Agrociencia

INTERNET OF THINGS SENSOR DATA AND META-ALGORITHMIC APPROACHES FOR ADVANCED CLIMATE CHANGE-RELATED NATURAL DISASTER PREDICTION

Thirumalaimuthu **Babu**[1], Karuppaiya Sathaiah **Balamurugan**[2],
Kanmani Pappa **Chandramohan**[3], Surendran **Rajendran**[4*]

[1]St. Joseph's College of Engineering. Department of Electrical and Electronic Engineering. Chennai, Tamil Nadu, India. 600119.
[2]Karpaga Vinayaga College of Engineering and Technology. Department of Electronics and Communication Engineering. Chengalpattu, Tamil Nadu, India. 603308.
[3]Vel Tech Rangarajan Dr. Sagunthala R&D Institute of Science and Technology. Department of Electronics and Communication Engineering. Chennai, Tamil Nadu, India. 600062.
[4]Saveetha Institute of Medical and Technical Sciences. Department of Computer Science and Engineering. Chennai, Tamil Nadu, India. 602105.
* Author for correspondence: dr.surendran.cse@gmail.com

ABSTRACT

Natural disasters are catastrophic events caused by natural phenomena such as earthquakes, floods, hurricanes, and wildfires, resulting in severe damage, destruction, and human suffering. Existing methods for predicting natural disasters often lack precision because of the complexity and variability of natural processes, limited data availability, and the unpredictable cascading effects of initial events such as earthquakes. In this work, a range of machine learning techniques was applied for predicting various natural disasters under climate change: Long Short-Term Memory (LSTM) networks for earthquakes, Support Vector Machines (SVM) for tsunamis, Convolutional Neural Networks (CNN) for cyclones, and Random Forest (RF) for extreme temperatures. These models were integrated into a comprehensive meta-algorithm, enhanced by the Internet of Things (IoT) for real-time data collection and analysis. Performance evaluation against traditional models, including the Ensemble Decision Tree model and the Logistic Discriminant model, showed that the meta-algorithm achieved 5 % greater accuracy, highlighting its effectiveness in natural disaster prediction.

Keywords: machine learning, Meta Algorithm, data collection, data analysis.

Citation: Babu T, Balamurugan KS, Chandramohan KP, Rajendran S. 2025. Internet of things sensor data and meta-algorithmic approaches for advanced climate change-related natural disaster prediction.
Agrociencia. https://doi.org/ 10.47163/agrociencia.v59i6.3427

Editor in Chief:
Dr. Fernando C. Gómez Merino

Received: March 19, 2025.
Approved: September 12, 2025.
Published in Agrociencia:
September 23, 2025.

INTRODUCTION

The Internet of Things (IoT) incorporates several enabling technologies, including 5G connectivity, advanced sensors, intelligent actuators, and software applications, and has become a crucial component of modern information technology and smart communications (Balamurugan *et al*., 2024). These technologies together establish a framework for smart connectivity that addresses key challenges such as data transfer in cloud-edge services, functional performance, availability, and quality of service (QoS) metrics.

Intelligent sensors and actuators in the IoT environment are constrained by limited storage capacity and short battery life, which restrict their ability to retain large volumes of data for intelligent networking (Vinuja and Devi, 2024). Efficient information broadcasting is essential for transferring data from the IoT layer to the distributed computing layer for processing and storage. Several studies have shown that the IoT framework is increasingly applied to safety-critical situations, such as wildfires, earthquakes, floods, blizzards, hurricanes, seasonal tornadoes, and landslides, where it supports disaster management and response. Due to global warming, natural disasters occur with increasing frequency and intensity. It is essential to implement preventive measures prior to such events, collect and analyze information in real time during their occurrence, and conduct accurate post-event assessments. The fatalities and injuries from these disasters not only cause long-term societal trauma but also result in significant economic losses. Voosen et al (2025) discussed according to the National Oceanic and Atmospheric Administration (NOAA), the global average temperature increased by approximately 1 °C between 1901 and 2020. This rise has contributed to sea levels rising at nearly twice their historical rate, while melting glaciers faster.

The National Aeronautics and Space Administration (NASA) has reported that climate-related changes intensify the natural disasters experienced worldwide, increasing both their frequency and diversity. Natural disasters have a profound impact on people's lives, causing displacement, loss of housing, and disruption of livelihoods. They also cause physical injuries, emotional trauma, and severe psychological distress. According to the International Monetary Fund (IMF), the negative consequences are most severe for low-income countries. The lack of access to critical resources such as safe drinking water, food, and healthcare exacerbates these issues. Communities frequently face long-term economic hardship and recovery barriers as a result of severe damage to support systems and infrastructure.

