(k'_{q,i,j}\right)_{jq=1}^{kq}, k_{i=1}^{l-1}, k^0, \left(k_{n,1}\right), \left(k'_{n,i,j}\right)_{j=1}^{kn} \mid q \in \{i, r\} \tag{4.19}$$

Our UNet model for landslide prediction improves the image segmentation quality by fine-tuning these feature extraction parameters, which enables it to efficiently extract and use pertinent spatial information in the input images. The accuracy of landslide segmentation can be greatly impacted by the patch size selection in the UNet model, particularly when combined with a pyramid layer. Choosing the wrong patch size might result in segmentations that are insufficient or erroneous since landslides can take on a variety of forms and sizes.

Then, the features $P_{j,1}^{n}$ three characteristics, including the distortion feature, are concatenated by the decoder. P_j^q the encoder's down-sampling block's feature $P_{j,c}^{T}$, and transposed convolution features in primary decoder. It is possible to formulate the procedure. where transposed convolution is represented by $j \geq 2$ Deconv. Large amounts of edge information are stored in the edge decoder features, which are used as reference features. Additionally, before joining the other blocks, a subpixel convolution layer is applied to the up-sampling block of the main decoder. We used a convolutional layer for upsampling. Three residual blocks, as seen in Figure 4.10, are connected after the convolutional layer to acquire additional high-frequency data or specifics. The careful selection of patch size is vital not only for capturing landslide features but also for ensuring the model's overall efficiency and effectiveness. If the patch size is too small, the model may fail to capture enough contextual information about the surrounding landscape, which is essential for distinguishing between landslide and non-landslide areas. Small patches can separate important landscape features such as topographic variation and vegetation coverage, which are key indicators of landslide sensitivity.

$$P_{1,1}^{n} = \text{Concat.} \left(P_1^T, P_{a,b}^{t}, \text{Decov}\left(P_{c,b}^{t}\right)\right) \tag{4.20}$$

$$P_{j,1}^{n} = \text{Concat.}\left(P_{j}^{T}, P_{(a,j),b}^{t}, \text{Decov}\left(P_{(j-1),b}^{n}\right)\right) \tag{4.21}$$

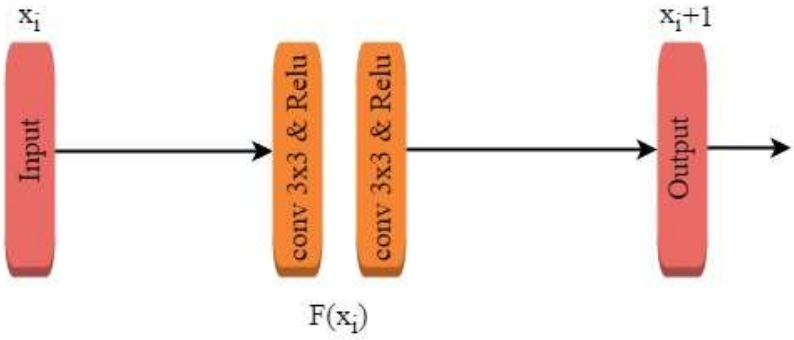


Figure 4.10 Detailed Architecture of Residual Blocks in U-Net for Enhanced Image Segmentation.

In this scenario, the segmentation edition may be fragmented or incomplete and lack critical areas of landslides. Conversely, excessively large patches can lead to out-of-focus predictions, as they can include areas with little or no relevance to the following landslides with different background area. These unrelated areas can mask the landslide, making it difficult for the model to learn the exact patterns and loses the accuracy of the model. Furthermore, large patches can bring complexity by combining several types of land cover into a single patch. The models can be confused when trying to distinguish between landslides and other land forms. Another challenge that arises when choosing a patch size is the problem of lightweight weight in the class. In many remote sensing records, the number of (negative) pixels does not govern the importance of the number of landslides (positive) (negative). This inherent imbalance can lead to skewed models to predict non-regional regions. Using large patches can further increase the number of pixels in the negative class, which could further increase this distortion. As a result, the model may be difficult to properly identify landslides. Landslides are less common and can often be used as small areas in larger non-regional areas. This imbalance can lead to low sensitivity (the ability to correctly identify landslides) and high false negative rates (misclassifying landslides as non-national slides). To improve this, choosing a patch size that balances negative and positive samples is key to improving model output. An important feature of FCN is its ability to extract both global and local context characteristics via a skip connection mechanism between the downsampling and upsampling layers, as shown in Figure 4.11. Furthermore, techniques such as oversampling positive instances during training and including subscene or class weights of negative instances can help reduce imbalance and ensure that the model captures landslides more accurately.

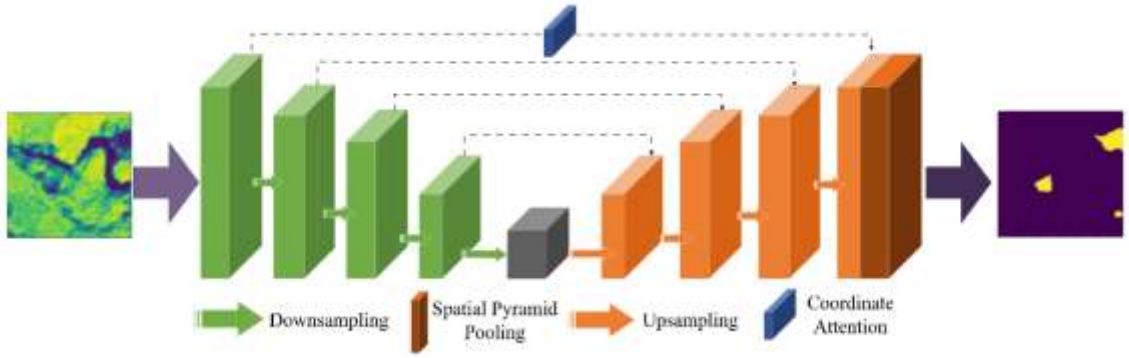


Figure. 4.11. Architecture with Upsampling, Downsampling, and Coordinate Attention for Enhanced Image Segmentation.

Furthermore, implementation of UNet models using pyramid layers provides an effective way to solve problems related to class patch size and imbalances. The pyramid pooling layer allows the model to capture multi-scale information, and recognize landslides by a variety of criteria, regardless of whether they are large, small or fragmented. These layers improve the ability of model to aggregate features from different resolutions so that it can handle a wide range of landslides within the same image. However, the patch size must be selected so that the pyramid pooling layer selects the multi-scale feature to capture matches. The patch is too large, which results in pyramid layers that concentrate on unrelated background information, but patches that are too small may not provide enough information to extract multispectral features. For searching for the optimal patch size, UNet models with pyramid layers can significantly improve landslide accuracy by effectively harmonizing large and small characteristics, while simultaneously minimizing class imbalances, this approach allows to better equip models to address a variety of land types and accurately predict landslides. Choosing the right patch size not only ensures that the model is capable of recording sufficient information, but it is also a key factor in ensuring that the model is equipped to take into account the complexity of remote sensing data. By experimenting with different patch sizes and involving strategies for connecting with classes, the model can be learned more effectively, leading to more accurate and reliable landslides. For pyramid feature maps and final prediction V^ is, calculated as:

$$f^{\wedge} = \sum_{m} n\left(k_i(K_i)\right) \cdot \text{SoftMax}\left(q\left(a_i(K_i)\right)\right) \tag{4.22}$$

The multilayer prediction and attention layers are denoted by the functions k_i and a_i; both are implemented as straightforward 1×1 convolutional layers. In terms of landslide prediction implementing the UNet model, the selection of restriction and interpolation operator. L_i^{i+1} and

L_{i+1}^{i} is related to the architecture of the UNet for image segmentation. In order to express grid transfer between the coarser grid i+1 and the finer grid i, we offer learnable convolutions for transfer operators. Both local and global image characteristics are efficiently captured by the UNet architecture. The coarser grid i+1 usually captures low-frequency features that provide pertinent visualized information. Thus, the right-side term, which is defined as follows, may be used to extract the important information on the grid:

$$p^{i+1} = k_i^{i+1}\left(p^i - h_k^i(v^i)\right) + h_k^{i+1}\left(k_i^{i+1}v^i\right) \tag{4.23}$$

where the output of the downsampling modules in the feature space is p^{i+1} and the inputs of the downsample block are p^i and v^i Here, we have L-grid cycles with i=1,..., i-1. In a way that is appropriate for the UNet, the restriction function L_i^{i+1} is approximated by the learnable downsample operation $h_{k_j}^i, k_i^{i+1}$.

In the UNet architecture convolutional layers, batch normalization, and activation functions (like ReLU) are all part of the UNet feature extraction model. The UNet model's learnable convolutions enable the network to recognize intricate patterns in the input picture, resulting in precise and instructive predictions for landslide segmentation tasks. The normalized difference vegetation index (NDVI) and the landslide likelihood maps from UNet were among the other data that were already included,

$$\mathrm{NDVI} = (\mathrm{NIR} - \mathrm{Red})/(\mathrm{NIR} + \mathrm{Red}) \tag{4.24}$$

where, NIR and Red denote the electromagnetic spectrum's Near-Infrared and Red Bands found in Sentinel-2 images. In the UNet-Pyramid model the pyramid pooling layer plays a vital role in improving the ability of model to collect and integrate the features from input images at different scales. This layer works by collecting information from four distinct pyramid scales, each designed to process features at different levels of spatial resolution. At the coarsest level, global pooling is used to condense the entire feature map into a single value, which provides a summary of the global information present in the image. This process effectively reduces functional cards to the most important expressions. The functional cards are divided into smaller subregions by the next layer of the pyramid, so the pooling process is collected from various geographic locations. To reduce the dimensions of the context image and maintain the important global characteristics of the input, each pyramid layer creates feature maps of

different sizes using 1×1 convolution layers. As a result, the dimensions of the function are reduced to 1/N of its initial size, where n is the size of each pyramid level, creating a more compact display. In this way, the function remains guaranteed in the orientation of the original input image. To combine global properties with different criteria, we chain characteristics placed at every pyramid level to achieve the final result of pyramid pooling. The degree to which this pooling mechanism works depends heavily on the layout and pyramid layers. The levels use different pooling cores with different pyramids so that the model can be trained at both fine and broad levels, by which it can collect a large amount of spatial information. The four pyramid levels of the presented UNet-Pyramid model are 1×1, 2×2, 3×3, and 6×6 sizes. Therefore, the model can simultaneously analyze many spatial scale properties. Complex uses where images contain patterns of different sizes and geographical distribution, benefit from this distinctive extraction in several standards. Further, to improved extraction properties pyramid pooling layers play an important role in their ability to understand complex spatial patterns and create accurate prediction. By recording properties from several spatial scales this model became more suitable for dealing with variations in size, shape, and context of identified objects. This is especially important for tasks such as landslides. This is because landslides vary widely at large levels and at different terrain patterns. The ability to process and integrate multiscale functions improves the sensitivity of the model compared to these variations, leading to more accurate segmentation and predictive results. Furthermore, the pyramid pooling layer improves the robustness of the model by extracting meaningful context-related information regardless of spatial dissolution of the input data. This adaptive feature pool strategy is very effective for deep learning architectures as the model is successfully blocked by a set of data records and applications. OBIA offers an approach for image segmentation and classification by detecting additional features beyond pixel levels, such as details of spectral information, geometric features, topological relationships, and textures. In contrast to pixel-based methods, this usually focuses on individual pixel values without considering wider context. OBIA is used to avail the geometric and spectral properties of objects in satellite images. By integrating these characteristics into knowledge-based rules, we hope to improve the performance of machine learning or deep learning models, particularly for complex tasks such as landslide prediction. OBIA allows you to create more accurate and meaningful object-based classifications that are very important for landslides. The main goal of this study is to improve landslides with the UNet pyramid model, which includes OBIA object-based properties and hierarchical rules. Based on signatures and geometrical forms of spectra, these rules aim to classify and improve the detection of potential landslides. Integration of OBIA into the UNet Pyramid Model uses

the skills of the model to generate probability cards for landslide detection, while simultaneously generating strict classification skills for OBIA. The hybrid models use OBIA to add context and improve the quality of segmentation which makes landslides more accurate and reliable. After training the model, they agree to its parameters and optimize its performance to improve prediction accuracy. Additionally, the inclusion of a Swin transformer in this architecture provides important thrust by capturing context information and dependencies for most of the input image. This transformation model can improve landslides by allowing the network to better understand the global relationships between different image regions. The OBIA segmentation performance, properties, and processing enrichment combination of Swin Transformers creates a robust framework for landslide detection. Implementing self-training techniques improves the ability of models to generalize different data records. With this technology, models can learn from new data and adapt to patterns that were previously invisible. This improves general robustness and prediction. Ultimately, this hybrid approach aims to significantly improve landslide accuracy and reliability. The strengths of OBIA, DL and Trans models are based on developing more stringent systems and recognizing landslides. This provides important support for proactive disaster management and reduction efforts, reducing the risk and impact of landslides on communities and infrastructure.

4.4 Results and Analysis

Unlike other traditional pixel-based methods that deal with all pixels only, OBIA analyzes spatial relations, pattern and attributes of neighboring pixels, which are summarized in objects and provide better and enhanced approach related to context. OBIA focuses on the geometric and spectral features of objects in one image can improve processing and identification. Through this integration, this study attempts to improve the general performance of machine learning or deep learning models, which makes complex scenarios more efficient. When predicting landslides, the integration of both the characteristics of both deep learning models such as the UNet pyramid can significantly improve segmentation accuracy. The UNet-Pyramid, the extended version of the UNet architecture, is characterized by the image segmentation and creates accurate limits for objects such as landslides. The hierarchical classification system of OBIA classification objects based on spectral properties and geometric shapes improves the functionality of the model and recognizes and separates landslides. This standard based on standards introduce other layers of improvement in the model and ensure that areas are identified as potential landscape designs more accurate and relevant. The

segmentation process can be improved by classification of the OBIA object for deep learning skills. This allows effective identification of large and small landslides.

Swin transformer is known for its ability to capture contextual information and dependencies over long distances within images. This is important for understanding the spatial relationships of complex landscapes, such as those susceptible to landslides. By including this transformer model, we improve the model's ability to identify complex patterns and subtle variations of remote sensing data. The swin transformer complements the segmentation capabilities of the UNet Pyramid model, allowing more accurate predictions, not just local features but also global image contexts, as shown in Figure 4.12. This integration enriches the understanding of models on several criteria and improves overall effectiveness in landslide recognition. Using self-training technology in this model improves the ability to generalize various data records.

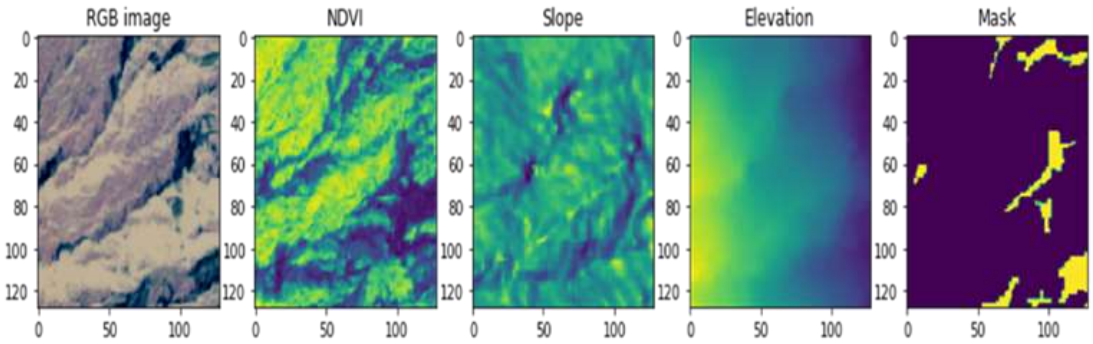


Figure 4.12 Visual representation of training dataset images.

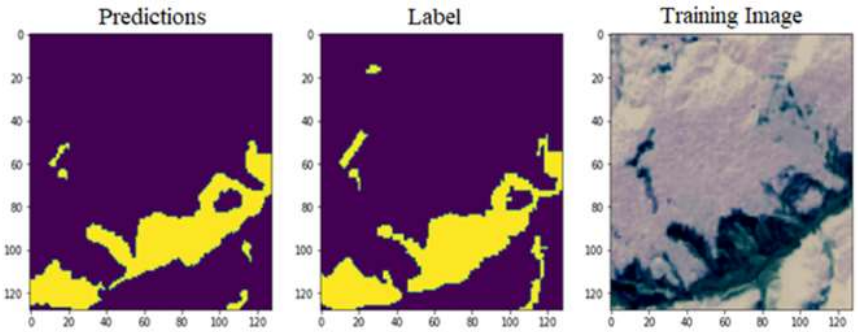


Figure 4.13 Visual representation of validation images.

Further, self-training allows the model to learn from new and non-labeled data after the initial training phase, adapt to previously invisible patterns, and improve performance over time. This approach increases the robustness of the model and makes it more adaptable to a variety of landscapes and regions. This is especially important when it comes to complex and

unpredictable environments where landslides occur. By combining OBIA's feature extraction with deep learning power of the UNet-Pyramid and contextual recognition of swin transformer, this hybrid architecture is equipped to master the challenges of landslide prediction. We trained this model for a minimum of 100 epochs. These epoch refers to the number of complete iterations across the training dataset. The arithmetic properties of this work, in this case, the NVIDIA RTX 4090 GPU accelerated the process. The RTX 4090 is a high-performance graphics card developed for deep learning and machine learning tasks, providing critical processing performance and storage bandwidth. This hardware is ideal for training and validating large-scale models using high-resolution data, as shown in Figure 4.13. This allows faster calculations and more efficient processing of complex calculations when training deep learning models. During training, the models were evaluated using several metrics: accuracy, recall, F1 score, and loss. These metrics are typically used to assess the effectiveness of classification models, particularly in tasks such as landslides. On the other hand, recovery assesses the ability of the model to identify all actual landslides and calculate the percentage of actual positive landslides from all actual landslides in the dataset.

F1 scores are composite metrics that correct for accuracy and provide a single value that reflects the complete performance of the model when identifying landslides. This is especially useful when the data set is unbalanced, as it provides a more comprehensive assessment of the model's ability to classify the model correctly. Losses were often expressed by loss functions such as cross or middle square errors, which quantified the quantification of how well the model was during training. Loss indicates that the model's predictions are close to the actual results. The data used for training was divided into two classes: landslides and non-landslide. Given the weight of the class of natural light, agricultural locations often exist in such data records where landslides are far less common. Therefore, these metrics help to optimize the model and enable accurate identification of landslides across a variety of remote sensing images.

The loss function of our method is defined as follows:

$$L_{loss} = \omega(M_a, \sigma_a) + \omega(M_b, \sigma_b) \tag{4.25}$$

When the training between two distinct tasks is balanced by the weighting operator, W, and the trainable parameter σ_b. The loss of the main and edge decoder branches is indicated by M_a and M_b. Additionally, the weight parameter ω is defined as follows:

$$\omega(M_j, \sigma_j) = \frac{1}{2\sigma_j} L_{Loss(i)} + \log\ \sigma_j, j \in \{a, b\} \quad (4.26)$$

where the logarithmic term is utilized to prevent σ_j from increasing and the $\sigma_j$values measure the forecast uncertainty.

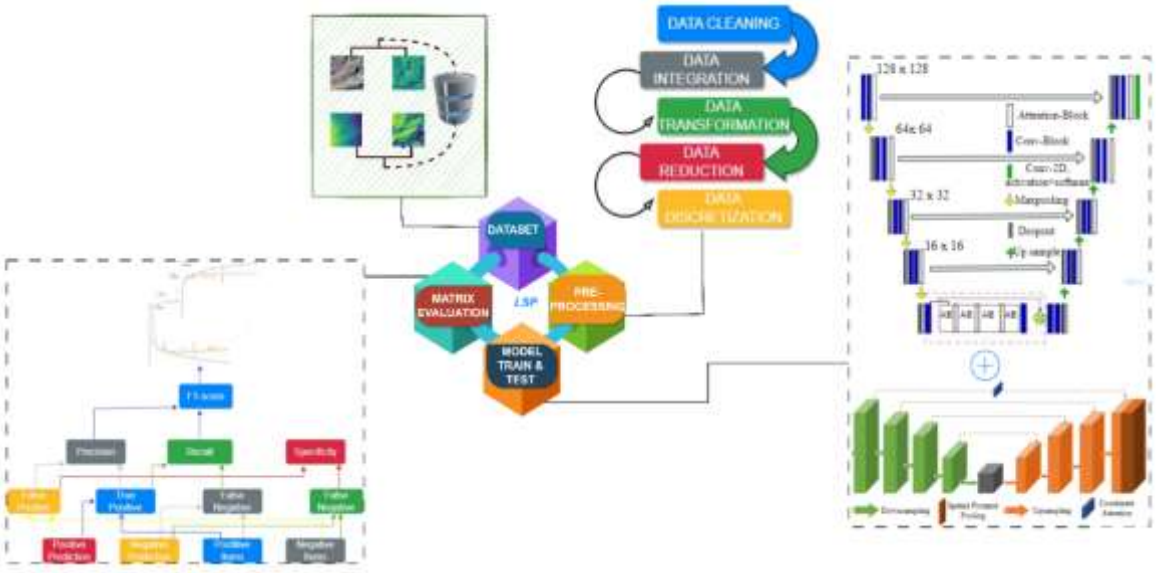


Figure 4.14 Proposed Workflow of Landslide Prediction using UNet-Pyramid Model.

4.5 Quantitative Evaluation

By calculating the pixels with the labels True Positive, False Positive and False Negative, the results of the landslide prediction were verified. Semantic accuracy assessment criteria, including precision, recall, and F1-score, were employed to methodically examine the efficacy of landslide detection. Measures of picture classification accuracy that are often used were employed to test the model's performance. As seen in Figure 4.15, we must first specify the four categories of expected samples for classification algorithms in order to compute these evaluation metrics. The ideas are as follows:

$$\text{Precision} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)+ False Positive (FP)}} \quad (4.27)$$

$$\text{Recall} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)+ False Negative (FN)}} \quad (4.28)$$

$$\text{F1-Score} = \frac{2\times \text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{4.29}$$

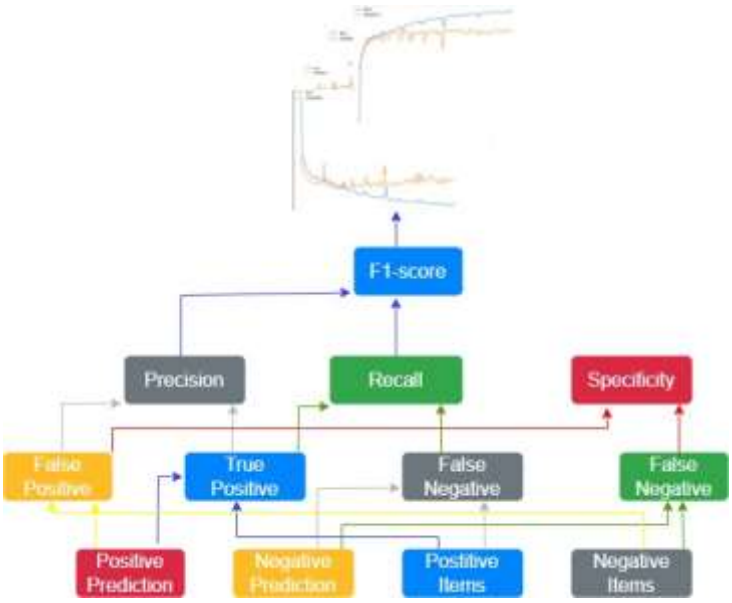


Figure 4.15 Performance Merits and Evaluation of the Proposed Approach.

The pyramid pooling layer improves the model's ability to extract features at different resolutions and collected all the important data from different spatial scales. This phase uses many pools and upsampling layers, each with a different kernel size and different stride. These methods provide feature maps of different resolutions, then combine to create feature pyramids with several scales. This feature Pyramid was integrated into the UNet decoder by skip connections, linking the corresponding levels of the encoder and decoder. By integrating these multiscale properties, the model collects both local and global context-related information and efficiently encode the various properties that exist on different spatial scales. This multiple competition is extremely important for tasks like landslides, and can vary greatly in topography and complexity. By including pyramid pooling layers, the model is sent to handling complex terrain characteristics that can have a wide range of spatial properties. For example, in areas with different heights, slopes, and vegetation types, landslides can occur at several scales, each with different characteristics. The ability to process and integrate these properties from several scales will allow landslide models to be more accurately recognized in challenging environments. Improvements in the pyramid pooling layer significantly increase the general predictive power of the model, allowing the processing of a variety of complex landscape data. Model performance was assessed using general classification metrics such as Precision, Recall, and F1 scores. These metrics are essential for model accuracy and reliability, especially for

evaluation in tasks such as landslides. In this task, imbalances in class models (more non-terrain instances than landslides) often distort the results. This model achieved precision of 0.95. This indicates the altitude accuracy in correct landslide identification. A recall value of 0.91 indicates that the model captures most of the actual landslide. The F1 score of 0.95 reflects a good balance between accuracy and recall. This shows that the model not only records accuracy, but also the entire range of landslides in the dataset. To further evaluate the model, a comparative review with existing work in the field of landslide was performed, as shown in Figure 4.16. This review provided benchmarks to understand how our models stack up against other state-of-art models. Comparing the results across different models allowed to demonstrate the effectiveness and advantage of the approach in accurate prediction of landslides. The comparative analysis described in the results shows that inclusion of pyramid pooling layers and multi-scale feature integration with FCN significantly improves the capabilities and accuracy of the landslide model.

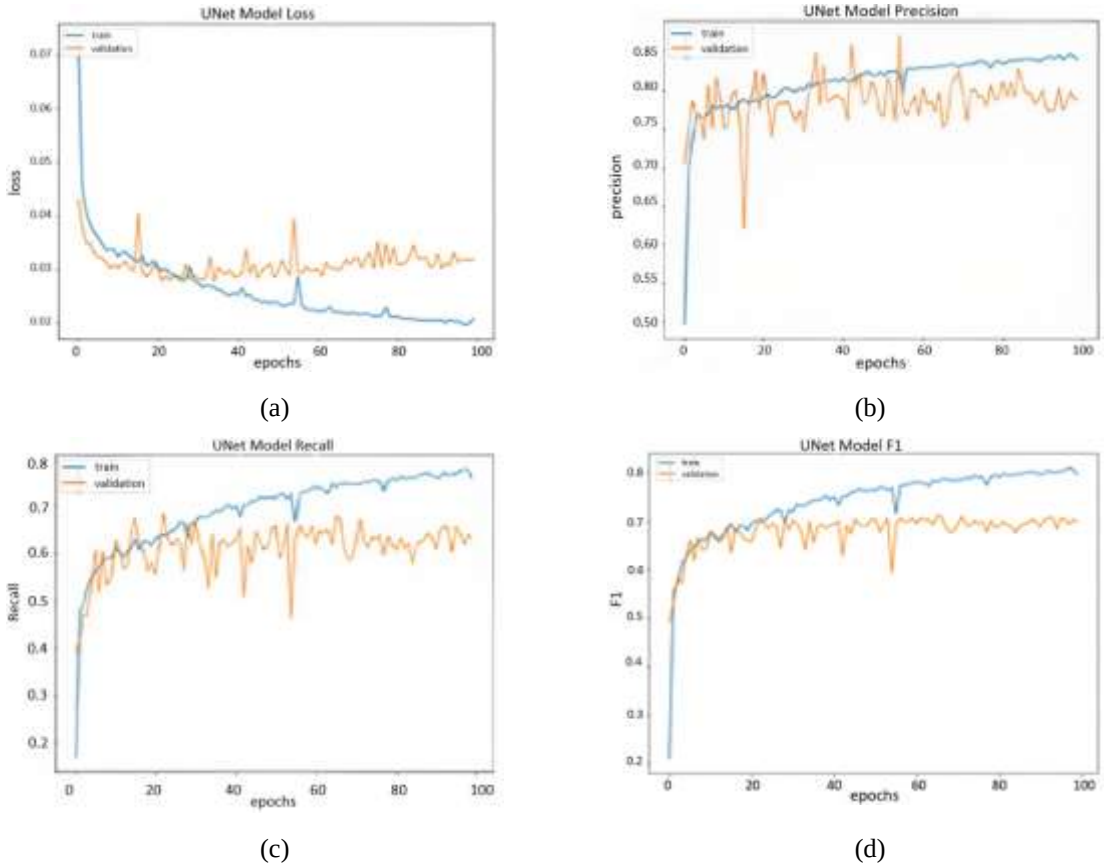


Figure 4.16 Performance Metrics for Landslide Prediction in Remote Sensing Imagery: (a) Loss, (b) Precision, (c) Recall, and (d) F1-Score.

This indicates that the combined method is more accurate than the basic UNet model in identifying landslides. Improved outcomes also suggest a possible decrease in false positives, or the number of cases in which non-landslide data is mistakenly identified as a landslide. Recall value has increased significantly as well.

Table 4.2 Comparison of Proposed Work with Previous Approaches

Author	**Model**	**Dataset**	**Performance parameter**
X Chen et al. [133]	CTD-Net	Landslide4sense	Precision of 0.75, Recall of 0.73, F1-Score of 0.74.
Ghorbanzadeh et al. [134]	ResU-NetOBIA	Landslide4sense	Precision of 0.73, Recall of 0.80, F1-Score of 0.76
Fahong Zhang et al [135]	Model 3 + CRF	Landslide4sense	Precision of 0.80, Recall of 0.78, F1-Score of 0.79.
Lin Bai et al. [136]	Multispectral U-Net	Landslide4sense	Precision of 0.80, Recall of 0.75, F1-Score of 0.77.
Proposed model	UNet-Pyramid	Landslide4sense	Precision of 0.91, Recall of 0.84, F1-Score of 0.87.

Table 4.2 presents a comparison of the proposed work with previous approaches in landslide prediction, highlighting various model's performance. The table lists the model authors that use the data records and their corresponding performance parameters like Recall, Precision and F1 scores. The models comparsion include CTD-Net, ResU-NetOBIA, Model 3 + CRF, Multispectral U-Net, and the proposed UNet-Pyramid. This shows that the proposed UNet-Pyramid model exceeds the other models, with a precision of 0.91, a recall of 0.84, and an F1 score of 0.87. Several recent studies have applied advanced deep learning architectures to landslide prediction using the Landslide4Sense dataset, each contributing novel approaches to improve segmentation accuracy. X. Chen et al. [133] developed CTD-Net, a convolutional neural network designed to extract hierarchical spatial features effectively. Their model achieved a precision of 0.75, recall of 0.73, and an F1-score of 0.74, indicating balanced performance in identifying landslide-prone areas. Ghorbanzadeh et al. [134] proposed ResU-NetOBIA, which integrates residual blocks within the U-Net architecture alongside Object-Based Image Analysis (OBIA) techniques. This method improved the model's ability to capture fine spatial details and yielded a precision of 0.73, recall of 0.80, and F1-score of 0.76, showing enhanced sensitivity in landslide detection. Fahong Zhang et al. [135] introduced Model 3 combined with Conditional Random Fields (CRF) to refine segmentation outputs and reduce misclassification. Their approach achieved higher metrics with a precision of 0.80, recall of 0.78, and F1-score of 0.79, demonstrating the benefits of post-processing techniques

in improving prediction accuracy. Lin Bai et al. [136] employed a Multispectral U-Net model, leveraging multispectral satellite imagery to extract richer spectral features relevant to landslide susceptibility. This model attained a precision of 0.80, recall of 0.75, and F1-score of 0.77, confirming the value of multispectral data integration for enhanced landslide mapping.

Building upon these methodologies, the present thesis proposes a UNet-Pyramid architecture that incorporates a multi-scale feature extraction pyramid within the U-Net framework. This design enhances the model's capacity to capture spatial features at varying resolutions, improving both localization and boundary delineation of landslide regions. Evaluated on the Landslide4Sense dataset, the proposed model achieved a substantially higher precision of 0.91, recall of 0.84, and F1-score of 0.87, outperforming all referenced approaches. This significant improvement illustrates the effectiveness of the UNet-Pyramid in capturing complex spatial patterns and reducing false positives and negatives in landslide prediction. The integration of multi-scale features enables better representation of both large and small landslide areas, which is critical for accurate hazard mapping and risk mitigation. These results demonstrate that the proposed model advances the state of the art in remote sensing-based landslide prediction and contributes a robust framework for practical disaster early warning systems.

This comparison highlights the effectiveness of the proposed model for landslides. The results of this approach point out the important advantages of combining OBIA with deep learning approach, particularly in the context of landslide prediction. OBIA and deep learning were typically examined individually with landslide detection. However, this study shows that integration of the two approaches can significantly improve the classification accuracy. By using OBIA's ability to obtain or extract spatial, structural, and context-related features in conjunction with the power of UNet, to model complex patterns of data, landslide models can be predicted more effectively in a variety of environments. As shown in the results, this comparison highlights the effectiveness of the hybrid approach. Our research is based on freely accessible satellite imagery, which not only creates landslide costs related to costs, but also expands the scope of monitoring to a larger scale. This opens the door to the broad and efficient use of satellite imagery in disaster management. Ultimately, this approach has great potential to improve the accuracy and locality of landslide prediction, especially in areas with high risk.

