IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received 29 March 2025, accepted 8 May 2025, date of publication 12 May 2025, date of current version 23 May 2025.

Digital Object Identifier 10.1109/ACCESS.2025.3569499

RESEARCH ARTICLE

Neural-XGBoost: A Hybrid Approach for Disaster Prediction and Management Using Machine Learning

MUHAMMAD ASIM SALEEM [1], **ASHIR JAVEED** [2], **WATIT BENJAPOLAKUL** [1], **WATTANASAK SRISIRI** [1], **SURACHAI CHAITUSANEY** [1], (Member, IEEE), **AND PASU KAEWPLUNG** [1]

[1]Center of Excellence in Artificial Intelligence, Machine Learning, and Smart Grid Technology, Department of Electrical Engineering, Faculty of Engineering, Chulalongkorn University, Bangkok 10330, Thailand
[2]Department of Computer Science, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden

Corresponding authors: Watit Benjapolakul (Watit.B@chula.ac.th), Pasu Kaewplung (Pasu.K@chula.ac.th), and Surachai Chaitusaney (surachai.c@chula.ac.th)

This work was supported in part by Thailand Science Research and Innovation Fund Chulalongkorn University under Grant 6641/2566 and Grant IND_FF_68_331_2100_042; and in part by the Ratchadapisek Somphot Fund for the Center of Excellence in Artificial Intelligence, Machine Learning, and Smart Grid Technology and by the Second Century Fund (C2F), Chulalongkorn University.

ABSTRACT Effective disaster prediction is essential for disaster management and mitigation. This study addresses a multi-classification problem and proposes the Neural-XGBoost disaster prediction model (N-XGB), a hybrid model that combines neural networks (NN) for feature extraction with XGBoost for classification. The NN component extracts high-level features, while XGBoost uses gradient-boosted decision trees for accurate predictions, combining the strengths of deep learning and boosting techniques for improved accuracy. The N-XGB model achieves an accuracy of 94.8% and an average F1 score of 0.95 on a real-world dataset that includes wildfires, floods and earthquakes, significantly outperforming baseline models such as random forest, Support vector machine and logistic regression 85% accuracy. The balanced F1 scores for wildfires 0.96, floods 0.93, and earthquakes 0.96 demonstrate the model's robustness in multi-class classification. The Synthetic Minority Oversampling Technique (SMOTE) balances datasets and improves model efficiency and capability. The proposed N-XGB model provides a reliable and accurate solution for predicting disasters and contributes to improving preparedness, resource allocation and risk management strategies.

INDEX TERMS Disaster prediction, feature extraction, XGBoost, SMOTE, machine learning.

I. INTRODUCTION

A. BACKGROUND

Natural disasters, including wildfires, floods and earthquakes, cause major damage to infrastructure and economic activity, as well as loss of life [1], [2]. In the last two decades, both the number of climate-related disasters and their dangerous extent have increased. Between 2000 and 2020, more than 3.3 billion people worldwide were affected by climate-related events [3]. Increasing floods, droughts and heatwaves are a consequence of climate change in Central America, East Africa and South Asia. Due to inadequate disaster risk reduction and limited infrastructure, developing countries are particularly vulnerable to destructive events and therefore need data-driven disaster management strategies [4], [5].

The demand for accurate disaster prediction models [6] is becoming increasingly urgent due to this growing trend, as they help communities to better prepare and reduce the impact of damage [7]. The growing number of destructive natural disasters recorded in the last two decades have caused billions of dollars of damage worldwide (CRED-EMDAT, 2023). Climate-related events such as wildfires, floods

The associate editor coordinating the review of this manuscript and approving it for publication was Cesar Vargas-Rosales.

and earthquakes are occurring with increasing frequency, underlining the urgency of predictive approaches that are able to effectively anticipate such hazards [8].

The traditional approach to disaster prediction, which strongly depends on historical patterns and meteorological data, is generally less able to make reliable predictions for multiple classes [9]. These limitations are both due to the inability to work with large amounts of imbalanced classes in the dataset. For this reason, today's research has gone further and tried to find a machine-learning approach to improve the accuracy of disaster classification. However, important challenges remain unsolved. These include imbalanced classes, missing measurements, and the inadequacy of typical versions to capture complicated, nonlinear emergencies.

Global disaster databases such as EM-DAT are useful for carrying out prediction analyses. However, major problems such as missing data lead to biased estimates, lower prediction accuracy and limited applicability of the models to smaller or underrepresented areas [10]. Furthermore, the disaster databases usually show a large imbalance between classes [11]. Floods are over-emphasized, while wildfires and earthquakes are under-emphasized, leading to biased machine learning models [12]. The Synthetic Minority Over-sampling Technique (SMOTE) has already shown that it can handle this problem by providing a balanced number of samples of all disaster types.

The numerous disaster datasets, which contain temporal elements in addition to spatial and numerical components, become a challenge for the analysis [13]. Support Vector Machines (SVM) and Random Forests have difficulties in processing complex data that differ from those of traditional machine learning models. Deep learning models can automatically extract features as well as create models between non-linear relationships [14]. Deep learning becomes more effective when combined with gradient-boosted decision trees through XGBoost, as the algorithm shows the highest performance when processing tabular datasets [15], [16].

SVM and other models, as well as logistic regression, are very popular in disaster prediction modeling, but each has several weaknesses. The model performs poorly when processing datasets with an imbalanced class distribution of disaster types, as floods outweigh wildfires and earthquakes, which leads to prediction biases. The modeling of non-linear interactions between meteorological, geological and socio-economic elements exceeds its predictive capabilities. EM-DAT performs poorly when necessary fields in the datasets, such as economic loss data or geographic coordinates, are incorrectly entered or misreported. The shortcomings of current prediction methods indicate that a better data prediction system needs to be developed that manages imbalanced data and analyzes complicated relationships to increase the efficiency of category recognition for different disaster events.

To address these limitations, the Neural-XGBoost (N-XGB) model is proposed, which combines deep learning methods for feature extraction with gradient-boosted decision tree classification. N-XGB utilizes neural network extraction techniques to discover high-dimensional nonlinear connections between disaster features for better prediction. XGBoost serves as a classification system for these extracted features and proves capable of processing databases of imbalanced data and missing values while maintaining an interpretable process. The model uses SMOTE as a technique to handle class imbalance problems so that different disaster types are adequately represented. The advanced features implemented in N-XGB provide better results than the baseline solutions and show high accuracy and recall in addition to the F1 score, suggesting the great potential of the model for disaster prediction in real-world scenarios.

In this study, a machine learning model is developed using SMOTE for oversampling to train the model on balance classes [17], a neural network for feature extraction, and XGBoost for disaster event classification. The proposed system improves disaster classification performance by addressing incomplete data, managing imbalance between classes, and considering different attribute types. Ultimately, this model provides important predictive insights for disaster response organizations and enables better preparation and resource allocation in emergencies.

The proposed model is evaluated on real disaster datasets and addresses key challenges, including missing data, class imbalance and feature complexity. The main contributions of this work are summarized as follows:

1. Development of a hybrid deep learning-based disaster prediction model (N-XGB) integrating neural networks for feature extraction and XGBoost for classification.
2. Addressing class imbalance and missing data problems through advanced preprocessing techniques, including SMOTE and feature imputation.
3. Comparing the performance of N-XGB with traditional machine learning models to demonstrate its effectiveness in multi-class disaster classification.
4. Evaluating the impact of deep feature extraction on prediction performance to confirm the importance of hybrid models in disaster analytics.

By achieving these objectives, this study contributes to the development of more accurate and interpretable disaster prediction models that can help disaster management agencies make data-driven decisions.

II. LITERATURE REVIEW

A. TRADITIONAL MACHINE LEARNING APPROACHES

Advances in machine learning have significantly improved the prediction and classification of natural disasters and overcome the challenges associated with complex and heterogeneous datasets. Traditional models such as decision trees, random forests and SVM have proven their worth with structured data, but reach their limits when it comes to capturing non-linear relationships and processing high-dimensional inputs [18], [19], [20]. Ensemble methods such as XGBoost have shown significant performance in disaster prediction due to their robustness in handling imbalanced datasets, efficiency in dealing with missing values, and strong interpretability [21].

