KESATRIA: Jurnal Penerapan Sistem Informasi (Komputer & Manajemen)
Terakreditasi Nomor 204/E/KPT/2022 | Vol. 6, No. 1, Januari (2025), pp. 302-311

Machine Learning for Tsunami Prediction: A Comparative Analysis of Ensemble and Deep Learning Models

Gregorius Airlangga
Department of Information Systems, Universitas Katolik Indonesia Atma Jaya, Jakarta, Indonesia
E-mail: gregorius.airlangga@atmajaya.ac.id

Abstract

Tsunamis, triggered by seismic activities, pose significant threats to coastal regions, necessitating accurate prediction models to mitigate their impact. This study explores the application of machine learning models, including ensemble methods (Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost) and deep learning (Neural Networks), for tsunami prediction based on seismic data. The dataset spans seismic events from 1995 to 2023, characterized by features such as magnitude, depth, and geographic location. A 10-fold cross-validation approach was employed to evaluate model performance using precision, recall, F1-score, accuracy, and ROC-AUC metrics. The results highlight that Gradient Boosting achieved the best balance between precision and recall, with an F1-score of 0.6544 and the highest ROC-AUC of 0.8606, demonstrating its strong discriminatory power. Random Forest excelled in precision (0.6920) and F1-score (0.6287), making it suitable for reducing false positives. Ensemble boosting models, such as CatBoost and LightGBM, offered consistent performance with low variability across folds. In contrast, Neural Networks underperformed, achieving an F1-score of 0.5497 and an ROC-AUC of 0.7936, indicating the need for further optimization. Despite promising results, challenges in recall scores underscore the need for enhanced detection of tsunami-triggering events. The findings establish ensemble methods, particularly Gradient Boosting and Random Forest, as robust tools for tsunami prediction, providing a foundation for early warning systems. Future work will focus on improving recall and exploring hybrid modeling techniques to optimize predictive accuracy and reliability.

Keywords: *Tsunami prediction, Machine learning, Ensemble models, Neural networks, Seismic data analysis.*

1. Introduction

Natural disasters remain one of the most significant challenges to human safety, infrastructure, and economic stability [1]–[3]. Among these, tsunamis, characterized by sudden and devastating waves often triggered by undersea earthquakes, pose a particularly grave threat to coastal regions globally [4]–[6]. Over the past three decades, advancements in geoscience have significantly improved our understanding of tsunami genesis, yet predicting their occurrence with precision remains a pressing challenge [7]–[9]. The complexity of seismic data, coupled with the inherent uncertainty in geological phenomena, necessitates innovative approaches to improve prediction accuracy and timeliness [10]–[12]. Recent studies have explored various computational techniques for tsunami prediction, including statistical modeling, physical simulations, and, more recently, machine learning (ML)[13]–[15]. ML models have demonstrated considerable potential due to their ability to analyze vast datasets and uncover patterns that traditional methods might overlook [16]. For instance, gradient boosting and random forest classifiers have been utilized to predict tsunami likelihood based on seismic parameters such as magnitude, depth, and geographic coordinates [17]. However, despite their promise, existing models often suffer from limitations such as overfitting, lack of

ISSN: 2720-992X

generalizability across different geographical regions, and challenges in handling missing or imbalanced data [18].

Moreover, the integration of deep learning models, including Convolutional Neural Networks (CNNs) and Neural Networks (NNs), has opened new frontiers in predictive modeling [19]. Deep learning excels in extracting high-level features from complex datasets, making it a promising candidate for seismic and tsunami prediction [20]. Nevertheless, there remains a research gap in evaluating the comparative performance of classical ML algorithms and advanced deep learning models within the context of tsunami prediction [21]. Specifically, the effectiveness of hybrid approaches that combine classical ensemble methods with deep neural architecture has not been adequately investigated [22]. The urgency of this research lies in the devastating impact of tsunamis, as exemplified by recent catastrophic events, such as the 2011 Tōhoku tsunami in Japan and the 2018 Sulawesi tsunami in Indonesia [23]. These events underscored the need for accurate and timely prediction systems to enable effective early warning and mitigation strategies [24]. Current state-of-the-art systems, such as the DART (Deep-ocean Assessment and Reporting of Tsunamis) buoy network, provide critical data on oceanic disturbances but require complementary predictive models for rapid risk assessment [25]. Integrating ML and deep learning models with such systems could revolutionize tsunami warning systems by providing probabilistic predictions based on real-time seismic data [24].