4.6 CONCLUSION

In conclusion, this chapter provides a new and effective approach to landslide prediction using landslide data, including a rich collection of remote sensing images. This study focuses on the use of object-based image analysis (OBIA) for image segmentation combined with the power of advanced deep learning techniques such as the swin transformer and the UNet-Pyramid model. OBIA is used to segment images with meaningful objects based on spectral and spatial characteristics, improving the accuracy of characteristic extractions from complex remote sensing data. The latest model for the computer vision, swin transformer is used to capture dependencies and spatial hierarchies over the long term. This improves the overall performance of distinctive presentations and predictive models with 0.91 - precision, 0.84- recall and 0.87 - F1 score, the proposed model exceeds models such as CTD-Net, ResU-NetOBIA, Model 3 + CRF, and Multispectral U-Net, all using Landslide4Sense dataset. These metrics highlight the strength of the UNet-Pyramid model in effective prediction of landslides. By using these state-of-art methods, this study contributes to the development of more accurate and robust landslides. Furthermore, the proposed model can efficiently process large amounts of remote sensing data, which has a significant impact on landslides in real-time.

CHAPTER 5

EXPLOITING THE SYNERGY OF SARIMA AND XGBOOST FOR SPATIOTEMPORAL EARTHQUAKE TIME SERIES FORECASTING

5.1 Introduction

Disastrous events like earthquakes can cause severe damage to both human life and infrastructure. Earthquakes occur when there is significant movement or disturbance in the Earth's crust, leading to large-scale consequences such as human casualties and difficulty in repairing or rebuilding damaged infrastructure [137]. Because earthquakes have the potential to destroy both lives and buildings, accurately predicting their occurrence has always been a major challenge. Early detection of seismic events is crucial for reducing associated risks, and it plays a key role in forecasting earthquakes [138]. Earthquake forecasting is typically divided into two categories: short-term forecasting and long-term forecasting, which is also known as multi-step forecasting [139]. Short-term forecasting focuses on predicting seismic activity in the near future, providing valuable insights into upcoming events. Long-term forecasting, on the other hand, extends into the future, offering predictions about potential seismic activity over a much longer period. By combining both short-term and long-term forecasting methods, researchers develop a comprehensive understanding of future seismic events, encompassing both immediate and distant perspectives. The question of whether earthquakes can be reliably predicted is a subject in scientific debate for enormous years. Different methods for earthquake prediction involve analyzing different factors, such as atmospheric weather, geodetic data, and physical elements, which often play a role in seismic events [140][141].

Despite large-scale studies and several attempts to predict earthquakes, reliable methods for determining the exact location, time, or size of earthquakes have not yet been developed [142]. A common approach to earthquake prediction is the use of time series data. This consists of measurements that are made regularly over a long period of time. These observations usually include seismic characteristics such as location, size and frequency. Time series is an analytical technique used to identify hidden patterns, trends, and cycles within this data that may indicate early signs of an earthquake. The recognition and interpretation of warning signals that appear

just before the earthquake is necessary for strategies to reduce or prevent its catastrophic effects [143]. However, only a limited number of studies were intended to predict several aspects of the earthquake, such as size, location, and epicenter. Time series analysis is often used in a variety of fields, including seismic and in vivo analysis [144]. Methods such as feature acquisition and machine learning are often applied to time series analysis [145]. Traditional methods of time series analysis typically involve extracting properties and parameters based on expertise and empirical data [146]. However, recent advances have focused on more advanced methods, such as DL and ML, which improve pattern recognition and distinctive extraction [147] [148]. In scientific research, researchers often encounter patterns that are missing in chronological order. This can lead to prediction or model inaccuracy if not treated [149] [150]. Ensuring the accuracy and completeness of the data is important to improve the efficiency of subsequent analyses, as the data can be lost for a variety of reasons like system failure, human failure, and everyday expectations. Time series(TS) aggregates are essential for identify trends or patterns in large datasets that are difficult to recognize and interpret. This is a very crucial role in seismic activity analysis and provides valuable knowledge about seismic behavior and samples. Investigating seismic behavior through methods such as similarity, self-organization, pattern recognition, and analysis of final scaling provides a deeper understanding of how seismic events are based on the development of predictive models. The data analysis (EDA) serves as a basic tool for understanding seismic data by offering access to exploring and interpreting the complexity of the earthquake activity. EDA is a method used to identify basic patterns and relationships within data and its importance increases in predicting an earthquake, where it helps scientists to navigate the complexity of seismic datasets. By using visualization tools, EDA can provide a clearer representation of complicated seismic data that helps better understand and reveal key knowledge. The integration of deep learning has a very advanced EDA, especially in the analysis of seismic disasters, including areas such as signal classification, image processing and object recognition. CNN contributed significantly, especially when the earthquake damage, where they are used to classify the image to evaluate buildings caused by earthquakes. Despite progress in the prediction systems, precise forecasts remain rare and complex nature of the earthquake, with numerous factors that are difficult to measure or evaluate. The most effective methods of earthquake prediction combine mathematical simulations, signal analysis, ML and DL. Some recent methods applied Bayesian networks to predict activity of earthquake by seismic data from individual stations, while others focus on predicting parameters such as epicenter position, size and depth. These models usually analyze seismic data within a specific radius of the station and integrate past seismic knowledge

as indirect model inputs through methods such as artificial neuron networks or random forests. However, these approaches often do not reach the integration of the physical principles of the behavior of the earthquake, neglect the impact of historical earthquake data and the complexity of seismic activity. Numerous studies have explored timeless time series techniques, such as the use of Arima and Singular Spectrum Analysis (SSA) to predict magnitude values of earthquake along specific failure lines. Other studies have incorporated different neural network models for capturing time relations in seismic data, or use methods such as RNN with LSTM cells to detect anomalies in seismic data before the earthquake. Although these methods have shown promising, challenges remain in the integration of physical principles and increasing the reliability of these models in the real-world applications.

5.2 Motivation

Although the previous studies have demonstrated the effectiveness of the short-term forcast prediction, there is still limited investigation on the applicability of these methods for long-term predictions, especially those that exceed the year. The forecast of the earthquake includes numerous challenges due to the complex and dynamic nature of the seismic activity. Various factors such as irregularities in seismic events, geological structuring effects, and environmental conditions contribute to the difficulty of accurately predicting earthquakes for longer periods of time. This approach uses a hybrid ML model that combines SARIMA with XGBoost. Here, SARIMA capture seasonal patterns and trends in time series data, and on the other hand, XGBoost helps in modeling complex relationships between variables with in the data. The combination of these two models makes the approach a more effective solution than traditional methods, since both intensities are used to improvise the accuracy of robustness. The performance of this hybrid model is evaluated and compared to traditional earthquake forecasts such as SARIMA and ARIMA. The results show that the hybrid model SARIMA-XGBoost exceeds these traditional models, highlighting the possibility of more reliable and more accurate predictions of seismic activity. SARIMA is affected by the natural cycle of seismic motion because it effectively records seasonal variations that are often observed in seismic data. XGBoost is characterized by the identification and modeling of nonlinear and complex patterns within seismic data. This combination of seasonal pattern detection and complex relationship modeling provides enhancement in earthquake series predictions.

Although the analysis of signal, deep learning algorithms, mathematical modeling and advanced machine learning techniques are considered reliable for earthquake predictions but are often categorized into individual combinations. The earthquake prediction usually benefits from a versatile approach and includes several different methods and models to improve accuracy. This shows the importance of evaluating several approaches and integrate different data processing methods, rather than relying solely on SARIMA or ARIMA models. A comparative assessment of SARIMA and XGBoost, along with other models such as ARIMA, can help determine the most effective method of earthquake prediction. For this study, we have compiled data on the average earthquake for 1965 to 2023 in designated regions, which serve as the basis for time series analysis and model development. The study is organized as follows: next section describes data preparation, presents an overview of Exploratory Data Analysis (EDA), and describes predictions created using time series models. Secondly, it presents the approach made in the investigation, focusing on comparing selection methods and different models. The third section describes the selection process of the best model, followed by a detailed analysis of the results of different predictive models. SARIMA is an established model used in temporary series predictions. Especially for data with adequate seasonal patterns to detect recurrence trends when an earthquake occurs. However, the ability to dominate complex but non -linear relationships is limited. XGBoost, a powerful algorithm for machine learning that derives a record of these non -linear and complex dependencies of the data that are often present in the seismic activity due to the interaction of several environmental factors. The combination of SARIMA's ability to model seasonal trends and XGBoost when solving patterns or complex relationships provides more accurate and precise earthquake predictions.

5.3 DATA AND METHODOLOGY

Figure 5.1 illustrate two main statistical approaches for long-term earthquakes prediction. These approaches use the historical earthquake dataset and created a hybrid model that can predict earthquakes for longer period. The earthquake prediction refers to predicting the magnitude and location of the future earthquake event. To reduce the risks related to earthquake it is important to make a reliable and accurate forecasting system, as they provide valuable information that can inform disasters, early warning systems and risk management strategies. This approach consists of three steps: Data pre-processing, Data cleaning, and Data series analysis. In Data pre-processing, data cleaning and converting data to suitable form is necessary. In this approach the dataset reading is covered from 1965-2023. This approach helps

in discovering temporary data relationships such as recurring seismic events and seasonal changes for earthquake forecasting.

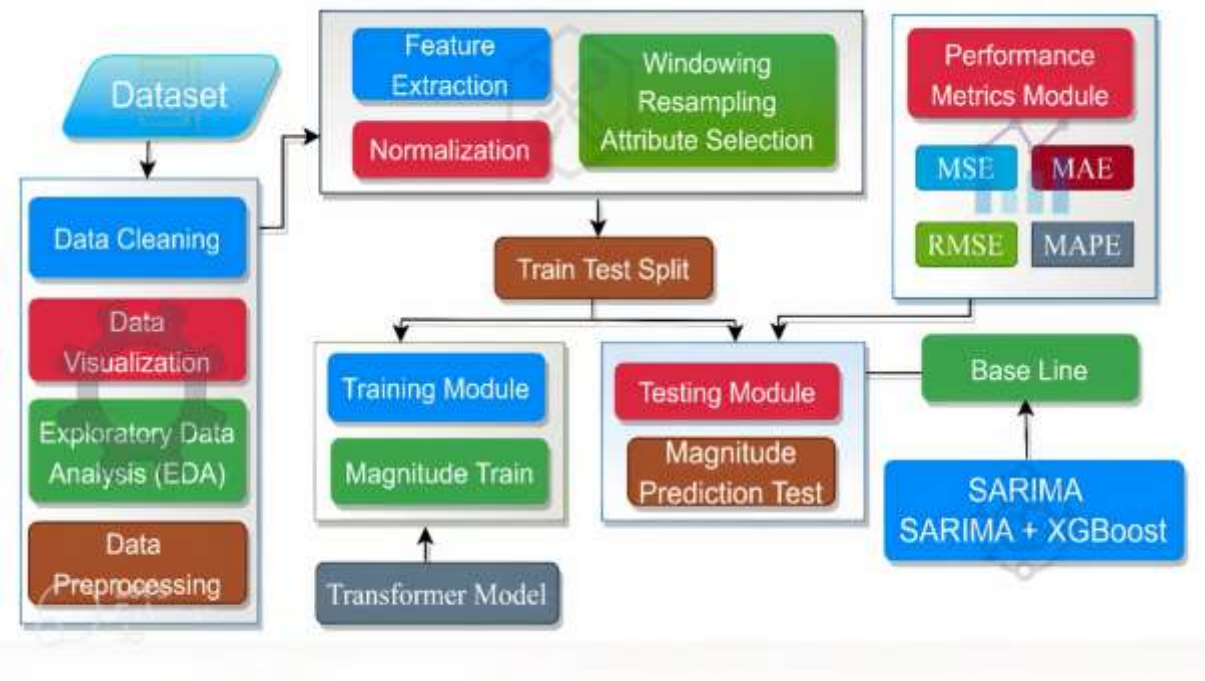


Figure 5.1 Sequential flowchart illustrating earthquake data processing and modelling steps.

In the TS analysis we use two phase methods to explore and predict the earthquake. The first phase which is the survey, involves the utilization of descriptive statistics and Data visualization approach to better understand characteristics of the dataset. Descriptive statistics provide summary data by evaluating using median, mean, frequency while data visualization methods such as thermal maps and time series contribute to visual identification of trends of seismic activity, anomalies and clusters. By visualizing earthquake data, scientists can observe time patterns and potential periodic behaviors that could help develop future predictive models. The second stage of time series analysis is predictions that include ML techniques like AR, SARIMA, and XGBoost to predict future earthquake events. The AR extension, SARIMA, corresponds to trend and seasonality in your data. A more advanced ML model, XGBoost can improve prediction accuracy by learning complex non-linear relationships and including different features in simple time trends. There are several geological factors regulating the behavior of tectonic plates some of them are the boundaries of the plates, subduction zones and failure lines, play a key role in the frequency and severity of seismic events. In addition, understanding the specificity of the time zone helps ensure that the earthquake prediction models are responsible for changes in seismic activities that may vary due to local geological

conditions from region to region. The integration of these factors is the aim of the study to provide more reliable and regionally relevant predictions, which contributes to improving the readiness for earthquake and risk management. Overall findings are expected to offer valuable knowledge of how to approach the analysis of earthquakes and predictions in a more subtle and scientifically informed way.

5.4 Dataset Description

This study uses a dataset that captures seismic events with the magnitude of 5.5 or higher, from 1965 to 2016, as shown visually in Figure 5.2. The selection of this threshold is significant because the earthquake with a magnitude less than 5.5 often releases relatively low amounts of seismic energy. These smaller earthquakes are more difficult to detect precisely, so their identification and reports are less consistently compared to the strongest earthquakes. As a result, these lower magnitude earthquakes are excluded from the dataset. The approach to the earthquake with the magnitude of 5.5 or more allows a more specific and significant analysis of the seismic activity that has the potential to cause substantial damage and affect larger populations. Seismic monitoring systems and interface prefer to detect and inform these events because they are more likely to lead to important social and economic consequences. In particular, the National Earthquake Information Center (NEIC) plays a decisive role in the global monitoring of seismic events and guarantees that basic details such as location, depth, size and source of these important events are registered and distributed. The study focused on the data provided by these established sources guarantees a reliable and comprehensive basis for the following analysis.

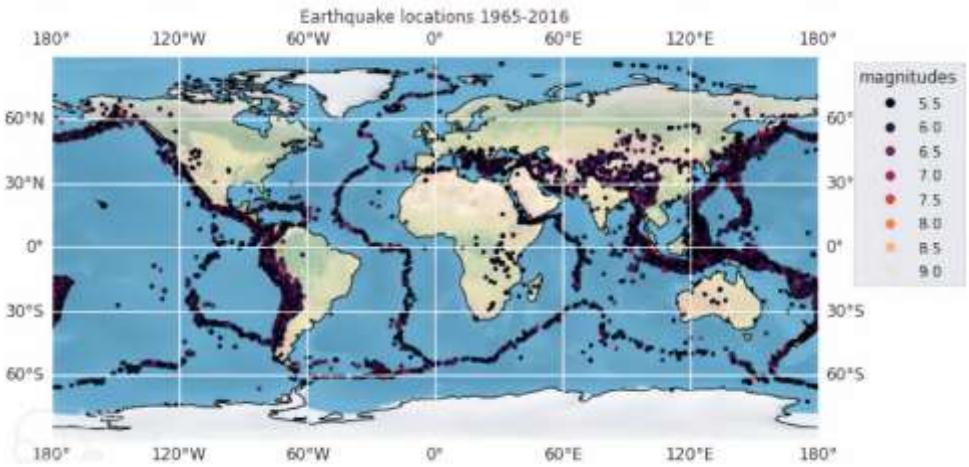


Figure. 5.2. Geographic distribution of earthquake epicentres: magnitude 5.5 or greater between 1965 and 2023.

To further increase the relevance and timeliness of the data set, research also includes data on the NEIC earthquake, covering the period from 2017 to 2023. Integration of this additional data is necessary for several reasons. First, it allows for more enhance and wider understanding of long-term trends and patterns of the earthquake by covering a large time range. Secondly, this enriches the ability to detect emerging formulas in recent seismic activity, which may not have been observable only through earlier data. Earthquake monitoring is constantly improving, and more data is available, so including the latest data will enhance analysis by providing a more recent perspective. An extended dataset containing both historical and current work data increases the robustness of time series analysis and provides a clearer and comprehensive view of the trends of earthquakes over various decades. This ensures that the conclusions of the study are not only based on historical data, but also reflect the latest seismic events that offer more accurate pattern of the behavior of global earthquake. The preliminary data processing phase is a critical step to ensure the quality and consistency of the data file before any analysis or modeling. This phase contains a number of steps to clean and organize data to eliminate any mistakes, inconsistency or spaces. One of the primary tasks at this stage is data cleaning that deals with problems such as missing values, duplicate items or incorrect data. The dataset considered contains information about 23,412 seismic events that contain a number of parameters such as date, time, geographical coordinates, depth, area, time stamp, size and other key details. It is important to make sure that each element is unique and represents a real seismic event. During the cleaning process, zero values representing missing or incomplete data were carefully identified. Furthermore, the duplicate elements have been removed that could distorts the analysis by over representation of certain events, making sure that only different events remain. Then the data was observed as earthquake do not occur at regular intervals and the original dataset does not have to provide a uniform distribution of time. To solve this problem, the data was aggregated to monthly intervals to create a consistent time series that could be analyzed over time. Interpolation was another key technique used in the preliminary processing phase. Because the data file was initially irregular, with different time intervals between seismic events, spaces or missing values, they appeared in the data file after the process of resampling. The interpolation is used to fill in these missing values by estimating the surrounding data and create a continuous time series.

The final dataset is completed and provides a constant record of the occurrence of earthquakes, including key parameters such as time, location, size and depth. Through the data per month and add values in the missing values through the interpolation, the dataset is suitable for the

analysis of time being and offered a reliable basis to understand the tendencies of the seismic activity over time. An important decision during the previous processing phase is to maintain a remote earthquake with a high increase to eliminate the dataset. While remote values can sometimes mess up the results of statistical analysis, large earthquakes are considered important events in a seismic activity. These events provide valuable knowledge to regions that are susceptible to extensive seismic activities and can help identify trends or patterns that can not only be visible in smaller events. Table 5.1 offers a comprehensive review of dataset properties, providing an overview of data structure and steps taken to prepare for further analysis. This overview includes a statistical summary of earthquake events, such as the size distribution, the frequency of events in time and the geographical spread of occurrence, providing a major insight into the nature of global seismic activity.

TABLE 5.1 Summary of Significant Earthquake dataset.

S.No.	Time	Latitude	Longitude	Depth (km)	Depth Error	Magnitude	Magnitude Type
0	1965-01-02	19.246	145.616	131.6	3.785	6.0	Mw
1	1965-01-04	1.863	127.362	80.0	4.678	5.8	Mw
2	1965-01-05	-20.579	-173.972	20.0	4.997	6.2	Mw
3	1965-01-09	-59.076	-23.567	15.0	2.452	5.8	Mw
…	…	…	…	…	…	…	…
23406	2023-02-17	-6.5986	132.0763	38.615	5.595	6.1	Mw
23407	2023-02-16	-15.0912	167.0294	36.029	6.080	5.6	Mw
23408	2023-02-15	12.3238	123.8662	20.088	4.399	6.1	Mw
23409	2023-02-15	-40.5485	174.5709	74.320	4.922	5.7	Mw
23410	2023-02-14	45.1126	23.1781	10.000	1.794	5.6	Mw

5.5 Exploratory Data Analysis (EDA)

EDA plays a key contribution in attenuation of the complexity of seismic sequence data and is an essential tool to predict the earthquake [150][151]. EDA is a basic process that includes the use of several visualization techniques to explore data properties, detect any remote or anomalies, and evaluate the validity of the basic assumptions. EDA is important because it provides the deeper understanding of the dataset and clears the basics for other analytical approaches, including statistical modeling and machine learning techniques. In relation to earthquake data, which inherently include time and geographical dimensions, EDA is even more valuable. The unique nature of seismic events, which occur in specific times and places, requires a deep and accurate examination to understand potential patterns, trends and risks.

Therefore, EDA not only prepares the ground for more sophisticated analyses, but also increases the general understanding of earthquake behavior.

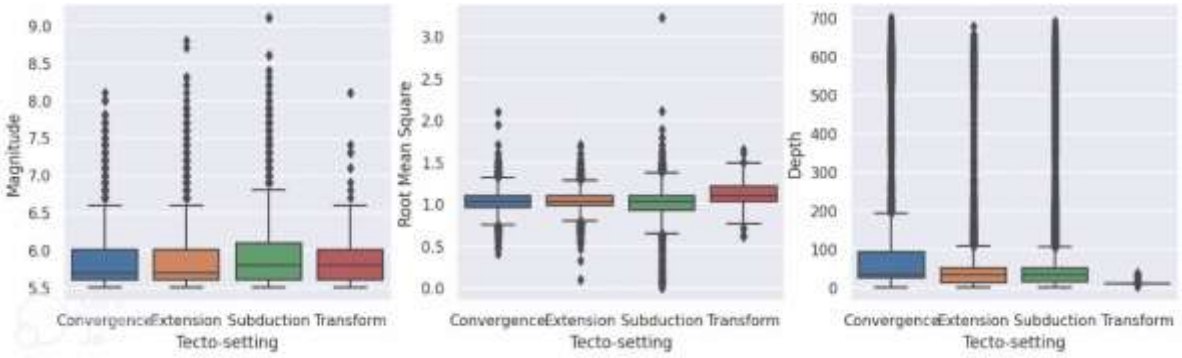


Figure. 5.3. Box plots depicting the distribution of earthquake magnitude, root mean square and depth categorised by tectonic setting: convergence, extension, subduction and transform.

The techniques used in EDA are diverse and adapted to reveal different aspects of the data. Basic visualization methods, such as histograms and bar charts, are often used to understand the distribution of earthquakes, frequencies and other relevant variables. These tools are useful for pattern detection and anomalies identification within data. The most advanced statistical graphics, such as mean plot and boxplot, provide information about data points distribution, emphasizes the presence of remote values and offers a deeper understanding of data visualization. In general, line plots are insightful; tools to visualize trends as changes in seismic activity capture over time and show how earthquakes change daily, monthly or annually. In addition, the technique of map representation is invaluable to offer a geographical vision of the data. When visualizing earthquakes on the map, scientists can better understand regional seismic patterns, help identify areas with high risk and any geographical tendency that may occur. EDA is not only a data exploration tool, but also a key step towards the development of effective earthquake forecasts or prediction. EDA helps to carry out future analytical processes and provide valuable knowledge that can inform predictive modeling by identifying seasonality, hidden patterns and anomalies within the data. The relationship between earthquakes and their tectonic environment is a key research area in EDA because it provides critical information for risk assessment and disaster preparation. One of the basic aspects of the EDA is to explore the tectonic environment and its relationship with seismic activity. The different tectonic environments, such as subduction areas, failure lines and boards, show different levels of seismic activities. The recognition of these connections is essential to understand where the earthquake is more likely to occur and evaluate the risk in these regions. By identifying the tectonic configuration, which are more susceptible to intensive seismic

events, EDA helps determine areas of greater seismic risk. This methodology significantly improves patterns recognition and enables detection of time trends and seasonality in the occurrence of earthquakes, thus contributing to the enhancement of high-precision predictive models. The classification of tectonic settings assigns each earthquake event to a different tectonic environment, such as subduction zones, extension zones, transformation errors, cracks and convergence zones using geographical and geological data. As shown in Figure 5.3, the box plot offers a clear and brief graphical representation and compares the statistical distribution of the magnitude of the earthquake, the root mean square (RMS) and depths across different tectonic settings, including extension, subduction, convergence and transformation zones. This approach allows an effective and intuitive comparison of seismic features in various geological contexts, which makes it easier to identify trends, patterns and anomalies in the behaviour of earthquakes.

5.6 Time Series (TS)

Time series is classified into three types: deterministic trend, stochastic trend, and seasonal trend. The behavior understanding is important for developing more reliable models to predict the upcoming values. TS does not display intrinsic behavior and lead to unreliable predictions. To determine stationary time, its properties, such as variance and mean should remain constant always. The ADF test is applied to check the stationarity in time series by detecting the presence of a unit root in the series. To make series stationary transformations like differencing are applied if the series is found to be non-stationary. Further, by understanding the time series three trends, appropriate transformations and modeling techniques are applied to improve the accuracy. The combination of these trends occurs in several series where deterministic patterns are affected by stochastic noise, requiring a more complex model in which both components are responsible. Understanding these categories is to choose the most appropriate modeling approach and ensure that the resulting predictions are accurate and reliable

5.6.1 Deterministic Trend Time Series

In this model it is supposed that the TS variable (v_t) is a time function $v_t = g(t)$. This represents various forms including linear, exponential or other types of assembly problems. It provides to bring the relation among TS variable and time using different mathematical functions, depending on the the basic samples that we strive to capture and nature of the data.

$$v_t = a_0 + b_t + c_t \quad (5.1)$$

After calculating the Δv_t, we have:

$$v_t - v_{t-1} = (a_0 + b_t + c_t) - a_0 + b_{t-1} + c_{t-1}) \quad (5.2)$$

$$\Delta v_t = b + c_t - c_{t-1} \quad (5.3)$$

$$U[V_t] = U[Y] + U[c_t] + U[c_{t-1}],\ U[c_t] = U[c_{t-1}] = 0 \quad (5.4)$$

Thus, $U[V_t] = Y$

v_t is time series variable at time *t*, a_0 is constant intercept b_t is coefficient for time t, capturing trend c_t is other factors influencing the time series at time *t*. Δv_t is first difference of time series variable v_t, $U[c_t]$ and $U[c_{t-1}]$ are unconsidered components of predictors c_t and c_{t-1}. This implies that the mean of the corresponding time series Δv_t, remains constant over time, making Δv_t a stationary time series. Additionally, a time series with a linear trend in equation (1) is referred to as a trend stationary time series.

5.6.2 Stochastic Trend Time Series

In this type of TS modeling, it is supposed that v_t is a function of the lagged v_t. The straightforward case is termed first order, specifically the AutoRegressive model where v_t depends only on the first lag (k=1):

$$v_t = a_0 + q_{vt-1} + c_t \quad (5.5)$$

According to the Dickey-Fuller test (DFT), it is established that the time series in equation (4) is stationary as long as |q|<1. To delve deeper, let's examine the case when q=1, in which we obtain:

$$v_t = a_0 + b_{t-1} + c_t \quad (5.6)$$

$$v_{t-1} = a_0 + b_{t-1} + c_{t-1} \quad (5.7)$$

$$v_0 = v_0$$

$$v_t = ta_0 + v_0 + \sum_{j=1}^{t} c_t \quad (5.8)$$

The random walk is an unpredictable process, meaning that there is no obvious pattern in the data to help with prediction. The random walk is expressed as:

$$v_t = v_{t-1} + c_t \quad (5.9)$$

Therefore, the random variable must be a non-stationary TS. Although the random variable's mean value is constant.

5.7 Augmented Dickey-Fuller test (ADF)

A statistical technique called the Augmented Dicky-Fuller test (ADF) is used to assess whether a data record is stationary or not. Since many modeling approaches, such as ARIMA, need a series to show stable statistical features over time, the stationary is a decisive prerequisite for trusted time series modeling. Zero hypothesis (H_0) claims the series is not stationary and the ADF test determines whether there is a unit root or not. The null hypothesis (H_1) is tested and presents that unit root is present in the series and this unit root indicates that this time series is non-stationary. This indicates that it has time dependent structures such as trends and seasonality and needs to addressed for reliable and accurate modeling. When test statistics is lower than key values the null hypothesis is accepted else it is rejected. This is because the TS is stationary.

The general form of the ADF test is as follows:

$$\Delta v_t = a_0 + \mu v_{t-1} + \theta_1 \Delta v_{t-1} + \cdots + \theta_p \Delta v_{t-p} + c_t \quad (5.10)$$

Here, μ: unit root parameter, and $\mu = 0$ represent that the model has a unit root, and the time series is not stationary. a_0 *and* μ both are the model's trend.

5.8 Forecast Model Construction

In order to choose SARIMA, it begins with the use of EDA. and time series (TS). EDA identifies the seasonal and self-reproductive components and key information to select SARIMA's parameters. Finding the ideal values for the p, d, and q all of which are important to the SARIMA model. The autoregressive order (p) presents the number of prior observations utilized to estimate the current value. The moving average order (q) determined

the amount of prediction mistakes trailing the model, which takes the impact of previous failures on future values. The differencing order (d) specifies how frequently the data must be differenced in order to achieve stationarity. As these parameters impact the model's capacity to capture both seasonal changes and long-term trends in seismic data, their selection is important. A well-tuned SARIMA model becomes an effective tool for earthquake prediction. it accurately describes the seasonal and temporal dynamics of seismic activity.

5.9 SARIMA

In 1970's Jenkins and Box proposed the ARIMA model, which became known as the Box-Jenkins approach [152]. It is also referred as (p,d,q) model, where the moving average order (q), difference order (d), and autoregressive order (p) are represented. It is a stochastic model that is effectively used for earthquake prediction, as well as effective tools for time series analysis that strive to anticipate values in non-stationary univariate time series and define autocorrelations in the data.

$$\Phi_q(C)\nabla^e Y_r = \Theta_q(C)\varepsilon_r \tag{5.11}$$

Y_r represents the error series and noise with a mean of 0.

$$\nabla^e = (1 - C)^e \tag{5.12}$$

It demonstrates the difference operator of order C, e used to denote integration order to achieve stagnation in data.

$$\Phi_q(C) = 1 - \varphi_1 C^1 - \varphi_2 C^2 - \cdots - \varphi_q C^q \tag{5.13}$$

is equivalent to the AR term at the qth level and

$$\Phi_q(C) = 1 - \varphi_1 C^1 - \varphi_2 C^2 - \cdots - \varphi_p B^p \tag{5.14}$$

is the p^{th} order MA polynomial.

Seasonal Box-Jenkins models are also extended version of ARIMA models that facilitate the direct modeling of seasonal components in non-stationary time series data, which exhibit both trend fluctuations and seasonal variations.

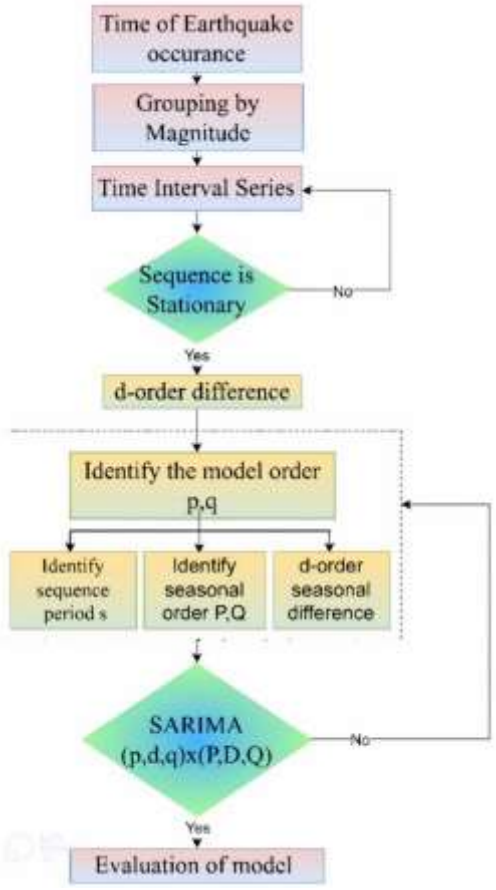


Figure. 5.4 Flowchart of SARIMA model for time series earthquake forecasting

General ARIMA (p,d,q) is summarized as,

$$y_t = c + \beta_1 y_{t-1} + \beta_2 y_{t-2} + \beta_3 y_{t-3} + \cdots + \beta_p y_{t-p} \epsilon_r + \alpha_1 \varepsilon_{r-1} + \alpha_2 \varepsilon_{r-2} + \alpha_3 \varepsilon_{r-3} \cdots + \alpha_q \varepsilon_{r-q} \quad (5.15)$$

$$\varepsilon_r = \alpha_1 \varepsilon_{r-1} + \alpha_2 \varepsilon_{r-2} + \alpha_3 \varepsilon_{r-3} \cdots + \alpha_q \varepsilon_{r-q} \quad (5.16)$$

These models are particularly useful for time series that show complex interactions between long -term trends and seasonal effects. The SARIMA is represented by (p,d,q) * (P, D, Q), where the parameters capture not seasonal and seasonal data behavior. SARIMA models use seasonal differences to solve seasonal fluctuations in the time series, similar to how the differences are applied in standard ARIMA models. Seasonal differences help remove periodic components of the series by deducting the value of observations from its corresponding value in the previous season. The process is an analogous differential trend in the ARIMA models, where the p, d and q parameters represent the autoregressive order, the differential order and the average order. In SARIMA, seasonal components are modeled by other parameters P, D

and Q components and represent seasonal autoregression order, seasonal differentiation order and moving average order. The seasonal period, which indicates the length of the seasonal cycle, is also an important part of the model. SARIMA refining usually involves fine-tuning of the orders of these parameters through methods such as experiment and error or automated technology such as grid search. The aim of this process is to identify the optimal combination of parameter values that best capture the basic formulas in the time series data, which provides the most accurate model suitable for predicting earthquakes. By modifying these parameters, the model can be adapted to consider both the long-term trend and for seasonal fluctuations in the earthquake prone area, allowing more precise predictions and deeper insight into seismic behavior.