Despite these strengths, traditional machine learning models are highly dependent on the manual selection of features, which can result in the underlying patterns in the disaster data not being captured. The analysis faces the problem that data on floods receives more attention than other hazards such as wildfires and earthquakes. The disparity in the number of examples means that prediction results favor one set of outcomes at the expense of prediction in other contexts. Scientific research shows that the classification accuracy of random forest and SVM models remains average, but their success varies from one dataset to another, suggesting the need for more adaptive methods [20].

B. DEEP LEARNING-BASED MODELS

Deep learning architectures of the convolutional neural network (CNN), recurrent neural network (RNN), and Long Short-Term Memory (LSTM) types help to recognize disaster patterns better than before, as they automatically extract features [22], [23]. These models are ideal for processing satellite imagery and sequential data required for real-time disaster monitoring [24]. Research in the field of self-supervised learning is now helping neural networks to use different types of data, such as electronic records and geospatial, to predict better outcomes from complex input sets [16], [25]. Due to their resource requirements, deep learning models pose great difficulties when used in real-world environments. The internal structure of deep learning networks makes them difficult to understand for decision-makers in emergency response organizations. The models tend to favor the larger class group and get stuck on training patterns, especially if the data classes are non-uniform.

C. HYBRID APPROACHES: DEEP LEARNING + XGBoost

To mitigate the limitations of traditional and deep learning models, researchers have explored hybrid architectures that integrate deep learning-based feature extraction with ensemble classification techniques such as XGBoost [26], [27]. These hybrid models aim to capture non-linear relationships, manage heterogeneous features, and improve scalability without compromising computational efficiency and interoperability [28].

CNN-XGBoost, CNNs for extracting spatial features, which are then classified by XGBoost. However these models are less suitable for electronic disaster record (EDR) disaster data [26]. Similarly, LSTM-XGBoost models are designed for time series data but are limited by the need for large labelled datasets, which limits their applicability in different scenarios [27]. These hybrid methods help with classification, but they do not deal effectively with problems such as missing data and class imbalance.

D. ADDRESSING CLASS IMBALANCE AND MISSING DATA

The major problem in predicting disasters is the imbalance between classes. Events such as floods are disproportionately represented in the datasets, while others, such as wildfires and earthquakes, are missing [29]. The problem of this imbalance affects the performance of the model, especially when predicting multiple classes. The Synthetic Minority Oversampling Technique (SMOTE) is effective in mitigating this problem by generating synthetic examples of minority classes, thereby improving the generalization of machine learning models [30].

Another pressing problem is missing data, which is common in global disaster databases such as EM-DAT [31]. Incomplete datasets missing critical attributes such as geographic coordinates, economic damage, or extent can significantly impact model performance. Common imputation methods, including mean, median and mode substitution, vary in effectiveness depending on the characteristics of the dataset. Therefore, robust preprocessing and tailored imputation strategies are crucial to minimize bias and ensure accurate predictions [32].

Although techniques exist to address these issues, few hybrid models explicitly integrate both SMOTE and extended imputation into their pipelines. In particular, SMOTE is rarely used in conjunction with deep learning and XGBoost, leading to biased results even for otherwise advanced models.

E. RATIONALE FOR AN INTEGRATED FRAMEWORK

Although previous studies have made significant progress, there is no existing approach that fully integrates deep learning, XGBoost and SMOTE into a unified disaster prediction model. To address this gap, the current study introduces the Neural-XGBoost (N-XGB) framework. N-XGB utilizes neural networks for automatic feature extraction, applies XGBoost for robust and interpretable classification, and integrates SMOTE to address the imbalance between classes and ensure a fair representation of all disaster types.

By combining these techniques, the N-XGB model provides a more accurate, efficient and interpretable framework for disaster prediction compared to existing models. It also incorporates advanced data imputation strategies to effectively manage missing information and further increase the reliability of the model.

This study presents N-XGB, a hybrid model for disaster prediction that integrates neural network-based feature extraction and XGBoost classification. Through SMOTE oversampling and refined data preprocessing, N-XGB provides robust multi-class predictions and significantly improves accuracy for underrepresented disaster types. The model combines the strengths of deep learning and ensemble methods and supports data-driven preparation, optimized resource allocation and effective disaster response.

III. METHODOLOGY

A. DATA COLLECTION AND DESCRIPTION

This study includes 45 features and 5,091 data samples with various disaster records that are publicly available [33]. The most important attributes are the type of disaster, location, extent, and associated damage. This includes attributes that capture the spatial and temporal dimensions of disasters as well as the socio-economic impacts.

This study uses the dataset of the Center for Research on the Epidemiology of Disasters (CRED) EM-DAT: The International Disaster Database, known as a global and widely used disaster database. EM-DAT offers a comprehensive, high-quality dataset and is, therefore, one of the most reliable sources for disaster prediction research.EM-DAT was chosen because it is robust, comprehensive, and credible, as the quality and completeness of disaster datasets often vary widely. This ensures that our results are based on accurate and well-documented disaster records.

Table 1 provides a summary of the dataset, including its key columns.

B. DATA PRE-PROCESSING

The dataset contains significant missing values that must be carefully considered during preprocessing. More than 80% of the entries are missing longitude and latitude information, making it difficult to include precise geographic information. Critical metrics such as the total number of fatalities and the extent of damage are also missing around 29% and 68% of values, respectively. Financial data, including insured losses and reconstruction costs, is even sparser, with almost 90% of entries missing. These significant gaps necessitated the implementation of robust data processing strategies to ensure the usability of the dataset and maintain its integrity for machine learning tasks. The dataset was initially filtered to include only relevant disaster types: wildfires, floods and earthquakes. Missing values in numerical features were handled using median imputation, while categorical features were imputed with the most frequent value. After imputation, all categorical variables were coded with label encoding to ensure compatibility with both the neural network and the XGBoost components.

To ensure feature uniformity, the numeric variables were normalized using the StandardScaler, which centres and scales the data to unit variance. This step improves convergence in gradient-based models. Given the imbalance of classes in the dataset, in particular, the overrepresentation of flood events the Synthetic Minority Oversampling Technique (SMOTE) was applied to the training set to equalize the class distributions and thus improve the sensitivity of the model to underrepresented classes such as wildfires and earthquakes.

A detailed pre-processing process was carried out to prepare the data. The first step focused on the selection of three specific disaster types for a more targeted and meaningful analysis: wildfire, flood and earthquake. This filtering process can be represented mathematically as follows:

$$\mathcal{S} = r \in \mathcal{R}, |, r_{\text{category}} \in \text{Wildfire, Flood, Earthquake} \quad (1)$$

where R is the original dataset, r represents a single record, and r_{category} denotes the disaster type.

1) HANDLING MISSING VALUES

Missing values were systematically addressed using imputation techniques. For numerical columns, the missing values were replaced by their median:

$$\hat{v}j = \text{median}\left(vij \mid v_{ij} \neq \text{NaN}, , i = 1, \ldots, M\right) \quad (2)$$

where $\hat{v}j$ is the imputed value for feature j,
v_{ij} is the value of feature j for record i, and M is the total number of samples. For categorical features, missing values were replaced by their mode:

$$\hat{v}_k = \text{mode}\left(v_{ik} \mid v_{ik} \neq \text{NaN}, i = 1, \ldots, M\right) \quad (3)$$

2) ENCODING CATEGORICAL VARIABLES

Categorical variables were converted into numerical formats using Label Encoder. Let $K = \{\kappa_1, \kappa_2, \ldots, \kappa_L\}$ denote the set of unique categories for a feature. Each category κ_i is mapped to a unique integer via the encoding function:

$$\phi(\kappa_i) = i, \quad \forall \kappa_i \in \mathcal{K} \quad (4)$$

Thus, each categorical feature v_k is transformed into a numerical vector v'_k suitable for machine learning models.

3) NORMALIZING NUMERICAL FEATURES

StandardScaler was applied to normalize numerical features. For a numerical feature v_j, normalization was achieved as:

$$z_{ij} = \frac{v_{ij} - \bar{v}_j}{s_j} \quad (5)$$

where: v_{ij}: Original value for feature j in record i,

$\bar{v}_j = \frac{1}{M} \sum_{i=1}^{M} v_{ij}$: Mean of feature j,

$s_j = \sqrt{\frac{1}{M} \sum_{i=1}^{M} (v_{ij} - \bar{v}_j)^2}$: Standard deviation of feature j, z_{ij}:

Normalized value. This ensured all numerical features had zero mean and unit variance, enhancing model performance by mitigating scale-related biases.

