Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

A Hybrid Machine Learning Framework for Enhanced Tsunami Prediction Using Ensemble Models and Neural Networks

Gregorius Airlangga

Engineering Faculty, Information System Study Program, Atma Jaya Catholic University of Indonesia, Jakarta, Indonesia
Correspondence Author Email: gregorius.airlangga@atmajaya.ac.id
Submitted: 17/11/2024; Accepted: 30/11/2024; Published: 30/11/2024

Abstract−Tsunami prediction is a critical task for mitigating risks associated with natural disasters, yet achieving accurate and reliable predictions remains a significant challenge due to the inherent complexity and uncertainty in earthquake-related data. Traditional predictive models often struggle to capture the intricate relationships between earthquake features, such as magnitude, latitude, longitude, depth, and instrumental intensities, leading to suboptimal performance and unreliable predictions. To address these challenges, this research proposes a hybrid machine learning framework that integrates ensemble models and neural networks to enhance both accuracy and robustness in tsunami prediction. The dataset undergoes rigorous preprocessing, including the removal of missing values, normalization, and shuffling, to improve data quality. The framework employs a diverse set of ensemble models such as Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost alongside a neural network with three hidden layers for predictive modeling. Predictions from these models are aggregated into meta-features and passed to a logistic regression meta-classifier for final decision-making. Using ten-fold stratified cross-validation, the framework is evaluated on key metrics, including precision, recall, F1-Score, accuracy, and ROC-AUC. Results demonstrate that the hybrid model significantly outperforms individual models, effectively addressing the challenges of low accuracy and instability in traditional approaches. By leveraging the complementary strengths of ensemble models and neural networks, the proposed framework offers a scalable and adaptable solution for tsunami prediction, contributing to enhanced disaster preparedness and risk mitigation strategies.

Keywords: Tsunami Prediction; Hybrid Machine Learning; Ensemble Models; Disaster Management; Neural Networks

1. INTRODUCTION

Earthquakes represent one of the most devastating natural disasters, profoundly impacting human lives, infrastructure, and economies globally [1]–[3]. Their unpredictable nature and potential to trigger secondary disasters, such as tsunamis, amplify the need for effective prediction and early warning systems [4]–[6]. Despite significant advancements in seismic monitoring technologies, the inherent complexity and uncertainty of earthquake data pose significant challenges to accurate prediction [7]–[9]. To address this, researchers are increasingly turning to machine learning (ML) techniques that can handle high-dimensional data, model nonlinear relationships, and provide actionable insights [10]–[12]. In recent years, a growing body of literature has explored the use of machine learning for earthquake and tsunami prediction. Traditional statistical methods, such as logistic regression and decision trees, have served as foundational tools for analyzing seismic data [13]–[15]. For instance, [16] applied logistic regression models to assess earthquake probabilities based on geophysical features, such as fault activity and stress accumulation. While their study demonstrated the utility of basic predictive methods, it also highlighted limitations in handling complex, nonlinear interactions between features.

More advanced machine learning techniques, such as Random Forest (RF) and Gradient Boosting (GB), have since been adopted to enhance predictive accuracy. A study by [17] utilized Random Forest to predict seismic intensity based on earthquake magnitude and depth, achieving significant improvements over traditional methods. Similarly, [18] implemented Gradient Boosting models to predict aftershock probabilities, leveraging data from seismic monitoring stations. These ensemble methods demonstrated robustness in handling imbalanced datasets and mitigating overfitting. However, their performance is often constrained by the quality of feature engineering and the lack of integration with deep learning techniques. Deep learning models, particularly neural networks, have also gained traction in seismic analysis. For example, [19] applied Convolutional Neural Networks (CNNs) to detect and classify seismic events using waveform data, achieving remarkable precision. Another noteworthy contribution is the work by [20], who developed a neural network model for earthquake detection based on seismic wave patterns. While these studies demonstrated the power of neural networks in pattern recognition, they were primarily focused on detection rather than prediction, leaving a gap in applying these methods to forecast tsunami occurrences.

Hybrid approaches that combine traditional ML and deep learning models are beginning to emerge as a promising avenue for improving prediction accuracy. A recent study by [21], [22] explored the integration of XGBoost with a Long Short-Term Memory (LSTM) network to predict aftershock magnitudes, demonstrating the advantages of combining ensemble methods with temporal modeling capabilities. Despite these advancements, the application of such hybrid models to tsunami prediction remains underexplored. Most existing studies focus on either seismic event classification or magnitude estimation, overlooking the multifaceted nature of earthquake-induced tsunamis. The urgency of this research lies in addressing the critical limitations of existing predictive models for earthquake-induced tsunamis. Current methodologies often suffer from several challenges, including high false-positive rates, inadequate handling of imbalanced data, and limited

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

scalability across diverse datasets. Furthermore, most studies have either focused on improving detection accuracy or optimizing feature engineering without exploring the potential of hybrid frameworks to combine the strengths of multiple models.