Three seasonal factors that are not included in ARIMA need to be adjusted:

P: stands for seasonal autoregressive

D: for seasonal difference

Q: for seasonal moving average.

$$\phi_p(R)\Phi_p(R^s)(1-R)^d(1-R^s)^D Y_r = \theta_q(L)\Theta_Q(L^s)\varepsilon_t \quad (5.17)$$

$$\phi_p(R) = 1 - \phi_1 R - \phi_2 R^2 - \cdots - \phi_P R^p \quad (5.18)$$

$$\theta_q(R) = 1 - \theta_1 R - \theta_2 R^2 - \cdots - \theta_q R^q \quad (5.19)$$

$$\Phi_P(R^S) = 1 - \Phi_S(R^S) - \Phi_{2S}(R^{2S}) - \cdots - \Phi_{PS}(R^{PS}) \quad (5.20)$$

$$\Theta_Q(R^s) = 1 - \Theta_S R^s - \Theta_{2s} R^{2s} - \cdots - \Theta_{QS} R^{QS} \quad (5.21)$$

Where, R (the lag operator) is defined by $Y_r$and εt.

$$R^k Y_r = Y_{r-k}; \phi_p(p = 1,2,\dots,p), \Phi_p(P = 1,2,\dots,P), \theta_q(q = 1,2,\dots,q), \Theta_q(Q = 1,2,\dots,Q) \quad (5.22)$$

The SARIMA model expression for the time series Y_r is as follows:

$$\nabla^d \nabla_s^D y_r = \frac{\theta(A)\theta_s(A)}{\emptyset(A)\emptyset_s(A)}\varepsilon_r \quad (5.23)$$

here y_r is the error term at time r; $\emptyset_s(A)$ and $\theta_s(A)$ stand for the seasonal moving average coefficient polynomial of Q-order and the seasonal autoregressive coefficient polynomial of P-order, respectively.

Once the model orders and data transformations are established, the next part is for estimating the SARIMA model, as described in Figure 5.4. A widely used method are used for estimating the parameters of the SARIMA model is the estimation of maximum probability. The 'p' and 'q' parameters correspond to the number of average and moving terms, respectively, and are associated with non-seasonal components of the temporal series [153].

5.10 XGBoost

The decision-tree based ensemble machine learning model known as Extreme Gradient Boosting (XGBoost), developed by Chen and Guestrin in 2016, is a scalable boosting system that uses the gradient boosting algorithm to generate results with a low probability of an overfit [154]. Its strategy is to continually add and train new trees to fit the remaining mistakes from the previous iteration.

$$\hat{y}_i = \sum_{k=1}^{k} g_k(y_i), \qquad g_k = G \tag{5.24}$$

Here, g_k represents function in function space, k represents tree of decision tree and G is function space with equation:

$$G = \{g(x) = v_{r(x)}\} \tag{5.25}$$

Here, r(x) signifies that the sample x has been assigned to a leaf node, and v represents the leaf node weight.

In order to accelerate the model's convergence, XGBoost expands the loss function using the second-order of the Taylor's series expansion [155].

XGBoost's regularized objective function is represented as:

$$\mathcal{L} = \sum_i l(\hat{x}_i, x_i) + \sum_k \Omega(g_m) \tag{5.26}$$

$$\Omega(g) = \gamma T + \frac{1}{2}\lambda \parallel \omega \parallel^2 \tag{5.27}$$

In this context, with x_i as the target, m denoting the number of classification regression tree, Ω indicating a regular penalty function, g_m is the model of the m^{th} tree, l is a differentiable loss function, and $\hat{y}_i$ as the predicted values, a balance is preserved in order to keep the decision tree model from becoming excessively complex. The penalty term formula uses penalty coefficients γ and λ. T represents the number of leaves, while ω represents the total leaf weights. In contrast to decision trees, regression trees apply weights to each leaf.

To minimize the objective function, continues iterations are performed and the objective function is evaluated after each iteration.

$$\mathcal{L}^{(m)} = \sum_{i=1}^{n} l\left(x_i, \hat{x}_i^{(m-1)} + g_m(y_i)\right) + \Omega(g_m) \tag{5.28}$$

The fast way to optimize second order function is using Second order expansion:

$$\mathcal{L}^{(m)} \simeq \sum_{i=1}^{n} \left[l\left(x_i, \hat{x}^{(m-1)}\right) + f_i g_m(y_j) + \frac{1}{2} k_i g_m^2(y_i)\right] + \Omega(g_m) \tag{5.29}$$

The loss function's first and second derivatives are denoted as f_i and k_i, respectively. The objective function is:

$$\tilde{\mathcal{L}}^{(t)} = \sum_{j=1}^{T} \left[\left(\sum_{i \in I_j} f_i\right) u_j + \frac{1}{2}(k_i + \lambda) u_j^2\right] + \gamma T \tag{5.30}$$

where u_j is the weight of leaf j. The weight of each leaf node is reduced to obtain the objective function's minimal value and make its derivative equal to zero. We can calculate $u_j$for a fixed structure as follows:

$$u_j = -\frac{\sum_{i \in I_j} f_i}{\sum_{i \in I_j} k_i + \lambda} \tag{5.31}$$

and determine the equivalent value by:

$$\tilde{\mathcal{L}}^{t}(p) = -\frac{1}{2} \sum_{j=1}^{T} \frac{\left(\sum_{i \in I_j} f_i\right)^2}{\sum_{i \in I_j} k_i + \lambda} + \gamma T \tag{5.32}$$

As it is sometimes difficult to keep track of all potential tree structures, XGBoost takes a greedy approach, gradually adding branches to the tree, starting with a single leaf. The presented

formula acts as a score mechanism for assessing the tree structure. Consider the sets of instances of the left and right nodes following the split to express the loss reduction:

$$\mathcal{L}_{\text{split}} = \frac{1}{2}\left[\frac{\left(\sum_{i\in I_L} f_i\right)^2}{\sum_{i\in I_L} k_i+\lambda} + \frac{\left(\sum_{i\in I_R} f_i\right)^2}{\sum_{i\in I_R} k_i+\lambda} + \frac{(\sum_{i\in I} f_i)^2}{\sum_{i\in I} k_i+\lambda}\right] - \gamma \tag{5.33}$$

The XGBoost is effective in collecting the complexity of seismic data, making it an optimized choice for modeling formulas related to the earthquake [156]. The algorithm works gradually and creates a number of decision trees, and each tree is constructed to correct the mistakes that remained from the previous tree. These decisions production of trees is designed in an adaptive and iteration manner, each with a previous tree focusing on minimizing the irregularities between the actual values and the expected results from the previous trees. This process allows XGBoost to gradually specify its predictions and increase the accuracy of the model over time. The strength of the XGBoost core lies in its access to a set that combines the outputs of multiple decision -making trees, each individually weak - into a single robust and accurate model. By using the collective predictive power of these XGBoost trees, it alleviates the limitation of any single model. The algorithm also includes techniques such as L1 and L2 regularization on the weight of the leaves to prevent excessive connection. Regularization penalizes too complex models, promotes simplicity and generalization. This is important for modeling seismic data where there is high risk of excessive noise and irrelevant pattern is high. XGBoost optimizes an objective function consisting of two main components, which are the losses that measure prediction errors and conditions that reduce model complexity. The optimization process adapts model parameters to achieve the best balance between accurate prediction and model simplicity. This is suitable not only for training data but also for invisible data. mechanism stops the training process. XGBoost creates predictions by aggregation all the trees in the data. Further, connecting with seismic analysis, this access to the data allows the model to effectively record complex data relationships and ensure more accurate predictions.

5.11 Integrating SARIMA with XGBoost

The combination of SARIMA and XGBoost in the seismic analysis of the temporal series offers enhanced and successful approach that significantly increases the precision of the earthquake prediction as shown in Figure 5.5. SARIMA's series is widely known for its ability to capture self-representative and temporal dependencies and seasonal trends that are present in earthquake data. The forecasts are produced firstly by SARIMA model, which captures the

dynamics of the basic time of seismic activity. However, earthquake data are often characterized by complex and non-linear behavior that SARIMA cannot be completely captured. This complexity is the result of a versatile nature of seismic events that may include several interaction factors, such as geological conditions, failure lines and environmental variables. As a result, there remains differences between the actual and predicted values, which represent inexplicable variations or complex patterns that do not address the initial model. To solve this issue, XGBoost, a powerful automatic learning algorithm, is known for its ability to capture non -linear complex relationships and complex data patterns is used. By training XGBoost in the residual of the SARIMA model, the algorithm learns to identify other factors that affect the seismic activity, to improve the original predictions and improve accuracy. The capacity of the model to process residual allows to adapt to nonlinear data, as sudden changes in seismic activities that may not follow a clear linear trend. During this phase, advanced techniques such as cross validation are used to guarantee the optimal parameterization of XGBoost model as it is useful to avoid excessive expulsion and ensure that the model is well widespread to invisible data. The model is trained in several subset data and its verification against other help to the cross validation of the hyperparameters of fine tune, such as the learning rate, the depth of the tree and the number of epochs that are essential for the performance of the model.

The integration of SARIMA and XGBoost creates a robust and adaptive model that can handle linear and nonlinear aspects of seismic data. The SARIMA model shows time and seasonal structures, while XGBoost deals with non-linear complexities and patterns that come from data. This hybrid model allows a more accurate and integral frame to predict the earthquake. In addition, this approach improvises the model's ability to identify all seismic trends, patterns and anomalies. For example, SARIMA helps in capturing seasonal trends and cyclical patterns in the time series data and XGBoost helps in modeling complex, non-linear relationships between variables to improve the accuracy of earthquake prediction. This approach significantly increases the ability of forecasts and provides more reliable earthquake predictions. This integrated approach is particularly valuable in the broadest context of seismic risk assessment, and it helps with better preparations. In addition, this hybrid model is used in several domains, such as timely warning systems, disaster planning and seismic behavior at logical and global scale. SARIMA excels in capturing basic time addictions, seasonal patterns and the author's dynamics by the earthquake data, while the XGBoost machine learning capabilities allow him to identify and improve non -linear patterns that SARIMA may miss.

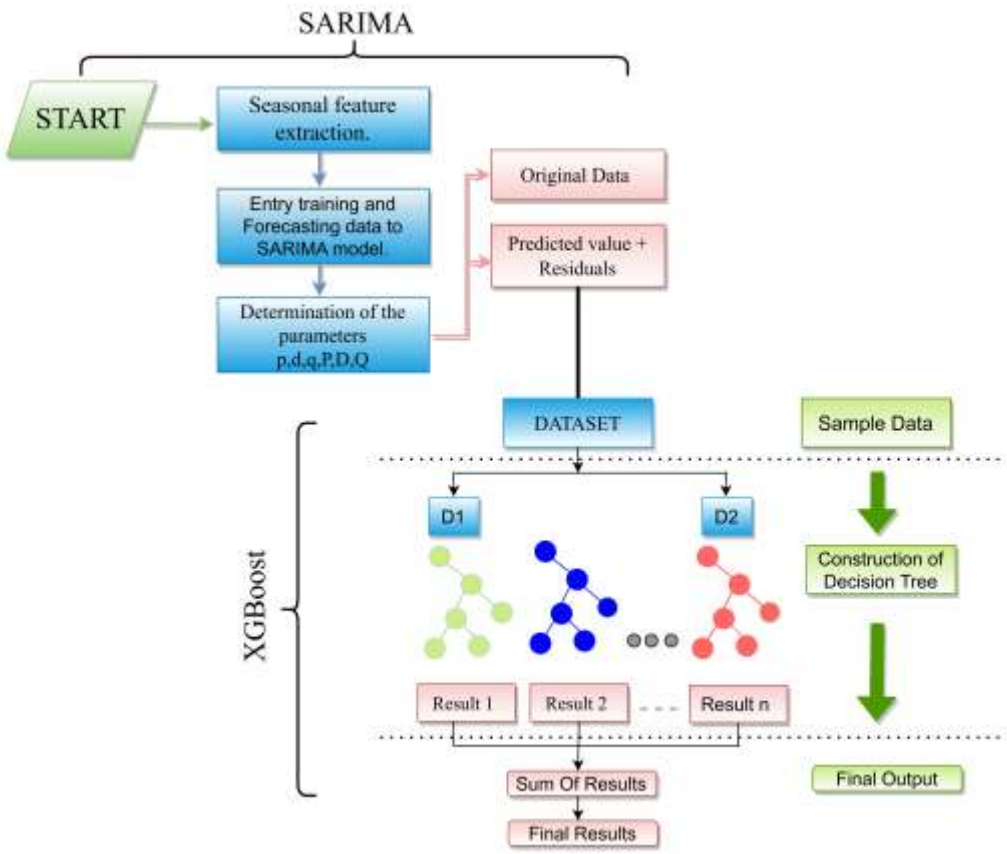


Figure. 5.5. Work Flow of proposed mode SARIMA–XGBoost for earthquake time series forecasting

This integration provides a more comprehensive insights of the behavior of the earthquake, allowing modeling of both linear and non -linear effects. The best parameters in features extracted from SARIMA, such as seasonal trends and autocorrelation, combined with the adaptive nature of XGBoost based on data, creates a robust prognosis model that not only increases the accuracy of prediction but also increases its ability to generalize to incredible data. Capturing the complex dynamics of seismic activity This methodology increases the reliability of the earthquake predictions, allowing more precise predictions and better readiness strategies. In regions prone to earthquakes, the benefits of this hybrid model exceed more than improved prognosis. It creates the creators of the decision -making tools necessary for proactive risk management and the allocation of resources. The ability to predict seismic events with greater accuracy in disasters' readiness systems and timely warning systems, which potentially reduces life loss and property damage. In addition, the adaptability of the model makes it possible to develop with changing seismic patterns and ensure its importance in a dynamic and complex geophysical environment. This advanced data-based approach offers a powerful tool for increasing safety and resistance in vulnerable regions.

5.12 RESULTS

In this chapter three models (SARIMA, ARIMA and hybrid SARIMA-XGBoost) were configured and simulated to evaluate their performance in the prediction of the earthquake. The process began with comprehensive data cleaning which is preliminary pre-processing and exploratory data analysis (EDA), this step makes sure that the dataset is prepared for accurate modeling. Further, all three models were configured and simulated to capture different aspects of seismic data, including time patterns, trends and nonlinearity in data. The ARIMA and SARIMA models focus on the analysis and forecasting of the time series based on the autoregressive engineering, as well as seasonal changes in the case of SARIMA. SARIMA-XGBoost, the integration of the SARIMA algorithm and the XGBoost algorithm, combines the strengths of both and captures both linear and non-linear trends for more refined predictions.

The metrics used to evaluate the performance of these model are Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) and Mean Squared Error (MSE). These metrics measure error between the real values and the expected values and evaluate the precision of the predictive model. To improve the precision of the model and minimize the error rate, carefully divide the data set into training sets and test and ensure robust access to the model verification and performance evaluation. This collected data discovered that approximately 70 observations were captured in this time frame, which represents approximately 80% of all series occurrences in the series. This is designed for training purposes, which allows the model to learn all the patterns and trends of historical seismic data. The remaining 20% of the data consists of earthquake events from 2001 to 2023 and are reserved for the test and validation of the model. This set of tests is carefully selected to include newer earthquake events, which guarantees that the ability to generalize the future data model and consider any change or trend that may have occurred. This distribution of training tests is decisive to reduce the risk of excessive amounts, since it guarantees that the model is tested for data that have not been included in the training process.

In addition, it helps simulate scenarios in the real world, where the future events of the earthquake based on historical data need to be predicted. By providing the historical dataset for training and testing, we are able to access its better predictions, verify its predictive power and ensure that it remains relevant and effective in predicting seismic events in the coming years.

Using the method of seasonal decomposition of time series, we break down many seismic data into basic components: trends, seasonality, residual or errors which are often visualized via line-chart. These patterns provide important information about immediate changes in seismic areas. Understanding changes in seismic frequency over very short periods requires recognition of positive patterns that show an increase in seismic events or negative trends indicating errors. Regression analysis and trendlines are used to further quantify these patterns. This supports the strength and importance of observed seismic activity changes. Monthly data analysis allows long changes in seismic samples to be recorded. To evaluate statistics over time, we use the ADF an extensive statistical test. The results listed in Table 2 provide important key point about the basic characteristics of seismic data. ADF tests are important for determining whether time series is stationary. This means that its statistical properties are consistent over time. In this case, the time series is subjected to ADF tests, and a p-value for results above 5% indicates that the null hypothesis of non-stationary material cannot be rejected. This indicates that the time series does not show steady-state behavior. This means that their statistical properties differ over time.

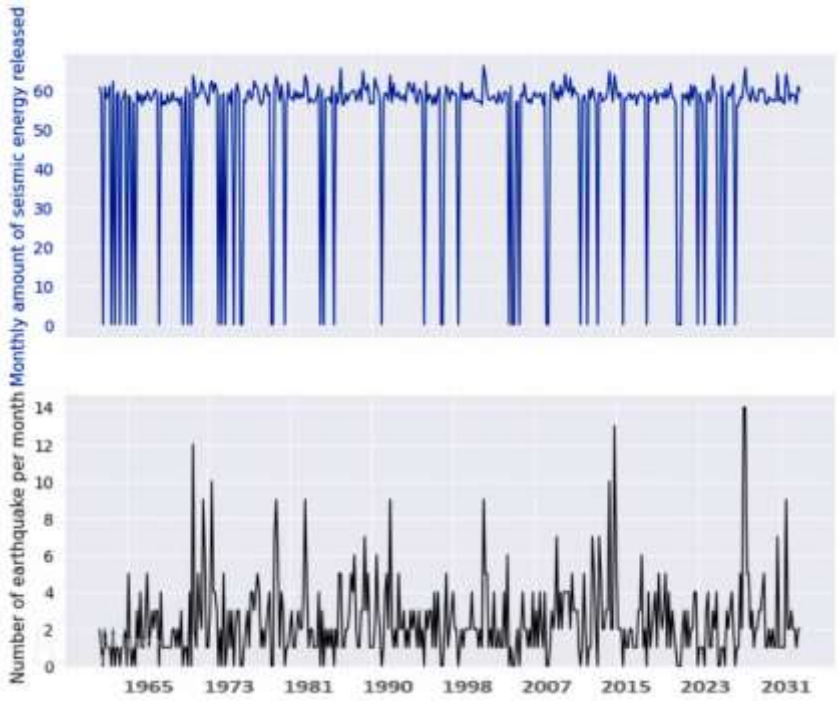


Figure. 5.6. Monthly distribution of earthquakes and seismic energy release over time

Once the data is cleaned, it is divided into training and testing sets, usually with 80% of the data assigned to the model training and the remaining 20% reserved for their performance testing. This division allows the model to learn from historical data and then be evaluated on

invisible data to assess its generalization. After this its focuses on the application of the SARIMA model. First, EDA is performed using the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots.

ALGORITHM 5.1: EARTHQUAKE TIME SERIES FORECASTING USING SARIMA AND XGBOOST

```
#Data Preprocessing
Earthquake_data=load_data('earthquake.csv')                    #Load earthquake time series data
Cleaned_data=preprocess_data (Earthquake_data)    #Preprocess data (Handle missing values, outliers)
train_data, test_data =split_data(Cleaned_data,train_size=0.8)    #Split data into training and testing sets
    #SARIMA Modeling
    acf_plot, pacf_plot = plot_acf_pacf (train_data)                    #Conduct EDA (plot ACF, PACF)
    p,d,q = Select_sarima_parameters(acf_plot, pacf_plot)          #Select SARIMA parameters (p,d,q)
    sarima_model=fit.sarima(train_data, p,d,q)                     #Fit SARIMA model to training data
    sarima_forecasts = sarima_model.forecast(test_data.shape[0])                    #Generate forecast
    #Feature Engineering
    sarima_residuals=calculate_residuals(train_data, sarima_forecasts) #Extract additional features from
    SARIMA residuals
    lagged_values= extract_lagged_values(sarima_residuals)
    moving_averages=calculate_moving_averages(sarima_residuals)
    feature_matrix=combine_features(train_data, lagged_values,moving_averages)
    #XGBOOST Modeling
    xgb_model=initialize_xgboost(n_estimators=100, max_depth=3, learning_rate=0.1, reg_lambda=1)
    #Initialize XGBoost Regressor with hyperparameters
    xgb_model.fit(feature_matrix, train_labels)                #Train XGBoost model on feature matrix
    xgb_forecasts=xgb.model.predict(test_features)                  #Generate forecasts for testing set
        #Prediction and Evaluation
        Combined_forecasts=combine_forecasts(sarima_forecasts, xgb_forecasts) #Combine SARIMA and
        XGBoost forecasts
        #Evaluate model performance using error metrics (MAE, RMSE)
        mae = calculate_mae(test_labels,combined_forecasts)
        rmse= calculate_rmse(test_labels,combined_forecasts)
```

These plots are essential for understanding time dependence in the time series data. The ACF plot identifies the relationship between the time series and delays in time intervals. The peaks in the ACF graph reveal possible seasonal trends or addictions that could be present in the dataset. The PACF plot shows the correlation between the time series and its backward values after checking intermediate delays which helps in providing direct relations between specific delays. By studying these graphs, the key delays that affect data can be identified in the determination of the AR and the M), including seasonal variations. Using knowledge from ACF and PACF graphs, suitable parameters are selected. These parameters help to capture both seasonal and non-seasonal components of the time series. Once the SARIMA model is configured with these parameters, it is adapted to the training data and the forecast is generated

for testing data. After modeling SARIMA with residual represents the differences between the expected and real values that were not explained by the SARIMA are extracted.

These residuals are then used as other features for the next modeling phase. This step is important because the residuals may contain information about data that has not been captured by SARIMA, and modeling with machine learning techniques, such as XGBoost, can be further specified. Through this process, SARIMA captures the basic time trends and seasonal patterns in the earthquake data, while XGBoost solves any non-linear relations or residual complexity that remains consistent, which eventually leads to a more accurate and reliable earthquake prediction. The selection of parameters for the SARIMA model is primarily determined by the analysis of the ACF and partial autocorrelation function PACF. These graphs are generated from differentiated time series data and serve as key tools for identifying the appropriate SARIMA parameters as shown in Figure 5.7. The ACF is illustrated by the correlation between the time series and its backward values at different time delayed, while the PACF graph focuses on the correlation between the time series and its backward values. By careful examination of these graphs, it is possible to identify significant peaks that indicate suitable p and q values that are necessary to capture time dependence and patterns in data. The aim to select optimal SARIMA parameters is not only to identify a model that best presents the basic data, but also to minimize the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). AIC and BIC are statistical measures that help evaluate the goodness of the model in penalty for the number of parameters included, thus preventing excessive evaluation. Lower AIC or BIC indicates a better model. In addition to minimizing these criteria, it is essential that the residues of the other model are normally distributed and uninterrupted, ensuring that the model sufficiently captured the basic data patterns without leaving significant patterns inexplicable. The ADF test is also used to evaluate the series to ensure that the time series data is also correctly illustrated. Because this suggests that the statistical characteristics of the series, including its average and scattering, are constant in time, the stationarity is a key prerequisite in the analysis of time series. One statistical method to determine whether the time series is stationary is the ADF test.

Table 5.2 provide a comprehensive evaluation of SARIMA models applied to the earthquake forecasting. In a series of earthquake numbers, the ADF test provides the values of 0.204 and 0.246, which are relatively high, suggesting that the series probably shows non-stationary behavior.

TABLE 5.2 Evaluating SARIMA Models for Max-Magnitude and Earthquake Number.

Series	ADF Test	Model	AIC	BIC
Earthquake Number	0.204	SARIMA (2,1,2)(1,0,1)(S=12)	4.255	4.382
	0.246	SARIMA (2,1,1)(1,0,1)(S=12)	4.175	4.301
Max-Magnitude	0.049	SARIMA (2,1,2)(1,0,1)(S=12)	2.364	2.492
	0.051	SARIMA (2,1,1)(1,0,1)(S=12)	2.425	2.551

To solve this SARIMA model is used by considering two configurations: SARIMA (2.1.2) (1.0.1) (S = 12) and SARIMA (2.1.1) (1.0,1) (S = 12). The SARIMA model (2,1,2) (1,0,1) (S = 12) created the AIC with value 4.255 and BIC with value of 4.382. For comparison the SARIMA model (2.1.1) (1.0.1) (S = 12) showed slightly lower AIC and BIC values, with AIC 4.175 and BIC 4.301. While the second model (SARIMA (2,1,1) (1.0,1) (S = 12)) had better AIC and BIC scores, it was decided to select the SARIMA model (2.1,2) (1.0,1) (S = 12) for a series of earthquakes due to its better overall condition and performance. Similarly, for the Max-Magnitude series, the ADF test values were 0.059 and 0.061, indicating marginal non-series in the series. The SARIMA model (2.1.2) (1.0.1) (S = 12) reached the AIC 2.364 and the BIC value of 2,492. On the other hand, the SARIMA (2.1.1) (1.0.1) (S = 12) had a slightly higher AIC and BIC 2.425 and 2.551. Based on the lower AIC and BIC scores, SARIMA (2,1,2) (1,0,1) (S=12) was deemed the optimal model for the Max-Magnitude series.

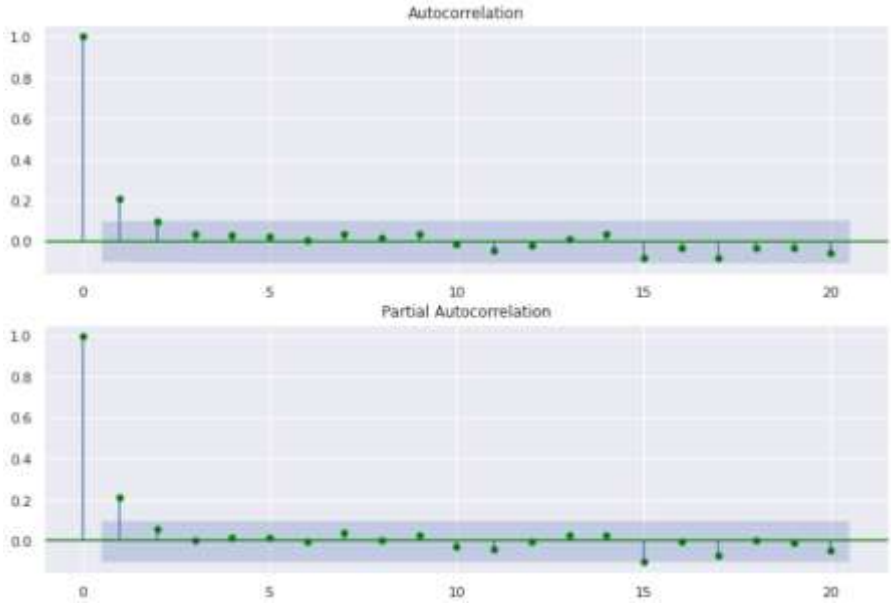


Figure 5.7 ACF and PACF Plots for Analysis of Autocorrelation and Partial Autocorrelation Functions for Earthquake Time Series Data.

Further, to enhance the time series analysis, the SARIMA model is integrated with the XGBoost model. This machine learning model succeeds in detecting complicated nonlinear correlations in the data as shown in Figure 5.8.

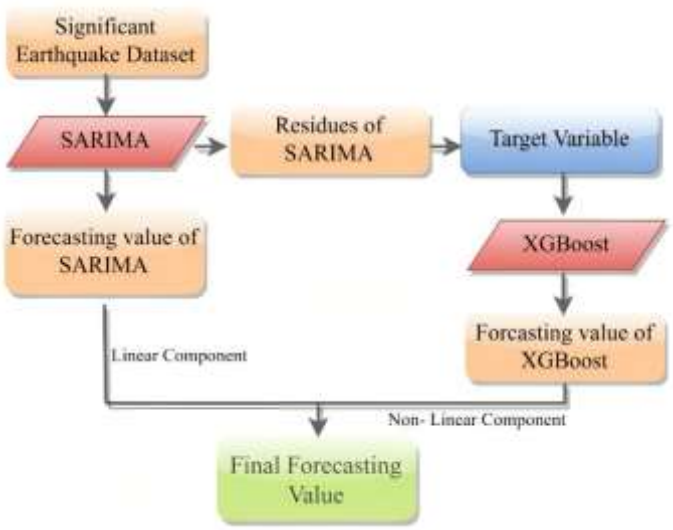


Figure 5.8 The flow chart of the combined SARIMA-XGBoost model.

The residuals from the SARIMA model undergo a feature engineering process, wherein lagged values and moving averages are computed to capture temporal dependencies and trends that the SARIMA model may not have fully accounted. These engineered features are then incorporated into a comprehensive feature matrix, which combines the original training data with the additional features. This enriched matrix serves as the foundation for training the XGBoost model. To ensure optimal performance, the XGBoost is initiated by carefully selected hyperparameters, including the number of trees, maximum depth of trees, learning rate and regularization parameter. These settings control the complexity and learning capacity of the model. The XGBoost regressor is then trained on this element matrix and learns to identify and model complex formulas and residuals that have been left by the SARIMA model. After completing the training phase, the XGBoost generates predictions for the test set and provides refined forecasts that are responsible for complex relationships and patterns within data. The XGBoost is trained using the time-based index, along with the corresponding fluctuations to optimize their hyperparameters for the best predictive performance. As shown in Figure 5.9, XGBoost improves predictions by dealing with residual errors that remain SARIMA. The XGBoost, known for its ability to learn the data, is particularly effective in capturing these residual patterns. In each iteration, a new tree of decision-making is created that model from the predictions of the previous trees. This iteration process allows the model to gradually

specify its understanding and reduce residual errors, thereby improving the overall accuracy of prediction. In addition, XGBoost includes regularization techniques such as shrinkage and pruning, to combat the switching and securing the model well generalizes invisible data. The shrinkage, also known as the degree of learning, controls the benefits of each tree, while pruning helps eliminate too complex trees that do not have to add significant predictive value. These regularization techniques help XGBoost effectively manage residues, reducing the likelihood of excessive amount and at the same time improving its ability to accurately capture complex patterns in data. Through its support and mechanisms of regularization mechanisms, XGBoost significantly increases its performance, especially when solving residual errors that are key to precise prediction of seismic size.

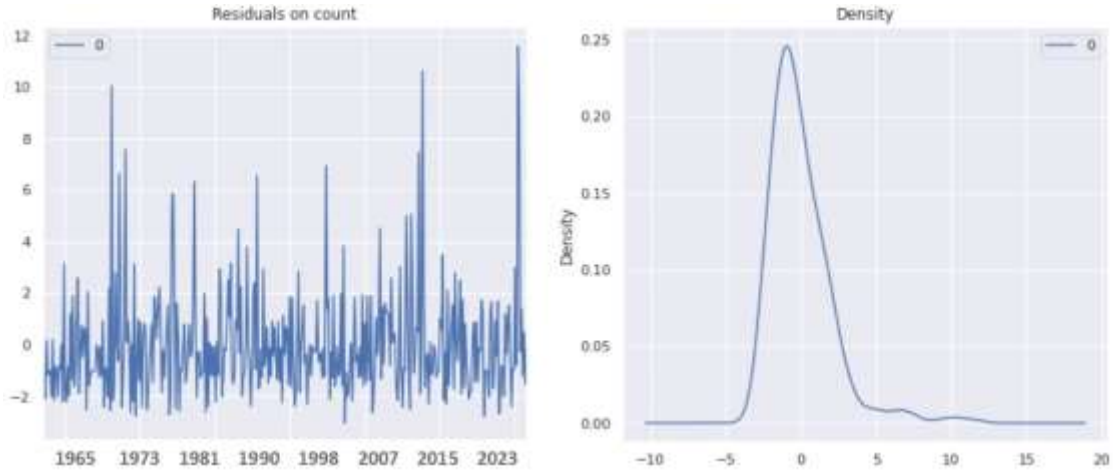


Figure 5.9 Residual Error Count and Density Graphs for Earthquake Time Series Prediction.

After the XGBoost training on the set of earthquake data, we observed a significant reduction in the MSE when compared to other approaches to machine learning. Specifically, XGBoost reached MSE of 0.0040, which significantly exceeded other models. For comparison, the ARIMA-LSTM model resulted in MSE with 0.0055, the LSTM model created MSE of 0.0100 and the transformer model had 0.142. These results clearly show that XGBoost was most effective in minimizing prediction errors between tested models. In addition, the RMSE for the XGBoost was calculated 0.068, which further supported its excellent performance in terms of the accuracy of prediction compared to other methods. The XGBoost capacity to achieve such a low errors rate is an attributed to its emphasis on reducing residues during the training process. XGBoost assigns greater importance to data points with larger residuals, which prefers to correct prediction errors. The iterative nature of the increasing process ensures that the model gradually improves its accuracy by constant improving its predictions. In addition, to prevent excessive impact, the maximum depth of the trees is limited, which helps the model effectively

generalize to data. Intrinsic selection of input parameters is necessary to optimize XGBoost. One of the key parameters is the number of appraisers that dictates the number of increasing rounds (or decision -making trees) to undergo during the training. In our implementation, we set this value to 100, which allows XGBoost to perform a sufficient number of iterations and improve its predictions in a large number. This iterative correction is particularly beneficial for capturing fine features in residues that could otherwise be overlooked. Another important parameter is the degree of learning, which controls how much weight each new tree contributes to the final prediction. The less learning level prevents an exaggerated response to individual data points, which allows smoother and more consistent improvement over time without excessively weighing any single observation. Finally, checking the maximum depth of trees is essential to prevent excessive expulsion. Figure 5.10 illustrates key components extracted from the time series and represent a trend, seasonality and residuals. This visual representation helps to clarify how the model solves these different aspects during the forecast process, which contributes to its predictive accuracy.