IoT and artificial intelligence (AI) are increasingly intertwined, transforming various sectors. IoT devices collect massive amounts of data, which AI algorithms analyze to produce actionable insights and support autonomous decision-making (Umamaheswari *et al*., 2025). This combination improves efficiency in industries like manufacturing, healthcare, and transportation by enabling predictive maintenance, personalized medicine, and self-driving vehicles. AI strengthens IoT by increasing data processing speed, accuracy, and security, while IoT extends AI's reach by collecting real-time data from interconnected devices. Together, these technologies foster smarter, more adaptive, and responsive systems across diverse applications. Despite significant advances in AI and IoT, predicting specific natural disasters, particularly earthquakes, remains a major scientific challenge. Their inherently unpredictable nature, absence of reliable precursors, and the complexity of subsurface processes make long-term forecasting extremely difficult. In seismically active regions such as Mexico, traditional prediction approaches remain virtually unfeasible. Suárez *et al*.

(2022) developed the early warning systems, such as the Mexican Seismic Alert System (SASMEX), provide only seconds to a few minutes of lead time by detecting initial seismic waves. This emphasizes the need to develop hybrid frameworks that integrate real-time sensing, rapid alerts, and efficient communication with communities, rather than relying exclusively on long-term predictions.

Urban floods, caused by the unpredictable intensity of rainfall, can cause significant damage and are highly nonlinear events that are difficult to analyze. Keum *et al*. (2020) proposed a real-time flood prediction system based on classification by utilizing linear and nonlinear regression techniques. The system incorporated screening methods such as Probabilistic Neural Networks (PNN) and Latin Hypercube Sampling (LHS) to enhance forecasting accuracy, achieving 85 % accuracy with a processing time of 1 min and 12 s. Additionally, instruments designed according to the disaster management framework were developed to ensure alignment with actual emergency response procedures (Elshami *et al*., 2025). The integration of expert evaluations and statistical validation techniques, including factor analysis and consistency assessments, enhances the reliability and practical applicability of these tools across diverse healthcare settings. Suhardono *et al*. (2025) highlighted the importance of integrating community education, early warning systems, and infrastructure planning to strengthen disaster resilience in coastal regions. Using a choice experiment methodology, researchers assessed public preferences and the economic valuation of preparedness strategies. Multi-scenario frameworks that integrate educational, logistical, and ecological components have proven to be effective models for comprehensive disaster management.

Duraisamy and Natarajan (2024) applied an enhanced Bidirectional Encoder Representations from Transformers (BERT) model with Twitter data for disaster prediction, showing superior performance compared to Gated Recurrent Unit (GRU) and Memory-Augmented Neural Networks (MANN). Khan *et al*. (2024) introduced a non-intrusive, cost-effective geophysical approach that integrates regional-scale geology for detecting and forecasting rock bursts in coal mines in northwestern China. Yang *et al*. (2023) proposed a hybrid model combining Long Short-Term Memory (LSTM), Random Forest (RF), and Multitask Deep Belief Networks (MDBN) for landslide prediction. Overall, the combined model achieved high predictive reliability for landslide vulnerability. Sathianarayanan *et al.* (2024) introduced a Convolutional Neural Network (CNN) detection model along with multi-digit phone number data and street view house numbers, and achieved over 79 % of average precision, 82 % aggregated average precision, and an overlap coefficient of 0.5. Zheng *et al*. (2024) proposed a multi-factor geographic game restoration approach for forecasting mineral water inrush catastrophes, using Principal Component Analysis (PCA), a Whale Optimization Algorithm-Random Forest-Geographic Information System (WOA-RF-GIS), and Collaborative Kriging Interpolation to address data dimensionality. Özen and Souri (2024) introduced a method combining Analysis of Variance (ANOVA)-based feature selection with the Optimized Ensemble Bagged Tree (OEBT) algorithm, integrated with cloud-based prediction and 5G technology. Their approach achieved a reliability of 97.9 %, precision of 98.7 %, recall of 78.9 %, and an F1-score of 78.3 %. Park and Lee (2024) developed new hybrid fuzzy-Deep Neural Network (DNN) method for urban floods with high accuracy. Chai and Wu (2023) proposed a method with AI and multi-criteria decision-making (MCDM) for the early prediction of earth quakes through continuous monitoring. Powers *et al*. (2023) introduced a method for sending emergency messages on social media during natural disasters using deep learning techniques, including CNNs, BERT, and XLNet, with CNNs demonstrating the highest accuracy. Similarly, Wu *et al*. (2023) developed a conceptual model based on Data Envelopment Analysis (DEA) to assess regional disaster susceptibility across mainland China, using annual data from official Chinese statistics spanning 2006 to 2021. The model incorporated five input factors: population density, gross domestic product (GDP) per square kilometer, per capita GDP, total regional population, and regional investments in water conservation, environmental management, and public services. Zhu *et al*. (2021) proposed a flood forecasting technique that integrates the Ensemble Decision Tree model with Geographic Information System (GIS) technologies, demonstrating that the combination of the RF algorithm and GIS is effective for analyzing the spatial distribution and underlying patterns of flood risk. Anbarasan *et al*. (2020) developed a flood detection approach using IoT, Big Data, and Convolutional Deep Neural Networks (CDNNs), which outperformed established algorithms such as Feedforward Neural Networks (FNNs) and Hierarchical Neural Networks (HNNs) in terms of accuracy. These studies highlight the value of integrating advanced computational methods and diverse data sources to improve natural disaster prediction under varying conditions.