This study addresses the research gaps by systematically comparing the performance of classical machine learning models, such as Random Forest, Gradient Boosting, and XGBoost, with deep learning models, including neural networks, for tsunami prediction [26], [27]. The research leverages a large dataset of seismic events from 1995 to 2023, incorporating features such as earthquake magnitude, latitude, longitude, depth, and tsunami occurrence [28]. Advanced preprocessing techniques, including scaling and cross-validation, ensure robust model evaluation. Additionally, the study explores the potential of hybrid ensemble approaches, such as majority voting, to improve prediction accuracy and reliability. The primary contributions of this research are threefold. First, it establishes a comprehensive benchmark for the performance of classical and deep learning models in tsunami prediction. Second, it demonstrates the efficacy of a hybrid ensemble method, integrating predictions from multiple models to achieve higher accuracy and robustness. Third, it provides actionable insights into the critical features influencing tsunami occurrence, thus aiding domain experts in refining predictive systems and early warning frameworks. The remainder of this article is structured as follows: The next section provides a detailed review of related work, highlighting recent advancements in ML and deep learning for disaster prediction. This is followed by the methodology section, which outlines the dataset, preprocessing steps, and model implementation. The results and discussion section presents a comparative analysis of model performance, emphasizing key findings and implications for tsunami prediction systems. Finally, the conclusion summarizes the contributions of this study and discusses potential directions for future research.

2. Research Methodology

Predicting natural disasters, particularly tsunamis, has been a focal point of research due to the devastating effects these phenomena can have on human lives and infrastructure. The application of machine learning (ML) and deep learning (DL) techniques in this domain has witnessed significant growth, with researchers striving to overcome the limitations of traditional statistical and simulation-based approaches. This section reviews recent advancements in ML and DL applied to disaster prediction, focusing on their evolution, methodologies, and identified gaps in the literature.

2.1. Machine Learning Approaches for Disaster Prediction

ML techniques have been widely used for disaster prediction due to their ability to analyze large datasets and identify complex patterns. Early studies employed traditional algorithms like Support Vector Machines (SVM) and Decision Trees to predict earthquakes and their associated risks, including tsunamis. For instance, [29] demonstrated the efficacy of Random Forest classifiers in predicting tsunami occurrences based on seismic parameters such as magnitude, depth, and location. Their study emphasized the importance of feature selection in improving model accuracy. Ensemble models, such as Gradient Boosting Machines (GBMs) and XGBoost, have since become popular due to their robust performance and ability to reduce overfitting. [30] employed GBMs to predict landslide susceptibility after seismic events, achieving higher accuracy than individual models. Similarly, [31] utilized XGBoost to assess tsunami risks in the Indian Ocean, demonstrating the model's capability to handle imbalanced datasets effectively through techniques like Synthetic Minority Oversampling Technique (SMOTE). Despite these advancements, ML models often face challenges such as limited generalizability across regions, sensitivity to noisy data, and reliance on manual feature engineering. These limitations underscore the need for more adaptive and automated methods, such as deep learning.