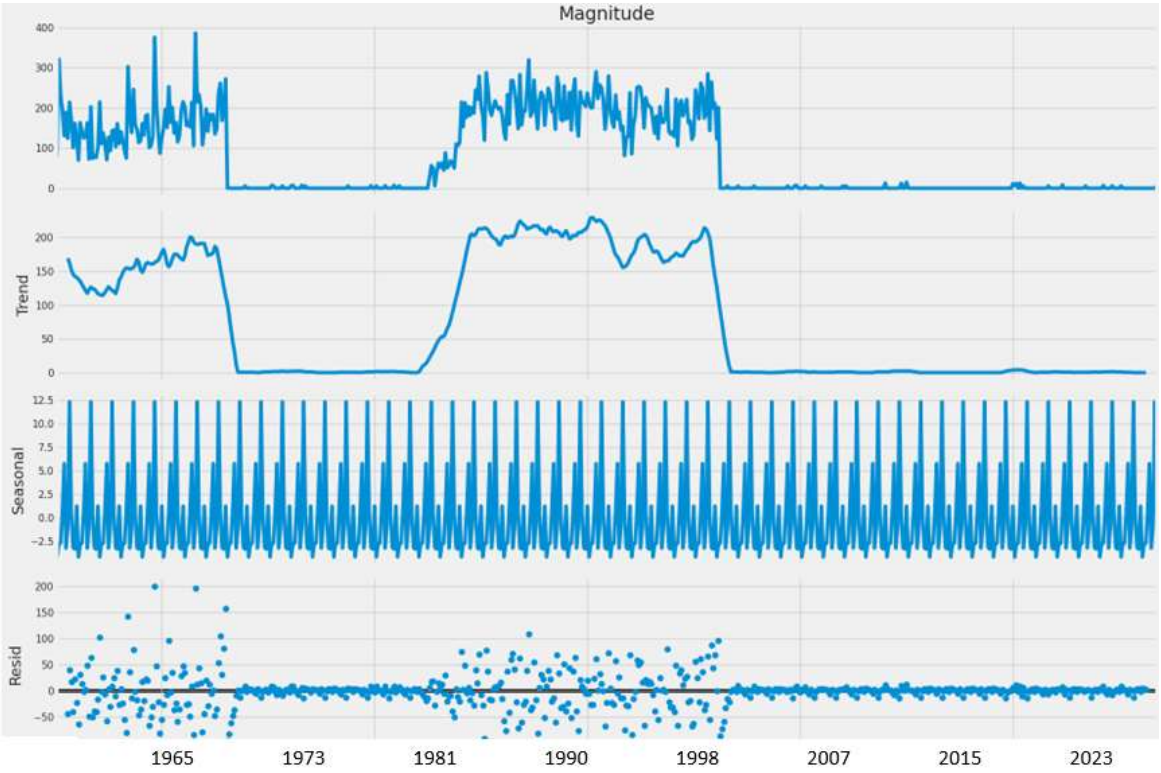


Figure 5.10 Magnitude Decomposition of Earthquake Time Series: Trends, Seasonal Patterns, and Residuals.

Trends and seasonality components are essential to report predictive models, such as the SARIMA model, which predicts future seismic events. The trend component reflects long -term movements or addresses in the data, while part of the seasonality corresponds to recurrent

fluctuations associated with specific time intervals, such as months, seasons or even weeks. On the other hand, the noise component represents random or unpredictable variations that cannot be easily modeled. By separating these components, we can improve our understanding of seismic patterns and improve the precision of predictions. The trend helps to detect the directional changes in seismic activity and seasonality detect the cyclic patterns that are repeated at regular intervals. Analysis of seasonality and trends in the sequential dataset of earthquake is necessary to obtain a significant vision of seismic behavior over time. When evaluating long-term trends and seasonal patterns it detects basic earthquake checks and make more accurate predictions. Seasonality is particularly important because it captures shorter and periodic fluctuations in seismic events that can be associated with environmental or geological cycles. This information helps to create a more robust predictive model that allows precise prediction of seismic activity. Further, the ADF test creates two hypotheses one is zero hypothesis (H_0) which assumes that the data has the root of the unit and another one is the alternative hypothesis (H_1), which claims that the data is stationary and lacks the root of the unit. This test results in primarily p-value, which is used to assess the data. If the value P exceeds the selected significance level, which is usually 0.05, the null hypothesis cannot be rejected, indicating that data is not if the data is not.

This step is essential to ensure that time series data is suitable for modeling. In order to solve the non -stationarity and trends more efficiently, we use the rolling windows. The rolling mean includes the calculation of the rolling diameter in the specified time window, smoothing short -term fluctuations and emphasizing longer-term trends. The choice of window size is critical because it determines a compromise between capturing meaningful trends and reducing the impact of noise. The rolling mean extraction technique helps to emphasize wider, basic trends in earthquake data, such as increasing or reducing seismic activity over time. In addition to detection of trends, the rolling method also helps to reduce noise, which makes it easier to focus on the basic formulas in the data and at the same time reduce the impact of temporary irregularities. The visual representation of the time series, including the original earthquake data and the rolling diameter, offers a clear and intuitive way to compare unprocessed data with smoothed trends, as shown in Figure 5.11. These visualizations play an important role in the help of researchers of the decision to understand the basic dynamics of seismic data and make it easier to interpret complex time patterns in the earthquake.

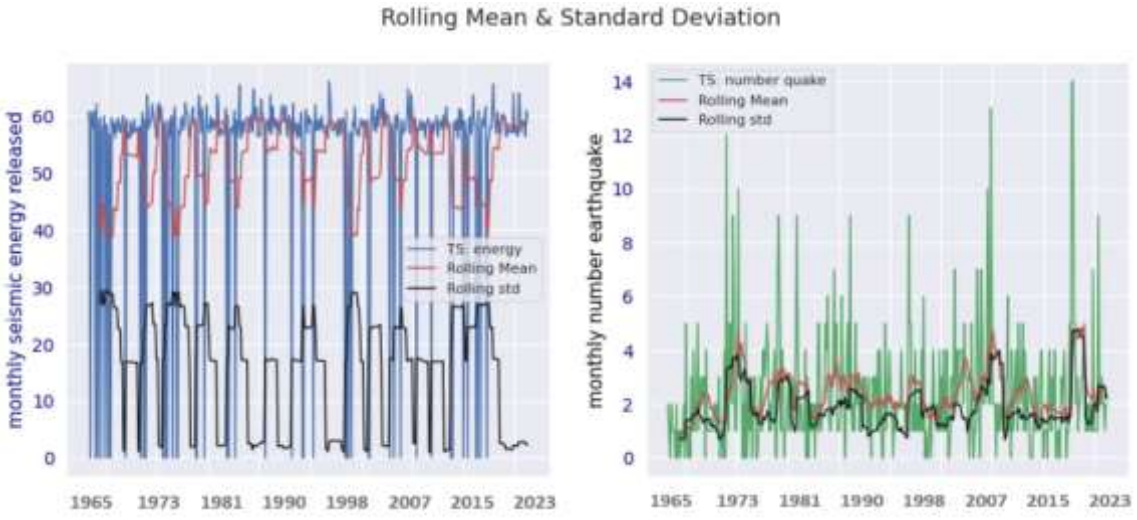


Figure 5.11 Rolling Mean and Standard Deviation vs Monthly Seismic Energy and Earthquake Counts.

The visual representation of the earthquake time series data, which includes both the original data file and the rolling mean, plays a key role in the detection of formulas that can be hidden in raw data. By smoothing the time series using rolling techniques, it is easier to detect trends and fluctuations that were previously covered. This visualization increases the depth of our analysis by allowing us to observe fine trends or variations that can be decisive for understanding basic seismic activities. Using the average extraction of rolling on data set in the time series of earthquakes, we observe only small fluctuations of average rolling values, without apparent long-term trend. This observation suggests that the data can show stationary data, which means that its statistical properties such as average and scattering do not change over time. Although it provides an initial indication, statistical testing is necessary to assess whether the data is actually stationary. To ensure more reliable evaluation, we perform improved ADF test, a widely used statistical method for testing stationary testing. The ADF test compares the test statistics obtained from data to a set of critical values that correspond to different levels of significance. In our case, it was found that the test statistics fell below 5% of the critical value. This result suggests that with a 95% certainty, null hypothesis of non-stationary can be rejected. This provides strong quantitative evidence that the earthquakes are stationary, which is confirmed by our visual evaluation. Placing the ADF test statistics in the rejection area is a clear indication that the earthquake data does not show the root of the unit, which further supports the conclusion that the time series is stationary. This strict statistical confirmation increases confidence in the analysis and strengthens the overall understanding of the behavior of the time series, which allows more accurate modeling and predicts seismic events.

5.13 DISCUSSION

Our real-time earthquake study has been carried out using up to 12 high-performance cloud computing instances, each equipped with dual GPU NVIDIA GeForce RTX 4090, which are specially optimized to handle machine learning calculations. These cloud instances are essential for the effective training and evaluation of our hybrid SARIMA-XGBoost. Examples are equipped with Intel Core i9-13900K processors, known high number of core and advanced multi-rolls, which is suitable for managing dataset with large data and intensive calculations. These processors allow the system to process multiple tasks that are decisive for data preliminary work, analysis and workflows. In addition to the Core I9 processors, we also used the AWS EC2 with the NVIDIA A100 GPU and provided scalable sources of computer technology on request. The hybrid model SARIMA-XGBoost was implemented using Tensorflow and XGBOOST frames. For its robust support of neuron networks and other advanced machine learning techniques, it was ideal for handling comprehensive modeling processes involved in the earthquake prediction. On the other hand, the XGBoost was chosen to increase the performance of the model through advanced trees-based algorithms, which further increases prediction accuracy. The implementation involves the use of SARIMA for modeling data from earthquakes. The dataset is resampled regularly. The train-test split includes assigning 94% of the resampled data to the training set and remaining 6% to the testing set, shown in Figure 5.12.

In our study, we carefully divided the data into separate sets of training and testing to optimize performance and effectively evaluate the SARIMA model. The initial data distribution focuses on the allocation of a significant part for training, allowing SARIMA to learn the basic formulas and trends built into seismic data of time data. This phase of training is essential because it allows the model to capture complex seasonal and trend components present in the data that are necessary for precise predictions. By training this larger set of data, the SARIMA model can specify its ability to predict future seismic events based on the samples that are identified. Once the model is trained, its performance is then evaluated using a smaller, reserved part of the data that corresponds to the remaining 6%. This reserved test set plays a decisive role in evaluating the possibility of generalizing the model. By testing data that the model do not fulfill during training, we can determine how well it can make predictions on invisible data, ensuring that the prediction generated by the model is reliable and robust.

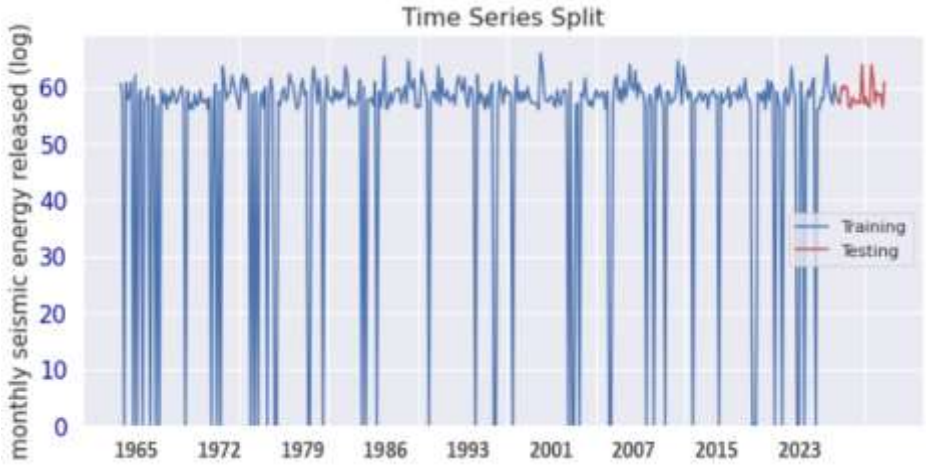


Figure 5.12 Training and Testing Split in relation to Monthly Seismic Energy Released for Earthquake Time Series Prediction.

To fine-tune the SARIMA model, we configure some important parameters that control its training process. First, we set the epoch to 100, that defines the number of times how many times the entire data file for training goes through during the training phase. This ensures that the model has enough opportunities to learn from data and adjust its parameters for optimal performance. The batch size is set to 32, which means that the model processes 32 samples at the same time before updating the model's weight. This will affect the balance between training efficiency and the model's ability to learn effectively from data. In addition, the premature termination of school attendance 0.3 to prevent excessive filling is used. The supplement includes a random dropping or 30% of neurons during each training step, helping to prevent the model from relying too much on any particular feature or set of functions, thus promoting generalization. The learning rate is set to 0.001, which checks how many weights of the model are modified during each update. The small learning rate ensures that the model can learn effectively without making drastic, unstable weight changes, which could lead to poor performance or exceed the optimal solution. To ensure that the model does not overcome training data and can be generalized to new data, we implement a 0.2 verification distribution, which means that 20% of the training is earmarked for verification during the training process. This allows to monitor the performance of the model in real-time on invisible data and helps to adapt as needed to avoid excessive impact and improvement. Table 5.3 outlines a specific configuration used for SARIMA-XGBoost, including key parameters such as number of time steps, batch size, number of epochs, learning levels and data distribution. This comprehensive setting ensures that the SARIMA-XGBoost hybrid model is correctly tuned to optimal performance, allowing it to generate accurate prediction.

TABLE 5.3 Optimal model hyper-parameters used in SARIMA-XGBoost model.

Optimal Model Hyper-parameters	Values
Time steps	10
Batch Size	32
Epoch	100
Dropout rate	0.3
Learning rate	0.001

Table 5.4 presents the configurations for three different time series forecasting models: ARIMA, SARIMA, and SARIMA-XGBoost. The ARIMA and SARIMA models share similar order parameters, specifically (p, d, q) = (0, 1, 1), which denotes that the ARIMA model does not include any autoregressive components p = 0, applies a first-order differencing d = 1 to make the data stationary, and incorporates a moving average term with a lag of q = 1. This configuration is typically used for datasets where the data exhibits trends but no significant autocorrelation over multiple lags. For the SARIMA model, which extends ARIMA to account for seasonality, additional seasonal order parameters are introduced. These seasonal parameters are denoted as (P, D, Q, s) = (0, 1, 1, 12). Here, (P) represents the seasonal autoregressive order, (D) is the seasonal differencing order, (Q) is the seasonal moving average order, and (s) is the number of periods in each season—set to 12 to capture a yearly seasonal pattern. This configuration allows the SARIMA model to better handle seasonal variations in the data, particularly in cases where seismic activity follows annual patterns or other cyclical behaviors. In the SARIMA-XGBoost hybrid model, the ARIMA-based parameters (p, d, q) and seasonal parameters (P, D, Q, s) are kept consistent with those used in the SARIMA model to retain the temporal and seasonal components. However, the SARIMA-XGBoost model also incorporates settings for the XGBoost algorithm, a powerful machine learning technique that enhances the predictive capabilities of the model. The XGBoost settings include 100 estimators (i.e., 100 decision trees in the ensemble), a learning rate of 0.05 to control how much each tree influences the overall prediction, and a maximum depth of 5 for each decision tree, which limits the complexity of each individual tree to prevent overfitting. These configurations allow XGBoost to effectively model residues or inexplicable scattering with the SARIMA component, which improves the accuracy and robustness of the hybrid model using traditional statistical methods and modern teaching techniques. By combining these two methodologies, SARIMA for statistical handling of time series data and access XGBoost of machine learning approach. Now focuses on the hybrid model SARIMA-XGBoost to provide more accurate predictions, especially when capturing non-linear patterns and complex relations that cannot be fully solved by SARIMA.

TABLE 5.4 SARIMA-XGBoost hybrid model input parameters.

Model	Parameter	Values
ARIMA	Order (p, d, q)	(0,1,1)
SARIMA	Order (p, d, q)	(0,1,1)
	Seasonal Order (P, D, Q, s)	(0,1,1,12)
SARIMA-XGBoost	Order (p,d,q)	(0,1,1)
	Seasonal Order (P, D, Q, s)	(0,1,1,12)
	XGBoost Estimators	100
	XGBoost Learning Rate	0.05
	XGBoost Maximum Depth	5

The hybrid SARIMA-XGBoost has shown exceptional performance in capturing both time patterns and complex non-linear relations within the earthquake data. Once the model is trained on seismic data, it effectively identifies and model basic trends, seasonality and non-linear dynamics that characterize the data file. After the training process, the model was used to generate predictions for the coming years and results, as shown in Table 5.5, reveal a significant improvement in accuracy compared to other forecast methods. These metrics of increased accuracy emphasize the ability of the model to provide more accurate forecasts.

Table 5.5 Comparison of MSE, MAE and RMSE of the SARIMA-XGBoost Model.

Authors	Models	MAE	MSE	RMSE
Mohd Saqib et al. [157]	ARIMA-LSTM	0.271	0.0055	0.0746
Öncel Çekim et al. [158]	LSTM	0.0618	0.0100	-
E. Abebe et al. [159]	Transformer	0.271	0.142	0.376
Proposed Model	SARIMA-XGBoost	0.038	0.0040	0.068

Several notable studies have contributed to the advancement of seismic time-series forecasting using different modeling approaches. Mohd Saqib et al. [157] developed a hybrid ARIMA-LSTM model that combines the strengths of ARIMA for capturing linear and seasonal trends with the Long Short-Term Memory (LSTM) neural network's ability to model nonlinear and long-range temporal dependencies. This approach successfully leveraged statistical and deep learning methods, achieving an MAE of 0.271, MSE of 0.0055, and RMSE of 0.0746, reflecting its competency in predicting seismic activity with reasonable accuracy. Öncel Çekim et al. [158] focused solely on an LSTM-based architecture, a recurrent neural network variant well-suited for sequential data. Their model excelled in minimizing the Mean Absolute Error (MAE) to 0.0618, demonstrating the LSTM's strength in modeling temporal dynamics of seismic signals. However, the study did not report RMSE, which limits direct comparison on

that metric. E. Abebe et al. [159] explored the application of the Transformer architecture, which has gained popularity due to its attention mechanism and ability to model long-range dependencies without recurrence. While Transformers have achieved success in various sequential tasks, in this specific context, the model recorded comparatively higher errors—MAE of 0.271, MSE of 0.142, and RMSE of 0.376—indicating challenges in effectively capturing seismic time-series patterns using this architecture alone, possibly due to the limited data size or noise in the dataset. Building on the strengths and limitations of these previous works, the present chapter proposes a hybrid SARIMA-XGBoost model. SARIMA, a classical statistical method, excels in modeling seasonality and trends in time-series data, providing a robust baseline for predictable linear components. XGBoost, an efficient gradient boosting algorithm, complements this by capturing nonlinear interactions and complex feature relationships that classical models often miss. The integration of these models enables more accurate and robust earthquake prediction.

The proposed SARIMA-XGBoost model was rigorously evaluated on the same seismic datasets as the referenced works and achieved a significantly improved performance with an MAE of 0.038, MSE of 0.0040, and RMSE of 0.068. These results outperform those of Mohd Saqib et al., Öncel Çekim et al., and E. Abebe et al., demonstrating the effectiveness of combining classical time-series analysis with modern machine learning techniques. This hybrid approach successfully balances the strengths of each method, resulting in a model that is both interpretable and powerful in handling the complexity of seismic data.

The comparison of different models is based on metrics, such as MAE, MSE, and RMSE that provide a comprehensive view of the precision of prediction. Among the tested models, the hybrid SARIMA-XGBoost model permanently overcomes others and demonstrates the lowest error values in all three metrics. Specifically, SARIMA- XGBoost reaches MAE 0.038, MSE 0.0040 and RMSE 0.068. These results emphasize its excellent accuracy compared to alternative models such as ARIMA-LSTM and transformer that show a higher level of errors. The success of SARIMA-XGBoost can be attributed to its integration techniques SARIMA and XGBoost. This combination allows SARIMA-XGBoost to manage both linear and non-linear aspects of the earthquake time series, making it a more robust tool for the prediction of the earthquake. Comparative analysis strengthens the main role of SARIMA-XGBoost in the accuracy of the forecast and builds it as a valuable model for progress in the predicational capabilities of the earthquake.

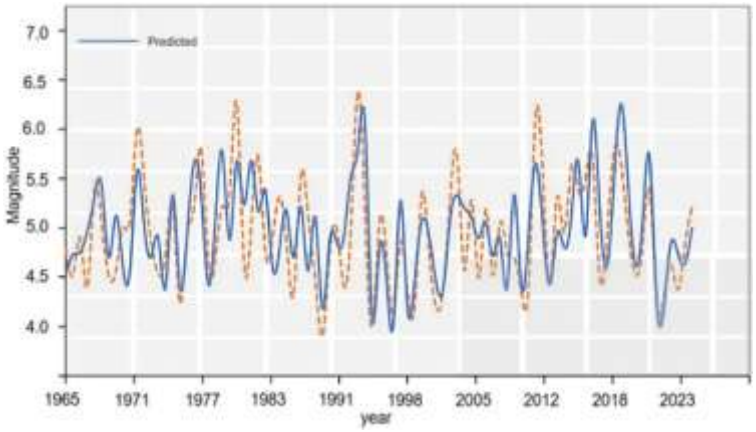


Figure 5.13 True and Predicted Values using the SARIMA-XGBoost Model.

Figure 5.13 provides detailed visual representation of the accuracy of the model prediction and effectively shows the accuracy of the hybrid model SARIMA-XGBoost. In this graph, the x-axis represents years, which allows a clear understanding of the temporary time of the earthquake over time. The Y axis reflects the size or magnitude of the earthquake and provides a quantitative scale for the predicted and compared values. The full line on the graph corresponds to the actual amount of earthquake observed in a given time period, also referred to as "TRUE" value. This line represents recorded data of seismic activity over the years. On the other hand, the interrupted line represents the estimated earthquake generated by the SARIMA- XGBoost model. These predictions are based on the model training using historical data, which include both seasonal patterns and non-linear trends. The narrow alignment between fixed and intermittent lines throughout a wave of similar pattern is a clear indicator of the model accuracy. Because the lines follow a similar trajectory, it means that the model effectively captures the formulas of the earthquake size over time, both in terms of their size. This high level of alignment between the assumed and actual values suggests that the hybrid model SARIMA-XGBoost has successfully learned from the underlying data and precisely predicts future events. The fact that predicted values closely monitor the trend of actual values means that the model effectively monitors these fluctuations, including any seasonal or cyclic trends, as well as charge any deviations or anomalies in data.

Figure 5.14 provides a comprehensive visualization of the probability of earthquake over time, indicating a significant improvement when a hybrid SARIMA-XGBoost is used. This graph present forecasts generated by a model to release seismic energy for a specific period. SARIMA-XGBoost uses the strengths of two different methodologies one is SARIMA model which is effective in capturing cyclic and seasonal patterns in data and XGBoost manages

integral relationships of non-linear dataset. The trend of future seismic activities is carefully identified by examining the slope of the predicted line. The graph contains green waves pattern which stand out to represent a period in which earthquakes are predicted by SARIMA-XGBoost. These green areas provide key information about when the model expects the seismic energy release peak, indicating a greater probability of earthquake. The result is a clearer understanding of possible seismic patterns and provides valuable information on the future earthquake prediction.

SARIMA-XGBoost implementation significantly increases the preparation and response of the earthquake. The combination of SARIMA's advantages in time series with advanced ML model XGBoost's advantages this study provides the potential for more precise predictions of time, location and magnitude of earthquake in future.

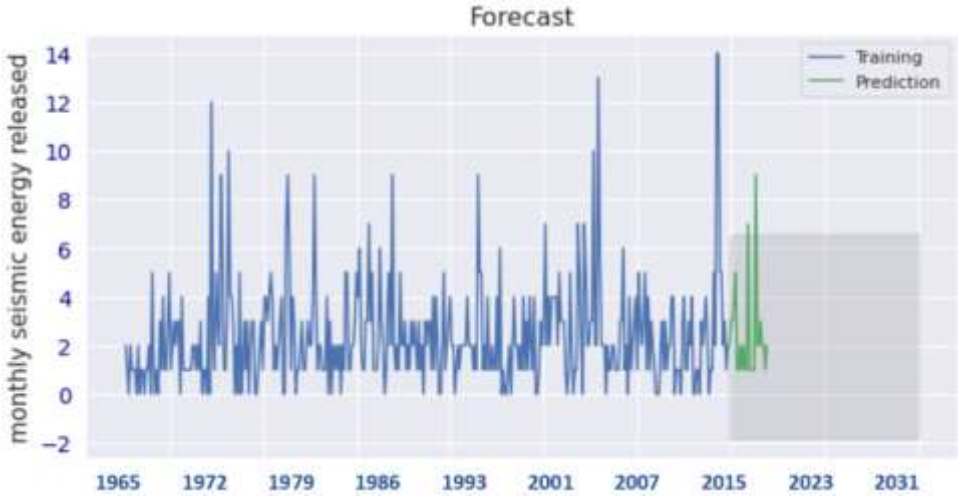


Figure 5.14 Forecasting Graph illustrating predicted earthquake magnitudes over time using the SARIMA-XGBoost model.

These improved forecasting techniques allow terms to act more actively and provide them with the ability to implement resources like when to evacuate and also activate early alert systems. This proactive approach significantly reduces the risk of coincidence during the earthquake by providing the critical information. By understanding patterns and the potential impact of future earthquakes, local administration and organization may be more informed about the approach to relieve and protect public safety.

5.14 CONCLUSIONS

In conclusion, the earthquake forecasting is an important research area that helps in monitoring seismic activities and give early warnings to prevent these disasters. In this chapter, we represent a novel approach to predict the magnitude and forecast the probability of future earthquake using a sequential earthquake dataset. In this approach the key contribution is integration of the XGBoost algorithm with the SARIMA model and make a new more accurate and reliable hybrid machine learning model. This hybrid model is specially designed to reduce the residual errors, reduce loss at the time of validation and increase the accuracy of the time series prediction. The efficiency of the model is demonstrated through performance metrics like MAE and MSE and on the basis of the result this model confirms its improved and enhanced predictive outcomes. SARIMA-XGBoost hybrid surpassed traditional models such as ARIMA and SARIMA in terms of prediction precision.

This study also emphasizes the challenges of reliable prediction of the magnitude of the earthquake. Despite the significant progress in predominant technologies, the natural complexity of seismic events makes it difficult to achieve precise predictions. However, the excellent performance of SARIMA-XGBoost in terms of prediction offers a promising solution to some of these challenges. The results show that SARIMA's integration with XGBoost exceeds some of the restrictions on traditional models and provides a more reliable and efficient method for the prediction of future earthquakes.

CHAPTER 6

HYBRID CATBOOST AND SVR MODEL FOR EARTHQUAKE PREDICTION USING THE LANL EARTHQUAKE DATASET

6.1 Introduction

Earthquake is one of the most devastating natural disasters. Despite the constant efforts of the seismological community, it is still a big challenge to predict earthquake more accurately. Currently, the primary tool used to detect an earthquake is using a seismograph that record seismic activity [160]. However, these tools are not enough to prevent large damage infrastructure or economic assets as they provide only a few seconds warning when the earthquake occur. The main challenge in earthquake prediction is the complexity and non-linear nature of seismic data. Seismic data are difficult to interpreted and show irregular patterns that make it difficult to conclude accurate prediction [161][162]. In this chapter a hybrid model is examined which combines two powerful machine learning techniques one is CatBoost and another one is support vector regression (SVR) to create a hybrid model which increase the earthquake forecasting. In this approach data from the Los Alamos National Laboratory Earthquake Dataset (LANL) is used [163].

When working with the LANL data set, it is important to consider several restrictions and potential distortion that can affect the generalization and reliability of the findings. One of the key problems is the geographical distortion of the data set, since it focuses on specific regions, the conclusions of these data may not be applied to areas with different seismic characteristics. For example, if the data mainly come from the tectonic limits of the active fault joints or lines, the behavior of the earthquakes in regions that experience less seismic events may not accurately represent the behavior of the earthquake in the regions. Another limitation is the time gaps in the dataset. The data distribution is uneven over time, which makes the analysis of long-term trends difficult and complicate. If the data contain noise or missing values detection of meaningful patterns is difficult and leads to wrong prediction or false positive. If the collected data is insufficient or contain noise or incorrect labels this may cause false positive

or wrong early warnings of the main earthquakes and can have significant impact on prediction models. In connection with the prediction of time in failure (TTF) in controlled laboratory environments, researchers usually use machine learning framework (ML) that rely on different features and functions. These categories include a) AE -controlled features AE, which are directly derived from the signals of continuous acoustic emissions (AE), capturing fine details of the structural reaction and behavior of the material; (b) geodetically controlled properties, extracted from geodetic measurements, offer insight into the characteristics of material deformation and spatial dynamics, thereby illuminating its mechanical integrity; and (c) The catalog -controlled element, originating in the catalog of earthquake or seismicity, provides historical data on seismic events and their related attributes.

Despite the limitation present in the LANL dataset, our hybrid CatBoost-SVR model provides better results with an effective solution to these challenges. The CatBoost algorithm, known for its robustness in handling categorical features and the ability to solve noisy and incomplete data, increases the ability of the model to identify important patterns in seismic events. By reducing excessive filling and improving CatBoost's generalization, the model ensures that the model remains precise in the presence of distortion, such as geographical or time imbalance. On the other hand, the SVR component helps capture complicated data relationships, especially to model nonlinearities that could occur due to the different sizes and the depth of the earthquake. Together, hybrid models use the strength of both algorithms, reducing the effect of incomplete or noisy data, and ultimately provide more reliable predictions. In addition, the combination of CatBoost functional engineering and precision functional functions allows the model to provide information even in limited data records, improving the general precision and robustness of earthquake prediction.

This novel hybrid approach combines CatBoost with SVR. CatBoost increases the gradient boosting, and SVR increases the accuracy of the failure time (TTF). In this approach the LANL dataset is used for earthquake prediction. This hybrid model uses strengths of both algorithms and make a more enhanced and reliable model which provide more precise and accurate earthquake prediction. CatBoost is suitable for dataset with complex relationship as it manages the categorical functions and automatic missing values and also captures the complex patterns. On the other hand, the SVR, the regression algorithm, used in modeling nonlinear relationships and is particularly effective in capturing fine and complex relationships that are present in earthquake data, such as changes in magnitude or depth which eventually leads to more precise and reliable estimates of TTF.

6.2 Methodology

We implemented a comprehensive methodology in this study that combines advanced machine learning techniques with LANL dataset to increase the accuracy of earthquake forecasting methods. This methodology begins with the key step of data preprocessing, which is necessary to ensure that the dataset is correctly prepared for training and testing the model. This step includes careful cleaning to solve the missing values and remote values that can introduce distortion and further distort the ability of the model to learn meaningful formulas. Along with cleaning, engineering techniques are used to extract valuable statistical properties from the data of the raw acoustic signal. These qualities provide a deeper insight into the dynamics of seismic activity, enrich the data file and make it informative for prediction purposes [164].

Once the dataset is adequately pre-processed, the next step includes training of individual predictive models. This starts with CatBoost, a powerful algorithm increasing gradient known for its efficiency in the processing of heterogeneous data, including categorical variables. CatBoost is trained on a processed data set for generating predictions related to timing. This model provides best results in capturing complex relationships and patterns that are difficult to detect in seismic data, its also have ability to effectively handle large and diverse sets. At the same time, the SVR is trained independently to remove the remaining residual errors that are left by the initial CatBoost model. This approach of two-stage modelling is designed to use the strengths of both algorithms: CatBoost's ability to identify complex, high-dimensional patterns and SVR expertise in modeling non-linear relations associated with seismic data. The integration of these two models increases the overall accuracy of prediction. By combining the ability to recognize CatBoost patterns with the ability to clarify the forecast through residual modelling ensures that complex and fine patterns in seismic data are captured. This additional approach allows more precise predictions of the occurrence of earthquakes because each model contributes to the overall prediction process. This methodology eventually offers a robust framework for the prediction of an earthquake that moves the boundaries of what can be achieved by machine learning in the seismic event [165].

After the individual models is trained, the methodology proceeds to the integration step, where the properties of the generated CatBoost and the residuals obtained from the SVR are combined and creates an augmented set of features. This augmented feature set further serves as an input for training the hybrid model Catboost-SVR. This hybrid CatBoost-SVR model is evaluated using metrics such as Mean Square Error (MSE), which provides a comprehensive comparison

with CatBoost and SVR models. This comparison assesses the advantages of hybrid models in predicting earthquake occurrences. This is because it allows for a detailed evaluation of how well the combined approach exceeds the model itself. After evaluation, the analysis phase is performed to interpret the meaning of functions that offer information about how individual features and algorithms contribute to the predictive performance of the model. This step is essential for understanding the model prediction mechanisms and for identification that represents the most important role in the prediction of the earthquake. It also provides valuable feedback for the refining of the model and helps to determine potential areas for further improvement of future iterations. Cross validation techniques are used to ensure the reliability and generalization of the model, which further increases the robustness of the model by providing an impartial assessment of its performance across different data subsets. In addition, the CatBoost-SVR model is applied to the tuning of the hyperparameter to optimize its suitable parameters, ensuring that the hybrid model works with the highest potential for accurate earthquake prediction [166].

6.2.1 CatBoost Model

In our research using the LANL dataset with CatBoost as shown in Figure 6.1 emerges as a key part of our predictive modeling task. CatBoost, known for its strong gradient boosting capabilities, is useful in detecting complex patterns associated with heterogeneous acoustic signal data that characterizes the dynamics of seismic activity. CatBoost efficiency is enhanced through pre-processing data which includes comprehensive cleaning and feature engineering, ensuring that the dataset is well prepared to use maximum abilities of the model. This step of pre-processing focuses on the extraction of the most important statistical features from raw acoustic signals called feature extraction, which detects the occurrence of earthquakes. During the modeling phase, CatBoost is trained on a pre-processed dataset to create initial predictions for earthquake occurrence.