4) BALANCING CLASS DISTRIBUTION

The class imbalance was addressed using SMOTE. For a minority class C_t, synthetic samples were generated by linear interpolation between a sample p and one of its k-nearest neighbors p_{nn}:

$$\mathbf{p}_{\text{new}} = \mathbf{p} + \lambda \cdot (\mathbf{p}_{\text{nn}} - \mathbf{p}), \quad \lambda \sim \mathcal{U}(0, 1) \quad (6)$$

After applying SMOTE, the size of all classes was balanced to match the size of the majority class:

$$|\mathcal{C}_t| = \max\left(|\mathcal{C}_1|, |\mathcal{C}_2|, \ldots, |\mathcal{C}_N|\right) \quad (7)$$

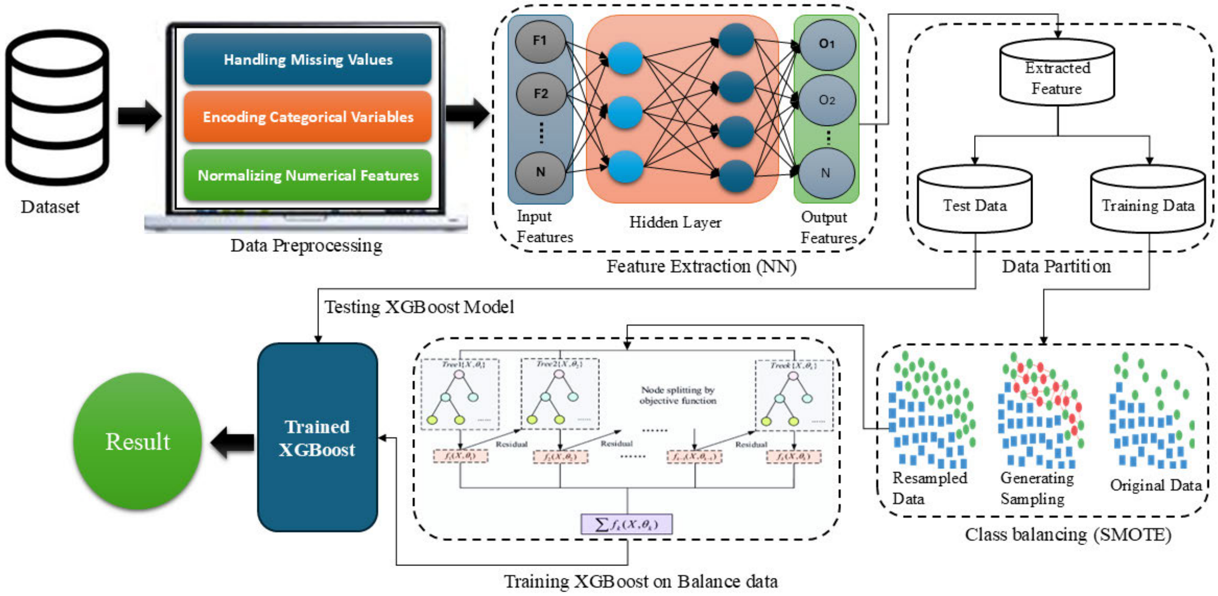

FIGURE 1. The working of the proposed model.

TABLE 1. Dataset description.

Feature Category	Data Type	Description
Disaster Type	Categorical	The primary classification of the disaster (e.g., Flood, Drought).
Country, Region, Subregion	Categorical	Geographic details about where the disaster occurred.
Dates	Temporal	Start and end dates of the disaster, broken into year, month, and day.
Impact Metrics	Numerical	Total deaths, number of injured, affected population, and economic damages.
Magnitude	Numerical	A numerical scale of the disaster's severity, accompanied by its unit.
Latitude/Longitude	Numerical	Geographic coordinates of the disaster (mostly missing).
Financial Contributions and Costs	Numerical	Aid and Reconstruction Costs, both raw and adjusted for inflation.

where:

$|C_t|$: Number of samples in class t,

N: Total number of classes.

To solve the problem of class imbalance in the disaster dataset, the synthetic minority oversampling technique (SMOTE) was used. SMOTE generates synthetic samples for the minority classes and thus ensures a balanced class distribution. Before applying SMOTE, the dataset had a significant imbalance, with floods being the dominant class and wildfires and earthquakes being significantly underrepresented. By using SMOTE, the class distributions were balanced so that the machine learning models work more reliably in all disaster categories. SMOTE was only applied to training data so that the model was trained on balanced data. The results of this balancing process are shown in Figure 2.

C. BIAS CONSIDERATIONS

Despite the use of the globally recognized EM-DAT dataset, distortions may occur, e.g. due to uneven geographical distribution, different reporting of events and varying completeness of characteristics. For example, underreporting of low-impact disasters in developing regions can lead to regional biases, while missing data on economic losses or magnitude characteristics can distort learning. To mitigate these issues, we used SMOTE to balance the classes and robust imputation techniques to reduce the impact of missing data. We also restricted our feature set to variables that are sufficiently represented in all disaster types to minimize bias due to the simplicity of the model.

D. HYBRID MODEL: NEURAL-XGBoost (N-XGB)

The hybrid model integrates neural networks for feature extraction and XGBoost for classification and utilizes the strengths of both approaches. The functionality of the Neural-XGBoost model (N-XGB) is shown in the algorithm 1.

1) FEATURE EXTRACTION USING NEURAL NETWORKS

A Neural Network extracts high-level features from the preprocessed dataset:

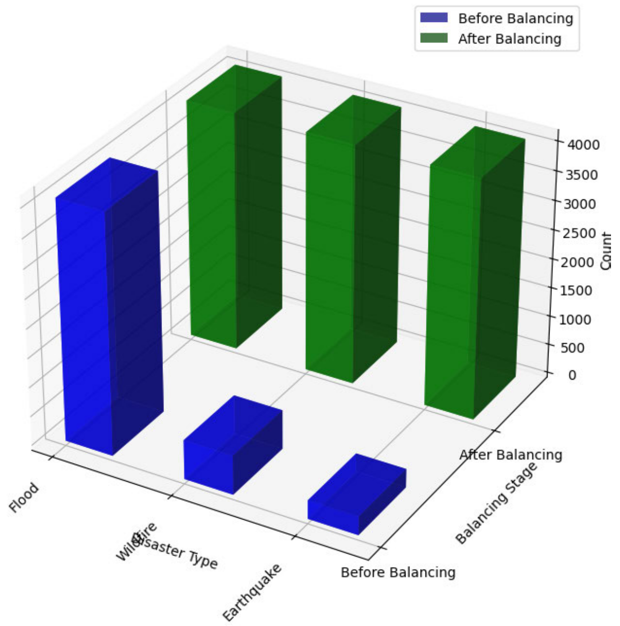

FIGURE 2. Class distribution before and after balancing using SMOTE.

Algorithm 1 Neural-XGBoost Model (N-XGB) for Disaster Prediction

---

Input: Disaster dataset $\mathcal{D} \in \mathbb{R}^{M\times N}$ with missing values and class imbalance.

Output: Predicted disaster categories $\mathbf{P}_{\text{pred}}$.

Step 1: Preprocessing:

1: Handle missing values:

2: For numerical features, impute using the median: $\hat{v}_j = \text{median}(\{v_{ij} \mid v_{ij} \neq \text{NaN}\})$.

3: For categorical features, impute using the mode: $\hat{v}_k = \text{mode}(\{v_{ik} \mid v_{ik} \neq \text{NaN}\})$.

4: Normalize features: $z_{ij} = \frac{v_{ij}-\mu_j}{\sigma_j}$, where μ_j and σ_j are the mean and standard deviation.

Step 2: Feature Extraction:

5: Extract embeddings $\mathbf{H}$ using a neural network f_{NN}, where layers h_1, h_2, h_3 process input features.

6: Balance classes using SMOTE:

$$\mathbf{p}_{\text{new}} = \mathbf{p} + \lambda(\mathbf{p}_{\text{nn}} - \mathbf{p}), \quad \lambda \sim \mathcal{U}(0, 1).$$

Step 3: Classification:

7: Predict disaster categories: $\mathbf{P}_{\text{pred}} = f_{\text{XGB}}(\mathbf{H})$, using the XGBoost classifier.

Step 4: Output:

8: Return predicted categories $\mathbf{P}_{\text{pred}}$.

9: **return** $\mathbf{P}_{\text{pred}}$

- **Input:** $X \in \mathbb{R}^{M\times N}$, where M is the number of samples and N is the number of features.
- **Architecture:**

$$h_1(X) \rightarrow h_2(h_1) \rightarrow h_3(h_2)$$

Fully connected layers with ReLU activation. Dropout layers for regularization.