MATERIALS AND METHODS

An integrated system for natural disaster prediction was developed using IoT sensors and advanced machine learning models (Figure 1). Sensors such as seismometers, buoys, barometers, and climate stations collect real-time environmental data, which is transmitted to a cloud-based infrastructure for processing. In the cloud, machine learning models such as LSTM for earthquakes, SVM for tsunamis, CNNs for cyclones, and RF for extreme temperatures analyze incoming data to generate predictive insights.

Data preparation

Sensors continuously collect environmental data to support the prediction of earthquakes, tsunamis, cyclones, and extreme temperature events. Seismometers detect ground vibrations for earthquake forecasting, while ocean buoys and pressure sensors monitor sea level changes for tsunami detection. Climate stations measure wind speed, atmospheric pressure, and temperature to track cyclones and extreme heat or cold events, enabling timely and informed disaster preparedness.

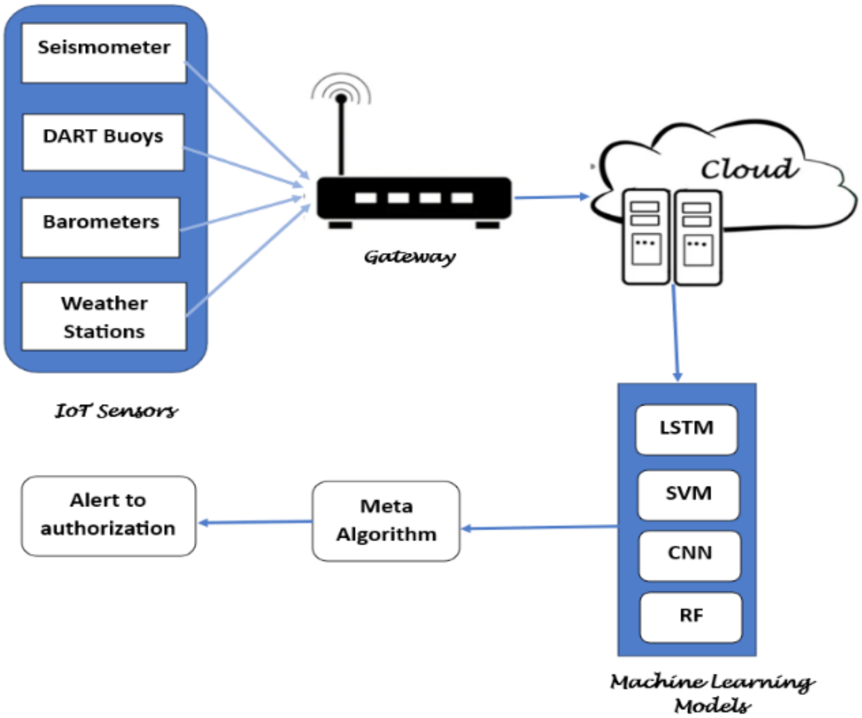

Figure 1. Proposed integrated Internet of Things (IoT) and machine learning model for real-time natural disaster prediction.

Datasets for this study were obtained from Kaggle, a platform providing curated data across multiple domains to support data-driven research and analysis. The earthquake dataset includes approximately three million records of seismic events worldwide from 1990 to 2023, with each entry detailing attributes such as timestamp, geographic coordinates (latitude and longitude), magnitude, epicenter depth, magnitude measurement type, affected area, and other relevant information.

The Global Historical Tsunami Database catalogs over 2400 tsunami events from 2100 BC to the present, covering the Atlantic, Indian, and Pacific Oceans, as well as the Mediterranean and Caribbean Seas. The National Hurricane Center (NHC) maintains the HURDAT historical tropical cyclone database, which includes Atlantic HURDAT2 and Northeast/North Central Pacific HURDAT2 datasets. These records, which are updated every six hours, include cyclone positions, peak wind speeds, central pressures, and storm dimensions for all tropical and subtropical cyclones.

For extreme temperatures, the Land Max Temperature and Land Min Temperature datasets were used, containing the highest and lowest average global land temperatures, respectively, each with associated uncertainty intervals. Additionally, the Land and Ocean Average Temperature dataset provides combined global land and ocean temperatures in Celsius, along with its corresponding uncertainty measurements.

Preprocessing

Data preprocessing was performed prior to transmission to the cloud to reduce bandwidth usage, improve data quality, minimize false alarms, enhance processing efficiency, and lower costs. Sensor data noise was mitigated using the moving average technique, which smooths measurements by averaging values within a sliding window:

$$MA_i = \frac{1}{\omega}\sum_{i=t-\omega+1}^{t} x_i$$

where MA_i represents moving average value at time t, ω is the size of window, and x_i represents data points within the window.

