LANDSLIDE DETECTION USING MACHINE LEARNING

ASEELA JAMALU

Department of Computer Engineering,
Mangalam college of Engineering
Jaseelapj0@gmail.com

SIJI MOL M S

Department of Computer Engineering,
Mangalam college of Engineering ``

Abstract- **Landslides and famine are major natural hazards affecting regions with complex topography and vulnerable populations. The study investigates the utilization of machine learning to predict landslides and food insecurity in the affected area. The system makes use of data like past weather patterns, soil and land use information, terrain, and socio-economic-related data to offer early warnings and other information. The proposed model for landslide prediction utilizes Convolutional Neural Networks (CNNs) which allows the model to recognize spatial patterns and risk zones from satellite images and DEMs. Time series analysis and regression methods monitor the shortage of food and predict food production. These models, which are trained and tested on empirical data from high-risk areas, have been found to be much more accurate than traditional forecasting. This paper also presents a hybrid prediction model based on CNNs and RNNs. While CNNs capture what you can see on the ground, RNNs focus on measuring various time-dependent phenomena like rainfall, soil moisture, and river water level. The system makes better predictions and gives alerts on time because both spatial and temporal data is integrated. The modified version was tested in the past for landslide-prone areas in India and shows strong potential to be used in actual disaster management and early warning system. Words- Landslide forecasting, Flood hazard map, Time level forecasting, Machine learning, CNN, RNN.**

INTRODUCTION

In previous years, natural calamities have increased significantly all around the world. The risk to safety is becoming ever more serious due to intensified natural disasters. Flood-prone areas are where natural hazards create the most danger to damage. Landslides are considered one of the most serious disaster which occur today. These incidents cause large-scale damage to infrastructures like roads and buildings. Destruction is in the nature of the area Now it is easier to combat natural disasters due to Machine Learning (ML). Machine learning uses data to enhance algorithms that automatically assess when and where floods and landslides are most likely to occur. Using this method makes people more prepared to respond quickly. Helps and preventing them.

1.EXISTING SYSTEM

To forecast floods and landslides one must determine how water propagates throughout the environment. The simulation of different water cycle aspects including runoff run off and groundwater infiltration falls under the responsibility of hydrological models. Scientists together with decision-makers depend on these models to observe water flow patterns in different locations and scenario settings. Traditional methods may prove inadequate in predicting water flow because various components such as climate change and human activities make the system intricate.

Machine learning (ML), which can evaluate vast amounts of historical and current data, spot underlying patterns, and produce better predictions about the likelihood of floods or landslides, is currently being incorporated into these models in an attempt to increase accuracy. Other than deep learning, neural networks, profound learning complex and shuffled data can be transformed to easier way to understand the patterns which will give more accurate prediction to us.

Another technique is using satellite drones and photograph to collect the data of the land of remote sensing area. Using such technology we can record important data about water levels, soil moisture content, and plant life elements which help us to predict floods and landslides. For example , a sudden declination of plants will lead to a landslide and it could be a sign of landslide. Keeping an eye on a river side plant life and other water bodies can help us to identify the flood prone areas.

It use satellite data with machine learning algorithm for powerful prediction. Machine learning algorithm can process images ,identify the changes of land, and evaluate the moisture content of the soil automatically. These factors aid in estimating the amount of rainfall that will result in surface runoff. Getting this detail right is essential for water resources management and flood risk assessment.

Doi:10.48047/ijiee.2025.15.6.9

Like any other complex system, rainfall-runoff models rely on empirical equations with a multitude of assumptions that do not always respond well to changes in weather or climate. This is where machine learning, ML, can help. ML models

can incrementally learn from new data and adapt to different climate conditions, improving their predictions over time.

Advanced ML techniques such as recurrent neural networks, random forests, and support vector machines, are particularly proficient in dissecting the intricate relationships within historical data and real-time weather data, and improve the accuracy and responsiveness of flood forecasting..