- **Output:** $H \in \mathbb{R}^{M\times P}$, where $P < N$.

2) CLASSIFICATION USING XGBOOST

The extracted features H are used as input to the XGBoost classifier, which combines multiple decision trees through gradient boosting:

$$\min_{w} \sum_{i=1}^{M} \ell(y_i, f(h_i; w)) + \Omega(f) \tag{8}$$

where:

- y_i: True labels,
- h_i: Extracted features for sample i,
- ℓ: Loss function (e.g., log loss),
- $\Omega(f)$: Regularization term.

E. DATA EXPLORATION AND FEATURE ANALYSIS

A correlation matrix was created to examine the key correlation relationships between the numerical features in the dataset, as shown in Figure 3. This visualization helps identify highly correlated variables that can affect the performance of machine learning models. The diagonal elements of the matrix represent self-correlation ($r = 1$), while the off-diagonal elements indicate the degree of correlation between different features. Understanding these relationships is crucial for feature selection and dimensionality reduction to ensure that multicollinearity does not affect the predictive accuracy of the models. The working of the hybrid prediction model is shown in Figure 1. The model integrates machine learning algorithms and geographical analysis to predict the occurrence of disasters, assess damage, and optimize response strategies.

IV. EXPERIMENTAL SETUP

The experimental setup comprises three important phases: Data preprocessing, model training and evaluation. This evaluates the performance of the proposed Neural-XGBoost disaster prediction model (N-XGB). These phases are described below:

A. MODEL TRAINING

The Neural-XGBoost disaster prediction model (N-XGB) integrates feature extraction via neural networks (NN) with XGBoost classification. The pre-processed dataset is fed through the neural network to generate high-level feature representations. The architecture includes fully connected layers with ReLU activation functions to capture nonlinear transformations and dropout layers for regularization to prevent overfitting. These extracted features, referred to as H, form a compressed and meaningful representation of the input data:

$$H = f_{\text{NN}}(D') \tag{9}$$

where D' is the preprocessed dataset.

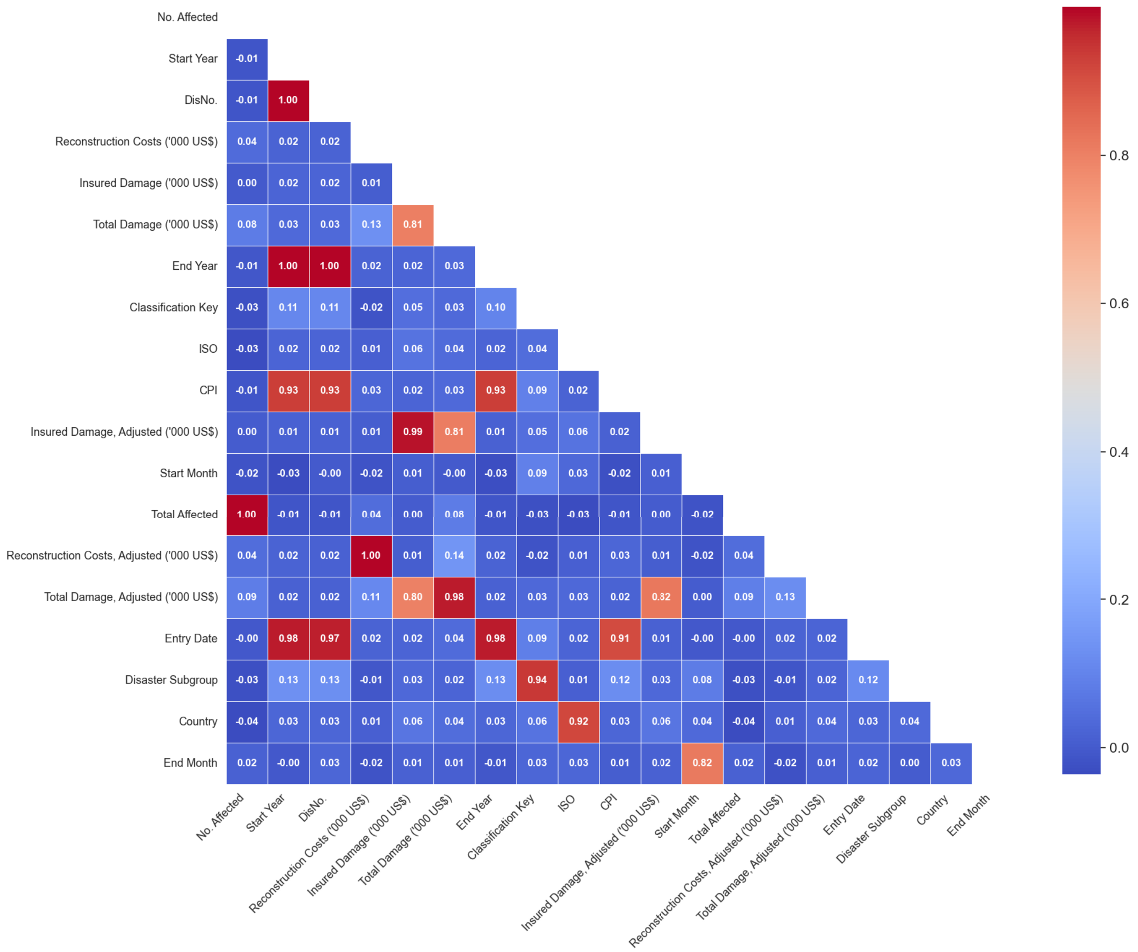

FIGURE 3. Correlation matrix heatmap with the most strongly correlated features in the disaster dataset.

The extracted features are then fed into the XGBoost classifier, which optimizes an ensemble of decision trees by gradient boosting. After feature extraction, SMOTE is performed to balance the classes in the training data so that the model is trained on balanced data. The model minimizes the following objective function:

$$\min_{w} \sum_{i=1}^{M} \ell(y_i, f(h_i; w)) + \Omega(f) \tag{10}$$

where:

- ℓ: Loss function (log-loss for multi-class classification),
- y_i: True labels,
- h_i: Feature vector for sample i,
- $\Omega(f)$: Regularization term to control overfitting.

To improve the performance of the model, the optimization of the hyperparameters is performed using a grid search, where combinations of learning rate, tree depth and number of estimators are systematically investigated. Table 2 summarizes the hyperparameter space investigated for the XGBoost optimization.

TABLE 2. Hyperparameter for XGBoost optimization.

Hyperparameter	Values Tested
Learning Rate (α)	0.01, 0.05, 0.1
Maximum Tree Depth (d)	5, 7, 9
Number of Estimators (n)	100, 200, 300

This setup ensures that the model achieves optimal accuracy, efficiency and robustness.

B. EVALUATION SETUP

The performance of the model is evaluated using various metrics using the holdout validation scheme (70:30) to assess the performance of the model and ensure its accuracy and robustness.The dataset is split into training and test subsets for learning and validation. Key metrics include accuracy (correct classification rate), F1 score (balancing precision and recall for class imbalances) and confusion matrix (detailed error breakdown by disaster type). The Receiver Operating Characteristic (ROC), Area Under the Curve (AUC), evaluates the model's ability to distinguish between disaster types and shows the trade-off between true positive and false positive rates.

To validate the efficiency of the Neural-XGBoost disaster prediction model (N-XGB), its performance is compared with the baseline classifiers, including random forest, support vector machines (SVM), logistic regression and k-nearest neighbors. Table 3 contains the formulas for these evaluation metrics and illustrates the mathematical basis underlying the model evaluation framework.

1) SYSTEM SPECIFICATIONS

The Neural-XGBoost disaster prediction model (N-XGB) was implemented on a system equipped with an Intel Core i9 processor, 32 GB RAM and an NVIDIA GPU, enabling accelerated computations. The environment was built on Windows 11 (64-bit) and Python 3.9 and includes libraries such as NumPy, Pandas, scikit-learn for preprocessing, PyTorch for feature extraction, XGBoost for classification and Matplotlib/Seaborn for visualization. The neural network used a batch size of 64, a learning rate of 0.001 and ran for 50 epochs, with the XGBoost parameters tuned via GridSearchCV. This configuration optimized the hybrid pipeline and fully utilized the hardware for deep learning and classification.

TABLE 3. Evaluation metrics.