To reduce the volume of data transmitted to the cloud, downsampling was applied, decreasing the frequency of data points by selecting a representative subset. For a time series $X + \{x_1, x_2, \ldots x_n\}$ and a down sampling factor k:

$$X' = \{x_1, x_{1+k}, x_{1+2k}, \ldots\}$$

where X' is the down sampling dataset, and k is the down sampling factor.

The processed data is transmitted to the cloud via a cloud gateway, which serves as an intermediary between the sensors and the cloud storage platform. In the cloud, data is securely stored in a scalable infrastructure, enabling efficient access, management, and analysis. Machine learning models are then applied to process the data, train predictive models, and generate actionable insights for different types of natural disasters (Santhanaraj *et al*., 2023). This integration of cloud storage and machine learning allows large-scale processing of sensor data, supporting real-time decision-making and analytics.

Hyperparameter tuning

Every machine learning model requires careful parameter tuning to achieve optimal performance. The systematic search technique was employed for this purpose, specifically using Grid Search, which evaluates a predefined set of parameters to identify the configuration that maximizes model performance (Surendran *et al*., 2023). The process involves several sequential steps to determine the best hyperparameters. First, the machine learning model M must be defined. A decision tree classifier $M(\theta)$ uses a parameter set θ that includes the maximum tree depth (d), the minimum number of samples required to split an internal node (s), and the minimum number of samples required at a leaf node (l). Hyperparameters along with their range of values need to be defined next. For example, let θ={d,s,l} where:

$d \in \{3,5,7,10\};\ s \in \{2,5,10\};\ l \in \{1,2,4\}$

A parameter grid Θ is created as the Cartesian product of all hyperparameter values:

$$\Theta = \{ (d, s, l) \mid d \in \{3,5,7,10\}, s \in \{2,5,10\}, l \in \{1,2,4\} \}$$

To evaluate system performance, an appropriate scoring metric *S* is selected. For classification tasks, this is typically accuracy (*A*), calculated as:

$$A = \frac{\text{Number of correct predictions}}{\text{Number of test cases}}$$

The dataset *D* was partitioned into training and testing sets, and *k*-fold cross-validation was employed to ensure robust performance evaluation. The model is trained and validated *k* times, each time using a different subset as the validation set and the remaining *k*-1 subsets for training. The overall performance is calculated as the average across all *k* iterations:

$$S_{cv} = \frac{1}{k}\sum_{i=1}^{k} S_i$$

Grid search is implemented by iterating over each combination of hyperparameters *θ* ∈ Θ, training the model M(*θ*) on the training dataset, and evaluating its performance using the selected scoring metric on the validation set. The optimal set of model parameters is denoted as *θ**:

$$\theta^* = \arg max_{\theta \in \Theta} S_{cv}(\theta)$$

After determining the optimal hyperparameters *θ**, the model M(*θ**) is retrained on the full training dataset and its performance is evaluated on an independent test set *T* to assess generalization capability:

$$S_{test} = S(M(\theta^*), T)$$

After completing hyperparameter tuning and cross-validation, each machine learning model was trained for validation. The selected algorithms for predicting each natural disaster are as follows: Earthquakes, Tsunami, Cyclone, and Extreme Temperature.

The likelihood and magnitude category of future earthquake events were predicted using time-series data, including seismic wave velocity, magnitude, and depth. A three-layer LSTM network with 64, 32, and 16 units was used, followed by a dense layer with sigmoid activation for binary classification. The input shape was *(batch_ size, time_steps, features)*, with features comprising magnitude, depth, and timestamp intervals. The model was compiled using binary cross-entropy loss and the Adam optimizer.

Model 1. Long Short-Term Memory (LSTM) earthquake prediction

The input to the LSTM model is a multivariate time series, where each sequence represents seismic activity over time. Each time step includes magnitude of the earthquake (float), depth of the event (float), latitude and longitude (float), and timestamp (converted to a numerical or cyclic format).

The final input shape for the LSTM model is:

$$X_{train} \in R^{(N,T,F)}$$

where N is the number of samples, T represents time steps, and F represents number of features per step.

The output is a single predicted magnitude of the next earthquake event:

$$y_{train} \in R^{(N,1)}$$

The model learns temporal patterns to forecast the magnitude of the next event based on previous sequences.

Model 2. Support Vector Machine (SVM) for tsunami prediction

A Support Vector Machine (SVM) classifier with a radial basis function (RBF) kernel was employed to determine whether a seismic or oceanic event would trigger a tsunami. Input features included wave height (1–30 m), sea-level pressure, water displacement, sea surface temperature, earthquake magnitude, and geographic coordinates (latitude and longitude). Hyperparameters C and γ were optimized using grid search. The target variable y was a binary label indicating tsunami occurrence (1 = tsunami, 0 = no tsunami).