2.2. Deep Learning in Disaster Prediction

Deep learning, characterized by its ability to automatically extract high-level features from raw data, has revolutionized disaster prediction. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) are among the most widely used architectures in this field. [32] demonstrated the effectiveness of CNNs in analyzing satellite imagery to predict flood-prone areas, highlighting the model's superior feature extraction capabilities. Similarly, [33] utilized RNNs, particularly Long Short-Term Memory (LSTM) networks, to model temporal dependencies in seismic data, achieving high accuracy in earthquake prediction. Hybrid models combining CNNs and RNNs have shown promising results in disaster prediction. For example, [34] proposed a CNN-LSTM model to predict tsunami wave heights based on real-time seismic data, achieving better performance than standalone CNN or LSTM models. The study highlighted the hybrid model's ability to capture both spatial and temporal features, which are critical for accurate tsunami prediction. Another emerging area is the application of attention mechanisms in disaster prediction. Attention-based models, such as Transformers, have been successfully applied to various tasks, including earthquake risk assessment. [35] incorporated attention mechanisms into their LSTM model to focus on critical time steps in seismic data, resulting in enhanced prediction accuracy and interpretability.

2.3. Hybrid and Ensemble Methods

Combining traditional ML models with DL architecture has emerged as a promising approach to leverage the strengths of both paradigms. Ensemble methods, such as stacking and majority voting, have been used to combine predictions from multiple models, improving robustness and accuracy. [36] developed a hybrid ensemble model integrating Random Forest, XGBoost, and a CNN-based deep learning model to predict tsunami occurrences. Their results showed that the hybrid model outperformed individual models, particularly in handling imbalanced datasets. Moreover, advanced ensemble strategies, such as weighted voting and meta-learning, have been explored. [37] proposed a meta-learning framework that dynamically selects the best-performing model based on the input data characteristics, achieving state-of-the-art results in earthquake prediction.

2.4. Challenges and Research Gaps

Despite these advancements, several challenges remain in applying ML and DL to disaster prediction. First, the availability of high-quality, labeled datasets is a significant bottleneck. Many existing datasets are region-specific, limiting the generalizability of trained models. Second, handling missing or noisy data is a persistent issue, particularly for seismic datasets with incomplete records of events. Third, the interpretability of complex models, especially deep learning architectures, poses challenges for their adoption in real-world applications, where transparency and trust are crucial. Furthermore, while hybrid models have shown promise, their computational complexity can be a barrier to real-time prediction. There is a need for optimized algorithms that balance accuracy and computational efficiency, particularly for deployment in early warning systems.

Emerging trends in disaster prediction research include the integration of multi-source data, such as combining seismic data with satellite imagery or oceanic sensor readings. The use of transfer learning to adapt pre-trained models to new regions or disaster types is also gaining traction. Additionally, the application of explainable AI (XAI) techniques is expected to address the interpretability challenges of complex models. In conclusion, ML and DL have made significant strides in disaster prediction, offering tools to improve the accuracy and timeliness of early warning systems. However, addressing the challenges of data quality, model interpretability, and computational efficiency remains critical for future research. This study builds on these advancements by systematically comparing classical ML and DL models, with a focus on their applicability to tsunami prediction, thereby contributing to the growing body of knowledge in this domain.

2.5. Proposed Methodology

The methodology employed in this study integrates comprehensive data preprocessing techniques, advanced predictive modeling approaches, and rigorous evaluation strategies to address the problem of tsunami prediction. Each stage is detailed to ensure the clarity and reproducibility of the research process, emphasizing the mathematical rigor underlying the applied methods.

2.5.1. Dataset Description

The data set used in this study comprises seismic events recorded globally from 1995 to 2023 and can be downloaded from [38]. Each seismic event is represented as a feature vector $X = \{x_1, x_2, \ldots, x_n\}$, where $(x_i \in R^d)$ includes key attributes such as earthquake magnitude $((m))$, latitude $((\phi))$, longitude $((\lambda))$, depth $((d))$, Modified Mercalli Intensity (MMI), Community Internet Intensity (CDI), and a binary target variable (y), indicating whether the event triggered a tsunami $((y = 1))$ or not $((y = 0))$. The dataset can be expressed as (1).

$$\mathcal{D} = \{(X, y)\} \tag{1}$$

where $(X \in R^{n \times d})$ and $(y \in \{0,1\}^n)$. Missing values in features such as MMI and CDI were addressed during preprocessing, and rows with critical missing values, particularly in (m), (ϕ), (λ), (d), and (y), were removed to ensure data integrity.