Metric	Formula
Accuracy	Accuracy $= \frac{TP+TN}{TP+TN+FP+FN}$
F1-Score	F1-Score $= 2 \cdot \frac{\text{Precision}\cdot\text{Recall}}{\text{Precision}+\text{Recall}}$
Confusion Matrix	Confusion Matrix = $\begin{bmatrix} TP & FP \\ FN & TN \end{bmatrix}$
ROC-AUC	True Positive Rate (TPR) vs. False Positive Rate (FPR) plotted as the area under the curve.

V. RESULTS AND DISCUSSION

This section examines the performance of the proposed model in comparison to baseline machine learning techniques. It is divided into three parts. The Neural-XGBoost disaster prediction model (N-XGB) is evaluated in the first experiment on a balanced dataset using SMOTE to remove class imbalance. The second experiment examines the performance of the model on an imbalanced dataset and emphasizes its robustness with different data distributions. Finally, in the third experiment, the performance of the N-XGB model is compared with other established machine learning algorithms to confirm its strength in disaster prediction.

A. DISASTER OCCURRENCE DISTRIBUTION

The disaster dataset shows remarkable differences in the frequency of disasters across different countries. Table 4 shows the distribution of disaster density and identifies the countries most affected by natural disasters. China and Indonesia report the most registered events, with 338 and 288, respectively. They are followed by India (212) and the United States (207).

In addition, Thailand is one of the most affected regions, with 74 recorded disasters, which underlines the country's vulnerability to frequent natural disasters. The data shows that significant disasters are concentrated in a small group of countries. This uneven distribution can be attributed to geographical, climatic and socio-economic factors that make certain regions more vulnerable to certain types of disasters. Table 4 shows that countries such as Brazil (117), Pakistan (107) and the Philippines (131) are also frequently hit by disasters, underlining the importance of developing targeted disaster prediction and management strategies for these high-risk areas.

TABLE 4. Disaster density across countries.

Country Name	Disaster Count	ISO Code
China	338	CHN
Indonesia	288	IDN
India	212	IND
United States of America	207	USA
Philippines	131	PHL
Brazil	117	BRA
Afghanistan	110	AFG
Pakistan	107	PAK
Russian Federation	81	RUS
Thailand	74	THA

B. GEOGRAPHICAL DISTRIBUTION OF DISASTERS

The geographical distribution of disasters shows significant differences between regions. Figure 4 shows the global density of disasters, including floods, wildfires, and earthquakes. Countries such as China, Indonesia and the United States have the highest density of disasters, which is represented by the darker shades on the map. These regions are particularly susceptible to recurring natural disasters due to their unique geographical and climatic conditions.

C. FIRST EXPERIMENT: PERFORMANCE OF THE MODEL ON THE BALANCED DATASET

The first experiment evaluates the performance of the proposed Neural-XGBoost disaster prediction model (N-XGB) on a balanced training dataset where class imbalance is handled using SMOTE. In this experiment, the prediction performance of the model is analyzed where the three disaster classes of wildfire, flood and earthquake are equally represented. Figure 5 shows the prediction results for each disaster type. The N-XGB model shows an excellent performance by correctly predicting 96% of wildfire cases (792 out of 826), 90% of flood cases (728 out of 808), and a near-perfect accuracy of 98% for earthquake cases (815 out of 828). The high diagonal values in the matrix represent the correctly classified samples. In contrast, the low off-diagonal values indicate very few misclassifications, especially for the highest accuracy earthquake predictions.

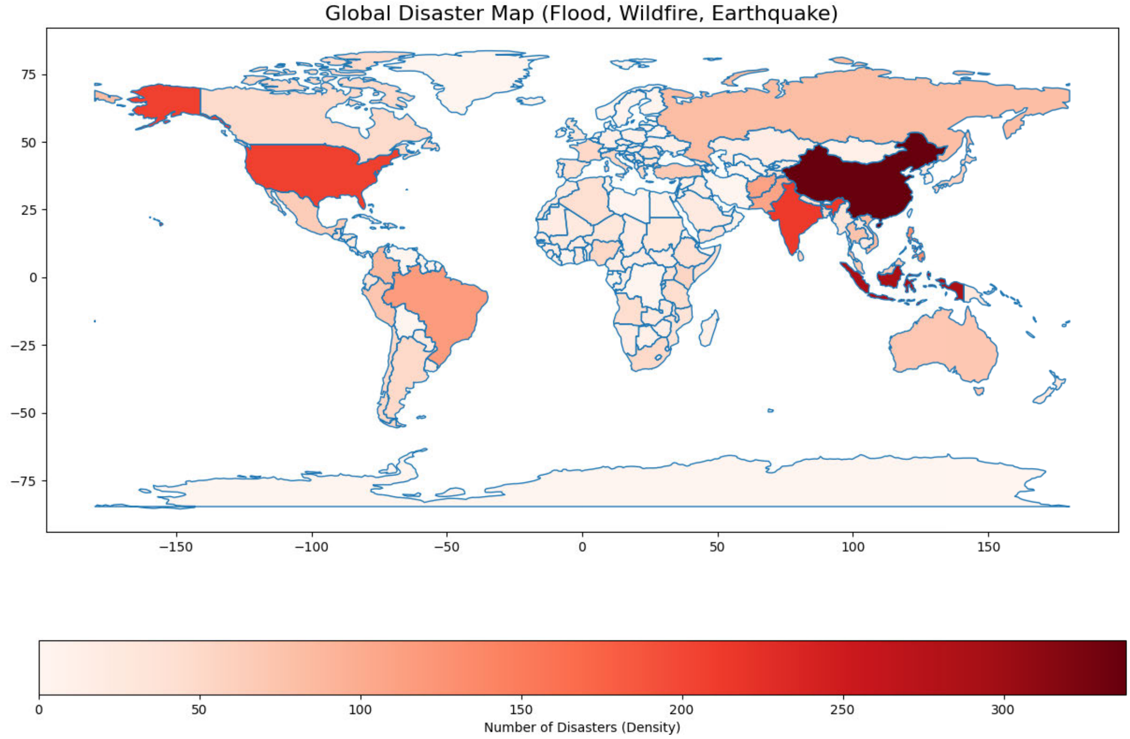

FIGURE 4. Map of global disaster density for the sources of floods, wildfires, and earthquakes.

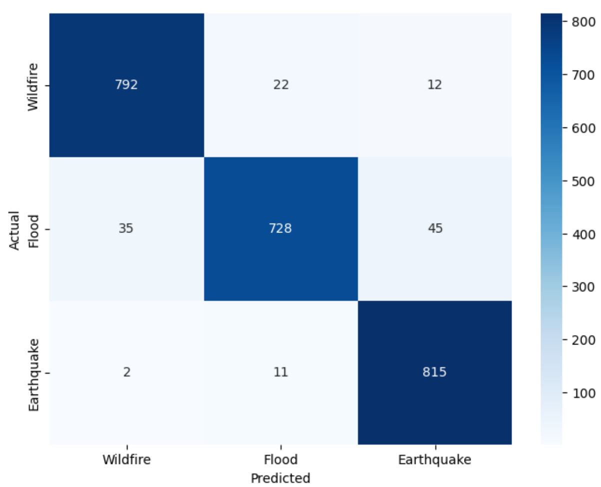

FIGURE 5. Confusion matrix for N-XGB on the balanced dataset.

Figure 6 compares the ROC-AUC curves of the N-XGB model with those of various baseline machine learning models, including Random Forest, Support Vector Machine (SVM), logistic regression and k-nearest Neighbors (k-NN), to further evaluate the classification performance. The area under the curve (AUC) values highlight the significant performance of the N-XGB model, which achieves an AUC of 0.99 for wildfires, 0.99 for floods and a perfect 1.00 for earthquakes. In contrast, the base models have lower AUC values, especially for floods and wildfires, indicating their relatively weaker performance. This comparison underscores the robust ability of the N-XGB model to accurately classify disaster events.