Model 3. Convolutional Neural Networks (CNN) for cyclone prediction

Convolutional Neural Networks (CNN) for cyclone prediction to classify the risk level of extreme temperature events (hot or cold) based on historical temperature, humidity, and pressure data, used a RF classifier with 100 estimators and gini criterion. Input features included land max temp, min temp, and surface pressure. Hyperparameters were tuned using GridSearchCV.

Model 4. Random Forest (RF) for prediction of extreme temperatures

Random Forest (RF) for prediction of extreme temperatures used after the completion of training and validation for each model using grid search, a meta-algorithm can be utilized to add the outputs of these individual models to achieve higher accuracy and robustness.

Meta Algorithm

Ensemble learning combines multiple models to achieve predictive performance superior to any single model (Zheng *et al*., 2024). By leveraging a diverse set of models, ensemble methods can correct individual errors, improving overall accuracy. In this study, the Adaptive Boosting (AdaBoost) algorithm was employed, a technique recently applied in cloud–IoT disaster management systems for enhanced reliability and precision (Sathianarayanan *et al*., 2024). Each model in the sequence is trained on a reweighted dataset, emphasizing previously misclassified instances, and the final prediction is obtained as a weighted aggregate of the individual model outputs.

AdaBoost is an ensemble learning method that combines multiple weak classifiers, typically decision stumps, into a strong predictive model. It trains models while reweighting misclassified instances, allowing subsequent classifiers to focus on harder cases. This iterative process reduces bias and improves the overall accuracy of the final model. Mathematically, the following is a description of the AdaBoost algorithm:

Initialization

All training instances are initially assigned equal weights. Given a training dataset $\{(x_1, y_1), (x_2, y_2), \dots (x_m, y_m)\}$, where x_i are the input features and $y_i \in \{-1,1\}$ are the class labels, the initial weight $w_1(i)$ for each instance i is:

$$w_1(i) = \frac{1}{m} \; for \; i = 1,2,\dots,m$$

During each boosting iteration $t = 1, 2, \dots, T$, a weak learner $h_t(x)$ is trained on the weighted dataset. The performance of each learner is evaluated by computing its weighted error rate ($\in_t$):

$$\epsilon_t = \sum_{i=1}^{m} w_t(i).I\big(y_i \neq h_t(x_i)\big)$$

where I is the indicator function, returning one if the condition is met and zero otherwise. Based on this error rate, the weight α_t of the weak learner is determined as:

$$\alpha_t = \frac{1}{2}\ln\left(\frac{1-\epsilon_t}{\epsilon_t}\right)$$

This weight reflects the contribution of each weak learner to the final ensemble, with lower-error learners receiving higher influence in the aggregated prediction.

Updating instance weights

For the following iteration, modify the training entities' sizes. Accurately identified cases are assigned a lower value than occurrences that are misclassified by the present weak learner:

$$w_{i+1}(i) \ = \ w_t(i). \exp\,(\alpha_t. I(y_i \neq h_t(x_i))$$

To maintain a valid probability distribution, the updated weights are normalized so that they sum to one:

$$w_{t+1}(i) = \frac{w_{t+1}(i)}{\sum_{j=1}^{m} w_{t+1}(j)}$$

Final classifier

After completing all iterations, the final classifier *H*(*x*) is obtained as a weighted sum of the individual weak learners:

$$H(x) = sign \sum_{t=1}^{T} \big(\alpha_t h_t(x)\big)$$

The class label is determined by the sign of the weighted sum, where the *sign* function assigns the predicted class based on whether the sum is positive or negative.

The ensemble algorithm integrates the outputs of multiple predictive models to determine the type of impending natural disaster, enhancing both accuracy and reliability. Once generated, predictions are transmitted through the Emergency Alert System (EAS) to authorities and emergency management agencies, enabling timely response, efficient resource allocation, and the issuance of public warnings. By continuously updating with new data, the system strengthens disaster preparedness and response.

Experimental Setup

The models were implemented in Python (version 3.9) with several open-source libraries. NumPy (v1.23.5) was used for numerical computation, matrix operations, and data manipulation. Keras (v2.11.0), a high-level neural network API, was used to build and train the LSTM and CNN models. Scikit-learn (v1.2.2) was used to implement the SVM, RF, and AdaBoost classifiers, as well as to compute evaluation metrics including accuracy, precision, and recall. Pandas (v1.5.3) facilitated dataset loading, preprocessing, and structuring. All tests were carried out on a system with an

Intel i5 processor and 64 GB of RAM to meet the computational demands of training and evaluation.

Given the scale and complexity of real-time disaster prediction, characterized by high-frequency sensor streams and large spatiotemporal datasets, advanced infrastructure is required. The proposed system is intended for deployment on cloud-based platforms that provide scalable computing resources, distributed storage, and real-time analytics capabilities. Future implementation will leverage services such as Amazon Web Services (AWS), Google Cloud Platform (GCP), or Microsoft Azure to ensure feasibility and robustness under production workloads.