2.5.2. Data Preprocessing

To ensure the data's suitability for predictive modeling, several preprocessing steps were applied. Missing values in continuous attributes were imputed using median imputation, defined as (2).

$$\hat{x}_j = \text{median}\big(\{x_j \mid x_j \neq \text{NA}\}\big) \tag{2}$$

for each feature (x_j) with incomplete observations. To ensure uniform feature scaling, each attribute was normalized to the range $([0,1])$ using the MinMaxScaler transformation as defined as (3).

$$x_{j,\text{scaled}} = \frac{x_j - \min(x_j)}{\max(x_j) - \min(x_j)} \quad (3)$$

Class imbalance was addressed using the Synthetic Minority Oversampling Technique (SMOTE), where synthetic samples ($\tilde{x}$) were generated as (4).

$$\tilde{x} = x_{\text{minority}} + \gamma \cdot (x_{\text{nearest neighbor}} - x_{\text{minority}}) \quad (4)$$

with ($\gamma \sim \text{Uniform}(0,1)$). The dataset was split into training and testing subsets using stratified (k)-fold cross-validation with ($k = 10$), preserving the class distribution in each fold. For a given fold, the dataset was divided into training ($\mathcal{D}_{train}$) and validation ($\mathcal{D}_{val}$) sets, where $\mathcal{D}_{val} = \mathcal{D}_i$ and $\mathcal{D}_{train} = \mathcal{D} \setminus \mathcal{D}_i$.

2.5.3. Model Implementation

The predictive modeling phase involved both classical machine learning algorithms and advanced deep learning architecture. Classical models included Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost. These models optimized the binary cross-entropy loss function, defined as $\mathcal{L}(y, \hat{y}) = -\frac{1}{n}\sum_{i=1}^{n}[y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)]$, where ($\hat{y}_i$) is the predicted probability for the (i)-th instance. Hyperparameters, such as the number of estimators ((T)), learning rate ((η)), and maximum tree depth, were tuned for optimal performance using grid search. In addition to classical models, a feedforward neural network was implemented with three hidden layers of sizes 256, 128, and 64 neurons, respectively. Each hidden layer applied the ReLU activation function, defined as (5).

$$h^{(l)} = \text{ReLU}\left(W^{(l)}h^{(l-1)} + b^{(l)}\right) \quad (5)$$

where ($h^{(l-1)}$) represents the activations from the previous layer, ($W^{(l)}$) denotes the weight matrix, and ($b^{(l)}$) is the bias vector. Dropout regularization with a rate of 0.3 was applied after each hidden layer to prevent overfitting. The output layer consisted of a single neuron with a sigmoid activation function $\sigma(z) = \frac{1}{1+e^{-z}}$, producing the probability ($\hat{y}$) of a tsunami-triggering event. The network was trained using the Adam optimizer with a learning rate of ($\alpha = 0.0005$) and binary cross-entropy loss for 50 epochs with a batch size of 32. An ensemble method combining predictions from all models was implemented using majority voting. For a given instance (x), the ensemble prediction ($\widehat{y_{\text{ensemble}}}$) was determined as (6).

$$\widehat{y_{\text{ensemble}}} = \text{mode}(\{\widehat{y_{\text{RF}}}, \widehat{y_{\text{GB}}}, \widehat{y_{\text{XGB}}}, \ldots\}) \quad (6)$$

where (mode) denotes the most frequent class among individual model predictions. For probabilistic evaluation, the ensemble averaged predicted probabilities as $P(y = 1 \mid x)_{\text{ensemble}} = \frac{1}{M}\sum_{m=1}^{M} P\left(y = 1 \mid x; f_\theta^{(m)}\right)$ where (M) is the number of models in the ensemble.