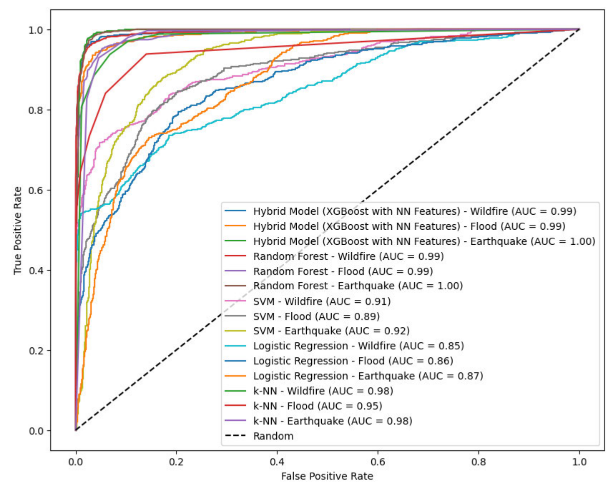

FIGURE 6. ROC curve comparison across models on the balanced dataset.

Figure 7 shows the performance of the N-XGB model in all disaster classes. The curve illustrates the model's exceptional ability to distinguish between wildfires, floods and earthquakes, with AUC values of 0.99 for wildfires, 0.99 for floods and a perfect value of 1.00 for earthquakes. This achieves near-perfect classification performance, which further strengthens confidence in the model and confirms its ability to classify balanced datasets well.

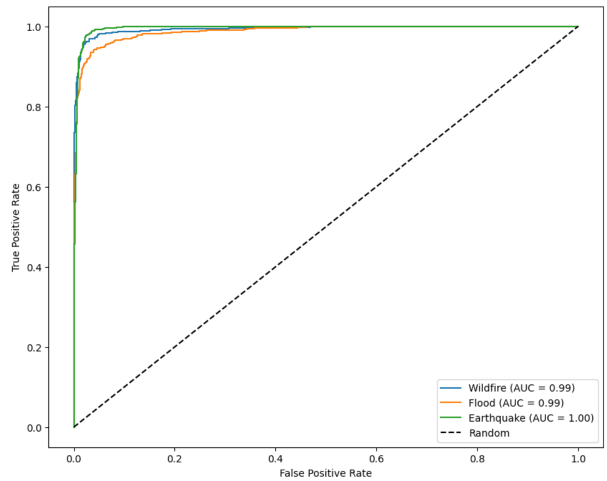

FIGURE 7. ROC curve for N-XGB model on the balanced dataset.

The results of the analysis show that the N-XGB model successfully predict a disaster. The comparisons show that N-XGB has high accuracy in recognizing different types of disasters compared to standard machine learning methods in ROC-AUC analysis. The model performs well on all classification tasks, including the prediction of earthquakes and wildfires, as well as on the flood dataset.

TABLE 5. Classification resutls.

Class	Precision	Recall	F1-Score	Support
Wildfire	0.96	0.96	0.96	826
Flood	0.96	0.90	0.93	808
Earthquake	0.93	0.98	0.96	828
Accuracy		0.95 (2462 total samples)		
Macro Avg	0.95	0.95	0.95	2462
Weighted Avg	0.95	0.95	0.95	2462

D. SECOND EXPERIMENT: PERFORMANCE OF THE MODEL ON THE IMBALANCED DATASET

The Neural-XGBoost disaster prediction model (N-XGB) was evaluated with the imbalanced dataset without applying class balancing techniques. This experiment evaluated the performance of the model in the presence of significant class imbalance and compared it to the results obtained with the balanced dataset. Figure 8 illustrates the prediction results for each disaster type on the imbalanced dataset. The model achieves relatively high accuracy for the majority class Flood but shows a significant drop in performance for the minority classes (Wildfire and Earthquake). Specifically, the model correctly predicts 118 cases of wildfires, while for floods, 814 out of 831 cases are correctly classified, with minor misclassifications. However, the performance for the earthquake class drops significantly as only 15 instances are predicted correctly. This highlights the challenges of handling minority classes in an imbalanced dataset.

Figure 9 shows a detailed evaluation of the AUC values for the N-XGB model compared to the basic machine learning techniques, including random forest, support vector machines (SVM), logistic regression and k-nearest neighbors (k-NN). In the case of SVM, the AUC values decrease to 0.86 for wildfire, 0.84 for floods and 0.73 for earthquakes, highlighting the better performance of the N-XGB model under imbalanced conditions. SVM, the AUC values decrease to 0.86 for wildfire, 0.84 for floods and 0.73 for earthquakes, highlighting the better performance of the N-XGB model under imbalanced conditions.

TABLE 6. Comparison of the proposed model and baseline models.

Model	Accuracy	F1-Score
Proposed Model	0.9824	0.9500
Random Forest	0.9222	0.5426
SVM	0.9160	0.5247
Logistic Regression	0.9139	0.5175
k-NN	0.9170	0.5404

The experiment shows that sampling imbalanced data makes it difficult to predict disasters in general and especially rare types of disasters. The results confirm the use of SMOTE to improve the class balance and thus improve the performance of the model and obtain reliable results.

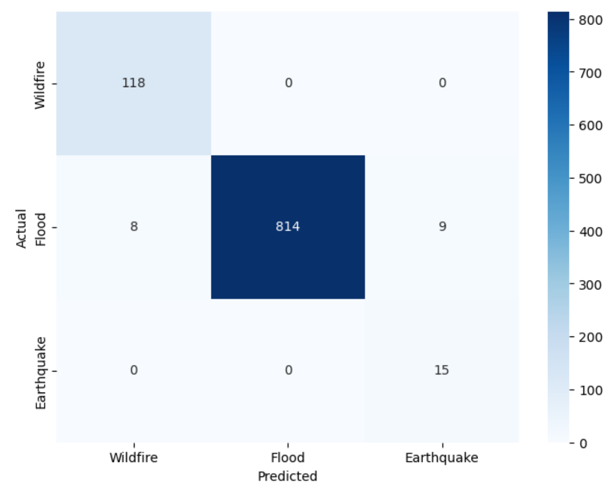

FIGURE 8. Confusion Matrix for N-XGB on imbalanced dataset.

E. THIRD EXPERIMENT: COMPARISON OF N-XGB WITH OTHER MACHINE LEARNING MODELS

To evaluate the performance of the Neural-XGBoost disaster prediction model (N-XGB), it was compared with some other basic machine learning techniques such as Random Forest, Support Vector Machines (SVM), logistic regression and k- Nearest Neighbors (kNN). In the evaluation, the standard performance metrics, i.e. accuracy, F1-score and ROC-AUC, were used for a comprehensive and robust comparison.

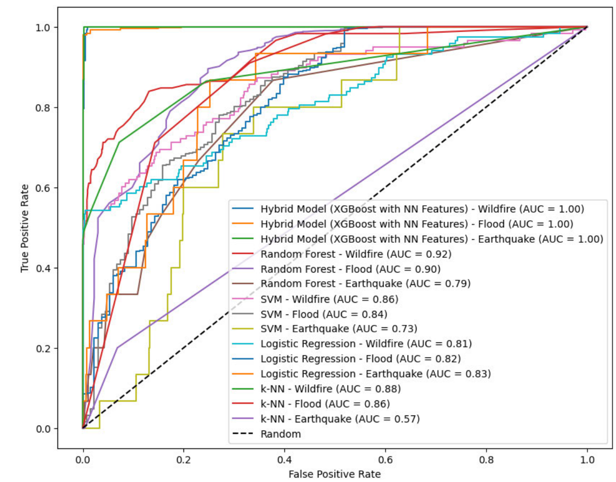

FIGURE 9. ROC-AUC curve comparison on the imbalanced dataset.

F. SENSITIVITY ANALYSIS OF HYPERPARAMETERS

The performance of the N-XGB model is significantly influenced by three important hyperparameters: learning rate (α), maximum tree depth (d) and the number of estimators (n). To systematically evaluate their influence, a sensitivity analysis was performed in which the classification accuracy and the F1 score of the model were evaluated for the test values listed in Table 2.

The results show that a learning rate (α) of 0.05 to 0.1 provides the best performance and ensures a balance between convergence speed and stability, while a lower value (0.01) slows down learning. For the maximum tree depth (d), increasing the depth improves classification accuracy up to d = 7, after which the performance stabilizes, indicating diminishing returns and potential overfitting at d = 9. The optimal number of estimators (n) is between 200 and 300, as increasing beyond 200 brings only marginal accuracy improvements while increasing computational costs, as shown in Figure 10.

These results emphasize the importance of the strategic selection of hyperparameters for improving the robustness, accuracy and efficiency of disaster classification models.

