International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue5
Sep-Oct- 2026, PP 01-07

Predicting Natural Disasters Using Artificial Intelligence And Machine Learning: A Review Of Methods, Data Sources And Early Warning Architectures

Bhalerao Shruti Santosh[1] Prof. M. A. Auti[2] Prof. M. C. Kolse[3] Prof. S. K. Said[4]
Department of Artificial Intelligence and Data Science
Jaihind College of Engineering
Pune, Maharashtra, India

Abstract—Floods, earthquakes, cyclones, landslides and droughts cause heavy loss of life and property, and their impact depends strongly on how early communities are warned. Conventional physics-based forecasting remains essential, but it struggles to exploit the volume and variety of data now produced by satellites, ground sensors and historical archives. This paper reviews how Artificial Intelligence (AI) and Machine Learning (ML) are being used to estimate disaster risk and support early warning. It surveys data sources, including rainfall, river levels, seismic waveforms, reanalysis fields and satellite imagery, and examines models ranging from random forests and support vector machines for susceptibility mapping to long short-term memory networks for flood forecasting, convolutional and transformer networks for earthquake detection, and graph-based models for medium-range weather prediction. A five-layer architecture for an AI-assisted early warning system is proposed, linking data ingestion, preprocessing, hazard-specific models, risk fusion and alert dissemination. The paper identifies research gaps such as data scarcity in ungauged regions, limited uncertainty reporting and the weak connection between forecasts and community action. It concludes that AI predictions should be treated as decision support that complements, rather than replaces, physical models and expert judgement.

Index Terms—Natural disaster prediction, early warning system, machine learning, deep learning, flood forecasting, earthquake detection, remote sensing, risk assessment

I. Introduction

India is exposed to almost every major natural hazard: monsoon floods in the Ganga and Brahmaputra basins, cyclones on both coasts, earthquakes along the Himalayan arc, landslides in the Western Ghats and recurrent droughts in the Deccan. The National Disaster Management Plan [1] and the global Sendai Framework [2] both emphasize that investment in early warning and preparedness saves far more lives than post-disaster relief.

Traditional forecasting relies on numerical models derived from physical laws. These models are indispensable but computationally expensive, and they depend on accurate initial conditions that are not always available. Reichstein *et al.* [3] argued that deep learning can complement physical models by learning spatial and temporal patterns directly from Earth-observation data, provided that physical consistency is respected.

This paper reviews AI and ML techniques for predicting different natural hazards, proposes an architecture for an AI-assisted early warning system, and highlights the gaps that limit operational use, particularly in developing regions.

II. Theoretical Background

2.1 Prediction, Forecasting and Susceptibility

Three related tasks should be distinguished. Susceptibility mapping estimates where a hazard is likely to occur, for example which slopes are prone to landslides. Forecasting estimates when and how intensely a hazard will occur, such as the river

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue5
Sep-Oct- 2026, PP 01-07

level at a gauge over the next three days. Detection identifies a hazard as it begins, such as the first seismic waves of an earthquake. Fig. 1 classifies the AI methods applied to each hazard.

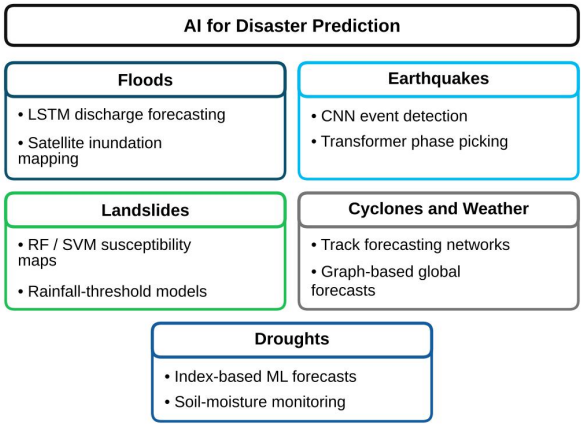


Fig. 1. Hazards and representative AI methods discussed in this review.

2.2 Data Sources

Inputs include rain-gauge and river-gauge time series, Doppler radar, satellite products for precipitation, soil moisture and land cover, digital elevation models, seismic waveforms from station networks, reanalysis datasets and historical disaster records. These sources differ in resolution, latency and reliability, so data fusion is a central design concern.