2.5.4. Evaluation Metrics

Model performance was evaluated using precision ((P)), recall ((R)), F1-score ((F_1)), accuracy ((A)), and the area under the receiver operating characteristic curve (AUC). These metrics were mathematically defined as (7) – (11).

$$P = \frac{\text{TP}}{\text{TP+FP}} \quad (7)$$

$$R = \frac{\text{TP}}{\text{TP+FN}} \quad (8)$$

$$F_1 = \frac{2PR}{P+R} \quad (9)$$

$$A = \frac{\text{TP+TN}}{\text{TP+TN+FP+FN}} \quad (10)$$

$$\text{AUC} = \int_0^1 \text{TPR}(\text{FPR})\, d(\text{FPR}) \quad (11)$$

where (TP), (FP), (TN), and (FN) denote true positives, false positives, true negatives, and false negatives, respectively. Results were averaged across all folds of cross-validation to ensure robustness and reliability. This methodology establishes a rigorous and comprehensive framework for evaluating tsunami prediction models, emphasizing the integration of classical and deep learning methods to advance predictive accuracy and reliability.

3. Results and Discussion

The results of the study demonstrate the performance of various machine learning models across 10-fold cross-validation, focusing on five key evaluation metrics: precision, recall, F1-score, accuracy, and ROC-AUC. The models evaluated include CatBoost, Gradient Boosting, LightGBM, Neural Network, Random Forest, and XGBoost. Additionally, the performance of a hybrid ensemble approach using majority voting is presented. The average performance metrics for the models are summarized in the table. Among the models, Random Forest achieved the highest precision (0.692) and demonstrated strong overall performance with an F1-score of 0.629, an accuracy of 0.779, and an ROC-AUC of 0.859. Similarly, Gradient Boosting exhibited balanced performance across metrics, with the highest recall (0.637) and a competitive F1-score (0.654).

CatBoost and LightGBM showed comparable performance, achieving F1-scores of 0.619 and 0.639, respectively, while maintaining high ROC-AUC values above 0.85. Neural Network lagged behind in all metrics, with the lowest F1-score (0.550) and accuracy (0.729), suggesting potential limitations in its ability to handle the dataset or the need for additional optimization. The standard deviation values indicate the stability of model performance across the 10 folds. Gradient Boosting demonstrated the most consistent performance, with low standard deviations across all metrics (e.g., accuracy: 0.044, ROC-AUC: 0.049). Conversely, CatBoost and Neural Network showed higher variability, especially in precision and recall, indicating less stable performance across the folds.

In addition, the hybrid ensemble approach, which combines the predictions of the individual models using majority voting, achieved an average precision of 0.661 (±0.088), recall of 0.646 (±0.107), and F1-score of 0.652 (±0.093). The ensemble also demonstrated competitive accuracy (0.777 ± 0.056) and ROC-AUC (0.795 ± 0.078). These results suggest that the ensemble effectively leverages the strengths of individual models to produce more robust predictions, although the improvement in metrics compared to the best-performing individual models is modest.

The results highlight the strengths and limitations of different machine learning models in the context of the problem. Random Forest and Gradient Boosting consistently outperformed other models in terms of precision, recall, and ROC-AUC, indicating their capability to capture complex relationships in the data. The strong performance of these models can be attributed to their ensemble nature and ability to handle diverse feature interactions. Neural Network, on the other hand, performed poorly compared to the ensemble methods. This may be due to insufficient tuning or the need for a more complex architecture to effectively model the data. It also suggests that deep learning methods may require more extensive feature engineering or larger datasets to match the performance of tree-based methods in this specific task. The hybrid ensemble approach provided a stable and competitive performance, demonstrating that combining multiple models can mitigate individual weaknesses and enhance overall robustness. However, the lack of substantial improvement over the best-performing individual models indicates that further exploration of ensemble techniques, such as weighted voting or stacking, may be necessary to fully exploit the potential of the hybrid approach. In terms of stability, Gradient Boosting showed the most consistent performance, which is a desirable attribute in practical applications where reliability is critical. High variability in CatBoost and Neural Network suggests these models are more sensitive to the specific training data splits, which might limit their applicability in real-world scenarios. Overall, the results emphasize the effectiveness of tree-based ensemble methods for this dataset, while also highlighting areas for improvement in neural network models and hybrid ensemble techniques. Future work could explore hyperparameter tuning, feature selection, and advanced ensemble methods to further enhance performance.