The dataset covers a broad spectrum of natural disasters and their associated parameters (Table 1). The established thresholds reveal the wide range of excessive and unpredictable conditions that natural disasters impose on different global regions. Threshold values for each natural disaster were defined using historical records, official monitoring agency datasets such as the United States Geological Survey (USGS), the National Oceanic and Atmospheric Administration (NOAA), and the World Meteorological Organization (WMO), and extreme events reported in the literature. These values represent the lower limits where impacts become significant and the upper limits corresponding to the most severe documented cases.

Table 1. Minimum and maximum threshold values of various natural disasters.

Natural Disaster	Minimum Threshold	Maximum Threshold
Earthquake	Magnitude 2.5 on Richter Scale	Magnitude 9.5 on Richter Scale
Tsunami	Wave height 0.5 m	Wave height 30 m
Cyclone	Wind speed 74 mph	Wind speed 157+ mph
Extreme Temperature	Below -30 °C (Extreme Cold)	Above 45 °C (extreme hot)

Machine learning models require performance metrics to assess their effectiveness across different tasks. Forecasting accuracy is evaluated using metrics such as the F1-score, recall, precision, and overall reliability. The general consistency of predictions is primarily measured through accuracy, calculated as the proportion of correct predictions to total predictions:

$$Accurateness = \frac{Correct\ predictions}{Total\ predictions}$$

Precision measures the percentage of accurate positive projections:

$$Precision = \frac{True\ positives}{True\ positives + False\ positives}$$

Recall assesses the fraction of actual positives correctly identified:

$$Recall = \frac{True\ positives}{True\ positives + False\ negatives}$$

The F1-score provides an equilibrium across precision and recall by taking the harmonic mean of the two criteria:

$$F1\ score = 2 \times \frac{Precision \times recall}{Precision + recall}$$

RESULTS AND DISCUSSION

Five machine learning algorithms were used, including LSTM combined with SVM, CNN, and RF for natural disaster prediction models, as well as the AdaBoost Ensemble method to bring them together. The AdaBoost ensemble model demonstrates a balanced improvement across all disaster types, effectively leveraging the strengths of individual models (Table 2). This confirms the advantage of using a meta-algorithm for integrated disaster prediction.

Table 2. Performance evaluation of the machine learning models for natural disaster classification.

Model	Disaster type	Accuracy	Precision	Recall	F1-score
LSTM	Earthquake	0.94	0.91	0.90	0.96
SVM	Tsunami	0.95	0.95	0.96	0.91
CNN	Cyclone	0.93	0.98	0.92	0.95
RF	Extreme temperatures	0.97	0.93	0.94	0.96
AdaBoost Ensemble	Combined disaster	0.96	0.97	0.95	0.97

LSTM: Long Short-Term Memory; SVM: Support Vector Machine; CNN: Convolutional Neural Network; RF: Random Forest.

This research evaluated the performance of LSTM, SVM, CNN, RF, and the Ensemble Method using the metrics of accuracy, precision, recall, and F1-score (Figure 2). The LSTM model achieved strong accuracy and F1-score but showed comparatively lower precision and recall. The SVM model excelled in precision but had a lower recall, while accuracy and F1-score remained balanced. The CNN model showed high recall but slightly lower precision and accuracy than the others. The RF model achieved the highest precision with balanced accuracy, though recall was lower. Finally, the Ensemble Method produced consistently strong results across all metrics, demonstrating its effectiveness in capitalizing on the strengths of individual models.

From 1990 to 2003, global disaster data (Figure 3) indicates that small-scale and large-

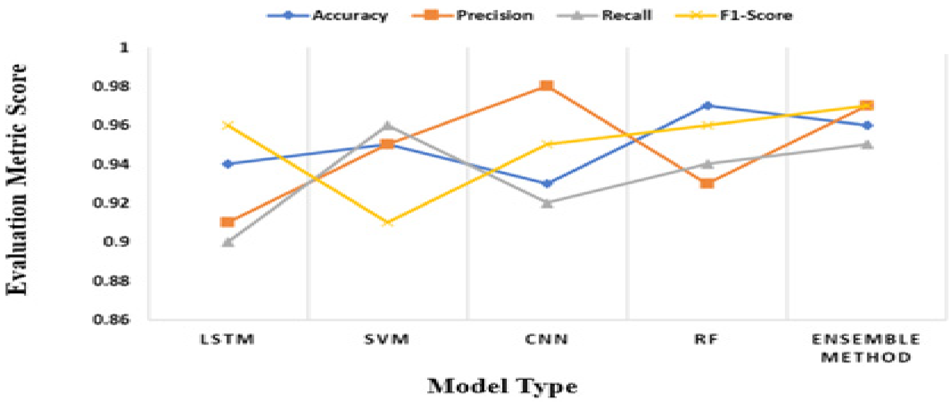

Figure 2. Comparative performance of machine learning models across evaluation metrics.