III. Survey of Existing Approaches

For floods, Kratzert *et al.* [4] showed that an LSTM network [5] trained on meteorological inputs can simulate daily river discharge across hundreds of catchments with accuracy comparable to calibrated hydrological models. Nevo *et al.* [6] described an operational system that combined LSTM-based stage forecasting with inundation mapping and delivered alerts to people in India and Bangladesh. Nearing *et al.* [7] later demonstrated that a globally trained model could provide reliable extreme-flood forecasts even in ungauged watersheds, extending early warning to regions without local data.

In seismology, Perol *et al.* [8] introduced ConvNetQuake, a convolutional network that detects and locates earthquakes from single-station waveforms far faster than template matching. Mousavi *et al.* [9] proposed the Earthquake Transformer, which uses attention to detect events and pick P- and S-wave arrivals simultaneously. DeVries *et al.* [10] used deep learning to predict aftershock locations from stress-change features, although later analyses showed that much simpler models achieve similar skill, a reminder that complexity must be justified.

For landslides, Pradhan [11] compared decision trees, support vector machines [12] and neuro-fuzzy models for GIS-based susceptibility mapping and found that data-driven models outperform traditional weighting. For cyclones, Giffard-Roisin *et al.* [13] fused historical tracks with reanalysis fields in a deep network to forecast storm trajectories.

At the scale of weather prediction, ConvLSTM [14] and deep generative nowcasting [15] improved short-term rainfall prediction, while GraphCast [16] showed that a graph neural network can produce ten-day global forecasts more accurately than a leading physics-based system on many variables. Table I summarizes these contributions.

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue5
Sep-Oct- 2026, PP 01-07

TABLE I Summary of Literature Survey

Ref.	Authors	Year	Technique Used	Merits	Demerits
[11]	Pradhan	2013	Decision tree, SVM and neuro-fuzzy in GIS	Data-driven landslide susceptibility maps	Static maps; no timing information
[4]	Kratzert et al.	2018	LSTM rainfall–runoff modelling	Matched calibrated hydrological models across many basins	Requires long discharge records
[8]	Perol et al.	2018	ConvNetQuake CNN	Fast detection and location of earthquakes	Single region; small catalogue
[10]	DeVries et al.	2018	Deep network on stress-change features	Improved aftershock location prediction	Simpler models later shown comparable
[9]	Mousavi et al.	2020	Earthquake Transformer	Joint detection and P/S phase picking	Does not forecast future quakes
[13]	Giffard-Roisin et al.	2020	Fused deep learning on tracks and reanalysis	Improved cyclone track forecasts	Intensity not addressed
[6]	Nevo et al.	2022	Operational ML flood forecasting and inundation mapping	Real alerts delivered in India and Bangladesh	Dependent on gauge availability
[16]	Lam et al.	2023	GraphCast graph neural network	Skilful 10-day global weather forecasts	Coarse resolution for local extremes
[7]	Nearing et al.	2024	Global AI model for extreme floods	Reliable forecasts in ungauged watersheds	Coarse for very small catchments

IV. Comparative Study

Table II compares the main model families. Tree ensembles are effective for susceptibility mapping with modest data, recurrent networks suit hydrological time series, convolutional and transformer networks excel on images and waveforms, and graph networks scale to global fields

.**TABLE II Comparison of Model Families**

Model	Best Suited For	Data Need	Interpretability
RF / SVM / XGBoost	Susceptibility mapping	Low–Medium	Medium (feature importance)
LSTM	Flood and discharge time series	Medium	Low
CNN	Satellite images, seismic waveforms	High	Low (saliency maps)
Transformer	Seismic phase picking	High	Low–Medium
Graph neural network	Global weather fields	Very high	Low

V. Research Gaps

The following gaps emerge from the survey:

- 1. Training data are scarce for rare extreme events and for regions with sparse monitoring networks.
- 2. Most models report point predictions; calibrated uncertainty estimates needed for decision-making are often missing.

- 3. Studies typically address a single hazard, whereas real events are often compound, such as cyclone-induced floods and landslides.
- 4. Explainability of predictions is limited, which reduces trust among disaster managers.
- 5. The link between a forecast and timely, understandable alerts for communities is rarely evaluated.