Table 1. Average Performance Across 10 Folds

Model	Precision	Recall	F1-Score	Accuracy	ROC-AUC
CatBoost	0.679585	0.572633	0.619368	0.774	0.855036
Gradient Boosting	0.675067	0.636932	0.654406	0.781	0.860639
LightGBM	0.661988	0.621496	0.639242	0.772	0.856496
Neural Network	0.60272	0.514489	0.549748	0.729	0.793574
Random Forest	0.691984	0.578693	0.628691	0.779	0.85888
XGBoost	0.65392	0.633996	0.642051	0.772	0.8555

Table 2. Standard Deviation Across 10 Folds

Model	Precision	Recall	F1-Score	Accuracy	ROC-AUC
CatBoost	0.111853	0.121187	0.116457	0.063105	0.060714
Gradient Boosting	0.076145	0.06412	0.064803	0.043576	0.049171
LightGBM	0.107307	0.104697	0.100581	0.066466	0.054485
Neural Network	0.111692	0.117946	0.097572	0.055668	0.062734
Random Forest	0.110516	0.109541	0.106873	0.062619	0.057015
XGBoost	0.081215	0.106209	0.089952	0.054324	0.05525

Table 3. Hybrid Ensemble Performance

Metric	Average	Standard Deviation
Precision	0.661	0.088
Recall	0.6463	0.1065
F1-Score	0.6517	0.0925
Accuracy	0.777	0.0562
ROC-AUC	0.795	0.0775

4. Conclusion

This study evaluated the performance of various machine learning models, including CatBoost, Gradient Boosting, LightGBM, Neural Network, Random Forest, and XGBoost, on a dataset using 10-fold cross-validation. The findings highlight the effectiveness of ensemble methods, with Random Forest and Gradient Boosting emerging as the top-performing models based on metrics such as precision, recall, F1-score, accuracy, and ROC-AUC. These models demonstrated a robust ability to capture complex patterns in the data, achieving high performance with relatively low variability across folds. The hybrid ensemble approach using majority voting achieved competitive results, combining the strengths of individual models to produce stable and reliable predictions. While the hybrid model did not significantly outperform the best single models, it showed potential for robustness and generalization, particularly in scenarios involving varied data distributions.

The Neural Network underperformed compared to tree-based methods, suggesting the need for further optimization or the application of more sophisticated deep learning architectures to improve its utility for this dataset. Furthermore, the variability in performance metrics for certain models, such as CatBoost and Neural

Network, indicates sensitivity to data splits, which could be addressed through enhanced feature engineering or hyperparameter tuning. In conclusion, ensemble methods, particularly Random Forest and Gradient Boosting, are well-suited for this problem due to their consistent and high performance across evaluation metrics. The hybrid ensemble approach offers a promising avenue for improving robustness, while neural networks require further exploration and refinement. Future research could focus on advanced ensemble techniques, feature engineering, and hyperparameter optimization to further enhance model performance and applicability. Additionally, integrating domain-specific insights into the model development process could improve the interpretability and utility of the results in real-world scenarios.

References

[1] M. T. I. Khan, S. Anwar, S. A. Sarkodie, M. R. Yaseen, and A. M. Nadeem, "Do natural disasters affect economic growth? The role of human capital, foreign direct investment, and infrastructure dynamics," *Heliyon*, vol. 9, no. 1, 2023.

[2] P. Cui *et al.*, "Scientific challenges of research on natural hazards and disaster risk," *Geogr. Sustain.*, vol. 2, no. 3, pp. 216–223, 2021.