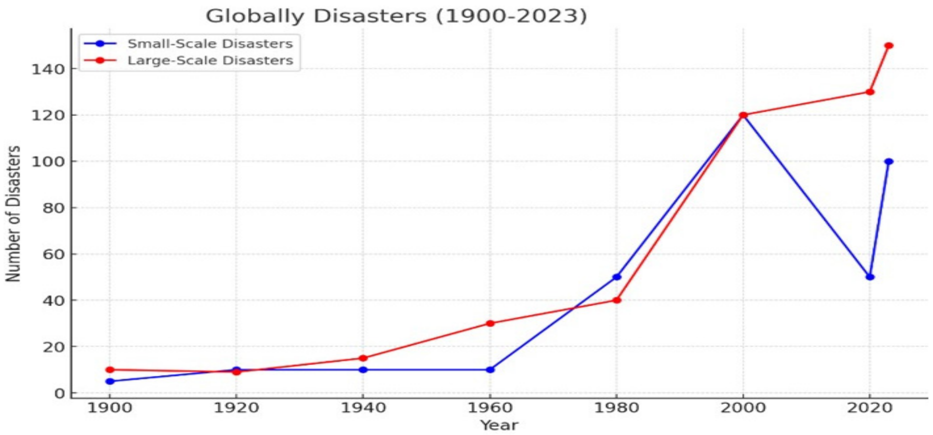

Figure 3. Global number of reported disasters by size, 1900 to 2023 (Small and Large -scale disasters events).

scale disasters events are characterized by a specific set of criteria. Small-scale disasters events involve five or fewer fatalities, affecting fewer than 1500 people, or resulting in economic losses below USD 13 million. Large-scale disasters events involving at least 50 fatalities, affecting more than 150 000 individuals, or causing economic damages of at least USD 320 million. This categorization provides a framework for understanding the severity and impact of such events over the analyzed period.

The study presents a comparative analysis of various earthquake prediction methods and their reported performance (Table 3). Keum *et al*. (2020) and Yang *et al*. (2023) both achieved 85 % accuracy using linear and non-linear regression techniques and the LSTM-RF-MDBN model, respectively. Sathianarayana *et al*. (2024) used CNNs, reaching 79 % accuracy. Özen and Souri (2024) used Bagged Tree and ANOVA techniques, achieving 95.5 % accuracy. Park and Lee (2024) implemented the Hybrid fuzzy and DNN model achieved 96 % accuracy. The proposed Ensemble AdaBoost method demonstrated the highest performance, reaching 96 % accuracy.

Table 3. Comparative performance of earthquake prediction methods.

Author	Methods	Performance
Keum *et al*. (2020)	Linear and Non-Linear Regression Techniques	85 %
Yang *et al*. (2023)	LSTM-RF-MDBN	85 %
Sathianarayana *et al*. (2024)	Convolutional Neural Networks	79 %
Özen and Souri (2024)	Bagged Tree and ANOVA techniques	95.5 %
Park and Lee (2024)	Hybrid fuzzy and DNN	95 %
Proposed method	Ensemble AdaBoost	96 %

CONCLUSIONS

A meta-algorithm-based framework was developed for predicting multiple natural disasters by integrating IoT sensor data with machine learning models. Combined using the AdaBoost ensemble technique, the system achieved high accuracy, precision, recall, and F1-scores, demonstrating its effectiveness for disaster forecasting and real-time alert generation. Despite the promising results, the proposed system has certain limitations. The model's performance heavily depends on the availability and quality of historical datasets, which may be sparse or inconsistent for certain regions or disaster types. Moreover, it has so far been validated only using historical data in a simulation environment. Real-time field deployment and integration with government-level emergency response systems remain untested. Future work includes enabling real-time data ingestion from IoT networks, implementing geospatial visualization and decision-support dashboards, incorporating deep learning-based anomaly detection, and conducting field trials with disaster management authorities to validate the system under operational conditions.

ACKNOWLEDGEMENTS

Data availability

https://www.kaggle.com/datasets/alessandrolobello/the-ultimate-earthquake-dataset-from-1990-2023

https://www.kaggle.com/datasets/andrewmvd/tsunamidataset?select= tsunami_dataset.csv

https://www.kaggle.com/datasets/noaa/hurricane-database

https://www.kaggle.com/datasets/berkeleyearth/climate-change-earth-surface-temperature-data

REFERENCES

Anbarasan M, Muthu B, Sivaparthipan CB, Sundarasekar R, Kadry S, Krishnamoorthy S, Samuel RDJ, Dasel AA. 2020. Detection of flood disaster system based on IoT, big data and convolutional deep neural network. Computer Communications 150: 150–157. https://doi.org/10.1016/j.comcom.2019.11.022

Balamurugan M, Sathesh M, Ramakrishnan K, Raja M, Kalaiarasi K. 2024. Integrating Remote sensing technologies for real-time disaster management, mitigation, and decision support systems. *In* 2024 International Conference on Recent Advances in Science and Engineering Technology. Institute of Electrical and Electronics Engineers: Nagara, India. https://doi.org/10.1109/icraset63057.2024.10895763