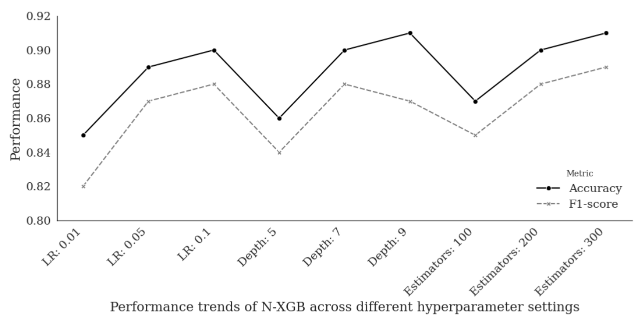

FIGURE 10. Sensitivity analysis of the proposed Neural-XGBoost (N-XGB) model for three important hyperparameters: Learning rate (α), maximum tree depth (d) and number of estimators (n).

VI. DISCUSSION

The results show that the proposed Neural-XGBoost (N-XGB) model significantly improves the accuracy of multi-class disaster classification compared to conventional models. It achieved an overall accuracy of 94.8%, with F1-scores of 0.96 for wildfires, 0.93 for floods and 0.96 for earthquakes outperforming models such as Random Forest, SVM and Logistic Regression. This performance is due to the deep feature extraction via neural networks that capture complex disaster patterns, combined with XGBoost's robustness in handling imbalanced data and missing values. SMOTE further improves classification by enhancing the representation of minority disaster types. These results confirm the effectiveness of hybrid models and are consistent with previous research while offering notable improvements.

Previous studies using models such as Random Forest, SVM and other ensemble techniques have reported accuracies between 80–89%. Deep learning models such as CNNs and LSTMs have achieved better results but often face challenges in terms of computational cost and data scarcity. The N-XGB model not only outperforms previous benchmarks with an accuracy of 94.8%, but also effectively addresses the imbalance between classes with SMOTE without sacrificing efficiency, making it well-suited for real-world applications.

In terms of computational complexity, N-XGB maintains a practical balance between performance and resource consumption. It was implemented in Python 3.9 on a Windows 11 system with an Intel Core i9 processor, 32 GB RAM, and an NVIDIA GPU. The neural network used a batch size of 64, a learning rate of 0.001 and 50 training epochs, while the XGBoost parameters were tuned using GridSearchCV. Although deep learning introduces a moderate overhead, the flat architecture and optimized batch processing keep the training process efficient. XGBoost incurs minimal computational costs during classification. The entire training cycle was completed within a reasonable time, with peak GPU memory consumption below 3 GB. This confirms the suitability of the model for use in high-performance disaster prediction systems that require both accuracy and computing power.

N-XGB offers automatic learning of key features, effective handling of imbalanced datasets through XGBoost and fairer classification through SMOTE. However, there are some limitations. The model depends on labeled data, which can limit its use in data-scarce regions. Also, it is optimized for tabular datasets and currently does not consider real-time sources such as satellite images or sensor feeds. Although training time increases with larger datasets, it remains more efficient than models based purely on deep learning.

N-XGB has great potential for practical use in disaster preparedness, emergency response, and financial risk assessment. It can support government agencies, first responders, and insurers with accurate, data-driven predictions. Future enhancements will include the integration of real-time data streams, such as satellite and IoT sensor inputs, to extend predictive capabilities. Validation with multiple datasets beyond EM-DAT is also planned to assess generalizability across different regions. The inclusion of explainable AI (XAI) methods will further improve interpretability for decision-makers. Despite pre-processing efforts, biases within the EM-DAT dataset, such as regional differences or differences in reporting, may affect generalizability, especially for underrepresented regions or disaster types.

Other limitations include the risk of overfitting due to the neural network component. Although dropout and batch normalization were used to reduce this, overfitting is still possible, especially for minority classes with limited samples. Since the model was trained and tested on a single dataset, its generalization across different standards and regions has yet to be evaluated. In addition, the model's reliance on high-quality labeled data may affect its performance in real-world conditions, where data is often incomplete or lagging. The inclusion of multimodal and partially labeled inputs in future work will help to address this issue. These limitations are part of the future development plan, which aims to improve the robustness, fairness, and scalability of the N-XGB framework for real-world disaster prediction.

VII. CONCLUSION

This study presented the Neural-XGBoost (N-XGB) disaster prediction model, a hybrid system that combines the deep feature extraction capabilities of neural networks with the accurate classification performance of XGBoost. The proposed model achieved impressive prediction performance by correctly classifying 94.8% of cases and achieving an average F1-score of 0.95. It consistently outperformed traditional machine learning algorithms such as random forest, support vector machine, logistic regression and k-nearest neighbors and achieved perfect AUC values of 1.00 on different disaster types. These results demonstrate the robustness and effectiveness of the hybrid approach on complex classification tasks. However, the analysis also showed that an imbalance between the classes has a negative impact on performance. This underlines the importance of appropriate pre-processing of the data and a balanced distribution. The sensitivity analysis also showed the critical impact of key hyperparameters, including learning rate, tree depth and number of estimators, on the model results. Future research will investigate automated tuning methods, such as Bayesian optimization, to streamline and improve model calibration. To ensure broader applicability and scalability, we plan to validate the N-XGB model on various datasets, including the NOAA Storm Events Database, the NASA Earth Observatory Disaster Archive, and the Kaggle Disaster Response Dataset, which covers different types of disasters and geographic regions. Overall, the N-XGB model offers a promising, interpretable and efficient solution for data-driven disaster prediction and management.

REFERENCES

[1] M. A. Saleem, W. Benjapolakul, W. Srisiri, S. Chaitusaney, and P. Kaewplung, "A hybrid prediction model integrating artificial intelligence and geospatial analysis for disaster management," *IEEE Access*, vol. 13, pp. 43716–43727, 2025.

[2] J. T. Thirukrishna, "An enhanced discovery of multiple natural disasters using machine learning model," *Earth Sci. Informat.*, vol. 18, no. 3, p. 324, Sep. 2025.

[3] Universit Catholique de Louvain. (2024). *Cred: Center of Research on the Epidemiology of Disasters*. Accessed: Oct. 15, 2024. [Online]. Available: https://uclouvain.be/en/research-institutes/irss/cred-center-of-research-on-the-epidemiology-of-disasters-0.html

[4] C. I. Donatti, K. Nicholas, G. Fedele, D. Delforge, N. Speybroeck, P. Moraga, J. Blatter, R. Below, and A. Zvoleff, "Global hotspots of climate-related disasters," *Int. J. Disaster Risk Reduction*, vol. 108, Jun. 2024, Art. no. 104488.

[5] E. Tung, A. Mostafavi, M. Li, S. Li, Z. Rasheed, and K. Shafique, "A system for community lifeline and resource optimization in disaster management and planning," *Int. J. Disaster Risk Reduction*, vol. 119, Mar. 2025, Art. no. 105269.

[6] X. Sun, X. Qin, Z. Liu, and Y. Yin, "Damaging effect of fine grinding treatment on the microstructure of polyurea elastomer modifier used in asphalt binder," *Measurement*, vol. 242, Jan. 2025, Art. no. 115984.

[7] G. Airlangga, "Machine learning for tsunami prediction: A comparative analysis of ensemble and deep learning models," *Kesatria : Jurnal Penerapan Sistem Informasi (Komputer dan Manajemen)*, vol. 6, no. 1, pp. 302–311, Jan. 2025.

[8] T. Lan, Y. Hu, L. Cheng, L. Chen, X. Guan, Y. Yang, Y. Guo, and J. Pan, "Floods and diarrheal morbidity: Evidence on the relationship, effect modifiers, and attributable risk from Sichuan Province, China," *J. Global Health*, vol. 12, p. 11007, Jul. 2022.

[9] S. Palaniappan, R. Logeswaran, S. Khanam, and Y. Zhang, "Machine learning model for predicting net environmental effects," *J. Informat. Web Eng.*, vol. 4, no. 1, pp. 243–253, Feb. 2025.

[10] R. L. Jones, A. Kharb, and S. Tubeuf, "The untold story of missing data in disaster research: A systematic review of the empirical literature utilising the Emergency Events Database (EM-DAT)," *Environ. Res. Lett.*, vol. 18, no. 10, Oct. 2023, Art. no. 103006.

[11] S. Sharma and A. Gosain, "Addressing class imbalance in remote sensing using deep learning approaches: A systematic literature review," *Evol. Intell.*, vol. 18, no. 1, pp. 1–28, Feb. 2025.

[12] N. V. Chawla, A. Lazarević, L. Hall, and K. W. Bowyer, "SMOTEBoost: Improving prediction of the minority class in boosting," in *Proc. Eur. Conf. Princ. Data Mining Knowl. Discovery*, Cavtat-Dubrovnik, Croatia, Jan. 2003, pp. 107–119.