[3] M. T. Chaudhary and A. Piracha, "Natural disasters—origins, impacts, management," *Encyclopedia*, vol. 1, no. 4, pp. 1101–1131, 2021.

[4] L. Velotti, "Natural Hazards: Tsunamis," in *Encyclopedia of Security and Emergency Management*, Springer, 2021, pp. 693–696.

[5] H. A. Bachay, A. H. Aldefae, S. L. Zubaidi, W. H. Humaish, and E. K. Sinichenko, "Tsunami Waves, Causes and Its Implications: A Review," *Geotech. Eng. Sustain. Constr. Sustain. Geotech. Eng.*, pp. 63–77, 2022.

[6] J. S. Becker *et al.*, "Earthquakes and tsunami," in *Routledge handbook of environmental hazards and society*, Routledge, 2022, pp. 13–32.

[7] T. Srinivasa Kumar and S. Manneela, "A review of the progress, challenges and future trends in tsunami early warning systems," *J. Geol. Soc. India*, vol. 97, pp. 1533–1544, 2021.

[8] J. Behrens *et al.*, "Probabilistic tsunami hazard and risk analysis: A review of research gaps," *Front. Earth Sci.*, vol. 9, p. 628772, 2021.

[9] P. M. De Martini *et al.*, "The Mediterranean Sea and the Gulf of Cadiz as a natural laboratory for paleotsunami research: Recent advancements," *Earth-Science Rev.*, vol. 216, p. 103578, 2021.

[10] X. Cao, Z. Liu, C. Hu, X. Song, J. A. Quaye, and N. Lu, "Three-Dimensional Geological Modelling in Earth Science Research: An In-Depth Review and Perspective Analysis," *Minerals*, vol. 14, no. 7, p. 686, 2024.

[11] Z. Ma and G. Mei, "Deep learning for geological hazards analysis: Data, models, applications, and opportunities," *Earth-Science Rev.*, vol. 223, p. 103858, 2021.

[12] L. Xu, N. Chen, Z. Chen, C. Zhang, and H. Yu, "Spatiotemporal forecasting in earth system science: Methods, uncertainties, predictability and future directions," *Earth-Science Rev.*, vol. 222, p. 103828, 2021.

[13] C. M. Liu, D. Rim, R. Baraldi, and R. J. LeVeque, "Comparison of machine learning approaches for tsunami forecasting from sparse observations," *Pure Appl. Geophys.*, vol. 178, no. 12, pp. 5129–5153, 2021.

[14] E. Cesario, S. Giampá, E. Baglione, L. Cordrie, J. Selva, and D. Talia, "A parallel machine learning-based approach for tsunami waves forecasting using regression trees," *Comput. Commun.*, vol. 225, pp. 217–228, 2024.

[15] A. Pourzangbar, M. Jalali, and M. Brocchini, "Machine learning application in modelling marine and coastal phenomena: a critical review," *Front. Environ. Eng.*, vol. 2, p. 1235557, 2023.