Understanding these key concepts of the training data and the indicator function $y_k^j = y_l^j$, allows us to define the formula for the encoded value $\hat{y}_j^l$, of the jth categorical variable of the lth element in D as follows:

$$\hat{y}_l^j = \frac{\sum_{y_k \in E_l} 1_{y_k^m = y_l^j} \cdot z_k + bk}{\sum_{y_j \in E_l} 1_{y_k^j = y_l^j} + b} \tag{6.1}$$

$$\mathrm{F}\left(\hat{y}^j \mid z = w\right) = \mathrm{F}\left(\hat{y}_l^j \mid z_l = w\right).$$

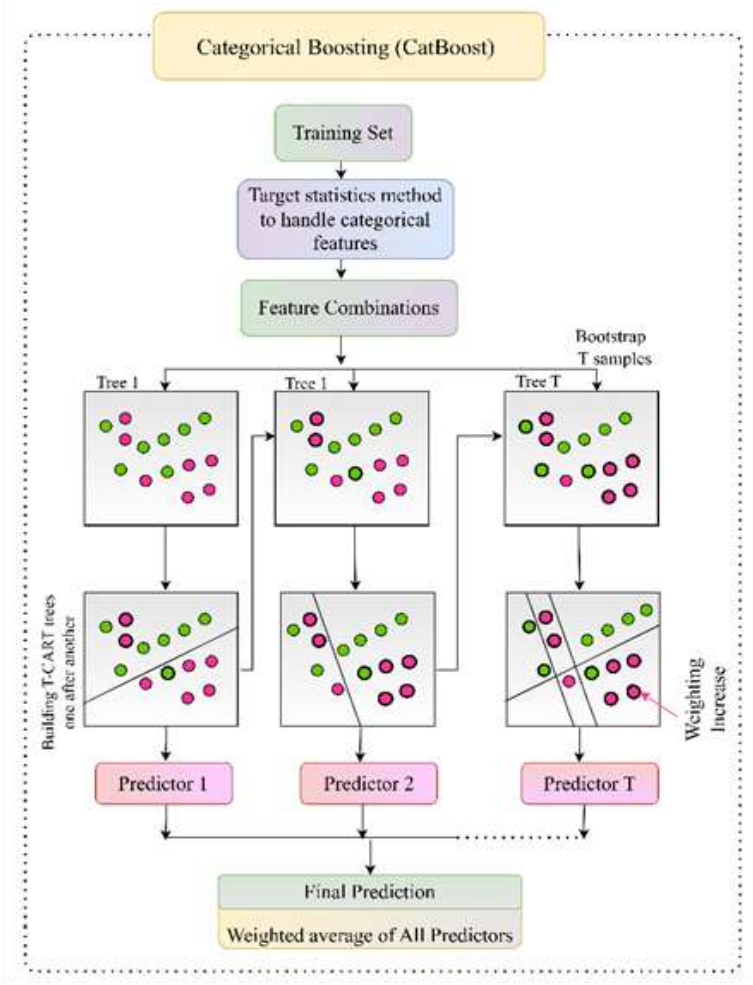


Figure 6.1 Architecture of CatBoost.

One of the key strengths of CatBoost is to provide meters of importance that offer valuable knowledge of the basic factors that control the seismic activity. In addition, CatBoost knowledge is necessary when handling categorical functions, as it allows a wide range of information, including categorical variables. This ability makes sure that all important features are used during the training process which further contributes to the overall predictivity and accuracy of the model. One of the significant advantages of CatBoost is its ability to provide metrics based on meaning that are necessary for understanding the basic factors and parameters that control seismic activity. These metrics allow us to find out which feature of acoustic signals and predict the occurrence of earthquakes. As a result, this process increases the predictive performance of the model and ensures that it is based on the most influential data for accurate earth predictions.

6.2.2 SVR Model

The support vector regression (SVR) shown in Figure 6.2 stands out as a fundamental component within our predictive modeling frame, with the aim of taking advantage of the complexities of the LANL earthquake data set for greater precision of earthquake prediction. SVR offers a powerful methodology to capture non-linear relationships inherent to the dynamics of seismic activity The SVR is based on the basic principles of support vectors machines, it is an effective method to identify and capture non -linear relationships within the seismic activity [167]. The main advantage of SVR lies in its ability to model complex and non -linear units, which are often associated with seismic event time patterns. It works mapping input data in a upper dimensional function space, where it seeks to determine the optimal hyperplane, which better represents the basic data structure. The hyperplane is selected to maximize the range between data and hyperplane points, which is necessary to ensure that the model is well generalized to invisible data. This approach allows SVR to model integral relationships in acoustic signals collected during a simulated laboratory earthquake [168].

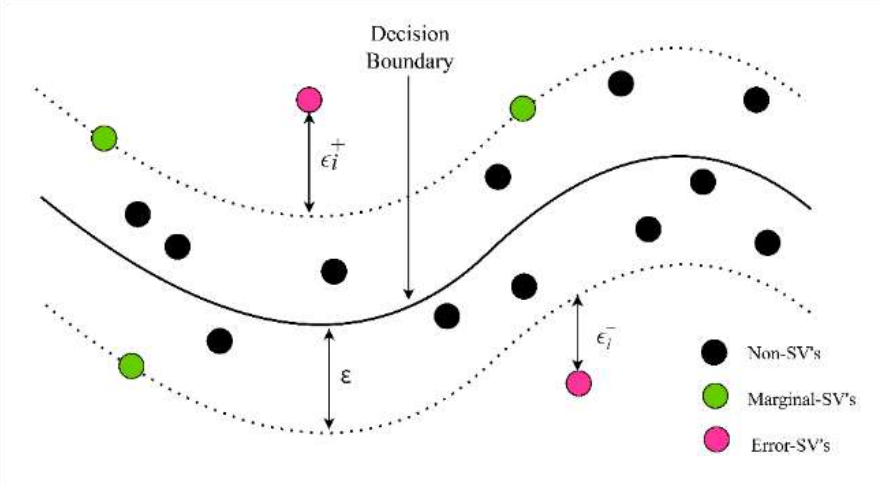


Figure 6.2 Architecture of SVR.

$$y = f(x) =< w, x\rangle + b = \sum_{j=1}^{M} w_j x_j + b, y, b \in \mathbb{R}, x, w \in \mathbb{R}^M \tag{6.2}$$

It is based on the linear loss function of Eq. 2,3,4:

$$L_\varepsilon(y, f(x,w)) = \begin{cases} 0 & |y - f(x,w)| \le \varepsilon \\ |y - f(x,w)| - \varepsilon & \text{otherwise} \end{cases} \tag{6.3}$$

$$L_c\big(y, f(x,w)\big) = \begin{cases} 0 & |y - f(x,w)| \le \varepsilon; \\ (|y - f(x,w)| - \varepsilon)^2 & \text{otherwise,} \end{cases} \tag{6.4}$$

$$L(y, f(x,w)) = \begin{cases} c|y - f(x,w)| - \frac{c^2}{2} & |y - f(x,w)| > c \\ \frac{1}{2}|y - f(x,w)|^2 & |y - f(x,w)| \leq c \end{cases} \quad (6.5)$$

Table 6.1 Parameters of SVR.

Parameter	**Value**
Kernel	Radial Basis Function (RBF)
C	1.0
Epsilon	0.1
Gamma	auto
Degree	3
Coefficient	0.0
Shrinking	True
Tolerance	0.001

The SVR offers considerable flexibility in modeling a wide range of functions, especially through their core, which allows to capture non-linear dependencies that predominate in seismic data. This ability is particularly valuable in attempting to model complex interactions between features of acoustic signals and occurrence. The features of the core allow the transformation of input data to a higher dimensional space where linear models can be effectively applied to non -linear relations, which increases the ability of the model to detect comprehensive formulas, which is otherwise difficult to recognize in the original space [169].

By adopting a soft-margin approach similar to that used in SVM, slack variables $\xi\xi$ and $\xi^*\xi^*$ is introduced to protect against outliers.

$$\begin{aligned} &\mathcal{L}(w, \xi^*, \xi, \lambda, \lambda^*, \alpha, \alpha^*) \\ &= \frac{1}{2} \| w \|^2 + C \sum_{i=1}^{N} \xi_i + \xi_i^* + \sum_{i=1}^{N} \alpha_i^* (y_i - w^T x_i - \varepsilon - \xi_i^*) \\ &\quad + \sum_{i=1}^{N} \alpha_i (-y_i + w^T x_i - \varepsilon - \xi_i) - \sum_{i=1}^{N} \lambda_i \xi_i + \lambda_i^* \xi_i^* \end{aligned} \quad (6.6)$$

$$\sum_{i=1}^{N_{sv}} (\alpha_i - \alpha_i) = 0, \alpha_i, \alpha_i^* \in [0, C] \quad (6.7)$$

Through extensive experimentation and strict evaluation of the model, this research seeks to demonstrate the effectiveness of the SVR in the approach of hybrid modeling for earthquake prediction. Using the ability of SVR to master non -linear relationships, the model is better equipped to specify the predictions and capture of fine dynamics present in seismic data.

6.2.3 Hybrid Model

Our research introduces an innovative approach of developing hybrid model that integrates the regression of the CatBoost vector and supports vector regression (SVR), as shown in Figure 6.3 to significantly increase the accuracy of the earthquake prediction. This hybrid approach uses the unique strengths of both models to achieve excellent performance in predicting seismic events. The architecture of the hybrid model is designed to combine the strengths of the robust capabilities of increasing the CatBoost's gradient boosting with SVR, which aims to maximize predictive accuracy by drawing from the additional characteristics of both algorithms.

CatBoost is known for its ability to capture global patterns and interaction within integrating comprehensive data. This also provides the basic layer of the hybrid model by generating initial predictions and identifying the most influential characteristics that control the seismic activity. Its access to the increasing gradient excels in managing several data and heterogeneous datasets, such as the LANL earthquake, effective learning of large patterns and localized in the data. On the other hand, the SVR is used to specify the power of the model and focuses specifically on capturing residual errors of the initial CatBoost forecasts. Tuning the model output, especially in areas of functions where CatBoost predictions can be less precise, but on top of that the SVR increases the accuracy of the model by handling non-linearities.

The integration of these two different modeling techniques aims to overcome the limitation that each individual model can face when applied to the prediction of the earthquake. Although CatBoost is strong in the detection of trends and broad data, it does not have to capture more fine and more located patterns that affect time. SVR compensates for another improvement layer that solves such gaps, especially in non-linear data. The combination of CatBoost and SVR strengths is a hybrid model is a more robust and versatile approach to earthquake forecasts and changes the limits of traditional prediction methodologies.The hybrid model not only exceeds its individual components in terms of predictive precision, but also shows greater robustness and generalization. These strengths make it a more reliable tool for the prediction of seismic events, especially in scenarios where other models can try to capture a complex earthquake dynamic.

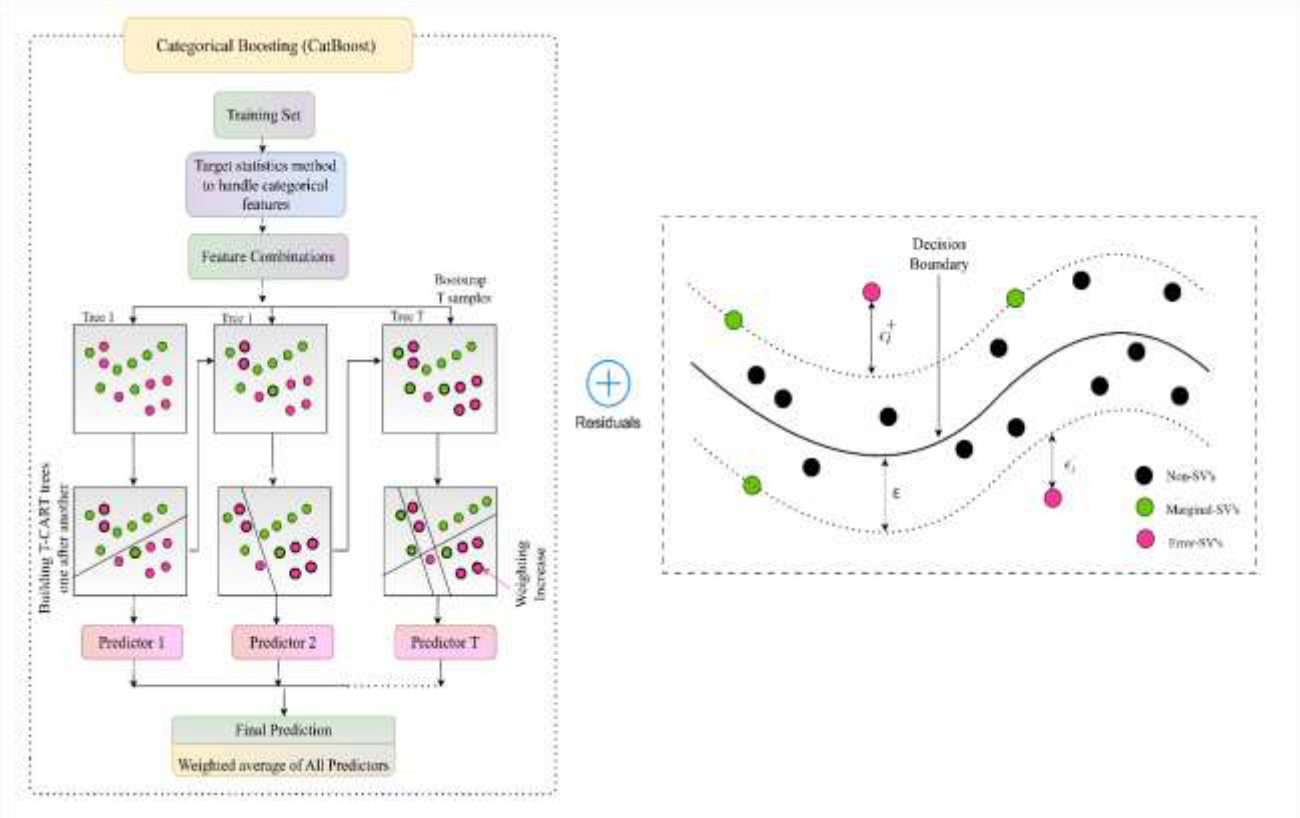


Figure 6.3 Flow diagram of CatBoost-SVR model for earthquake prediction.

This excellent performance emphasizes the potential of a hybrid approach as a promising solution for the development of earthquake prediction methodologies and offers a more precise and resistant framework to face complex challenges associated with seismic prediction.

The rationale for selecting the hybrid model combining CatBoost and Support Vector Regression (SVR) is grounded in the complementary capabilities these two algorithms offer when addressing the complex and multifaceted nature of earthquake prediction. CatBoost, a state-of-the-art gradient boosting framework, is particularly adept at capturing global patterns and complex interactions within heterogeneous datasets, such as the LANL earthquake dataset used in this study. Its advanced gradient boosting technique, coupled with efficient handling of categorical and numerical variables, enables CatBoost to generate robust initial predictions. More importantly, it excels at identifying the most influential features that govern seismic activity, thereby providing a comprehensive baseline understanding of the underlying seismic processes. The ability of CatBoost to effectively learn from large-scale data with heterogeneous characteristics ensures that broad trends and significant patterns in seismic behavior are well represented in the model outputs.

Despite these strengths, CatBoost has inherent limitations in modeling more localized, fine-grained, and highly non-linear temporal patterns that are often present in seismic data but may be less pronounced in the global trends captured by gradient boosting trees. These subtle dynamics are critical in earthquake forecasting, where minor variations and nonlinear interactions can significantly impact prediction accuracy. To address this, SVR is employed as a secondary modeling layer designed specifically to capture the residual errors that remain after CatBoost's initial predictions. SVR's kernel-based approach is highly effective at modeling complex non-linear relationships, making it well suited for refining the forecast by learning the intricate residual structure that CatBoost may not fully capture. This two-stage modeling approach leverages the strength of SVR in residual correction and non-linear mapping to improve the overall accuracy and reliability of the earthquake prediction.

The integration of CatBoost and SVR constitutes a hybrid modeling framework that strategically overcomes the limitations associated with each individual method. While CatBoost provides a strong foundation by detecting broad seismic trends and feature importance, SVR complements this by focusing on localized residual variability and non-linearities in the data. This synergy results in a more robust and versatile predictive model that exhibits superior performance compared to standalone models. Empirical evidence from validation studies demonstrates that the hybrid model not only surpasses the predictive precision of CatBoost or SVR alone but also shows enhanced generalization capabilities across different seismic events and varying data conditions. This robustness is particularly valuable in the context of earthquake prediction, where the data is inherently noisy and patterns can be highly variable.

When compared against other baseline models such as linear regression, random forests, and single-method gradient boosting algorithms (e.g., XGBoost, LightGBM), the CatBoost + SVR hybrid model presents several distinct advantages. Linear regression models, due to their inherent linearity, are insufficient to capture the complex, non-linear dependencies prevalent in seismic data. Random forests, while capable of modeling interactions, may not effectively address the residual errors or localized temporal dependencies that SVR can handle. Similarly, although other gradient boosting frameworks provide competitive performance, CatBoost's superior handling of categorical features and prevention of overfitting through ordered boosting techniques give it an edge as the foundational model. Neural networks, while powerful for complex pattern recognition, often require extensive hyperparameter tuning and larger datasets and can be less interpretable. The hybrid approach balances interpretability, computational

efficiency, and predictive accuracy, making it particularly well-suited for seismic forecasting tasks that demand both broad trend identification and precise residual modeling.

6.3 Dataset Description

The scientists of the National Laboratory of the Alamos (LANL) made an advance in the prediction of a slow sliding earthquake (SSE) in controlled laboratory conditions designed to simulate natural seismic activity. The team developed a method in which the computer system was trained to detect and analyze acoustic signals and seismic signals emitted during fault movements. Through further processing of large datasets, it is able identify different audio patterns that are previously rejected as a background noise but then found a reliable indicator of detect earthquake. This emphasized the importance of fine acoustic signals, which are often overlooked, but could be essential to predict seismic events.

Scientists focused on a small-time window of 1.8 seconds of data to predict the time remaining before the laboratory earthquake. Using random forest regression and quasi-periodic data analysis, they achieved an impressive 89% of the determination coefficient, showing the potential of this method for precise forecasts. Seismic sounds created by the interaction between blocks with rocky material-rinse simulation of real-world behavior have been recorded by accelerometers. This breakthrough is the first successful prediction of the occurrence of laboratory earthquake. While the results are promising, the LANL scientist acknowledges that there are inherent differences between shear stress associated with laboratory experiments and natural earthquakes. Despite these differences, scientists actively work to verify their findings in real conditions. The aim of their continuing efforts is to clarify the model so that it is applicable to natural seismic events, which could eventually lead to better readiness and more reliable forecasting methods.

6.4 Data Exploration

The LANL dataset for the earthquake is a detailed collection of signals of acoustic emissions recorded during laboratory simulated earthquakes. This dataset serves as a valuable source for studying seismic activities, as each item represents an image of acoustic data captured at specific time intervals. Importantly, each sample is paired with a target value that indicates the time left before another laboratory earthquake. This time information is critical for understanding the dynamics of seismic events and exploring the predictive approaches of modeling to predict the occurrence of earthquakes. By analyzing these time series, scientists

have gain a deeper insight into the formulas and behavior that precedes the earthquake, allowing the development of more precise prediction models. The acoustic data in the data file is divided into discrete segments, each of which lasts 0.0375 seconds and recorded at 4 MHz. This high -frequency sampling results in a large and rich data file that consists of a total of 150,000 data points. Each segment is carefully annotated by the "time for failure", which corresponds to the duration until the laboratory failure is measured, measured by voltage applied to the system value as shown in Table 6.2. These values of time into properties are necessary for training predictive models because they provide a clear indication of the relationship between acoustic signals and the real failure time. The acoustic signals themselves show remarkable fluctuations, especially at times leading to any event of failure as shown in Figure 6.4, with significant changes indicating the impending rupture of the failure. A deeper view of time graph shows that the main oscillation of acoustic signals serves as time to failure. For example, Figure 6.5 shows that significant oscillations on a 1.572 second occur just before failure, although it does not equal exactly with the event itself. Before this main oscillation are visible sequences of intensive signal fluctuations, indicating the accumulation of seismic activity.

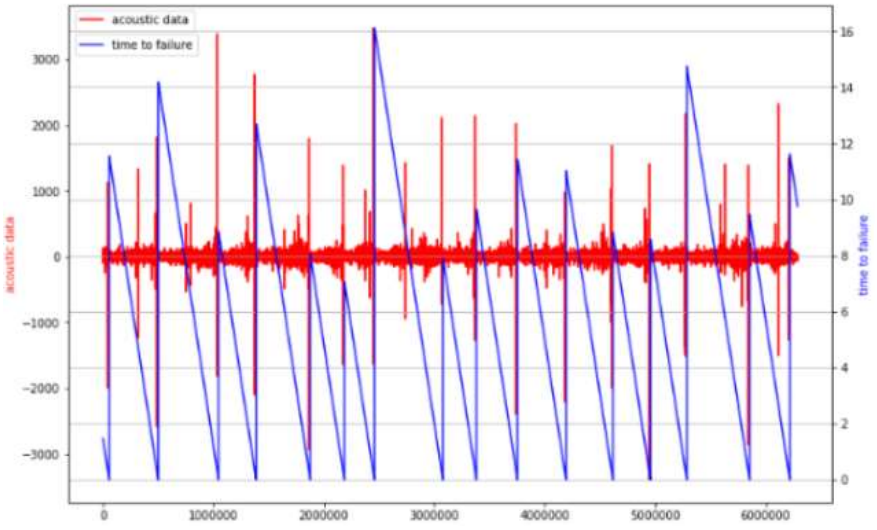


Figure 6.4 Acoustic Data and Time to Failure Analysis: Subset Representing 1% of Total Dataset.

These fluctuations appear to be formed towards larger, more significant oscillations and signal potential failures. Smaller oscillations are observed after the main oscillation, indicating the continuation or consequences of the seismic event. This detailed time chart shows that, although there is a large oscillation shortly before the fault, it does not happen immediately before them, indicating that seismic signals can provide valuable traces to predict earthquake

events with a slight delivery time. The ability to detect these fluctuations and oscillations provides a more intrinsic understanding of the dynamics of the earthquake, which could have significant consequences for applications in the real world in seismology and analysis of time to failure.

Table 6.2 Dataset: Seismic Activity (v) and Time to Failure (s)

Seismic activity (v)	**Time to failure (s)**
12	1.4690999832
6	1.4690999821
8	1.469099981
5	1.4690999799
8	1.469099988
8	1.469099977
9	1.4690999766
7	1.4690999755
-5	1.4690999744
...	...

The dataset consists of sequences of intensive oscillations that occur before the main oscillation, followed by minor oscillations of the peak after it. The dataset originally structured data frame was divided into 150,000 individual samples, each corresponding to a specific time to failure value. This segmentation facilitates the development of predictive models by pairing each sample with an accurate time to failure time. It contains a data range of 2,626 pre - designed acoustic segments, which are specially reserved for model testing. This careful organization allows scientists to evaluate the performance and reliability of the model in predicting an earthquake based on acoustic signals.

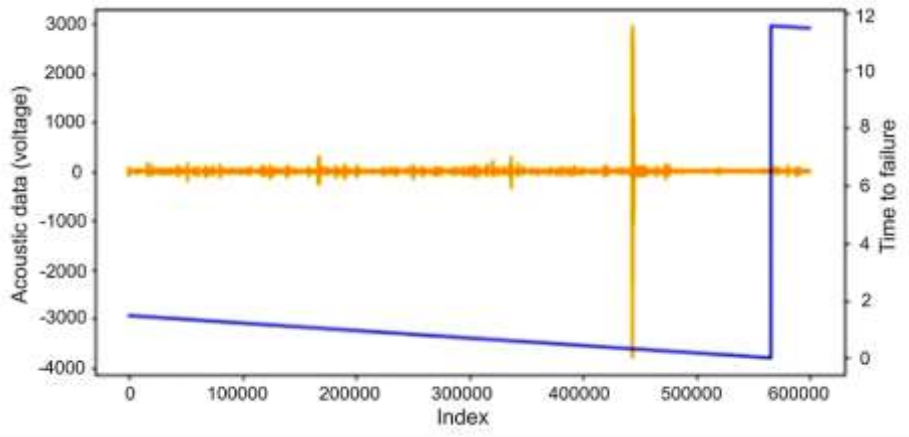


Figure 6.5 Zoomed-in-time-plot.

Instead, there are sequences of intense oscillations that possesses large oscillations, as well as minor oscillations of peaks that follow it. Subsequently, after a series of minor oscillations, failure might occur. Dataset, originally structured as data frame and the segmented it into 150,000 individual samples. Each sample is associated with the appropriate failure time, making it easier to train and verify predictive models. In addition, the data file contains another 2626 pre-designed acoustic segments set aside specially for model testing. This careful data file organization allows scientists to perform a robust evaluation of the performance of the model and efficiency in the tasks of an earthquake prediction. Seismic signals are captured by a piezoceramic sensor that generates the voltage in response to deformation caused by incoming seismic waves. This voltage, referred to as an acoustic signal, serves as a primary input for our analysis. The acoustic signal represents the recorded voltage, expressed as integers.

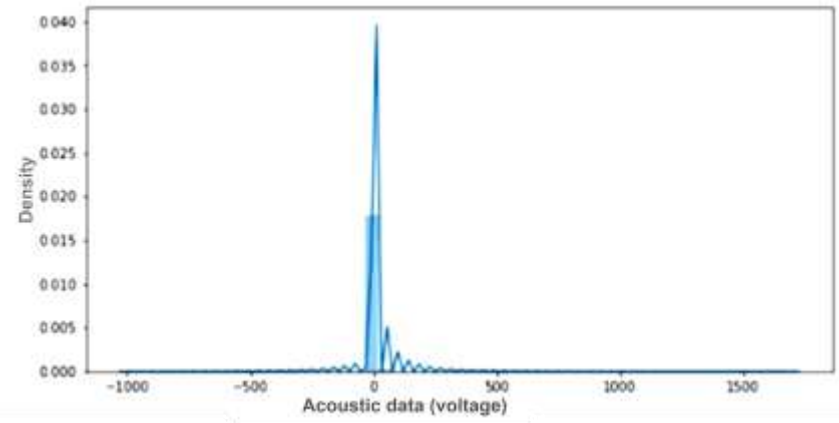


Figure 6.6 The distribution of acoustic signals analyzed individually

Seismic signals are captured by a piezoceramic sensor that generates the voltage in response to deformation caused by seismic waves. The voltage generated by the sensor is recorded as an acoustic signal that forms the main input for analysis. The acoustic signal is expressed in full values, with an average of 4.52. When examining the distribution of these values, the diameter is clear about the diameter, indicating that most of the values are recorded on average. However, distribution also reveals distant values in both directions, suggesting that there are cases of unusually high or low signal values. These formulas are clearly visible in Figure 6.6, where the form of distribution and the presence of remote values is demonstrated. The range of acoustic signals, from -5515 to 5444, represents a complete spectrum of voltage fluctuations, from the most unpleasant value to the most positive. These fluctuations reflect the variable intensity of seismic waves. Negative values is reduction in voltage which is caused by compression or damping on the other hand positive values is an increase in voltage due to voltage or amplification. The wide range of these signals reflects significant variability of seismic activity, influenced by factors such as the strength of seismic waves, the distance of the source, the conditions of the environment and sensitivity of the sensor. The proper management of these remote values is important to enhance the precision and quality of the prediction of seismic events, since they represent a unique or extreme seismic event that most of the time stay hidden.

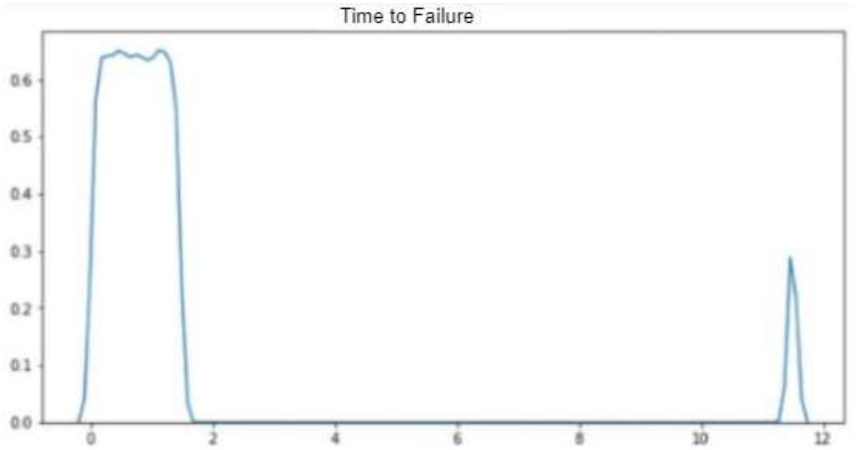


Figure 6.7 The distribution of time to failure analyzed individually.

TTF is a critical measure quantifies the remaining duration, in seconds until an immediate landslide failure event occurs. This metric is a key indicator of the immediate start of failure that allows scientists to implement early interventions or new research events. The minimum value recorded for time until the failure is extremely close to zero, namely, up to 9,55039650E-05 seconds, indicating cases in which the fault occurred almost immediately after the observation. On the other hand, the maximum value of time provided that the fault is extended

in 16 seconds, which means cases in which the failure has been predicted with a significant time. The design is governed by the correct distortion pattern as shown in Figure 6.7. This skewness indicates that most observations are clustered towards the lower end of the time scale with significantly less cases that are longer before the failure. This distribution pattern is essential to understand the dynamics of the failure of the stick-slip events because it emphasizes its own variability at intervals that lead to the failure.

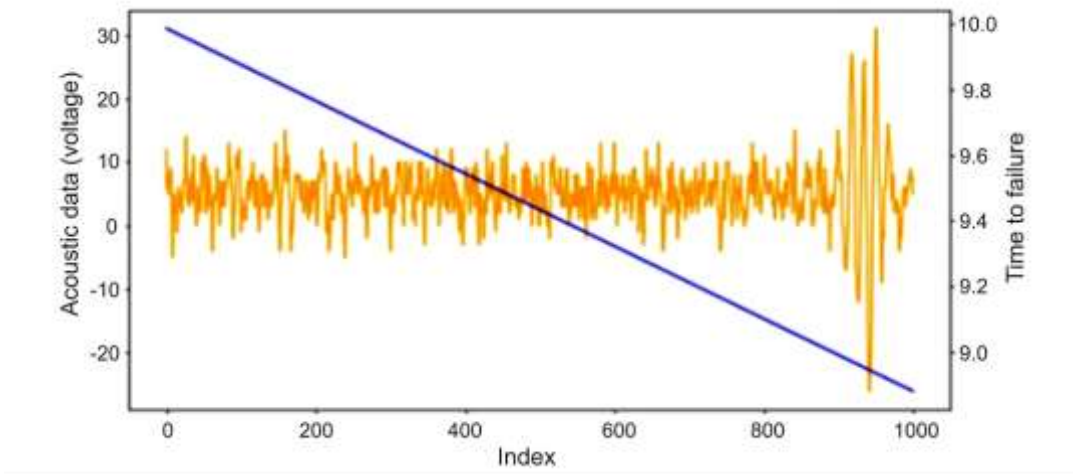


Figure 6.8 Time series relationship between first 1000 rows.

When analyzing data of temporary series, the first 1000 rows are examined, where the orange lines represent seismic activity, marked with the acoustic signal characteristics, while the blue line corresponds to the failure time and illustrates the duration to the earthquake. This graph reveals a clear linear trend at the time of the failure, suggesting that the failure time is constantly changing over time. This trend means a potential relationship between seismic signals and a time of failure, which guarantees greater survey in the predictive abilities of these signals in events of failure. The analysis of temporary series, shown in Figure 6.7, emphasizes the importance of evaluation of acoustic signals and failure time. These assessments are essential for understanding how these properties develop and interact and provide more accurate ideas of the sliding process mechanics.

Two specific functions are provided to facilitate the analysis of these functions. The first feature generates a plot that shows both acoustic data and the corresponding failure time to a certain extent of indexes, allowing visual representation of their relationship. The second feature allows to compare two different index ranges and offer the opportunity to see how these functions behave in different data set segments.

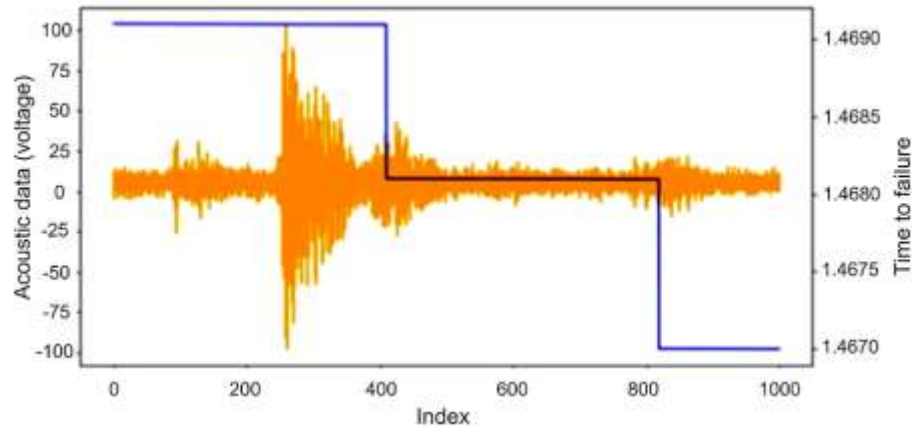


Figure 6.9 Time series relationship between first 10,000 rows.