VI. Aims and Objectives

The objectives of this review are:

- **O1:** To survey AI and ML methods used for prediction and detection of major natural hazards.
- **O2:** To identify the data sources and preprocessing steps required for each hazard.
- **O3:** To propose an architecture for a multi-hazard, AI-assisted early warning system.
- **O4:** To compare model families by data needs, lead time and interpretability.
- **O5:** To discuss limitations and future directions for operational deployment in India.

VII. Proposed Architecture

The proposed architecture, shown in Fig. 2, follows the data-to-decision chain of an early warning system.

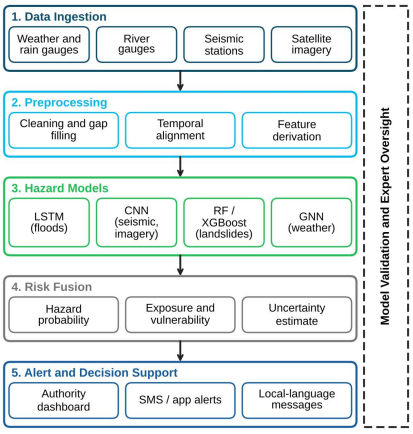


Fig. 2. Proposed architecture of an AI-assisted multi-hazard early warning system.

7.1 Data Ingestion Layer

This layer collects real-time feeds from weather stations, river gauges, seismometers and satellites, together with historical event catalogues and static layers such as elevation and land use.

7.2 Preprocessing Layer

Data are cleaned, gaps are filled by interpolation, time series are aligned to a common interval and satellite tiles are reprojected to a common grid. Features such as antecedent rainfall, slope and soil type are derived.

International Journal of Computational Research in Science and Technology
E-ISSN: 3108-0871, Volume 2 Issue5
Sep-Oct- 2026, PP 01-07

7.3 Hazard Model Layer

Separate models serve each hazard: LSTM networks for river levels, convolutional networks for satellite imagery and seismic waveforms, gradient-boosted trees [17] or random forests [18] for susceptibility maps, and graph models for large-scale weather fields.

7.4 Risk Fusion Layer

Hazard probabilities are combined with exposure and vulnerability data to estimate risk for each district or village. Uncertainty is propagated so that decision-makers see both the expected impact and its confidence range.

7.5 Alert and Decision Support Layer

Risk levels are shown on map dashboards for authorities and converted into colour-coded, local-language alerts sent by SMS, mobile apps and community radio.

VIII. Process Flow and Evaluation

The workflow in Fig. 3 describes the forecasting cycle. Model accuracy for continuous targets such as river level is commonly measured by the Nash–Sutcliffe Efficiency

$$NSE = 1 - \Sigma(Q_o - Q_s)^2 / \Sigma(Q_o - \bar{Q}_o)^2 \quad (1)$$

where Q_o is observed and Q_s simulated discharge. An NSE of one is a perfect match, while values below zero indicate that the model is worse than the observed mean. For categorical warnings, the probability of detection and the false-alarm ratio are reported together, because both missed events and excessive false alarms erode public trust.

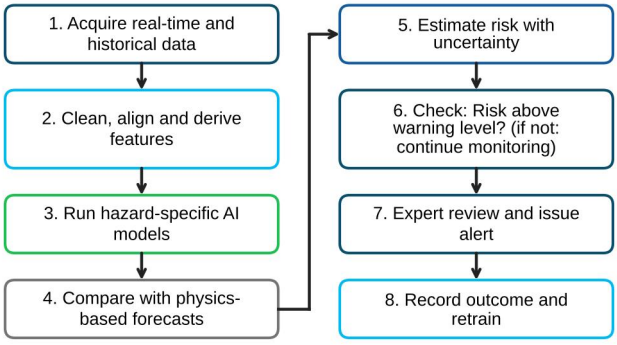


Fig. 3. Forecasting and warning workflow.

Models are retrained periodically as new events are recorded, and forecasts are compared with physics-based outputs; large disagreements trigger expert review before warnings are issued.

IX. Limitations and Future Scope

Table III lists key challenges. Promising directions include physics-informed neural networks that respect conservation laws, transfer learning from data-rich to data-poor basins, multi-hazard models that capture cascading events, and impact-based

forecasting that estimates damage rather than only the physical hazard. Integration with national systems and community feedback loops will determine whether improved forecasts translate into saved lives.