[16] X. Shu and Y. Ye, "Knowledge Discovery: Methods from data mining and machine learning," *Soc. Sci. Res.*, vol. 110, p. 102817, 2023.
[17] Y. Zhao and D. Gorse, "Earthquake prediction from seismic indicators using tree-based ensemble learning," *Nat. Hazards*, vol. 120, no. 3, pp. 2283–2309, 2024.
[18] 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, 2023.
[19] W. Khan, A. Daud, K. Khan, S. Muhammad, and R. Haq, "Exploring the frontiers of deep learning and natural language processing: A comprehensive overview of key challenges and emerging trends," *Nat. Lang. Process. J.*, p. 100026, 2023.
[20] Z. Ma, G. Mei, and N. Xu, "Generative deep learning for data generation in natural hazard analysis: motivations, advances, challenges, and opportunities," *Artif. Intell. Rev.*, vol. 57, no. 6, p. 160, 2024.
[21] T. Sharma, A. Singhal, K. Kundu, and N. Agarwal, "Machine Learning/Deep Learning for Natural Disasters," in *Applications of Artificial Intelligence, Big Data and Internet of Things in Sustainable Development*, CRC Press, 2022, pp. 91–121.
[22] M. A. Ganaie, M. Hu, A. K. Malik, M. Tanveer, and P. N. Suganthan, "Ensemble deep learning: A review," *Eng. Appl. Artif. Intell.*, vol. 115, p. 105151, 2022.
[23] R. Triatmadja and others, "Vegetation-based approached for tsunami risk reduction: Insights and challenges," *Prog. Disaster Sci.*, p. 100352, 2024.
[24] M. Esposito, L. Palma, A. Belli, L. Sabbatini, and P. Pierleoni, "Recent advances in internet of things solutions for early warning systems: A review," *Sensors*, vol. 22, no. 6, p. 2124, 2022.
[25] H. Li, "Deep-Ocean Assessment and Reporting of Tsunamis (DART) BUOY," in *Encyclopedia of Ocean Engineering*, Springer, 2022, pp. 306–314.
[26] Y. Zhao, S. Lv, and P. Liu, "Advances in earthquake prevention and reduction based on machine learning: a scoping review (July 2024)," *IEEE Access*, 2024.
[27] S. K. Nezhad, M. Barooni, D. Velioglu Sogut, and R. J. Weaver, "Ensemble neural networks for the development of storm surge flood modeling: A comprehensive review," *J. Mar. Sci. Eng.*, vol. 11, no. 11, p. 2154, 2023.
[28] P. Govindarajan, N. Venkatanathan, and others, "Towards real-time earthquake forecasting in Chile: Integrating intelligent technologies and machine learning," *Comput. Electr. Eng.*, vol. 117, p. 109285, 2024.
[29] Y. Li and K. Goda, "Risk-based tsunami early warning using random forest," *Comput. \& Geosci.*, vol. 179, p. 105423, 2023.
[30] E. K. Sahin, "Comparative analysis of gradient boosting algorithms for landslide susceptibility mapping," *Geocarto Int.*, vol. 37, no. 9, pp. 2441–2465, 2022.
[31] E. Karampotsis, K. Kioskli, A. Tsirimpa, G. Dounias, and A. Polydoropoulou, "Understanding evacuation behavior for effective disaster preparedness: a hybrid machine learning approach," *Nat. Hazards*, pp. 1–39, 2024.
[32] E. CHERIF, "Deep learning-enhanced satellite-based Flood detection for early Warning and prevention," University Larbi Tébessi--Tébessa, 2024.
[33] M. H. Al Banna *et al.*, "Attention-based bi-directional long-short term memory network for earthquake prediction," *IEEE Access*, vol. 9, pp. 56589–56603, 2021.
[34] G. Xu *et al.*, "Machine learning in coastal bridge hydrodynamics: a state-of-the-art review," *Appl. Ocean Res.*, vol. 134, p. 103511, 2023.
[35] C. Zhou, M. Ye, Z. Xia, W. Wang, C. Luo, and J.-P. Muller, "An interpretable attention-based deep learning method for landslide prediction based on multi-temporal InSAR time series: A case study of Xinpu landslide in the TGRA," *Remote Sens. Environ.*, vol. 318, p. 114580, 2025.
[36] A. S. Albahri *et al.*, "A systematic review of trustworthy artificial intelligence applications in natural disasters," *Comput. Electr. Eng.*, vol. 118, p. 109409, 2024.

[37] J. Deng, G. Gong, G. Zhou, M. Yan, and L. Li, "A gated feature fusion network with meta-encoder and self-calibrating cross module for building change detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, 2024.

[38] P. L. Ford, "Earthquakes Historical Data and Live Data." Kaggle, 2023. https://www.kaggle.com/datasets/patricklford/earthquakes-historical-data-and-live-data/data