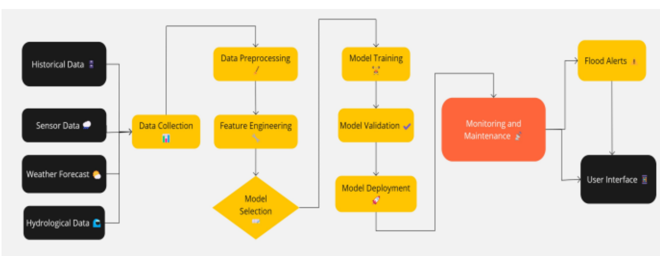
Incorporating thermal and remote sensing, satellite hydrology, as well as hydrograph analysis into an ML framework along with hydrological modeling enables the construction of a sophisticated intelligent disaster prediction system. This system is not capable of providing early warning system so that people can't take sudden actions and will cause human lose as well as damages. Authorities also can't take quick action to save human life and can't make a good decision.

III.PROPOSED SYSTEM

By analysing a variety of meteorological and environmental data, the system implements machine learning to forecast floods and landslides. It collects data from a variety of sources like satellite photos, rainfall trends, topography, soil moisture, and river flow levels. The system can identify the different patterns in the data that can suggest a flood or landslide that is likely to occur by using sophisticated algorithms like Convolutional Neural Networks (CNNs). By continuous updating of data with new data, it will learn and get better next time and provide more accuracy for its predictions.

The system's repeated monitoring, which will detect dangerous areas and inform the authorities and provide messages to local communities. So it will help the authorities to take immediate action and can provide safety to people.

IV.ARCHITECTURE



The above diagram show the architecture of predicting flood and landslides. This combines the machine learning with conventional hydrological model. These are the interconnected layers from data collection to output prediction. Each layer has interconnection. One layer output is the input of the next layer.

Data Acquisition and Inputs

The diagram shows the sources of environmental data, like remote sensing and satellite imagery. These provide important information about the locations of water bodies, plants life, and land use. Weather related information also provided. The topographic details, moisture content, rain fall patterns are some inputs. Combining all these different data is important to analyzing the different elements that effect water flow and intensity of the natural disaster like landslide and flood.

Hydrological and Rainfall-Runoff Modelling

The right side of the diagram explains the application of hydrological models. The model use information below to help us to understand how waterflow occur in our environment:

- Rain falls
- Runoff from the surface

- The process water seeps into the earth
- Stream and river flow

The rain fall run off model estimate the amount of rain fall

Based on rainfall data that is also highlighted in our diagram. This step is very important because it used for determining whether and when the excessive amount of water will come and which will cause flooding

Machine Learning Integration

The machine learning algorithm is the important of the diagram. The weather data is analyzed using the Convolutional Neural Networks (CNNs). These model is will give more accuracy because it will continuously update the historical data contemporary data. The system will provide more accuracy for prediction and will provide warning indicators compared to other approaches.

Predictive Outputs and Decision Support

The right side of the diagram show the output of the integrated system. It explains how our system provide risk assessment and early warnings of flood and landslide using our predictive insights. Authorities and local communities can use this approaches to take effective and efficient actions using the forecasting.

By combining predictive machine learning with deep environmental analysis, the diagram shows a good approach to prevention of landslide and flood. This approach increases resource allocation and provide a high priority for safety of people. .In conclusion, the architecture explains a complicated system in which accurate data collection will provide advanced hydrological models.

When occurring natural disasters, our models will improve by continuous machine learning updation and will provide useful prediction to speed up the decision making and effective action for local communities and authorities

V. METHODOLOGY

A. DATA COLLECTION

1. Weather Information:

Precipitation levels, air temperature, humidity, and other meteorological factors that affect environmental patterns are all included in this category.

2. Geographical Data

Geographical data encompasses terrain characteristics, soil classification, land elevation, gradient (slope), and overall land cover distribution. These factors play a crucial role in understanding surface water flow and erosion risks

3. Hydrological Data

This includes details about water movement in natural systems, such as river discharge rates, fluctuations in water levels, and soil moisture content. These parameters are critical in flood and drought monitoring.