Challenge	Mitigation
Rare extreme events	Data augmentation, transfer learning, synthetic scenarios
Sparse monitoring	Satellite data and global models
Uncertainty	Ensembles and probabilistic outputs
Low trust	Explainable outputs; physics comparison
Last-mile alerting	Local-language, multi-channel dissemination

TABLE III Challenges and Mitigation

X. Conclusion

AI and ML have demonstrated clear value in forecasting floods, detecting earthquakes, mapping landslide susceptibility and predicting weather. The reviewed studies show that learning-based models can match or exceed traditional approaches in specific tasks and can extend warnings to regions without dense monitoring. Their effectiveness, however, depends on data quality, careful validation and transparent communication of uncertainty. Because natural systems are complex, AI outputs should be used as decision-support information within an early warning chain that also includes physical models, expert judgement and community preparedness.

10.1 Acknowledgment

The authors express their sincere gratitude to the Department of Artificial Intelligence and Data Science, Jaihind College of Engineering, Pune, for providing the facilities, guidance and encouragement needed to carry out this work. The authors also thank the faculty members and colleagues whose valuable suggestions helped improve the quality of this review.

References

1. National Disaster Management Authority, National Disaster Management Plan. New Delhi: Government of India, 2019.
2. United Nations Office for Disaster Risk Reduction, Sendai Framework for Disaster Risk Reduction 2015–2030. Geneva: UNDRR, 2015.
3. M. Reichstein et al., “Deep learning and process understanding for data-driven Earth system science,” Nature, vol. 566, pp. 195–204, 2019.
4. F. Kratzert, D. Klotz, C. Brenner, K. Schulz, and M. Herrnegger, “Rainfall–runoff modelling using long short-term memory (LSTM) networks,” Hydrol. Earth Syst. Sci., vol. 22, no. 11, pp. 6005–6022, 2018.
5. S. Hochreiter and J. Schmidhuber, “Long short-term memory,” Neural Comput., vol. 9, no. 8, pp. 1735–1780, 1997.
6. S. Nevo et al., “Flood forecasting with machine learning models in an operational framework,” Hydrol. Earth Syst. Sci., vol. 26, no. 15, pp. 4013–4032, 2022.
7. G. Nearing et al., “Global prediction of extreme floods in ungauged watersheds,” Nature, vol. 627, pp. 559–563, 2024.
8. T. Perol, M. Gharbi, and M. Denolle, “Convolutional neural network for earthquake detection and location,” Sci. Adv., vol. 4, no. 2, Art. no. e1700578, 2018.

9. S. M. Mousavi, W. L. Ellsworth, W. Zhu, L. Y. Chuang, and G. C. Beroza, “Earthquake transformer—An attentive deep-learning model for simultaneous earthquake detection and phase picking,” Nat. Commun., vol. 11, Art. no. 3952, 2020.
10. P. M. R. DeVries, F. Viégas, M. Wattenberg, and B. J. Meade, “Deep learning of aftershock patterns following large earthquakes,” Nature, vol. 560, pp. 632–634, 2018.
11. B. Pradhan, “A comparative study on the predictive ability of the decision tree, support vector machine and neuro-fuzzy models in landslide susceptibility mapping using GIS,” Comput. Geosci., vol. 51, pp. 350–365, 2013.
12. C. Cortes and V. Vapnik, “Support-vector networks,” Mach. Learn., vol. 20, no. 3, pp. 273–297, 1995.
13. S. Giffard-Roisin et al., “Tropical cyclone track forecasting using fused deep learning from aligned reanalysis data,” Front. Big Data, vol. 3, Art. no. 1, 2020.
14. X. Shi et al., “Convolutional LSTM network: A machine learning approach for precipitation nowcasting,” in Proc. NeurIPS, 2015, pp. 802–810.
15. S. Ravuri et al., “Skilful precipitation nowcasting using deep generative models of radar,” Nature, vol. 597, pp. 672–677, 2021.
16. R. Lam et al., “Learning skillful medium-range global weather forecasting,” Science, vol. 382, no. 6677, pp. 1416–1421, 2023.
17. T. Chen and C. Guestrin, “XGBoost: A scalable tree boosting system,” in Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min., 2016, pp. 785–794.
18. L. Breiman, “Random forests,” Mach. Learn., vol. 45, no. 1, pp. 5–32, 2001.