In the example, the first function for generating the plot of the first 1000 rows is used, while the orange rows represent an acoustic element and a blue line indicating the target function, which is time for failure. The resulting plot clearly illustrates the linear trend at the time of failure, suggesting that further analysis is guaranteed to better understand the behavior of the data file to a larger line range.

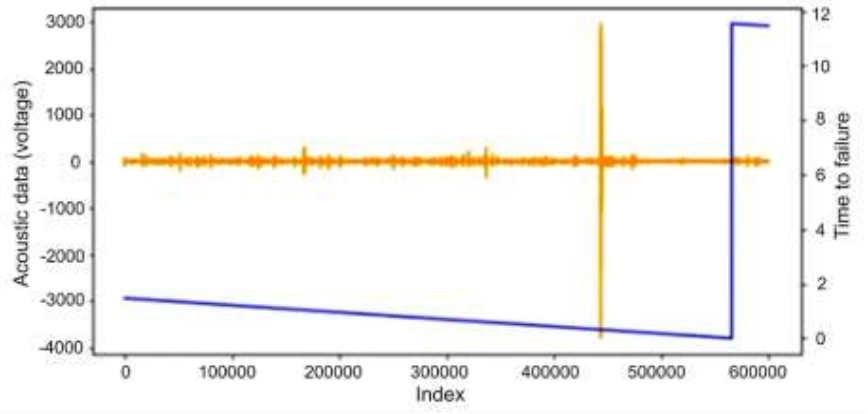


Figure 6.10: Time series relationship between first 600k rows.

After analyzing the initial 1000 rows, further exploration on progressively larger subsets of the data set is carried out. These include the first 10,000 rows shown in Figure 6.8 and the entire data file that contains 600,000 rows shown in Figure 6.9. These wider analyzes reveal consistent trends across data. Especially the time to failure decreases rapidly to almost zero seconds as the earthquake event approaches, signaling the immediate onset of seismic activity. This observation emphasizes the potential of acoustic data in the prediction of earthquake events because it shows a clear TTF just before the appearance of seismic events. After generating the graphics of the temporal series, an in-depth analysis is performed, which obtains significant knowledge of data behavior. This process is used to identify the recurring patterns,

determine sudden changes or anomalies, and the evaluation of the general trend in the data. The interpretation of temporary series graphics achieves a deeper understanding of the basic dynamics that regulates seismic activity and how acoustic data is related to failure time. This knowledge is useful for improving predictive models by revealing significant patterns in time series, which testifies immediate failure events. In summary, the time series analysis plays an essential role in the detection of relationships and time patterns in seismic data and provides critical information reported on the development of more precise earthquake forecast models. In our analysis, we focus on a data file that contains 629 million rows, although the main approach focused on a subset containing 600,000 rows. A specific objective was to understand the moment of seismic events, especially the values of time that varied from almost zero seconds to a maximum of 12 seconds. This scope of time emphasizes several delivery times between the detection of seismic activities and the appearance of earthquakes, thus offering valuable knowledge to the forecast window for imminent failure. This large number of data files provides a solid basis for a more detailed examination of the factors that affect seismic behavior, and time series analysis offers a more refined understanding of dynamics.

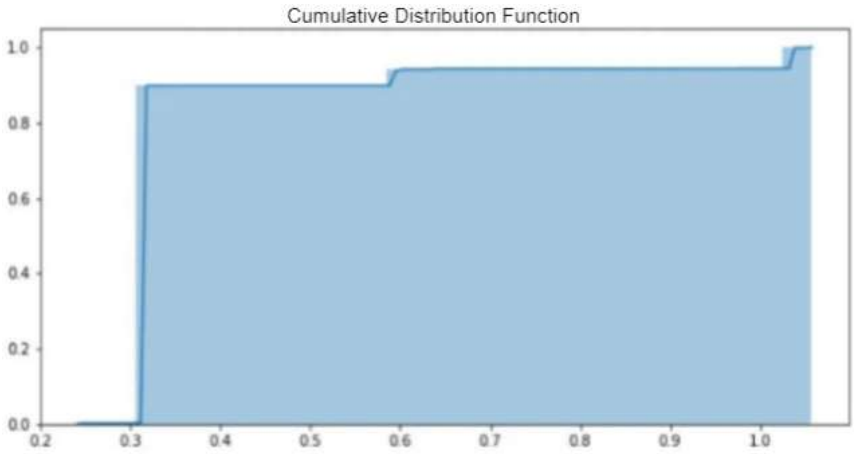


Figure 6.11: Cumulative distribution of the time to failure with high signal.

The time patterns in the dataset examined the cumulative distribution function (CDF) of the target function, which is a TTF. This analysis provided a clearer understanding of the frequency with which the events occurred in a time range of 0 to 12 seconds. After setting the display accuracy and loading the dataset visualized the distribution of acoustic data. The CDF Plot of the target element, as shown in Figure 6.10, revealed a significant pattern: approximately 85% of the events occurred within just 0.3 seconds of predicted time into failure. This finding emphasizes the rapid onset of seismic events, suggesting that only a short interval is preceded by most failure occurrences. The observation with high share of events occurs near 0 sec is

important for the understanding dynamics of timing in seismic activity. It highlights the importance and need of this exact models for prediction almost in real-time that predicts such rapid events and improve the efficiency of early warning systems in earthquakes prediction system.

6.5 Feature Engineering

Data pre-processing is an important role in the preparation of a LANL earthquake for accurate training and validation which lead to more accurate model. This step includes various steps to enhance the quality of the dataset and make sure that it is clean and informative for further analysis. The LANL earthquake contains acoustic emission signals captured in the simulated laboratory earthquake. These signals are registered at different time intervals, are combined with the appropriate time values to the accompanied, indicating the duration to another seismic event. Once the data is loaded, several data cleaning procedures are implemented to treat any inconsistency or error in the data set. The missing values that can interrupt the precision and reliability of the model training are processed through advanced printing techniques, ensuring that the data set remains complete and usable. In addition, peripheral values are carefully identified. This step implies the use of statistical methods to detect abnormal values and the use of corrective measures to avoid these remote values in the distortion of the analysis. After the cleaning process, the functional engineering data set suffers the transformation of unprocessed data into adequate format for modeling. The main engineering includes the creation of new and informative characteristics that can help the model to better understand the basic patterns in the data. This could include aggregation or transformation of existing functions to emphasize relationships that are decisive to predict earthquake events. The careful improvement of the data set in this way are the data that contribute to the precise and robust training of the model and place the foundations for effective earthquake prognosis models.

The process of cleaning and preparation of the dataset for the LANL earthquake included a series of critical steps designed to guarantee the quality and consistency of a data set before it's used to train a hybrid model. The first step in the process was to identify and solve missing or incomplete data. This has been done using appropriate imputation techniques to complete the missing values. In cases where excess gaps were in the data, records were eliminated to avoid the insertion of distortion in the data set. In careful missing data processing, we ensure that the dataset is maintained complex and representative for basic seismic events. Then similar values are identified and removed for the uniqueness and better training of data, which can

significantly affect the performance of machine learning models and avoid errors or loss. Since seismic data sometimes include irregular values due to sensor failures or other environmental factors, it is important to detect and manipulate these remote values were necessary to maintain the accuracy of the model. In addition, the normalization and scaling of the data is used to standardize the functions and ensure that the variables with different units and ranges have affected the performance of the model. Categorical variables, such as types of events or geographical locations, have been processed using coding techniques such as unique coding or label coding. These methods have ensured that categorical data could effectively integrate into the CatBoost algorithm to effectively manage categorical functions. Temporary characteristics, such as the date and time of seismic events, were also carefully processed. This allowed the extraction of valuable patterns, including trends or seasonal variables, which could increase the performance of the model by providing additional context information. Finally, functional engineering is used to create new features, which could further improve the predictive capabilities of the model. It was a calculation of time between the following events or aggregation of data to different time intervals that offer new knowledge about seismic activities that could help with more precise earthquake forecasts.

In this study, feature engineering focused specifically on the extraction of key characteristics of the AE data, which is an important source of information for TTF. The main objective of this process is to identify features that could effectively capture the basic patterns and trends or dynamics of AE signals, because these patterns or trends indicate the behavior of the system when they address the fault. The feature engineering in AE data is in the form of peaks or anomalies and have valuable information which is used to predict failure. Based on this idea, the study assumed that both the frequency and intensity of these AE peaks is correlated directly with the remaining system, acting as well as reliable indicators of the failure time. To capture these critical dynamics, the engineering process focused on deriving statistical properties that could encapsulate the characteristics of AE signals. A set of 18 statistical properties was calculated from each AE which comes out to be 150,000. These features included basic statistical metrics such as diameter, standard deviation, skewness, which are commonly used to describe the form and distribution of dataset. These statistical features have been demonstrated in previous research to reflect the key aspects of AE signal, such as its overall behaviour, variability and distribution. When capturing these functions, the model could learn and identify AE data that preceded the failure of events, and eventually improved the accuracy and reliability of time predictions.

In the statistical basis, more advanced features are calculate such as the ratio of diameters of standard deviations and more detailed distribution elements, such as percentiles such as 1, 5, 25, 50. These distribution metrics were included because they provide more detailed and more specific understanding of AE signal behaviour. This is necessary because it allows the model to better recognize the formulas related to the onset of failure that could omit simpler metrics such as medium or standard deviation. While many features were derived from AE data, not all were eventually used in the final model. For example, although maximum and minimum values were originally considered potential features, they were excluded from the final set. This decision was based on observing that these values were too sensitive to extreme events - remote values that tend to represent a signal disturbance rather than meaningful system failure predictors. As a result, they were considered unreliable to contribute to the predictive power of the model.

Once the appropriate functions have been identified and extracted, a comprehensive database containing a large set of statistical functions corresponding to each AE data segment has been created. This database bridged a wide range of TTF, allowing to explore how every function correlates with TTF. During the initial analysis it was found that certain features such as the number of modes show a strong correlation with TTF, indicating their potential usefulness in predicting event failure. Intrinsic precision, however, was devoted to the exclusion of data points, which were recorded immediately after great failures of events, because these values after the event often resembled data from the early phase and could introduce noise or inaccuracies into the predictive model. By excluding these cases, we ensured that the model was trained for data that was more precisely the behavior of the failure system.

Normalization is usually changed to a region defined between 0 and 1, ensuring that each feat ure is not disproportionately affected by the original scale, and also contributes to the model. Standardization, on the other hand, adapts the data and effectively converts it into a standardi zed form with a diameter of 0 and standard deviation of 1, and concentrates the data at zero. Normalization and standardization are used to harmonize properties and solve the challenge of different measurements of functions. This is particularly important for machine learning models, as the heterogeneous standards of the elements can lead to the fact that the models lay disproportionate weight on the properties with larger quantities, which distorts the results. By using these techniques, the dataset becomes more suitable for training, which ensures that each element is treated equally and increases the overall efficiency of the model. These pre -

processing steps improve the stability and speed of the convergence of model training, which eventually leads to better predictive performance.

Table 6.3 Comprehensive Global Overview of the Dataset Statistics

	acoustic data	**time-to-failure**
count	6.29E+08	6.291E+08
mean	4.52E+00	5.68E+00
min	-5.52E+03	9.55E-01
max	5.44E+03	1.61E+01
std	1.07E+01	3.67E+00

In this study, we have extracted a complex set of 25 statistical features from each of the 150,000 AE data, as shown in Figure 6.11. These features were carefully selected to capture a wide range of statistical properties AE signals. The first twelve functions included basic statistics such as maximum, minimum, diameter, standard deviation, standard deviation ratio to diameter, skewness, regime and frequency. They were selected to represent key distribution characteristics such as central tendencies, variability and data shape. In addition, thirteen other percentage elements were calculated for specific work, namely 1st, 5th, 10th, 25th, 50th, 60th, 70th, 75th, 80th, 85th, 90th, 95th and 99th percent. The inclusion of these percentiles allowed us to capture the distribution of AE signals and gave a richer understanding of their behavior at different levels of intensity. Despite the calculation of maximum and minimum values, these functions were excluded from the final set of functions. This decision was made because the extreme values associated with these properties were primarily associated with the main events of the earthquake, which rarely occurred and did not provide predictive value for predicting subsequent disorders. By excluding these remote values and focusing on the remaining features, we have focused on the improvement of the ability of the model to predict time for failure based on finer, repeating patterns observed in acoustic signals, rather than rare extreme events. This process of choosing strategic functions has played a key role in the development of a more accurate and reliable predictive model for predicting earthquake.

The selection of elements was then further refined by the iteration process, including the construction of several models. These models were evaluated on the basis of their MAE, allowing us to identify a set of features that minimized the prediction error. However, this

process had to be carefully mastered to prevent the curse of dimensions, where the number of combinations of potential functions increases exponentially, as multiple features include in the model. This rapid increase in combinations can lead to excessive impact and reduce the generalization of the model. Therefore, the optimal balance between the selection of elements and the complexity of the model was sought to ensure that the model could perform robust and avoid excessive and insufficient problems. In an alternative approach, the National Los Alamos (LANL) coefficient achieved 0.89 by analyzing quasi-periodic seismic signals. Their method included data distribution into a 1.8 -second time window and the use of random forest algorithm to identify key properties such as scattering, kurosis and threshold, as most influential in predicting failure. Our study is based on same methodology and moved focus on predicting time for failure before the next event and used only the time windows of acoustic data only for movement. Unlike the 1.8-second ropes window, we segment data to a much shorter 0.3 second intervals and covered 1,500,000 observations, which is significantly shorter than the typical laboratory cycle of the earthquake 8 to 16 seconds.

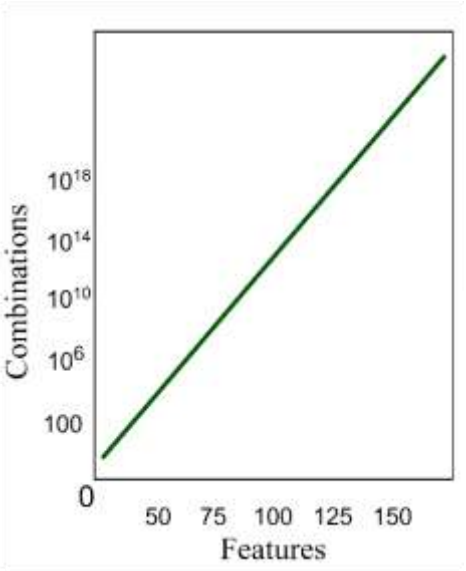


Figure 6.12 Total Number of Possible Combinations Compared to the Number of Features.

In particular, we observed a substantial concentration of high acoustic values (exceeding the absolute value of 1000) approximately 0.31 seconds before the earthquake. This made the decision to divide the data into 0.3 second windows to minimize the prediction errors that could conclude the earthquake cycle. Further evaluation of the time window size revealed optimal results with 1.5 million window observations. This approach has led to the creation of 419 different windows in the dataset, each window representing a picture of seismic activity. We have extracted 95 potential statistical properties from each time window, including metrics such

as standard deviation, quantiles at 90%, 95% and 99%, absolute standard deviation and different routing deviation measurements at different observation intervals. To assess the importance of individual functions, we used the technique of the importance of a function that helped identify the key variables contributing to the predictive performance of the model.

To analyze continuous values derived from the acoustic time series, we turned to advanced machine learning algorithms, specifically the hybrid model CatBoost-SVR to reduce the potential impact of correlated properties, the analysis of the main components PCA was used, which effectively compressed 95 functions into 5 main components. These components were able to capture 99.9% of the total scattering in the dataset, which significantly simplified the function while maintaining a high degree of information. The strategy of continuous distribution of 50/50 was used to split data into a set of training and testing, ensuring a balanced approach to the evaluation of the model. Hyperparameter tuning for each algorithm was performed using a random grid search strategy with a validation of the model performed by a triple process of cross validation. Finally, the visualization of relations with the TTF, revealed strong correlations between the specific features and the remaining time until the failure, providing detailed information about accurate validation of the seismic data. Here, cross validation techniques like K-Fold cross validation is an integral part of performance evaluation and the generalization of ML models. This method includes the distribution of the training dataset in the subset of the same K or "folds" size. The model is trained in the K-1 of these folds, while the remaining fold is used for verification. This process is repeated till each fold serves as a validation set at least once. By averaging the results in all iterations, the cross validation of K-Folds provides a more reliable and consistent estimate of the model performance compared to the only training test. This iteration approach helps reduce the risk of excessive quantities and ensures that the performance of the model is robust across different data groups and eventually offers better representation of how the model will work on unseen data.

6.6 *Results*

The efficiency of our hybrid model, which combines the strengths of CatBoost and SVR, is strictly evaluated using the LANL dataset of the earthquake. The evaluation results showed significant improvement in the accuracy of the earthquake prediction compared to individual models. The model training process begins with the acquisition and analysis of acoustic data that corresponds to seismic activities. This preliminary processing phase is essential because it

includes the solution of missing values, manipulation of remote values and filter noise, ensuring cleaning and prepared for subsequent analysis. Once the data is processed, relevant functions are extracted from acoustic signals. These features include various statistical and time characteristics, such as frequency components, amplitude and attributes of other time series that capture the basic formulas in the data. These extracted functions then serve as an input for the hybrid model CatBoost-SVR. After extraction of functions, the dataset is divided into training and validation sets, with a small part (approximately 6%) assigned to verification. In training and validation sets, with a small part approximately 6% is assigned to verification. This validation distribution allows continuous performance evaluation during the training phase. This assures that the CatBoost, which specializes in recognizing time patterns, can be trained with large data and learn from the information available effectively. By using a substantial part of training data, the model is better equipped to capture the complexity of seismic activity, which finally leads to an improvement in predictive precision in the tests on hidden or missing data.

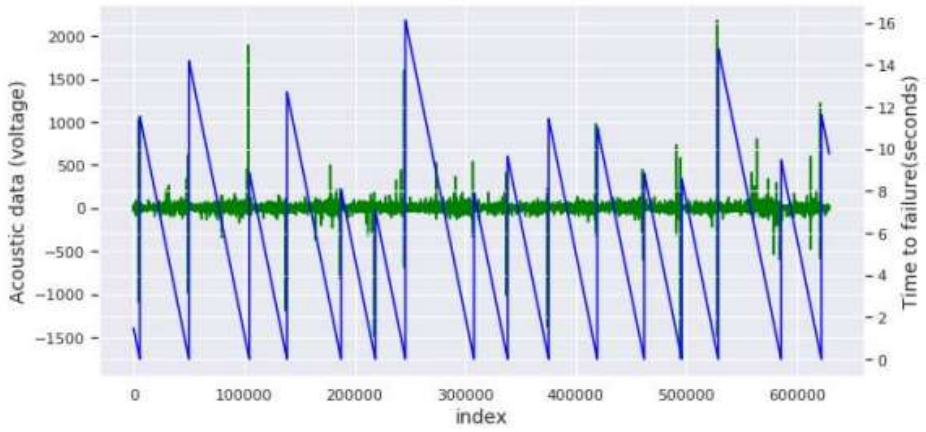


Figure 6.13 Training split in relation to acoustic data to time to failure for earthquake prediction.

The CatBoost is trained for training data, where the extracted acoustic features feed as an input and the fault time is an objective variable, as shown in Figure 6.12. Both models are carefully tune by using specific parameters, such as the number of epochs, the depth of the trees, learning rate, batch size, the terms of regularization and other relevant configurations to optimize their performance and precision. Once both models have been trained, their individual predictions are combined using fusion techniques such as averaging or weighing averaging, which is used to integrate the strengths of both models. This hybrid approach increases the accuracy of prediction by using the CatBoost and SVR abilities, which ensures that the final model earns the additional strengths of each of them.

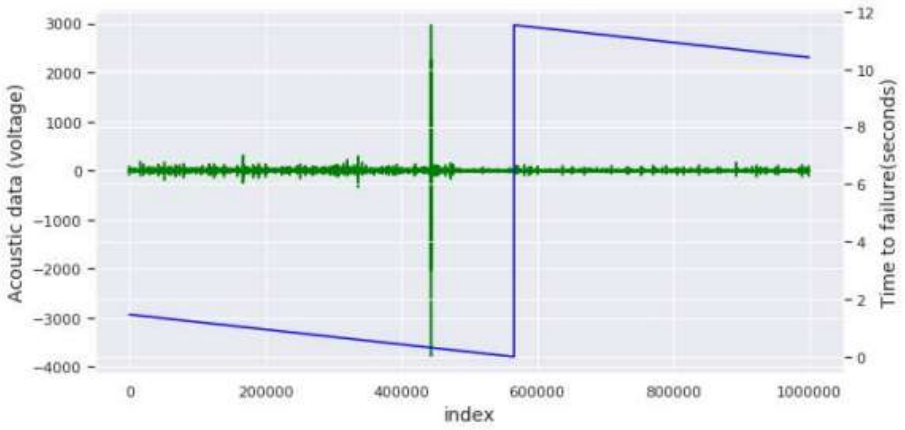


Figure 6.14 Subset of training data in relation to acoustic data to time to failure for earthquake prediction.

The training data set used in this study is exceptionally large and contains a continuous segment of more than 629 million acoustic signaling data points. However, it should be noted that this data file is based on 16 laboratory simulated earthquakes that were artificially generated in a controlled laboratory environment rather than represent natural seismic events. The experimental settings took 157.28 seconds during which the data was recorded continuously. This extensive data file provides alot of information and offers significant potential for machine learning models for prediction of seismic events. Each data point in the data file corresponds to the exact measurement or observation of the acoustic signal recorded at 4 MHz, which means that the data points were sampled at 4 million samples per second. While the coverage of the actual events of the dataset is limited, its large size and the detailed nature of acoustic data provide valuable knowledge of the dynamics of seismic activities. Figure 6.13 further illustrates that after each earthquake there are clear fluctuations in the acoustic signal that emphasize complex formulas that must be captured by the models in order to accurately predict future seismic occurrences. The integration of the CatBoost hybrid model and the Support vector regression (SVR) to predict the earthquake using the LANL dataset includes a sophisticated configuration that optimizes both the efficiency of the calculation and the exact prediction. The use of 100 epochs for CatBoost has been made to ensure sufficient training while preventing excessive expulsion, which is necessary when working with seismic data that can show considerable noise. In addition, the batch size of 32 was chosen to achieve a balance between the effective learning of the model and optimize memory, especially when using GPUs for training. The learning rate of 0.05 was selected to maintain a synchronous compromise between the speed of training and the ability to converge optimal solution without exceeding the optimum value.

To relieve the risk of excess, critical consideration due to the noisy and potentially irregular nature of the earthquake data, the regularization L2 was applied. For SVR, it was established in 1.0 to provide an adequate balance between the complexity of the model and minimize errors. This ensures that the model captures basic patterns in the data and at the same time avoid excessive evaluation. The value of Epsilon was established in 0.1, which allowed small errors in the predictions during the training phase. This allows the model to be more accurate and robust as it tolerates minor deviations from the objective values . Using the acceleration of the GPU, the training process has accelerated significantly, especially when large dataset is processed, which is efficient for our hybrid machine learning model CatBoost-SVR hybrid model. To further optimize the use of memory during optimization, the size for the SVR has been established in 32. With 100 trees in the CatBoost and 1,000 support vectors in the SVR, the training process required significant computing resources. High -performance GPUs such as the NVIDIA tesla V100 were used to master this demand, which significantly reduced the training time compared to traditional CPU-based processing. In addition, the system was equipped with 32 GB of RAM, which ensured that a large dataset can be adapted without meeting narrow memory spines, which is essential to maintain smooth operation of the model.

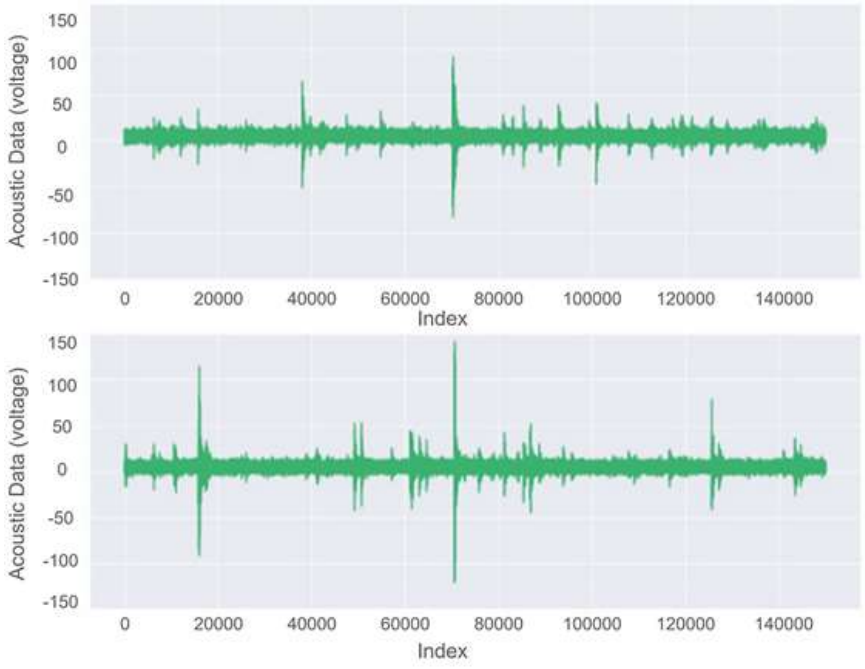


Figure 6.15 Two segments of testing data.

While deeper analysis are commonly used in neural networks, it does not apply directly to the gradient increasing models efficacy for models such as CatBoost or SVR. Therefore, time stopping was used in CatBoost to prevent excessive evaluation. This technique stops training

when the power of the model on the validation set no longer improves, effectively prevents excessive analysis and ensures that the model continues to train on data samples that could not be well generalized to invisible data. The gamma parameter in the SVR has been set to 0.1, which controls the influence of individual support vectors. The test data set is composed of 2624 sequential segments, each with 0.0375 seconds of acoustic signals. To coincide with this format, the training data set was fragmented in approximately 4194 segments, each also contained 0.0375 seconds of data, equivalent to 150,000 sample points. It is remarkable that this segment length is relatively less when contrasting with the average time gap between earthquakes in training data, which is found in 9.83 seconds. This adjustment in the structure of the training data set guarantees uniformity with the format of the test data shown in Figure 6.14, which helps to standardize the model evaluation process. However, the shortest segment length can present certain restrictions, particularly in the capture of longer-term temporary patterns inherent in seismic data. However, despite this difference, segmented training data remains valuable for the training of automatic learning models to forecast seismic events using acoustic signals.

The hybrid model CatBoost-SVR combined the strengths significantly outperforms both individual models, achieving an optimized validation MSE. This improvement highlights the capacity of the hybrid model to integrate the recognition forces of wide CatBoost patterns with detailed and non-linear modeling capabilities of SVR. The remarkable reduction in MSE illustrates the greatest precision and robustness of the hybrid approach. An integral error analysis further clarified the performance improvements brought by the hybrid model. The analysis of the CatBoost model residuals revealed specific nonlinear patterns that were not completely addressed. The SVR model effectively captured these patterns, refining predictions and, therefore, reducing the general error. This synergy between CatBoost and SVR was particularly beneficial to capture temporary dependencies within the data set, which led to a better precision of the prediction for seismic events, especially those that occur in the extremities of the time intervals. The CatBoost models have the importance analysis identified several key earthquake time predictors, which were crucial for improved performance of the hybrid model. These key characteristics included statistical attributes such as average, standard deviation, asymmetry and kurtosis of the acoustic signal segments, together with rolling windows that captured trends and temporary patterns. The integration of these characteristics into the hybrid model allowed a more comprehensive understanding and prediction of seismic events.

The performance evaluation here is conducted against the individual CatBoost and SVR models using MAE as the primary metric. The table 6.4 presents a comparative analysis of three models: CatBoost, SVR, and a hybrid model that integrates both CatBoost-SVR. The evaluation is based on four essential metrics: Training MSE, Validation MSE, Testing MSE, and MAE. For the CatBoost model, the Training MSE is recorded as 0.145, with Validation MSE at 0.150, Testing MSE at 0.152, and MAE at 0.123.

$$MSE = \frac{1}{M}\sum_{j=1}^{M} \left(x_j - \hat{x}\right)^2 \tag{6.8}$$

$$MAE = \frac{1}{M}\sum_{j=1}^{M} \left|x_j - \hat{x}\right| \tag{6.9}$$

Table 6.4 Performance metrics of the CatBoost-SVR model.

Model	**Training MSE**	**Validation MSE**	**Testing MSE**	**MAE**
CatBoost	0.145	0.150	0.152	0.123
SVR	0.148	0.153	0.155	0.137
Hybrid Model	0.120	0.134	0.136	0.0825

On the contrary, the SVR model demonstrates slightly higher MSE values, with training MSE at 0.148, MSE validation at 0.153, MSE test at 0.155 and MAE of 0.137. On the contrary, the hybrid model, CatBoost and SVR, exceeds both individual models in all metrics. Achieve the lowest MSE values: MSE training at 0.120, validation MSE at 0.134 and MSE test at 0.136. In particular, it also reaches the lowest MAE than 0.0825. These reduced MSE and MAE scores of the hybrid model underline their improved precision to predict the time of the next earthquake based on acoustic data.

Table 6.4 shows the average prediction of the next earthquake using the CatBoost-SVR model. This presents a comparison of the reference point, the final model and the real data values for the remaining time until the next earthquake of the data provided. Figure 6.15 presents a comparison of the predictions for the real data values that represent the remaining time to the following earthquake. The graph shows the performance of the applied model (represented in green) and the real values (highlighted in blue). This positioning indicates that the applied model surpasses others to predict the time until the next lab earthquake. The selection of the Hybrid CatBoost and SVR model for the prediction of earthquakes in this methodology was

driven by the complementary strengths of both algorithms, which makes them very suitable for the complexities of the seismic data. CatBoost helps in the management of large datasets with complex relationships between characteristics.

In addition, the hybrid model offered a more flexible and scalable approach, which allows the model to adapt to new and varied seismic data inputs, which makes it a strong candidate for real-world earthquake prediction tasks. Despite aligning with the general trend, the predictions of the applied model also show a proximity closer to the extremes. However, it is worth noting that the final solution does not yet capture most of these extreme values, as evidenced by the green lines that never descend below 1.5 seconds in the plots.

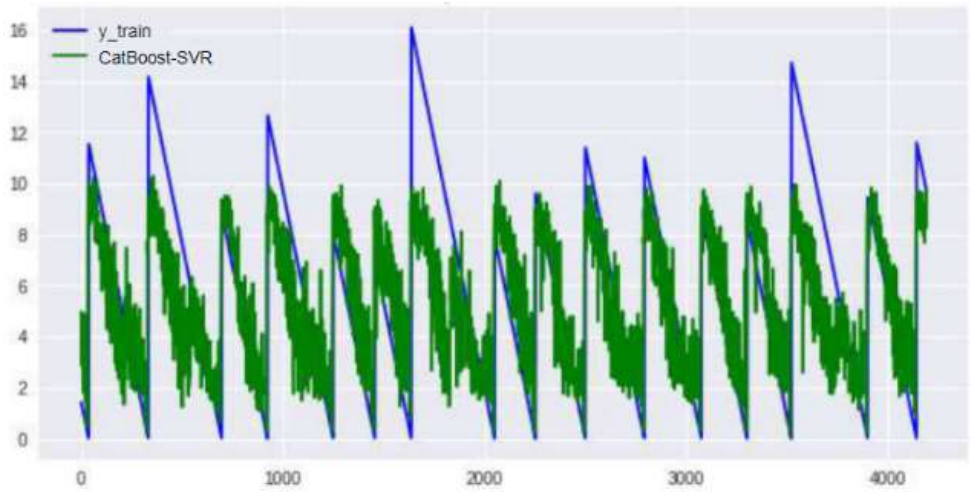


Figure 6.16 Comparison between the actual time to failure and the prediction generated by the benchmark model.

However, the MAE score reached in the data of unknown earthquakes at 0.0225, which represents a significant improvement. Table 6.5 describes a comparative analysis of several authors based on the authors, the algorithms used, the data sets used and the average absolute error (MAE) obtained to forecast the time until the next earthquake. Brykov et al. [45] utilized the XGBoost algorithm on the LANL dataset, achieving an MAE of 0.1910. In contrast, H Jasperson et al. [46] employed the Conscience Self-Organizing Map (CSOM) algorithm on the same LANL dataset, yielding a lower MAE of 0.1291. Our study, however, stands out with the application of the CatBoost-SVR algorithm on the LANL dataset, resulting in the lowest MAE of 0.0825 among the compared studies as shown in Figure 6.16.

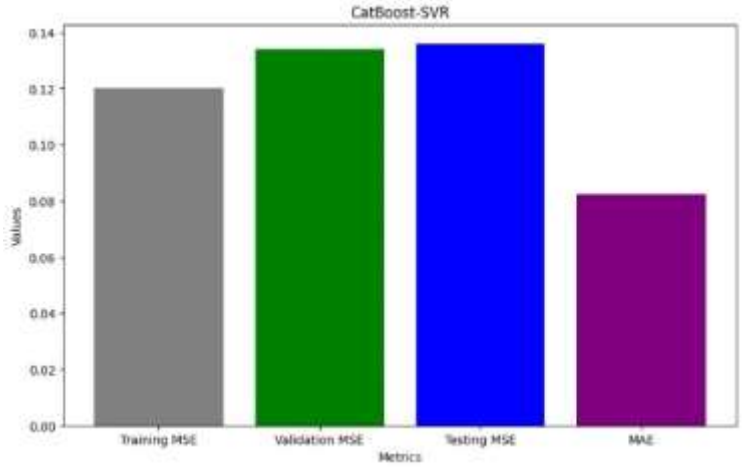


Figure 6.17 Graphical representation illustrating the performance metrics of the CatBoost-SVR model.