4. Historical Data

Historical records provide insights into previous flood and landslide occurrences, documenting their severity, duration, and consequences. These past events help improve predictive models for disaster risk assessment.

B. DATA PRE PROCESSING

Data preprocessing is a crucial step in machine learning and data analysis, ensuring that raw data is clean, structured, and suitable for modeling. It enhances the quality of data, leading to better performance and accuracy in predictive models. The key stages of data preprocessing include data cleaning, normalization, and feature engineering.

1. Data cleaning

Data cleaning is the process of finding and fixing errors, inconsistencies, and missing values in a dataset. This step ensures that the data used for analysis is accurate and trustworthy.

Handling Missing Values: Missing values can effect our system. It will decrease the performance of the system badly. The different methods to deal with missing values are

- Remove the data if it contain large amount of missing data.
- We can use mean, median , mode method to check missing values
- We can use predictive algorithm to predict the missing values.

Eliminating Duplicates: Analysis and model predictions may be skewed by duplicate entries. Data integrity is preserved by identifying and removing duplicate records.

Correcting Inconsistencies: Ensuring consistency in data formats (e.g., date formats, text case) and rectifying incorrect entries.

Filtering Outliers: Outliers, extreme values that can skew data distribution and impact model training, can be identified and managed using statistical methods like Z-score, IQR (Interquartile Range), and domain expertise.

2. Normalization

Normalization, known as feature scaling, involves adjusting numerical values to a consistent range. This process prevents features with large ranges from overpowering others, ensuring an equal contribution to the learning algorithm.

- **Min-Max Scaling:** This method transforms values to a fixed range, usually [0,1], utilizing the formula:

$$X_{standardized} = \frac{X - \mu}{\sigma}$$

3. Feature Engineering

Feature engineering involves creating new meaningful variables from raw data to improve model performance. It helps in uncovering hidden patterns and relationships in the dataset.

Feature Extraction: Deriving new features from existing ones (e.g., extracting 'day of the week' from a date column).

Polynomial Features: Creating higher-order interactions of variables to capture complex relationships.

Encoding Categorical Variables: Converting categorical data into numerical form using techniques like:

One-Hot Encoding: Creating binary columns for each category.

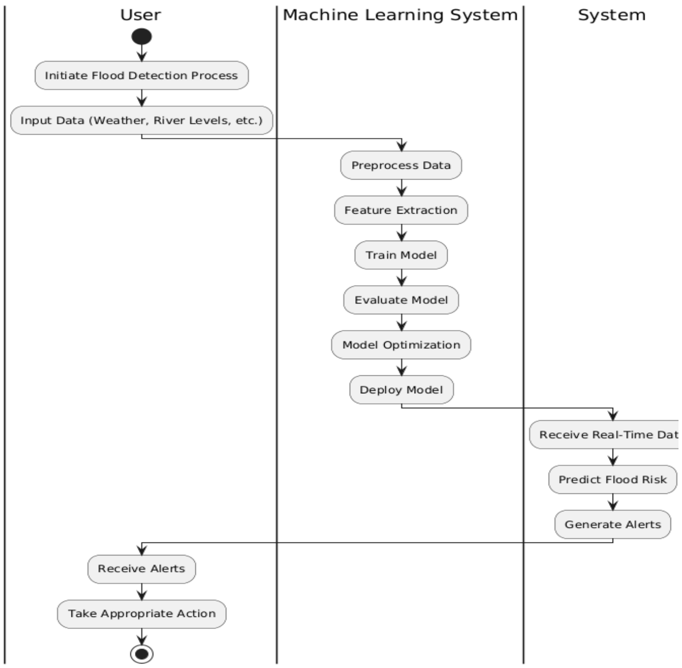
Label Encoding: Assigning numerical labels to

Categories .