Chai J, Wu HZ. 2023. Prevention/mitigation of natural disasters in urban areas. Smart Construction and Sustainable Cities 1 (1). https://doi.org/10.1007/s44268-023-00002-6

Duraisamy P, Natarajan Y. 2024. Twitter disaster prediction using different deep learning models. SN Computer Science 5 (1). https://doi.org/10.1007/s42979-023-02520-7

Elshami S, Ibrahim MIM, Abdel-Rahman ME, Rahim HA, Mukhalalati B. 2025. Developing and evaluating a disaster management assessment tool for health care practitioners. BMC Emergency Medicine 25 (1). https://doi.org/10.1186/s12873-025-01199-8

Keum HJ, Han KY, Kim HI. 2020. Real-time flood disaster prediction system by applying machine learning technique. KSCE Journal of Civil Engineering 24 (9): 2835–2848. https://doi.org/10.1007/s12205-020-1677-7

Khan M, He X, Guo J, Song D. 2024. Accurate prediction of indicators for engineering failures in complex mining environments. Engineering Failure Analysis 155: 107736. https://doi.org/10.1016/j.engfailanal.2023.107736

Özen F, Souri A. 2024. Cloud-based disaster management architecture using hybrid machine learning approach in IoT. Multimedia Tools and Applications 83 (29): 72357–72370. https://doi.org/10.1007/s11042-024-18333-6

Park K, Lee EH. 2024. Urban flood vulnerability analysis and prediction based on the land use using Deep Neural Network. International Journal of Disaster Risk Reduction, 101: 104231. https://doi.org/10.1016/j.ijdrr.2023.104231

Powers CJ, Devaraj A, Ashqeen K, Dontula A, Joshi A, Shenoy J, Murthy D. 2023. Using artificial intelligence to identify emergency messages on social media during a natural disaster: A deep learning approach. International Journal of Information Management Data Insights 3 (1): 100164. https://doi.org/10.1016/j.jjimei.2023.100164

Santhanaraj RK, Rajendran S, Tavera-Romero CA, Murugaraj SS. 2023. Internet of Things enabled energy aware metaheuristic clustering for real time disaster management. Computer Systems Science and Engineering 45 (2): 1561–1576. https://doi.org/10.32604/csse.2023.029463

Sathianarayanan M, Hsu PH, Chang CC. 2024. Extracting disaster location identification from social media images using deep learning. International Journal of Disaster Risk Reduction 104: 104352. https://doi.org/10.1016/j.ijdrr.2024.104352

Suhardono S, Lee CH, Suryawan IWK. 2025. Valuation of marine integrated disaster management amidst global warming in Southern Coast of Java, Indonesia. Marine Pollution Bulletin 211: 117446. https://doi.org/10.1016/j.marpolbul.2024.117446

Suárez, G., 2022. The seismic early warning system of Mexico (SASMEX): A retrospective view and future challenges. Frontiers in Earth Science, 10, p.827236.

Surendran R, Alotaibi Y, Subahi AF. 2023. Lens-oppositional wild geese optimization based clustering scheme for wireless sensor networks assists real time disaster management.

Computer Systems Science and Engineering 46 (1): 835–851. https://doi.org/10.32604/csse.2023.036757

Umamaheswari U, Suneel S, Ravikumar K, Baburao D, Upadhyay S, Swapna B. 2025. Design and development of a rescue robot to identify human presence in disaster scenarios by using artificial intelligence assisted sensor support. *In* 2025 6th International Conference on Mobile Computing and Sustainable Informatics. Institute of Electrical and Electronics Engineers: Goathgaun, Nepal. https://doi.org/10.1109/icmcsi64620.2025.10883140

Vinuja G, Devi NB. 2024. Leveraging 6G networks for disaster monitoring and management in remote sensing. *In* Development of 6G Networks and Technology, pp.115-143.

John Wiley and Sons: Hoboken, NJ, USA, pp: 115–143. https://doi.org/10.1002/9781394230686.ch5

Voosen, P., 2025. NASA, NOAA face major climate science cuts. Science (New York, NY), 388(6744), pp.237-238.

Wu L, Ma D, Li J. 2023. Assessment of the regional vulnerability to natural disasters in China based on DEA model. Sustainability 15 (14): 10936. https://doi.org/10.3390/su151410936

Yang X, Fan X, Wang K, Zhou Z. 2023. Research on landslide susceptibility prediction model based on LSTM-RF-MDBN. Environmental Science and Pollution Research 31 (1): 1504–1516. https://doi.org/10.1007/s11356-023-31232-x

Zheng Q, Wang C, Zhu Z. 2024. Research on the prediction of mine water inrush disasters based on multi-factor spatial game reconstruction. Geomechanics and Geophysics for Geo-Energy and Geo-Resources 10 (1). https://doi.org/10.1007/s40948-024-00761-1

Zhu Z, Zhang Y. 2021. Flood disaster risk assessment based on random forest algorithm. Neural Computing and Applications 34 (5): 3443–3455. https://doi.org/10.1007/s00521-021-05757-6