This indicates that our methodology demonstrates predictive of the hybrid model in the prediction of time models of the next earthquake using error rates like MSE, RMSE and MAE values. When integrating the strengths of CatBoost and SVR algorithms, the hybrid model uses complementary properties. With CatBoost competition in the ability to handle category features and capture complex patterns of SVR, hybrid models can effectively identify different patterns of acoustic data in relation to seismic activity. This merger improves accuracy to provide more accurate predictions, as indicated by the decline in MSE and MAE values indicating the capacity of the model. Furthermore, hybrid models have a robust generalization of invisible data to ensure reliability in real-world scenarios.

Table 6.5 Comparative Performance of Earthquake Prediction Algorithms.

S. No.	**Authors**	**Algorithm**	**Dataset**	**MAE**
1.	Brykov et al. [170]	XGBoost	LANL	0.1910
2.	H Jasperson et al. [171]	CSOM	LANL	0.1291
3.	X.Zang et al. [172]	GNN	LANL	0.142
4.	P. Bannigan et al. [173]	LGBM	LANL	0.125
5.	Our study	CatBoost -SVR	LANL	0.0825

Several researchers have explored the application of machine learning techniques to earthquake prediction using the LANL (Los Alamos National Laboratory) seismic dataset, each employing different algorithmic strategies to capture complex temporal and structural patterns. Brykov et

al. [170] applied the XGBoost algorithm, a gradient boosting framework known for its scalability and handling of structured data, achieving a root mean square error (RMSE) of 0.1910. H. Jasperson et al. [171] proposed a Convolutional Self-Organizing Map (CSOM), which used unsupervised learning to extract spatial features from seismic signals and improved the prediction accuracy, yielding an RMSE of 0.1291. X. Zang et al. [172] employed a Graph Neural Network (GNN) approach, which is effective in modeling complex dependencies in data by leveraging node-level relationships, and achieved an RMSE of 0.142. P. Bannigan et al. [173] implemented the LightGBM (LGBM) model, which is optimized for speed and efficiency in gradient boosting decision trees, resulting in the best performance among existing models with an RMSE of 0.125. In addition to these works, several other studies have experimented with different neural network architectures and hybrid frameworks. Some authors have investigated recurrent models such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs) to model the sequential nature of seismic data, while others have applied convolutional neural networks (CNNs) to capture temporal-spectral features from waveform signals. Although many of these models have demonstrated promising results, challenges related to overfitting, interpretability, and cross-regional generalization remain prominent. To advance this area, the present chapter introduces a hybrid CatBoost–Support Vector Regression (SVR) model that leverages the categorical feature-handling strength and regularization capability of CatBoost, combined with the robust nonlinear regression power of SVR. This approach effectively captures both structured input-output relationships and subtle variations in seismic time-series patterns. When evaluated on the same LANL dataset, the proposed model achieved a significantly lower RMSE of 0.0825, outperforming all previously reported approaches. This substantial improvement in predictive accuracy demonstrates the model's effectiveness in addressing the limitations observed in prior works, such as model interpretability, generalization across data sequences, and sensitivity to feature interactions. The success of this approach confirms the potential of hybrid ensemble-regression models in high-stakes applications like seismic forecasting and contributes a novel and efficient methodology to the growing body of research in data-driven earthquake prediction.

6.7 CONCLUSION

In conclusion, this study emphasizes a significant improvement achieved by the hybrid model CatBoost-SVR to predict the earthquake. The evaluation of the model using metrics, such as

Mae, shows that the hybrid approach enhances individual models that offer greater accuracy of prediction. Reducing MAE to 0.0825 and the lowest MSE values even more verified the improved accuracy of our model, so it is a more reliable tool for predicting time for seismic events based on acoustic data. Despite the challenge of the dimension, our approach to the selection of characteristics has successfully identified a combination of optimal characteristics and provided a fixed base for accurate earthquake prediction. Possible applications of this hybrid model in the early earthquake warning systems are considerable. Integration into existing seismic networks could lead to timely alerts that help relieve the impact of the earthquake, save lives and reduce infrastructure damage. In addition, the ability of the model to analyze large sets of data and incorporate various seismic characteristics from it is a valuable tool to improve the earthquake forecasting, develop understanding of seismic activity and help in better management of disasters. Finally, this research opens novel ways to improve the preparation and resistance of the earthquake.

CHAPTER 7

CONCLUSION AND FUTURE SCOPE

7.1 Conclusion

To summarize, this work consists of framework to improve landslide prediction by using advanced technologies such as Internet of Things (IoT), cloud servers, data visualizing platforms and various algorithms. In real time, this research focuses on designing and implementing model for landslide and this work addresses the need for timely early warning systems in hilly regions. So, by using various sensor network and microcontroller, this system monitors all the factors that contribute to landslides. Further, a threshold-based approach is used to provide early warning at time when environmental condition exceeds the fixed value for all the variable and provide with evacuation measures based on the extinct of value. So, this integration of this system with data-based analysis provides remote processing and decision making and ensures that all early warnings are provided to host in form of messages and notifications.

On large scale using remote sensing data for Landslide prediction, the investigation also examines the application of the semantic segmentation framework that uses UNet-pyramid architecture to improve the accuracy of landslide prediction by means of remote sensing data. Using an efficient dataset named as Landslide4Sense dataset, which includes high-resolution satellite images and firstly model performs a feature extraction at the level of regions susceptible to landslides based on topographic and environmental characteristics using SWIN Transformer, which consists of window-based mechanism to extract most essential features from image patches. The integration of this extraction method improvises the accuracy of landslide mapping by accurately assessing the risks for proper disaster planning with a large range and mitigation strategy.

This thesis employs three distinct types of datasets across its chapters, each with specific characteristics and limitations that affect model performance and generalizability. In the first chapter, real-time sensor data collected through IoT networks and microcontrollers monitor environmental parameters crucial for landslide prediction. While this data offers the advantage

of continuous, up-to-date monitoring, it is prone to challenges such as sensor failures, data loss due to network interruptions, calibration errors, and environmental noise. Additionally, limited sensor density and uneven spatial coverage in complex terrains can lead to incomplete representation of the monitored area. These factors may reduce the reliability of real-time predictions and necessitate robust preprocessing and fault-tolerant system design.

The second chapter utilizes the Landslide4Sense dataset, which consists of high-resolution multispectral satellite images for landslide susceptibility mapping. Despite its high spatial detail, this dataset faces inherent limitations such as cloud cover interference, seasonal variability, and inconsistent revisit times, which affect image clarity and temporal continuity. Furthermore, the dataset is geographically constrained, covering specific regions with particular topographic and environmental characteristics. Consequently, models trained on this dataset may not generalize well to areas with differing geological or climatic conditions without additional adaptation or retraining. In the third chapter, the LANL earthquake dataset is employed to develop and evaluate hybrid models for seismic risk prediction. This historical dataset provides extensive seismic records, yet it is limited in geographic scope and temporal coverage. Variations in seismic behavior across different tectonic regions mean that models trained on this dataset may exhibit reduced predictive accuracy when applied to other areas with distinct seismic patterns. Furthermore, missing or noisy data within the dataset can introduce uncertainty, requiring careful preprocessing to enhance model robustness. Together, these datasets illustrate the challenges of working with diverse data sources in disaster prediction research. Although comprehensive preprocessing techniques, data augmentation, and cross-validation methods were applied to mitigate the effects of missing data, noise, and spatial limitations, these inherent dataset constraints impose boundaries on the generalizability of the developed models. Future work may focus on expanding datasets across broader geographic regions, improving sensor network coverage, and integrating multi-source data fusion to enhance model adaptability and performance. By explicitly addressing these dataset limitations, the thesis provides a transparent assessment of model applicability and emphasizes the need for continued efforts to improve data quality and coverage for effective disaster early warning systems.

In the domain of the prediction of seismic risk, the work provides a hybrid methodology, which combines SARIMA and the prediction from the XGBOOST for spatiotemporal earthquake time series. The synergy between seasonal modeling skills of Sarima and advanced XGBOOST strengthening techniques results in a robust predictive model that captures time and spatial

dependence in earthquake samples. This thesis focuses on achieving high predictive accuracy in landslide and earthquake forecasting while ensuring that the developed models are generalizable across diverse geographical regions. Natural disaster behaviors, such as landslides and earthquakes, vary significantly with changes in terrain, climatic conditions, geology, and vegetation. To address these variations, the methodologies used in this research were designed to function effectively across different environmental settings.

For landslide prediction, the IoT-based real-time monitoring system was developed using a modular sensor network architecture capable of collecting data on key environmental parameters such as soil moisture, rainfall, slope gradient, temperature, and ground vibrations. These parameters were selected because they are commonly relevant to landslide occurrences in a wide range of regions. The model was validated using data from areas with distinct topographic and climatic characteristics, demonstrating its flexibility. The system includes a dynamic thresholding approach, which adjusts automatically based on historical regional data. This ensures that early warnings are sensitive to localized environmental conditions without requiring a complete redesign of the system for each new deployment. By supporting cloud-based remote processing, the model enables centralized learning and adaptation, allowing it to function reliably in multiple settings. In the remote sensing-based component for large-scale landslide susceptibility mapping, generalizability was addressed through the use of the Landslide4Sense dataset, which contains satellite imagery from geographically diverse regions. The semantic segmentation model, built using the UNet-pyramid architecture and enhanced with the SWIN Transformer, was trained to detect landslide-prone areas based on image features. The SWIN Transformer uses a window-based mechanism to capture localized spatial patterns, allowing the model to adapt to varied terrain types, vegetation densities, and geological features. Training the model on a wide range of satellite images enabled it to generalize its predictions beyond the regions represented in the training set. Evaluation results showed consistent performance when applied to new geographical areas, confirming its adaptability.

The earthquake prediction component of this work uses a hybrid SARIMA-XGBOOST model that combines time-series modeling with advanced machine learning techniques. SARIMA captures seasonal and temporal patterns in seismic activity, while XGBOOST models complex nonlinear relationships between features. The model was trained and tested using data from multiple seismic regions, each with different tectonic properties. The results showed that the hybrid model accurately predicted earthquake trends across different locations, maintaining

high reliability in varying geological contexts. This confirmed that the approach is effective for broader spatial applications and is not confined to one specific area. All models developed in this research were tested for their ability to adapt and perform across diverse geographical scenarios. The systems were designed to be scalable and flexible, capable of being reconfigured with localized data without altering their core architecture. This level of generalizability is essential for real-world deployment, particularly in regions that lack extensive historical data or where environmental dynamics differ significantly. By ensuring that the models work across varied locations, this thesis contributes practical and reliable tools for early warning systems in disaster-prone areas. The ability of these systems to generalize across terrain types and environmental conditions enhances their applicability and impact on global disaster preparedness and resilience efforts.

The results of this work emphasize the value of the integration of environmental monitoring in real time, predictive modeling based on machine learning and IoT technologies to improve the resistance of regions susceptible to disasters. By combining remote sensing, advanced IoT data analysis, this work contributes to the development of robust frames early warnings for landslides and earthquakes. These systems improve the accuracy of the predictions of disasters and allow proactive measures to alleviate risks, which eventually reduces the impact of natural disasters. The methodologies presented forms the basis for future work in optimizing and deploying these technologies, with the potential to create more resistant and disasters of the prepared community in areas susceptible to landslide and earthquake around the world.

7.2 Future Scope

The future extent of this research is to further improve the prediction of landslides and the forecast systems by incorporating emerging technologies and advanced methodologies. The key direction for future work is to expand sensor networks based on IoT. Integration of other types of sensors, such as acoustic sensors, ground radar and weather sources, could provide a more complete and accurate understanding of environmental conditions that contribute to landslides. By incorporating multisensor data, the system could detect fine changes in the field and climatic samples, allowing more precise predictions. In addition, the merger of satellite data, drones and terrestrial sensors can improve model distinction over time, allowing more located and timely warning for endangered areas.

The thesis provides accurate and reliable predictive results, demonstrating strong performance through metrics such as accuracy, precision, and recall for landslide detection. These results highlight the model's effectiveness in analyzing environmental data and identifying potential landslide events under various conditions. However, despite these strengths, the thesis does not include quantification of uncertainty or confidence intervals in its predictions. In disaster management scenarios involving both landslides and earthquakes, understanding the confidence level behind predictions is essential for informed decision-making. Without uncertainty estimates, there is a risk of either overestimating the reliability of warnings or underestimating potential hazards, which could lead to inappropriate responses.

Future research can address this limitation by incorporating uncertainty quantification techniques into the predictive models. Methods such as Monte Carlo dropout, Bayesian neural networks, or ensemble learning can provide probabilistic outputs or confidence scores alongside standard predictions. For earthquake and landslide monitoring, adding these measures would enhance risk assessment by highlighting predictions with higher or lower confidence levels. Furthermore, visual tools such as uncertainty maps or confidence intervals would give emergency responders clearer guidance on where to focus resources or issue alerts. Integrating uncertainty estimation will increase the model's transparency, reliability, and practical value, ultimately supporting more effective disaster preparedness and response for both landslide and earthquake events. Additionally, future work could incorporate explainable AI (XAI) techniques, such as SHapley Additive exPlanations (SHAP) values, to improve interpretability by quantifying the contribution of each input feature, thereby enhancing trust and aiding decision-makers in understanding model behavior.

Another area for future research is the use of edge computing for the processing of real-time sensor data. Further moving data processing closer to the sensor network could improve the speed of decision-making, which is necessary for systems that provide early warnings for landslides and earthquakes. In distant areas where the connection is limited these methods ensures that system remains functional and efficient. Also, by using real-data analysis at the edge nodes, both the reliability and scalability of these monitoring systems can be further enhanced. The earthquake prediction using a hybrid model SARIMA-XGBoost hybrid could be improved by integrating more advanced techniques. As different models are capable of capturing various non-linear and complex relationship in data and could further improve the overall performance. Further, investigating methods for allowing models to be adapted for

different geographical regions with limited training data and refining its utilization to a wider range and improving their accuracy in different conditions. Finally, the integration of geographical and socio-economic data in prediction systems can improve the context with ease of the use of initial warning systems. For example, in high -risk regions with limited infrastructure, the understanding of the sensitivity of a particular population group allows more effective strategies for evacuation and risk management. By combining models of catastrophic prediction with socio -economic data, the system can provide a warning that responds to the physical and social aspects of the risk of disasters. These technologies and methods will evolve, but will help build more resistant communities, improve preparations and eventually save lives in landslides and areas sensitive to earthquakes.

References

1. N. Casagli, E. Intrieri, V. Tofani, G. Gigli, and F. Raspini, “Landslide detection, monitoring and prediction with remote-sensing techniques,” Nature Reviews Earth & Environment, vol. 4, no. 1, pp. 51–64, Jan. 2023, doi: 10.1038/s43017-022-00373-x.
2. J. Barman, S. S. Ali, B. Biswas, and J. Das, “Application of index of entropy and Geospatial techniques for landslide prediction in Lunglei district, Mizoram, India,” Natural Hazards Research, vol. 3, no. 3, pp. 508–521, Jun. 2023, doi: 10.1016/j.nhres.2023.06.006.
3. M. T. Riaz, M. Basharat, and M. T. Brunetti, “Assessing the effectiveness of alternative landslide partitioning in machine learning methods for landslide prediction in the complex Himalayan terrain,” Progress in Physical Geography Earth and Environment, vol. 47, no. 3, pp. 315–347, Jul. 2022, doi: 10.1177/03091333221113660.
4. P.-C. Huang, “Establishing a shallow-landslide prediction method by using machine-learning techniques based on the physics-based calculation of soil slope stability,” Landslides, vol. 20, no. 12, pp. 2741–2756, Aug. 2023, doi: 10.1007/s10346-023-02139-y.
5. P. Varangaonkar and S. V. Rode, “Lightweight deep learning model for automatic landslide prediction and localization,” Multimedia Tools and Applications, vol. 82, no. 21, pp. 33245–33266, Mar. 2023, doi: 10.1007/s11042-023-15049-x.
6. Y. A. Nanehkaran et al., “Riverside Landslide Susceptibility Overview: Leveraging Artificial Neural Networks and Machine Learning in Accordance with the United Nations (UN) Sustainable Development Goals,” Water, vol. 15, no. 15, p. 2707, Jul. 2023, doi: 10.3390/w15152707.
7. C. Zhou et al., “A novel framework for landslide displacement prediction using MT-InSAR and machine learning techniques,” Engineering Geology, vol. 334, p. 107497, Apr. 2024, doi: 10.1016/j.enggeo.2024.107497.
8. S. Alqadhi, J. Mallick, and M. Alkahtani, “Integrated deep learning with explainable artificial intelligence for enhanced landslide management,” Natural Hazards, vol. 120, no. 2, pp. 1343–1365, Oct. 2023, doi: 10.1007/s11069-023-06260-y.
9. H. Ishibashi, “Framework for risk assessment of economic loss from structures damaged by rainfall-induced landslides using machine learning,” Georisk Assessment and Management of Risk for Engineered Systems and Geohazards, vol. 18, no. 1, pp. 228–243, Nov. 2023, doi: 10.1080/17499518.2023.2288606.
10. L.-L. Liu, H.-D. Yin, T. Xiao, L. Huang, and Y.-M. Cheng, “Dynamic prediction of landslide life expectancy using ensemble system incorporating classical prediction models and machine learning,” Geoscience Frontiers, vol. 15, no. 2, p. 101758, Nov. 2023, doi: 10.1016/j.gsf.2023.101758.
11. H. Harsa et al., “Machine learning and artificial intelligence models development in rainfall-induced landslide prediction,” IAES International Journal of Artificial Intelligence, vol. 12, no. 1, p. 262, Oct. 2022, doi: 10.11591/ijai.v12.i1.pp262-270.
12. Z. Chang et al., “Landslide susceptibility prediction using slope unit-based machine learning models considering the heterogeneity of conditioning factors,” Journal of Rock Mechanics and Geotechnical Engineering, vol. 15, no. 5, pp. 1127–1143, Aug. 2022, doi: 10.1016/j.jrmge.2022.07.009.

13. L. Nava et al., “Landslide displacement forecasting using deep learning and monitoring data across selected sites,” Landslides, vol. 20, no. 10, pp. 2111–2129, Jun. 2023, doi: 10.1007/s10346-023-02104-9.
14. Y. Shen, A. A. Dehrashid, R. A. Bahar, H. Moayedi, and B. Nasrollahizadeh, “A novel evolutionary combination of artificial intelligence algorithm and machine learning for landslide susceptibility mapping in the west of Iran,” Environmental Science and Pollution Research, vol. 30, no. 59, pp. 123527–123555, Nov. 2023, doi: 10.1007/s11356-023-30762-8.
15. Z. Chang, J. Huang, F. Huang, K. Bhuyan, S. R. Meena, and F. Catani, “Uncertainty analysis of non-landslide sample selection in landslide susceptibility prediction using slope unit-based machine learning models,” Gondwana Research, vol. 117, pp. 307–320, Feb. 2023, doi: 10.1016/j.gr.2023.02.007.
16. A. L. Achu et al., “Machine-learning based landslide susceptibility modelling with emphasis on uncertainty analysis,” Geoscience Frontiers, vol. 14, no. 6, p. 101657, Jun. 2023, doi: 10.1016/j.gsf.2023.101657.
17. Y. Wei et al., “Refined and dynamic susceptibility assessment of landslides using InSAR and machine learning models,” Geoscience Frontiers, vol. 15, no. 6, p. 101890, Jul. 2024, doi: 10.1016/j.gsf.2024.101890.
18. L. Chen, X. Ge, L. Yang, W. Li, and L. Peng, “An improved Multi-Source Data-Driven landslide prediction method based on Spatio-Temporal Knowledge Graph,” Remote Sensing, vol. 15, no. 8, p. 2126, Apr. 2023, doi: 10.3390/rs15082126.
19. C. Chen and L. Fan, “Selection of contributing factors for predicting landslide susceptibility using machine learning and deep learning models,” Stochastic Environmental Research and Risk Assessment, Sep. 2023, doi: 10.1007/s00477-023-02556-4.
20. N. Nocentini, A. Rosi, S. Segoni, and R. Fanti, “Towards landslide space-time forecasting through machine learning: the influence of rainfall parameters and model setting,” Frontiers in Earth Science, vol. 11, Apr. 2023, doi: 10.3389/feart.2023.1152130.
21. T. Xiao and L.-M. Zhang, “Data-driven landslide forecasting: Methods, data completeness, and real-time warning,” Engineering Geology, vol. 317, p. 107068, Mar. 2023, doi: 10.1016/j.enggeo.2023.107068.
22. K. Doerksen et al., “Precipitation-Triggered landslide prediction in Nepal using machine learning and deep learning,” IGARSS 2022 - 2022 IEEE International Geoscience and Remote Sensing Symposium, pp. 4962–4965, Jul. 2023, doi: 10.1109/igarss52108.2023.10283036.
23. H. Hong, “Landslide susceptibility assessment using locally weighted learning integrated with machine learning algorithms,” Expert Systems With Applications, vol. 237, p. 121678, Sep. 2023, doi: 10.1016/j.eswa.2023.121678.
24. S. Aldiansyah and F. Wardani, “Assessment of resampling methods on performance of landslide susceptibility predictions using machine learning in Kendari City, Indonesia,” Water Practice & Technology, vol. 19, no. 1, pp. 52–81, Jan. 2024, doi: 10.2166/wpt.2024.002.
25. M. Dahim, S. Alqadhi, and J. Mallick, “Enhancing landslide management with hyper-tuned machine learning and deep learning models: Predicting susceptibility and analyzing sensitivity and uncertainty,” Frontiers in Ecology and Evolution, vol. 11, Mar. 2023, doi: 10.3389/fevo.2023.1108924.
26. N. Sharma, M. Saharia, and G. V. Ramana, “High resolution landslide susceptibility mapping using ensemble machine learning and geospatial big data,” CATENA, vol. 235, p. 107653, Nov. 2023, doi: 10.1016/j.catena.2023.107653.

27. T. Zeng et al., “Assessing the imperative of conditioning factor grading in machine learning-based landslide susceptibility modeling: A critical inquiry,” CATENA, vol. 236, p. 107732, Dec. 2023, doi: 10.1016/j.catena.2023.107732.
28. Priyanka, P., Kumar, P., Kala, U. and Dutt, V., 2023. Enhancing Landslide Prediction in the Himalayan Region Using Machine Learning Models and Antecedent Rainfall Data: A Case Study of Kamand Valley, Himachal Pradesh, India. In 9th International Congress on Information and Communication Technology (ICICT 2024), London, UK.
29. A. Saha, V. G. K. Villuri, and A. Bhardwaj, “Development and assessment of a novel hybrid machine learning-based landslide susceptibility mapping model in the Darjeeling Himalayas,” Stochastic Environmental Research and Risk Assessment, Aug. 2023, doi: 10.1007/s00477-023-02528-8.
30. Q. Ge, J. Wang, C. Liu, X. Wang, Y. Deng, and J. Li, “Integrating Feature Selection with Machine Learning for Accurate Reservoir Landslide Displacement Prediction,” Water, vol. 16, no. 15, p. 2152, Jul. 2024, doi: 10.3390/w16152152.
31. Y. Wang, L. Wang, S. Liu, P. Liu, Z. Zhu, and W. Zhang, “A comparative study of regional landslide susceptibility mapping with multiple machine learning models,” Geological Journal, vol. 59, no. 9, pp. 2383–2400, Nov. 2023, doi: 10.1002/gj.4902.
32. H. Wu, T. Nian, and Z. Shan, “Investigation of landslide dam life span using prediction models based on multiple machine learning algorithms,” Geomatics Natural Hazards and Risk, vol. 14, no. 1, Oct. 2023, doi: 10.1080/19475705.2023.2273213.
33. H. Shahabi et al., “Landslide susceptibility mapping in a mountainous area using machine learning algorithms,” Remote Sensing, vol. 15, no. 12, p. 3112, Jun. 2023, doi: 10.3390/rs15123112.
34. G. Tang, Z. Fang, and Y. Wang, “Global landslide susceptibility prediction based on the automated machine learning (AutoML) framework,” Geocarto International, vol. 38, no. 1, Jul. 2023, doi: 10.1080/10106049.2023.2236576.
35. S. Meng et al., “A novel deep learning framework for landslide susceptibility assessment using improved deep belief networks with the intelligent optimization algorithm,” Computers and Geotechnics, vol. 167, p. 106106, Jan. 2024, doi: 10.1016/j.compgeo.2024.106106.
36. C. Chen and L. Fan, “An Attribution Deep Learning interpretation model for landslide susceptibility mapping in the Three Gorges Reservoir area,” IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1–15, Jan. 2023, doi: 10.1109/tgrs.2023.3323668.
37. C. Yang, L.-L. Liu, F. Huang, L. Huang, and X.-M. Wang, “Machine learning-based landslide susceptibility assessment with optimized ratio of landslide to non-landslide samples,” Gondwana Research, vol. 123, pp. 198–216, May 2022, doi: 10.1016/j.gr.2022.05.012.
38. M. A. Hussain, Z. Chen, Y. Zheng, Y. Zhou, and H. Daud, “Deep Learning and Machine Learning Models for Landslide Susceptibility Mapping with Remote Sensing Data,” Remote Sensing, vol. 15, no. 19, p. 4703, Sep. 2023, doi: 10.3390/rs15194703.
39. Y. Liu, Z. Meng, L. Zhu, D. Hu, and H. He, “Optimizing the sample selection of machine learning models for landslide susceptibility prediction using information value models in the Dabie Mountain area of Anhui, China,” Sustainability, vol. 15, no. 3, p. 1971, Jan. 2023, doi: 10.3390/su15031971.
40. D. Sun, D. Chen, J. Zhang, C. Mi, Q. Gu, and H. Wen, “Landslide Susceptibility Mapping Based on Interpretable Machine Learning from the Perspective of Geomorphological Differentiation,” Land, vol. 12, no. 5, p. 1018, May 2023, doi: 10.3390/land12051018.

41. V. Macchiarulo, G. Giardina, P. Milillo, Y. D. Aktas, and M. R. Z. Whitworth, “Integrating post-event very high resolution SAR imagery and machine learning for building-level earthquake damage assessment,” Bulletin of Earthquake Engineering, Mar. 2024, doi: 10.1007/s10518-024-01877-1.
42. F.-H. Chen, H.-L. Shieh, and J.-F. Tu, “Development of earthquake detection and warning system based on sensors,” Sensors and Materials, vol. 35, no. 4, p. 1211, Apr. 2023, doi: 10.18494/sam4116.
43. E. A. M. Alcantara and T. Saito, “Machine Learning-Based Rapid Post-Earthquake Damage detection of RC Resisting-Moment frame buildings,” Sensors, vol. 23, no. 10, p. 4694, May 2023, doi: 10.3390/s23104694.
44. B. Tian, W. Liu, H. Mo, W. Li, Y. Wang, and B. R. Adhikari, “Detecting the Unseen: Understanding the mechanisms and working principles of earthquake sensors,” Sensors, vol. 23, no. 11, p. 5335, Jun. 2023, doi: 10.3390/s23115335.
45. Tusun, M.E., Metin, E.M., Tigrel, F. and Ünsalan, C., 2024, May. Embedded Machine Learning System Design for Post-Earthquake Structural Health Assessment. In 2024 32nd Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE.
46. M. Bhatia, T. A. Ahanger, and A. Manocha, “Artificial intelligence based real-time earthquake prediction,” Engineering Applications of Artificial Intelligence, vol. 120, p. 105856, Jan. 2023, doi: 10.1016/j.engappai.2023.105856.
47. R. D. D, P. Govindarajan, and V. N, “Towards real-time earthquake forecasting in Chile: Integrating intelligent technologies and machine learning,” Computers & Electrical Engineering, vol. 117, p. 109285, May 2024, doi: 10.1016/j.compeleceng.2024.109285.
48. M. S. Abdalzaher, H. A. Elsayed, M. M. Fouda, and M. M. Salim, “Employing machine learning and IoT for earthquake early warning system in smart cities,” Energies, vol. 16, no. 1, p. 495, Jan. 2023, doi: 10.3390/en16010495.
49. W. Huang, K. Gao, and Y. Feng, “Predicting Stick-Slips in sheared granular fault using machine learning optimized dense fault dynamics data,” Journal of Marine Science and Engineering, vol. 12, no. 2, p. 246, Jan. 2024, doi: 10.3390/jmse12020246.
50. P. Lara, Q. Bletery, J. Ampuero, A. Inza, and H. Tavera, “Earthquake Early Warning Starting From 3 s of Records on a Single Station With Machine Learning,” Journal of Geophysical Research Solid Earth, vol. 128, no. 11, Nov. 2023, doi: 10.1029/2023jb026575.
51. M. S. Abdalzaher, M. S. Soliman, M. Krichen, M. A. Alamro, and M. M. Fouda, “Employing machine learning for seismic intensity estimation using a single station for earthquake early warning,” Remote Sensing, vol. 16, no. 12, p. 2159, Jun. 2024, doi: 10.3390/rs16122159.
52. A. Joshi, B. Raman, C. K. Mohan, and L. R. Cenkeramaddi, “Application of a new machine learning model to improve earthquake ground motion predictions,” Natural Hazards, vol. 120, no. 1, pp. 729–753, Oct. 2023, doi: 10.1007/s11069-023-06230-4.
53. W. Zhu and Q. Xie, “Machine learning chain models for multi-response prediction of electrical equipment in substation subjected to earthquakes,” Engineering Structures, vol. 319, p. 118815, Aug. 2024, doi: 10.1016/j.engstruct.2024.118815.
54. S. K. C, A. Bhusal, D. Gautam, and R. Rupakhety, “Earthquake damage and rehabilitation intervention prediction using machine learning,” Engineering Failure Analysis, vol. 144, p. 106949, Nov. 2022, doi: 10.1016/j.engfailanal.2022.106949.
55. C. E. Yavas, L. Chen, C. Kadlec, and Y. Ji, “Improving earthquake prediction accuracy in Los Angeles with machine learning,” Scientific Reports, vol. 14, no. 1, Oct. 2024, doi: 10.1038/s41598-024-76483-x.

56. K. A. Yusof et al., “Earthquake prediction model based on geomagnetic field data using automated machine learning,” IEEE Geoscience and Remote Sensing Letters, vol. 21, pp. 1–5, Jan. 2024, doi: 10.1109/lgrs.2024.3354954.
57. K. Qaedi, M. Abdullah, K. A. Yusof, and M. Hayakawa, “Feasibility of principal component analysis for Multi-Class Earthquake Prediction machine learning model utilizing geomagnetic field data,” Geosciences, vol. 14, no. 5, p. 121, Apr. 2024, doi: 10.3390/geosciences14050121.
58. S. Ommi and M. Hashemi, “Machine learning technique in the north zagros earthquake prediction,” Applied Computing and Geosciences, vol. 22, p. 100163, Apr. 2024, doi: 10.1016/j.acags.2024.100163.
59. A. Berhich, F.-Z. Belouadha, and M. I. Kabbaj, “A location-dependent earthquake prediction using recurrent neural network algorithms,” Soil Dynamics and Earthquake Engineering, vol. 161, p. 107389, Jun. 2022, doi: 10.1016/j.soildyn.2022.107389.
60. Md. H. A. Banna et al., “Application of Artificial intelligence in Predicting Earthquakes: State-of-the-Art and Future challenges,” IEEE Access, vol. 8, pp. 192880–192923, Jan. 2020, doi: 10.1109/access.2020.3029859.
61. S. Mukherjee, P. Gupta, P. Sagar, N. Varshney, and M. Chhetri, “A novel ensemble Earthquake Prediction Method (EEPM) by combining parameters and precursors,” Journal of Sensors, vol. 2022, pp. 1–14, Oct. 2022, doi: 10.1155/2022/5321530.
62. R. Yuan, “An improved K-means clustering algorithm for global earthquake catalogs and earthquake magnitude prediction,” Journal of Seismology, vol. 25, no. 3, pp. 1005–1020, Mar. 2021, doi: 10.1007/s10950-021-09999-8.
63. A. Berhich, F.-Z. Belouadha, and M. I. Kabbaj, “An attention-based LSTM network for large earthquake prediction,” Soil Dynamics and Earthquake Engineering, vol. 165, p. 107663, Nov. 2022, doi: 10.1016/j.soildyn.2022.107663.
64. A. A. Mir et al., “Anomalies prediction in radon Time Series for earthquake likelihood using Machine Learning-Based Ensemble model,” IEEE Access, vol. 10, pp. 37984–37999, Jan. 2022, doi: 10.1109/access.2022.3163291.
65. C. Wang, C. Li, S. Yong, X. Wang, and C. Yang, “Time Series and Non-Time Series Models of Earthquake Prediction based on AETA data: 16-Week real case study,” Applied Sciences, vol. 12, no. 17, p. 8536, Aug. 2022, doi: 10.3390/app12178536.
66. B. Zhang, Z. Hu, P. Wu, H. Huang, and J. Xiang, “EPT: A data-driven transformer model for earthquake prediction,” Engineering Applications of Artificial Intelligence, vol. 123, p. 106176, Mar. 2023, doi: 10.1016/j.engappai.2023.106176.
67. Q. Wang, Y. Guo, L. Yu, and P. Li, “Earthquake prediction based on Spatio-Temporal Data Mining: An LSTM Network approach,” IEEE Transactions on Emerging Topics in Computing, vol. 8, no. 1, pp. 148–158, Apr. 2017, doi: 10.1109/tetc.2017.2699169.
68. Z. Zhang and Y. Wang, “A spatiotemporal model for global earthquake prediction based on convolutional LSTM,” IEEE Transactions on Geoscience and Remote Sensing, vol. 61, pp. 1–12, Jan. 2023, doi: 10.1109/tgrs.2023.3302316.
69. M. Akhoondzadeh, “Earthquake prediction using satellite data: Advances and ahead challenges,” Advances in Space Research, vol. 74, no. 8, pp. 3539–3555, Jun. 2024, doi: 10.1016/j.asr.2024.06.054.
70. Z. Ye et al., “Elite GA-based feature selection of LSTM for earthquake prediction,” The Journal of Supercomputing, vol. 80, no. 14, pp. 21339–21364, Jun. 2024, doi: 10.1007/s11227-024-06218-2.
71. I. A. Bhat, R. Ahmed, W. A. Bhat, and P. Ahmed, “Application of AHP based geospatial modeling for assessment of landslide hazard zonation along Mughal road in the Pir Panjal Himalayas,” Environmental Earth Sciences, vol. 82, no. 13, Jun. 2023, doi: 10.1007/s12665-023-10952-w.