Domain-Specific Transformations: Applying domain knowledge to create meaningful features (e.g., calculating Body Mass Index (BMI) from height and weight data).

Doi:10.48047/ijiee.2025.15.6.9

VI .ACTIVITY DIAGRAM



This image represents a **flowchart** illustrating a **flood detection system** using **machine learning**. It consists of three key components: **User, Machine Learning System, and System**, each playing a crucial role in flood prediction and alert generation.

User Actions

The user **initiates the flood detection process** by providing **input data** (such as weather conditions, river water levels, etc.).

Once the system processes the data and generates alerts, the user **receives notifications** and can take appropriate action based on the risk level.

Machine Learning System Actions

The system **preprocesses the input data** to remove noise and inconsistencies.

It performs **feature extraction**, identifying key characteristics relevant to flood prediction.

A **machine learning model is trained** using historical data.

The trained model is **evaluated** for accuracy and performance.

If needed, the model undergoes **optimization** to enhance accuracy. The final, optimized model is then **deployed** for real-time predictions.

System Actions

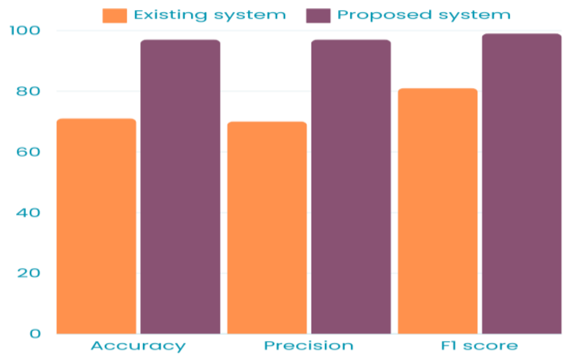
Sensors or other sources provide real-time data to the deployed model. According to the current situation the system will forecast the risk of flooding. It will inform the user with alert notification if it detects a possible flood.

Outcome

If there is a chance of flood or landslide the alert will send to the user. So the user can take appropriate action and take precautions.

Our methods and procedure guarantees an effective landslide detection that uses machine learning to evaluate the different data like environmental data, forecast hazards, and will provide alert in correct time.

VI. RESULT



	Proposed System	Exiting System
Accuracy	97	71
Precision	97	70
F1 score	99	81

VII.CONCLUSION

Our proposed system delivers enhanced performance capabilities to improve both accuracy and precision and increase F1 score compared to the current system. Our novel system delivers accuracy levels of 97% better than the existing system's 71% accuracy. The proposed system demonstrates increased accuracy by producing fewer wrong results that boosted accuracy from 70% up to 97%. The system achieves 99% F1 score performance through its ability to maintain equal precision and continuation. Our proposed system selection delivers enhanced security for implementation by showing improvement in accuracy and precision and operational efficiency. The research demonstrates that CNNs (Convolutional Neural Networks) enhance the capacity to predict natural disasters while helping prevent their consequences.

REFERENCES

- A. Wannachai, S. Aramkul, B. Suntaranont, Y. Somchit, and P. Champrasert introduced the HERO system, a hybrid and resilient operational station designed for early flash flood warnings. Their research appeared in *Sensors*, vol. 22, issue 11, article 4108, in 2022.
- S. Zhong, C. Wang, Z. Yu, Y. Yang, and Q. Huang explored spatiotemporal hazard mapping of tropical cyclones along China’s coastline, with their findings published in *Advances in Meteorology*, vol. 2018, in 2018.
- C. Hu, Q. Wu, H. Li, S. Jian, N. Li, and Z. Lou applied deep learning, specifically long short-term memory (LSTM) networks, to simulate rainfall-runoff processes. Their work was featured in *Water*, vol. 10, issue 11, article 1543, in 2018.
- Z. Fang, Y. Wang, L. Peng, and H. Hong utilized LSTM neural networks to predict flood susceptibility, as documented in *Journal of Hydrology*, vol. 594, article 125734, in 2021.