[13] G. Liu, Z. Zhong, T. Ye, J. Meng, S. Zhao, J. Liu, and S. Luo, "Impact failure and disaster processes associated with rockfalls based on three-dimensional discontinuous deformation analysis," *Earth Surf. Processes Landforms*, vol. 49, no. 11, pp. 3344–3366, Sep. 2024.

[14] Y. Tian, Y. Zhao, S.-W. Son, J.-J. Luo, S.-G. Oh, and Y. Wang, "A deep-learning ensemble method to detect atmospheric rivers and its application to projected changes in precipitation regime," *J. Geophys. Res., Atmos.*, vol. 128, no. 12, Jun. 2023, Art. no. e2022JD037041.

[15] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, Aug. 2016, pp. 785–794.

[16] C. Delgrange, O. Demler, S. Mora, B. Menze, E. de la Rosa, and N. Davoudi, "A self-supervised model for multi-modal stroke risk prediction," 2024, *arXiv:2411.09822*.

[17] G. Husain, D. Nasef, R. Jose, J. Mayer, M. Bekbolatova, T. Devine, and M. Toma, "SMOTE vs. SMOTEENN: A study on the performance of resampling algorithms for addressing class imbalance in regression models," *Algorithms*, vol. 18, no. 1, p. 37, Jan. 2025.

[18] P. Ghorpade, A. Gadge, A. Lende, H. Chordiya, G. Gosavi, A. Mishra, B. Hooli, Y. S. Ingle, and N. Shaikh, "Flood forecasting using machine learning: A review," in *Proc. 8th Int. Conf. Smart Comput. Commun. (ICSCC)*, Jul. 2021, pp. 32–36.

[19] A. Mosavi, P. Ozturk, and K.-W. Chau, "Flood prediction using machine learning models: Literature review," *Water*, vol. 10, no. 11, p. 1536, Oct. 2018.

[20] M. W. Ahmad, J. Reynolds, and Y. Rezgui, "Predictive modelling for solar thermal energy systems: A comparison of support vector regression, random forest, extra trees and regression trees," *J. Cleaner Prod.*, vol. 203, pp. 810–821, Dec. 2018.

[21] Y. Zhang and P. J. Thorburn, "Handling missing data in near real-time environmental monitoring: A system and a review of selected methods," *Future Gener. Comput. Syst.*, vol. 128, pp. 63–72, Mar. 2022.

[22] J. Jia and W. Ye, "Deep learning for earthquake disaster assessment: Objects, data, models, stages, challenges, and opportunities," *Remote Sens.*, vol. 15, no. 16, p. 4098, Aug. 2023.

[23] H. Xu, Y. Zhao, Z. Dajun, Y. Duan, and X. Xu, "Exploring the typhoon intensity forecasting through integrating AI weather forecasting with regional numerical weather model," *npj Climate Atmos. Sci.*, vol. 8, no. 1, p. 38, Feb. 2025.

[24] R. Cantini, C. Cosentino, F. Marozzo, D. Talia, and P. Trunfio, "Harnessing prompt-based large language models for disaster monitoring and automated reporting from social media feedback," *Online Social Netw. Media*, vol. 45, Jan. 2025, Art. no. 100295.

[25] W.-Y. Wang, W.-W. Du, D. Xu, W. Wang, and W.-C. Peng, "A survey on self-supervised learning for non-sequential tabular data," *Mach. Learn.*, vol. 114, no. 1, p. 16, Jan. 2025.

[26] E. Karampotsis, K. Kioskli, A. Tsirimpa, G. Dounias, and A. Polydoropoulou, "Understanding evacuation behavior for effective disaster preparedness: A hybrid machine learning approach," *Natural Hazards*, vol. 120, no. 14, pp. 13627–13665, Nov. 2024.

[27] M.-S. Stupariu, S. A. Cushman, A.-I. Pleşoianu, I. Pătru-Stupariu, and C. Fuerst, "Machine learning in landscape ecological analysis: A review of recent approaches," *Landscape Ecol.*, vol. 37, no. 5, pp. 1227–1250, May 2022.

[28] H. Qi, Z. Zhou, P. Manu, and N. Li, "Falling risk analysis at workplaces through an accident data-driven approach based upon hybrid artificial intelligence (AI) techniques," *Saf. Sci.*, vol. 185, May 2025, Art. no. 106814.

[29] Y. Wang, A. W. Z. Chew, and L. Zhang, "Building damage detection from satellite images after natural disasters on extremely imbalanced datasets," *Autom. Construct.*, vol. 140, Aug. 2022, Art. no. 104328.

[30] M. A. Saleem, A. Javeed, W. Akarathanawat, A. Chutinet, N. C. Suwanwela, P. Kaewplung, S. Chaitusaney, S. Deelertpaiboon, W. Srisiri, and W. Benjapolakul, "An intelligent learning system based on electronic health records for unbiased stroke prediction," *Sci. Rep.*, vol. 14, no. 1, p. 23052, Oct. 2024.

[31] R. L. Jones, A. Kharb, and S. Tubeuf, "The untold story of missing data in disaster research: A systematic review of the empirical literature utilising the emergency events database (EM-DAT)," *Environ. Res. Lett.*, vol. 18, no. 10, 2023, Art. no. 103006.

[32] J. M. Jerez, I. Molina, P. J. García-Laencina, E. Alba, N. Ribelles, M. Martín, and L. Franco, "Missing data imputation using statistical and machine learning methods in a real breast cancer problem," *Artif. Intell. Med.*, vol. 50, no. 2, pp. 105–115, Oct. 2010.

[33] EM-DAT: The International Disaster Database. (2024). *EM-DAT Data Repository*. Accessed: Jun. 17, 2024. [Online]. Available: https://public.emdat.be/data

MUHAMMAD ASIM SALEEM received the bachelor's degree and the M.S. degree in information technology from Government College University Faisalabad, Pakistan, in 2017, and the Ph.D. degree in software engineering from the School of Information and Software Engineering, University of Electronic Science and Technology, China. He is currently a Postdoctoral Researcher with the Faculty of Engineering, Chulalongkorn University, Thailand. He has served as a reviewer in many journals, including IEEE. His research interests include code optimization, machine learning, deep learning, health informatics, and biomedical science.

ASHIR JAVEED received the B.S. degree in software engineering and the M.S. degree in computer science, and the Ph.D. degree in software engineering from the University of Electronic Science and Technology, Chengdu, China. He has researched several areas, including artificial intelligence, data mining, and applications in the medical field to develop diagnostic systems for real-world problems.

WATIT BENJAPOLAKUL received the D.Eng. degree in electrical engineering from The University of Tokyo, in 1989. He is currently a Professor with the Department of Electrical Engineering, Chulalongkorn University, Thailand. He is the Head of the Center of Excellence in Artificial Intelligence, Machine Learning, and Smart Grid Technology, Faculty of Engineering, Chulalongkorn University. His current research interests include mobile communication systems, broadband networks, applications of artificial intelligence in communication systems, wireless networks, and applications of communication networks in smart grids.

WATTANASAK SRISIRI received the B.Eng. and M.Eng. degrees in electrical engineering from the Khon Kaen University of Thailand. His current research interests include artificial intelligence and the Internet of Things for medicine, smart grid, and renewable energy.

SURACHAI CHAITUSANEY (Member, IEEE) received the Ph.D. degree from The University of Tokyo, Japan, in 2007, with a JICA Scholarship. He is currently an Assistant Professor and the Head of Power System Research Laboratory (PSRL), Chulalongkorn University. His research interests include renewable energy, solar PV, distributed generation, power system planning and reliability, and smart grid regulation. In addition, he has served as a member of subcommittee, committee, and working group in several professional organizations, such as the Engineering Institute of Thailand (EIT), the Energy Policy and Planning Office (EPPO), the Energy Regulatory Commission (ERC), and the Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI) Association.

PASU KAEWPLUNG received the M.Eng. degree in electrical engineering from Yokohama National University, Japan, in 1998, and the Ph.D.Eng. degree in electrical engineering from Chulalongkorn University, Thailand, in 2004. He is currently an Assistant Professor with the Department of Electrical Engineering, Chulalongkorn University. His research interests include mobile communications, signal processing, applications to optical devices, long-haul fiber transmission, fiber access networks, and application of artificial intelligence.