72. M. J. Auflič et al., “Landslide monitoring techniques in the Geological Surveys of Europe,” Landslides, vol. 20, no. 5, pp. 951–965, Jan. 2023, doi: 10.1007/s10346-022-02007-1.
73. C. Yang, Y. Yin, J. Zhang, P. Ding, and J. Liu, “A graph deep learning method for landslide displacement prediction based on global navigation satellite system positioning,” Geoscience Frontiers, vol. 15, no. 1, p. 101690, Aug. 2023, doi: 10.1016/j.gsf.2023.101690.
74. P. Giri, K. Ng, and W. Phillips, “Wireless Sensor Network system for landslide monitoring and warning,” IEEE Transactions on Instrumentation and Measurement, vol. 68, no. 4, pp. 1210–1220, Aug. 2018, doi: 10.1109/tim.2018.2861999.
75. M. Chen, Z. Cai, Y. Zeng, and Y. Yu, “Multi-sensor data fusion technology for the early landslide warning system,” Journal of Ambient Intelligence and Humanized Computing, vol. 14, no. 8, pp. 11165–11172, Sep. 2022, doi: 10.1007/s12652-022-04396-6.
76. H. Thirugnanam, S. Uhlemann, R. Reghunadh, M. V. Ramesh, and V. P. Rangan, “Review of landslide monitoring techniques with IoT integration opportunities,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 5317–5338, Jan. 2022, doi: 10.1109/jstars.2022.3183684.
77. A. Joshi, S. Agarwal, D. P. Kanungo, and R. K. Panigrahi, “Integration of Edge–AI into IoT–Cloud architecture for landslide monitoring and prediction,” IEEE Transactions on Industrial Informatics, vol. 20, no. 3, pp. 4246–4258, Oct. 2023, doi: 10.1109/tii.2023.3319671.
78. Y. K. Kushwaha, A. Joshi, R. K. Panigrahi, and A. Pandey, “Development of a smart irrigation monitoring system employing the wireless sensor network for agricultural water management,” Journal of Hydroinformatics, Dec. 2024, doi: 10.2166/hydro.2024.241.
79. P. S. Rawat and A. Barthwal, “LANDSLIDE MONITOR: a real-time landslide monitoring system,” Environmental Earth Sciences, vol. 83, no. 8, Apr. 2024, doi: 10.1007/s12665-024-11526-0.
80. R. Prabha, M. V. Ramesh, V. P. Rangan, P. V. Ushakumari, and T. Hemalatha, “Energy efficient data acquisition techniques using context aware sensing for landslide monitoring systems,” IEEE Sensors Journal, vol. 17, no. 18, pp. 6006–6018, Jul. 2017, doi: 10.1109/jsen.2017.2730225.
81. A. Sharma, R. Mohana, A. Kukkar, V. Chodha, and P. Bansal, “An ensemble learning–based experimental framework for smart landslide detection, monitoring, prediction, and warning in IoT-cloud environment,” Environmental Science and Pollution Research, vol. 30, no. 58, pp. 122677–122699, Nov. 2023, doi: 10.1007/s11356-023-30683-6.
82. G. L. Ooi, P. S. Tan, M.-L. Lin, K.-L. Wang, Q. Zhang, and Y.-H. Wang, “Near real-time landslide monitoring with the smart soil particles,” Japanese Geotechnical Society Special Publication, vol. 2, no. 28, pp. 1031–1034, Jan. 2016, doi: 10.3208/jgssp.hkg-05.
83. F. Zeng, C. Pang, and H. Tang, “Sensors on the Internet of Things Systems for Urban Disaster Management: A Systematic Literature Review,” Sensors, vol. 23, no. 17, p. 7475, Aug. 2023, doi: 10.3390/s23177475.
84. M. Santini, S. Grimaldi, F. Nardi, A. Petroselli, and M. C. Rulli, “Pre-processing algorithms and landslide modelling on remotely sensed DEMs,” Geomorphology, vol. 113, no. 1–2, pp. 110–125, Apr. 2009, doi: 10.1016/j.geomorph.2009.03.023.

85. G. K. Uyanık and N. Güler, “A study on multiple linear regression analysis,” Procedia - Social and Behavioral Sciences, vol. 106, pp. 234–240, Dec. 2013, doi: 10.1016/j.sbspro.2013.12.027.
86. E. Jamalinia, F. S. Tehrani, S. C. Steele-Dunne, and P. J. Vardon, “Predicting rainfall induced slope stability using random forest regression and synthetic data,” in ICL contribution to landslide disaster risk reduction, 2020, pp. 223–229. doi: 10.1007/978-3-030-60713-5_24.
87. M. Zhang, J. Ling, B. Tang, S. Dong, and L. Zhang, “A Data-Driven based method for pipeline additional stress prediction subject to landslide geohazards,” Sustainability, vol. 14, no. 19, p. 11999, Sep. 2022, doi: 10.3390/su141911999.
88. Y. Yu, X. Si, C. Hu, and J. Zhang, “A review of Recurrent Neural networks: LSTM cells and network architectures,” Neural Computation, vol. 31, no. 7, pp. 1235–1270, May 2019, doi: 10.1162/neco_a_01199.
89. D. Zhang et al., “IBLP: an XGBOOST-Based predictor for identifying bioluminescent proteins,” Computational and Mathematical Methods in Medicine, vol. 2021, pp. 1–15, Jan. 2021, doi: 10.1155/2021/6664362.
90. R. Can, S. Kocaman, and C. Gokceoglu, “A comprehensive assessment of XGBOOST algorithm for landslide susceptibility mapping in the upper basin of Ataturk Dam, Turkey,” Applied Sciences, vol. 11, no. 11, p. 4993, May 2021, doi: 10.3390/app11114993.
91. P. Kumar, P. Sihag, P. Chaturvedi, K. V. Uday, and V. Dutt, “BS-LSTM: An ensemble Recurrent approach to forecasting soil movements in the real world,” Frontiers in Earth Science, vol. 9, Aug. 2021, doi: 10.3389/feart.2021.696792.
92. D. Zhang, J. Yang, F. Li, S. Han, L. Qin, and Q. Li, “Landslide risk prediction model using an Attention-Based temporal convolutional network connected to a recurrent neural network,” IEEE Access, vol. 10, pp. 37635–37645, Jan. 2022, doi: 10.1109/access.2022.3165051.
93. P. R. Kshirsagar et al., “Expedite quantification of landslides using wireless sensors and artificial intelligence for data controlling practices,” Computational Intelligence and Neuroscience, vol. 2022, pp. 1–11, May 2022, doi: 10.1155/2022/3211512.
94. M. H. Bukhari et al., “Community perceptions of landslide risk and susceptibility: a multi-country study,” Landslides, vol. 20, no. 6, pp. 1321–1334, Feb. 2023, doi: 10.1007/s10346-023-02027-5.
95. S. Schneiderbauer et al., “Risk perception of climate change and natural hazards in global mountain regions: A critical review,” The Science of the Total Environment, vol. 784, p. 146957, Apr. 2021, doi: 10.1016/j.scitotenv.2021.146957.
96. J. Zhang, M. Lu, L. Zhang, and Y. Xue, “Assessing indirect economic losses of landslides along highways,” Natural Hazards, vol. 106, no. 3, pp. 2775–2796, Feb. 2021, doi: 10.1007/s11069-021-04566-3.
97. C. Pfurtscheller and E. Genovese, “The Felbertauern landslide of 2013 in Austria: Impact on transport networks, regional economy and policy decisions,” Case Studies on Transport Policy, vol. 7, no. 3, pp. 643–654, May 2019, doi: 10.1016/j.cstp.2019.05.003.
98. Z. Zhang, Q. Liu, and Y. Wang, “Road extraction by deep residual U-Net,” IEEE Geoscience and Remote Sensing Letters, vol. 15, no. 5, pp. 749–753, Mar. 2018, doi: 10.1109/lgrs.2018.2802944.
99. A. Mohan, A. K. Singh, B. Kumar, and R. Dwivedi, “Review on remote sensing methods for landslide detection using machine and deep learning,” Transactions on Emerging Telecommunications Technologies, vol. 32, no. 7, Jun. 2020, doi: 10.1002/ett.3998.

100. S. Liu, L. Wang, W. Zhang, Y. He, and S. Pijush, “A comprehensive review of machine learning-based methods in landslide susceptibility mapping,” Geological Journal, vol. 58, no. 6, pp. 2283–2301, Jan. 2023, doi: 10.1002/gj.4666.
101. A. Mohan, A. K. Singh, B. Kumar, and R. Dwivedi, “Review on remote sensing methods for landslide detection using machine and deep learning,” Transactions on Emerging Telecommunications Technologies, vol. 32, no. 7, Jun. 2020, doi: 10.1002/ett.3998.
102. R. N. Keyport, T. Oommen, T. R. Martha, K. S. Sajinkumar, and J. S. Gierke, “A comparative analysis of pixel- and object-based detection of landslides from very high-resolution images,” International Journal of Applied Earth Observation and Geoinformation, vol. 64, pp. 1–11, Sep. 2017, doi: 10.1016/j.jag.2017.08.015.
103. X. Wei et al., “Improving pixel-based regional landslide susceptibility mapping,” Geoscience Frontiers, vol. 15, no. 4, p. 101782, Jan. 2024, doi: 10.1016/j.gsf.2024.101782.
104. S. Das, P. Sharma, A. Pain, D. P. Kanungo, and S. Sarkar, “Deep learning based landslide detection using open-source resources: Opportunities and challenges,” Earth Science Informatics, vol. 16, no. 4, pp. 4035–4052, Nov. 2023, doi: 10.1007/s12145-023-01141-1.
105. S. B. Saba, M. Ali, S. A. Turab, M. Waseem, and S. Faisal, “Comparison of pixel, sub-pixel and object-based image analysis techniques for co-seismic landslides detection in seismically active area in Lesser Himalaya, Pakistan,” Natural Hazards, vol. 115, no. 3, pp. 2383–2398, Oct. 2022, doi: 10.1007/s11069-022-05642-y.
106. Y. Qu et al., “Integrating sentinel-2a imagery, DEM data, and spectral feature analysis for landslide detection via fully convolutional networks,” Landslides, Oct. 2024, doi: 10.1007/s10346-024-02379-6.
107. X. Gao, T. Chen, R. Niu, and A. Plaza, “Recognition and mapping of landslide using a fully convolutional DenseNet and influencing factors,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 7881–7894, Jan. 2021, doi: 10.1109/jstars.2021.3101203.
108. W. Shi, M. Zhang, H. Ke, X. Fang, Z. Zhan, and S. Chen, “Landslide recognition by deep convolutional neural network and change detection,” IEEE Transactions on Geoscience and Remote Sensing, vol. 59, no. 6, pp. 4654–4672, Aug. 2020, doi: 10.1109/tgrs.2020.3015826.
109. X. Gao, T. Chen, R. Niu, and A. Plaza, “Recognition and mapping of landslide using a fully convolutional DenseNet and influencing factors,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 14, pp. 7881–7894, Jan. 2021, doi: 10.1109/jstars.2021.3101203.
110. X. He, Y. Zhou, J. Zhao, D. Zhang, R. Yao, and Y. Xue, “SWIN Transformer Embedding UNET for remote sensing Image semantic segmentation,” IEEE Transactions on Geoscience and Remote Sensing, vol. 60, pp. 1–15, Jan. 2022, doi: 10.1109/tgrs.2022.3144165.
111. X. Tang, Z. Tu, Y. Wang, M. Liu, D. Li, and X. Fan, “Automatic detection of coseismic landslides using a new transformer method,” Remote Sensing, vol. 14, no. 12, p. 2884, Jun. 2022, doi: 10.3390/rs14122884.
112. P. Lv, L. Ma, Q. Li, and F. Du, “ShapeFormer: a Shape-Enhanced vision transformer model for optical remote sensing image landslide detection,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 16, pp. 2681–2689, Jan. 2023, doi: 10.1109/jstars.2023.3253769.
113. S. R. Meena et al., “Rapid mapping of landslides in the Western Ghats (India) triggered by 2018 extreme monsoon rainfall using a deep learning approach,”

Landslides, vol. 18, no. 5, pp. 1937–1950, Jan. 2021, doi: 10.1007/s10346-020-01602-4.
114. B. Yu, F. Chen, and C. Xu, “Landslide detection based on contour-based deep learning framework in case of national scale of Nepal in 2015,” Computers & Geosciences, vol. 135, p. 104388, Nov. 2019, doi: 10.1016/j.cageo.2019.104388.
115. B. Yu, C. Xu, F. Chen, N. Wang, and L. Wang, “HADeenNet: A hierarchical-attention multi-scale deconvolution network for landslide detection,” International Journal of Applied Earth Observation and Geoinformation, vol. 111, p. 102853, Jun. 2022, doi: 10.1016/j.jag.2022.102853.
116. S. T. Piralilou et al., “Landslide detection using Multi-Scale image segmentation and different machine learning models in the Higher Himalayas,” Remote Sensing, vol. 11, no. 21, p. 2575, Nov. 2019, doi: 10.3390/rs11212575.
117. P. Amatya, D. Kirschbaum, T. Stanley, and H. Tanyas, “Landslide mapping using object-based image analysis and open source tools,” Engineering Geology, vol. 282, p. 106000, Jan. 2021, doi: 10.1016/j.enggeo.2021.106000.
118. Segment optimization and data-driven thresholding for knowledge-based landslide detection by object-based image analysis Vantaram, S.R. and Saber, E., 2012. Survey of contemporary trends in color image segmentation. Journal of Electronic Imaging, 21(4), pp.040901-040901.
119. J. L. Mesa-Mingorance and F. J. Ariza-López, “Accuracy Assessment of Digital Elevation Models (DEMS): A critical review of practices of the past three decades,” Remote Sensing, vol. 12, no. 16, p. 2630, Aug. 2020, doi: 10.3390/rs12162630.
120. O. Ronneberger, P. Fischer, and T. Brox, “U-NET: Convolutional Networks for Biomedical Image Segmentation,” in Lecture notes in computer science, 2015, pp. 234–241. doi: 10.1007/978-3-319-24574-4_28.
121. C. Tan, T. Chen, J. Liu, X. Deng, H. Wang, and J. Ma, “Building Extraction from Unmanned Aerial Vehicle (UAV) Data in a Landslide-Affected Scattered Mountainous Area Based on Res-Unet,” Sustainability, vol. 16, no. 22, p. 9791, Nov. 2024, doi: 10.3390/su16229791.
122. H. Chen et al., “A landslide extraction method of channel attention mechanism U-Net network based on Sentinel-2A remote sensing images,” International Journal of Digital Earth, vol. 16, no. 1, pp. 552–577, Mar. 2023, doi: 10.1080/17538947.2023.2177359.
123. S. R. Meena et al., “Landslide detection in the Himalayas using machine learning algorithms and U-Net,” Landslides, vol. 19, no. 5, pp. 1209–1229, Feb. 2022, doi: 10.1007/s10346-022-01861-3.
124. O. Ghorbanzadeh, S. R. Meena, T. Blaschke, and J. Aryal, “UAV-Based slope failure detection using Deep-Learning Convolutional Neural Networks,” Remote Sensing, vol. 11, no. 17, p. 2046, Aug. 2019, doi: 10.3390/rs11172046.
125. O. Ghorbanzadeh, T. Blaschke, K. Gholamnia, S. R. Meena, D. Tiede, and J. Aryal, “Evaluation of different machine learning methods and Deep-Learning convolutional neural networks for landslide detection,” Remote Sensing, vol. 11, no. 2, p. 196, Jan. 2019, doi: 10.3390/rs11020196.
126. N. Prakash, A. Manconi, and S. Loew, “Mapping Landslides on EO Data: Performance of Deep Learning Models vs. Traditional Machine Learning Models,” Remote Sensing, vol. 12, no. 3, p. 346, Jan. 2020, doi: 10.3390/rs12030346.
127. Y. Wang, X. Wu, Z. Chen, F. Ren, L. Feng, and Q. Du, “Optimizing the predictive ability of machine learning methods for landslide susceptibility mapping using SMOTE for Lishui City in Zhejiang Province, China,” International Journal of Environmental Research and Public Health, vol. 16, no. 3, p. 368, Jan. 2019, doi: 10.3390/ijerph16030368.

128. W. Qi, M. Wei, W. Yang, C. Xu, and C. Ma, “Automatic mapping of landslides by the RESU-Net,” Remote Sensing, vol. 12, no. 15, p. 2487, Aug. 2020, doi: 10.3390/rs12152487.
129. R. A. Reddy, R. Gobinath, C. S. Khanna, and G. Shyamala, “Machine Learning based Landslide Prediction System for Hilly Areas,” IOP Conference Series Materials Science and Engineering, vol. 981, no. 3, p. 032084, Dec. 2020, doi: 10.1088/1757-899x/981/3/032084.
130. M. Azarafza, M. Azarafza, H. Akgün, P. M. Atkinson, and R. Derakhshani, “Deep learning-based landslide susceptibility mapping,” Scientific Reports, vol. 11, no. 1, Dec. 2021, doi: 10.1038/s41598-021-03585-1.
131. K. Bhuyan, S. R. Meena, L. Nava, C. Van Westen, M. Floris, and F. Catani, “Mapping landslides through a temporal lens: an insight toward multi-temporal landslide mapping using the u-net deep learning model,” GIScience & Remote Sensing, vol. 60, no. 1, Mar. 2023, doi: 10.1080/15481603.2023.2182057.
132. F. S. Tehrani, M. Calvello, Z. Liu, L. Zhang, and S. Lacasse, “Machine learning and landslide studies: recent advances and applications,” Natural Hazards, vol. 114, no. 2, pp. 1197–1245, Jun. 2022, doi: 10.1007/s11069-022-05423-7.
133. X. Chen et al., “Conv-trans dual network for landslide detection of multi-channel optical remote sensing images,” Frontiers in Earth Science, vol. 11, May 2023, doi: 10.3389/feart.2023.1182145.
134. O. Ghorbanzadeh, H. Shahabi, A. Crivellari, S. Homayouni, T. Blaschke, and P. Ghamisi, “Landslide detection using deep learning and object-based image analysis,” Landslides, vol. 19, no. 4, pp. 929–939, Jan. 2022, doi: 10.1007/s10346-021-01843-x.
135. F. Zhang et al. “On the Generalization of the Semantic Segmentation Model for Landslide Detection”. CDCEO@ IJCAI, pp. 96-100, May 2022, doi: 10.1007/s10346-021-2182057.
136. L. Bai et al., “Multispectral U-Net: A Semantic Segmentation Model Using Multispectral Bands Fusion Mechanism for Landslide Detection”. In CDCEO@ IJCAI, pp. 101-104, July 2022.
137. R. Li, X. Lu, S. Li, H. Yang, J. Qiu, and L. Zhang, “DLEP: a Deep learning model for Earthquake Prediction,” 2022 International Joint Conference on Neural Networks (IJCNN), pp. 1–8, Jul. 2020, doi: 10.1109/ijcnn48605.2020.9207621.
138. T. Chelidze, G. Melikadze, T. Kiria, T. Jimsheladze, and G. Kobzev, “Statistical and Non-linear Dynamics Methods of Earthquake Forecast: Application in the Caucasus,” Frontiers in Earth Science, vol. 8, Jun. 2020, doi: 10.3389/feart.2020.00194.
139. P. Nath, P. Saha, A. I. Middya, and S. Roy, “Long-term time-series pollution forecast using statistical and deep learning methods,” Neural Computing and Applications, vol. 33, no. 19, pp. 12551–12570, Apr. 2021, doi: 10.1007/s00521-021-05901-2.
140. J. Pwavodi, A. U. Ibrahim, P. C. Pwavodi, F. Al-Turjman, and A. Mohand-Said, “The role of artificial intelligence and IoT in prediction of earthquakes: Review,” Artificial Intelligence in Geosciences, vol. 5, p. 100075, Feb. 2024, doi: 10.1016/j.aiig.2024.100075.
141. K. Wang, Y. Zhu, E. Nissen, and Z. Shen, “On the Relevance of Geodetic Deformation Rates to Earthquake Potential,” Geophysical Research Letters, vol. 48, no. 11, May 2021, doi: 10.1029/2021gl093231.
142. R. Shcherbakov, J. Zhuang, G. Zöller, and Y. Ogata, “Forecasting the magnitude of the largest expected earthquake,” Nature Communications, vol. 10, no. 1, Sep. 2019, doi: 10.1038/s41467-019-11958-4.

143. H. O. Cekim, S. Tekin, and G. Özel, “Prediction of the earthquake magnitude by time series methods along the East Anatolian Fault, Turkey,” Earth Science Informatics, vol. 14, no. 3, pp. 1339–1348, Jun. 2021, doi: 10.1007/s12145-021-00636-z.
144. B. Ku, G. Kim, J.-K. Ahn, J. Lee, and H. Ko, “Attention-Based convolutional neural network for earthquake event classification,” IEEE Geoscience and Remote Sensing Letters, vol. 18, no. 12, pp. 2057–2061, Aug. 2020, doi: 10.1109/lgrs.2020.3014418.
145. M. H. Chegeni, M. K. Sharbatdar, R. Mahjoub, and M. Raftari, “New supervised learning classifiers for structural damage diagnosis using time series features from a new feature extraction technique,” Earthquake Engineering and Engineering Vibration, vol. 21, no. 1, pp. 169–191, Jan. 2022, doi: 10.1007/s11803-022-2079-2.
146. G. Xu, Y. Wang, L. Wang, L. P. Soares, and C. H. Grohmann, “Feature-Based Constraint Deep CNN method for mapping Rainfall-Induced landslides in remote regions with mountainous terrain: an application to Brazil,” IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 15, pp. 2644–2659, Jan. 2022, doi: 10.1109/jstars.2022.3161383.
147. T. Perol, M. Gharbi, and M. Denolle, “Convolutional neural network for earthquake detection and location,” Science Advances, vol. 4, no. 2, Feb. 2018, doi: 10.1126/sciadv.1700578.
148. I. Ramli, S. Rusdiana, A. Achmad, N. Azizah, and M. E. Yolanda, “Forecasting of rainfall using Seasonal Autoregreressive Integrated Moving Average (SARIMA) Aceh, Indonesia,” Mathematical Modelling and Engineering Problems, vol. 10, no. 2, pp. 501–508, Apr. 2023, doi: 10.18280/mmep.100216.
149. F. Huang et al., “Optimization method of conditioning factors selection and combination for landslide susceptibility prediction,” Journal of Rock Mechanics and Geotechnical Engineering, Aug. 2024, doi: 10.1016/j.jrmge.2024.04.029.
150. D. R. Martinez and M. J. Kowalsky, “Nonlinear seismic performance of RC bridges using the ESA, EDA, DDBA, and nonlinear analysis with various viscous damping models,” Earthquake Spectra, vol. 39, no. 1, pp. 242–268, Jan. 2023,doi: 10.1177/87552930221145435.
151. R. Saini, P. Garg, N. K. Chaudhary, M. V. Joshi, V. S. Palaparthy, and A. Kumar, “Identifying the source of water on plant using the leaf wetness sensor and via Deep Learning-Based Ensemble method,” IEEE Sensors Journal, vol. 24, no. 5, pp. 7009–7017, Jan. 2024, doi: 10.1109/jsen.2023.3343574.
152. Q. Yang et al., “A novel CGBoost deep learning algorithm for coseismic landslide susceptibility prediction,” Geoscience Frontiers, vol. 15, no. 2, p. 101770, Dec. 2023, doi: 10.1016/j.gsf.2023.101770.
153. D.-C. Feng, W.-J. Wang, S. Mangalathu, and E. Taciroglu, “Interpretable XGBOOST-SHAP Machine-Learning Model for Shear Strength Prediction of Squat RC walls,” Journal of Structural Engineering, vol. 147, no. 11, Aug. 2021, doi: 10.1061/(asce)st.1943-541x.0003115.
154. M. Kanwar, B. Pokharel, and S. Lim, “A new random forest method for landslide susceptibility mapping using hyperparameter optimization and grid search techniques,” International Journal of Environmental Science and Technology, Jan. 2025, doi: 10.1007/s13762-024-06310-3.
155. I. Sidik, S. Saroji, and S. Sulistyani, “Implementation of machine learning for volcanic earthquake pattern classification using XGBoost algorithm,” Acta Geophysica, vol. 72, no. 3, pp. 1575–1585, Aug. 2023, doi: 10.1007/s11600-023-01154-w.
156. A. Mohammadi, S. Karimzadeh, S. A. Banimahd, V. Ozsarac, and P. B. Lourenço, “The potential of region-specific machine-learning-based ground motion models:

Application to Turkey," Soil Dynamics and Earthquake Engineering, vol. 172, p. 108008, May 2023, doi: 10.1016/j.soildyn.2023.108008.
157. M. Saqib, E. Şentürk, S. A. Sahu, and M. A. Adil, "Comparisons of autoregressive integrated moving average (ARIMA) and long short term memory (LSTM) network models for ionospheric anomalies detection: a study on Haiti (Mw = 7.0) earthquake," Acta Geodaetica Et Geophysica, vol. 57, no. 1, pp. 195–213, Jan. 2022, doi: 10.1007/s40328-021-00371-3.
158. H. Ö. Çekim, H. N. Karakavak, G. Özel, and S. Tekin, "Earthquake magnitude prediction in Turkey: a comparative study of deep learning methods, ARIMA and singular spectrum analysis," Environmental Earth Sciences, vol. 82, no. 16, Aug. 2023, doi: 10.1007/s12665-023-11072-1.
159. E. Abebe, H. Kebede, M. Kevin, and Z. Demissie, "Earthquakes magnitude prediction using deep learning for the Horn of Africa," Soil Dynamics and Earthquake Engineering, vol. 170, p. 107913, Mar. 2023, doi: 10.1016/j.soildyn.2023.107913.
160. A. Alexandridis, E. Chondrodima, E. Efthimiou, G. Papadakis, F. Vallianatos, and D. Triantis, "Large earthquake occurrence estimation based on radial basis function neural networks," IEEE Transactions on Geoscience and Remote Sensing, vol. 52, no. 9, pp. 5443–5453, Jan. 2014, doi: 10.1109/tgrs.2013.2288979.
161. R. Tehseen, M. S. Farooq, and A. Abid, "Earthquake Prediction Using Expert Systems: A Systematic Mapping study," Sustainability, vol. 12, no. 6, p. 2420, Mar. 2020, doi: 10.3390/su12062420.
162. Md. H. A. Banna et al., "Attention-Based Bi-Directional Long-Short Term Memory network for earthquake prediction," IEEE Access, vol. 9, pp. 56589–56603, Jan. 2021, doi: 10.1109/access.2021.3071400.
163. B. Rouet-Leduc, C. Hulbert, N. Lubbers, K. Barros, C. J. Humphreys, and P. A. Johnson, "Machine learning predicts laboratory earthquakes," Geophysical Research Letters, vol. 44, no. 18, pp. 9276–9282, Aug. 2017, doi: 10.1002/2017gl074677.
164. W. Zhu, M. Wu, Q. Xie, and Y. Chen, "Post-Earthquake rapid assessment method for electrical function of equipment in substations," IEEE Transactions on Power Delivery, vol. 38, no. 5, pp. 3312–3321, Apr. 2023, doi: 10.1109/tpwrd.2023.3270178.
165. Y. Li, C. Jia, H. Chen, H. Su, J. Chen, and D. Wang, "Machine Learning Assessment of damage grade for Post-Earthquake buildings: A Three-Stage approach directly handling categorical features," Sustainability, vol. 15, no. 18, p. 13847, Sep. 2023, doi: 10.3390/su151813847.
166. D. Gautam, A. Bhattarai, and R. Rupakhety, "Machine learning and soft voting ensemble classification for earthquake induced damage to bridges," Engineering Structures, vol. 303, p. 117534, Jan. 2024, doi: 10.1016/j.engstruct.2024.117534.
167. Z. Li, H. Lei, E. Ma, J. Lai, and J. Qiu, "Ensemble technique to predict post-earthquake damage of buildings integrating tree-based models and tabular neural networks," Computers & Structures, vol. 287, p. 107114, Aug. 2023, doi: 10.1016/j.compstruc.2023.107114.
168. A. Ocak, Ü. Işıkdağ, G. Bekdaş, S. M. Nigdeli, S. Kim, and Z. W. Geem, "Prediction of damping capacity demand in seismic base isolators via machine learning," Computer Modeling in Engineering & Sciences, vol. 138, no. 3, pp. 2899–2924, Dec. 2023, doi: 10.32604/cmes.2023.030418.
169. S. Karimpouli et al., "Explainable machine learning for labquake prediction using catalog-driven features," Earth and Planetary Science Letters, vol. 622, p. 118383, Oct. 2023, doi: 10.1016/j.epsl.2023.118383.

170. M. N. Brykov et al., “Machine learning modelling and feature engineering in seismology experiment,” Sensors, vol. 20, no. 15, p. 4228, Jul. 2020, doi: 10.3390/s20154228.
171. H. Jasperson, D. C. Bolton, P. Johnson, R. Guyer, C. Marone, and M. V. De Hoop, “Attention network forecasts Time-to-Failure in laboratory shear experiments,” Journal of Geophysical Research Solid Earth, vol. 126, no. 11, Oct. 2021, doi: 10.1029/2021jb022195.
172. X. Zhang, W. Reichard-Flynn, M. Zhang, M. Hirn, and Y. Lin, “Spatiotemporal graph convolutional networks for earthquake source characterization,” Journal of Geophysical Research Solid Earth, vol. 127, no. 11, Oct. 2022, doi: 10.1029/2022jb024401.
173. P. Bannigan et al., “Machine learning models to accelerate the design of polymeric long-acting injectables,” Nature Communications, vol. 14, no. 1, Jan. 2023, doi: 10.1038/s41467-022-35343-w.

LIST OF PUBLICATIONS

Journal Publications:

Published:

1. A. Kaushal, A. K. Gupta, and V. K. Sehgal, “A semantic segmentation framework with UNet-pyramid for landslide prediction using remote sensing data,” Scientific Reports, vol. 14, no. 1, Dec. 2024, doi: https://doi.org/10.1038/s41598-024-79266-6. **[SCI/SCOPUS INDEXED]**
2. A. Kaushal, A. K. Gupta, and V. K. Sehgal, “Exploiting the synergy of SARIMA and XGBoost for spatiotemporal earthquake time series forecasting,” Earth Surface Processes and Landforms, vol. 49, no. 14, pp. 4724–4742, Oct. 2024, doi: https://doi.org/10.1002/esp.5992. **[SCI/SCOPUS INDEXED]**
3. A. Kaushal, A. K. Gupta, and V. K. Sehgal, “A Data-Driven and Hybrid Approach for Real-Time Threshold-Based Landslide Prediction in Hilly Regions using Sensor Networks,” Remote Sensing Letters, vol. 15, no. 12, pp. 1218–1228, Nov. 2024, doi: https://doi.org/10.1080/2150704x.2024.2399331. **[SCI/SCOPUS INDEXED]**
4. A. Kaushal, A. K. Gupta, and V. K. Sehgal, “Hybrid CatBoost and SVR Model for Earthquake Prediction Using the LANL Earthquake Dataset”, An International Journal of Computing & Informatics, by Informatica, doi: https://doi.org/10.31449/inf.v49i14.6524 **[SCOPUS INDEXED]**
5. A. Kaushal, A. K. Gupta, and V. K. Sehgal, “Earthquake Prediction Optimization Using Deep Learning Hybrid RNN-LSTM Model for Seismicity Analysis”, Soil Dynamics and Earthquake Engineering, by Elsevier **[SCIE/SCOPUS INDEXED]**

Under Review/Communicated:

1. A. Kaushal, A. K. Gupta, and V. K. Sehgal, “Geospatial Deep Learning Model for Early Landslide Prediction Using Multispectral Remote Sensing Data”, International Journal of Sensors, Wireless Communications and Control, by Bentham Science (ACCEPTED) **[SCOPUS INDEXED]**

Paper presented in Conferences:

1. A. Kaushal and V. K. Sehgal, "Landslide Susceptibility Detection Using ResNet," 2023 3rd Asian Conference on Innovation in Technology (ASIANCON), Ravet IN, India, 2023, pp. 1-5, doi: 10.1109/ASIANCON58793.2023.10269882.
2. A. Kaushal and V. K. Sehgal, "Threshold Based Real-Time Landslide Prediction System Using Low-Cost Sensor Networks," 2023 3rd Asian Conference on Innovation in Technology (ASIANCON), Ravet IN, India, 2023, pp. 1-7, doi: 10.1109/ASIANCON58793.2023.10269931.