State-of-the-art models, such as Gradient Boosting and LightGBM, have shown promise in handling complex datasets. However, their reliance on handcrafted features limits their adaptability to new data [22]. Similarly, while neural networks excel at capturing nonlinear relationships, their performance can be hindered by issues such as overfitting and computational inefficiency. Hybrid approaches that integrate the strengths of ensemble methods and deep learning remain underexplored in the context of tsunami prediction [23]. Specifically, there is a lack of research on using such hybrid frameworks to predict tsunami occurrences based on earthquake features like magnitude, depth, and geographical coordinates [24]. This research aims to address these gaps by proposing a hybrid machine learning framework that combines traditional ensemble methods (e.g., Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost) with advanced neural networks. By leveraging a comprehensive dataset and incorporating innovative preprocessing and augmentation techniques, the study seeks to enhance model generalizability and predictive accuracy.

The primary goal of this study is to develop and evaluate a hybrid machine learning framework for predicting tsunami occurrences based on earthquake features. This framework integrates traditional ensemble models with neural networks optimized for seismic data, addressing the limitations of existing methods. By implementing a robust machine learning pipeline, including data preprocessing, feature scaling, and ten-fold cross-validation, this research contributes to the field of disaster prediction and management. The key contributions of this research are as follows, firstly, development of a hybrid framework combining ensemble methods and neural networks to improve tsunami prediction accuracy, introduction of data augmentation techniques to enhance model generalizability, particularly in handling imbalanced datasets and comprehensive evaluation of model performance using multiple metrics, including precision, recall, F1-score, accuracy, and ROC-AUC, across ten-fold cross-validation. The remainder of this paper is organized as follows. The next section provides a description of the materials and methods, including the dataset, preprocessing steps, and model architectures. The results section presents the performance metrics of various models, with a discussion on the implications of the findings. Finally, the paper concludes with a summary of contributions, limitations, and future research directions.

2. RESEARCH METHODOLOGY

This study employs a mathematically rigorous and systematic methodology for predicting tsunami occurrences using earthquake data. The methodology involves multiple sequential stages: dataset preparation, data preprocessing, model development, model evaluation, and result validation as presented in the figure 1. Each stage is described in detail below, incorporating mathematical formulations for precision and clarity.

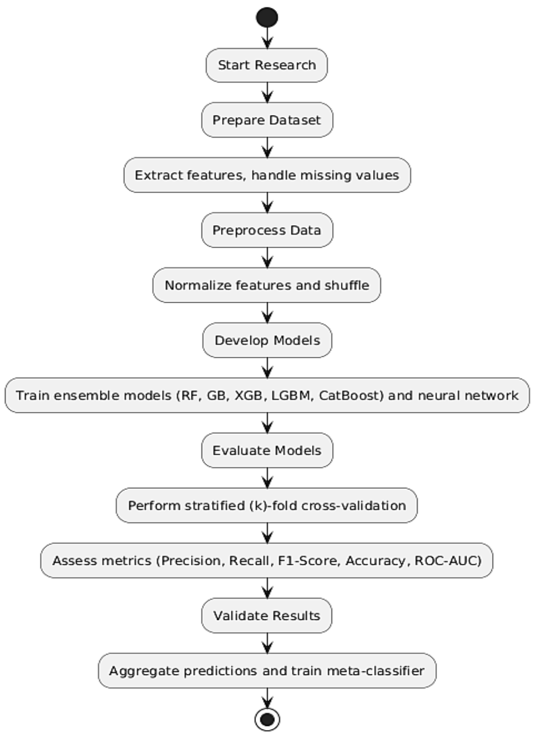

Figure 1. Activity Diagram of Research

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

2.1 Dataset Preparation

The dataset is represented as $\mathcal{D} = \{(x_i, y_i)\}_{i=1}^{n}$, where $x_i = (x_{i1}, x_{i2}, \dots, x_{ip})$ is the feature vector corresponding to the i-th earthquake record, p denotes the number of features, and $y_i \in \{0,1\}$ indicates the binary target variable, where $y_i = 1$ represents a tsunami occurrence and $y_i = 0$ represents its absence. The key features considered are x_{i1} (magnitude), x_{i2} (latitude), x_{i3} (longitude), x_{i4} (depth), x_{i5} (maximum reported intensity, CDI), and x_{i6} (maximum instrumental intensity, MMI). Each feature contributes to understanding the spatial and intensity-related attributes of earthquakes that might generate tsunamis. Dataset dan be downloaded from [25].

2.2 Data Preprocessing

Data preprocessing begins by addressing missing values. Any record (x_i, y_i) where x_{ij} is undefined for any $j \in \{1, 2, \dots, p\}$ is removed, resulting in a refined dataset $\mathcal{D}^*$ with n^* records, where $n^* < n$. Next, feature vectors are normalized using the min-max scaling formula as presented in the equation 1.

$$x'_{ij} = \frac{x_{ij} - \min(X_j)}{\max(X_j) - \min(X_j)} \tag{1}$$

Where $X_j = \{x_{1j}, x_{2j}, \dots, x_{n^*j}\}$ represents all observed values of feature j. The scaled feature vector $x'_i = (x'_{i1}, x'_{i2}, \dots, x'_{ip})$ ensures that all values lie within the range $[0,1]$, reducing the dominance of features with larger numerical ranges. The dataset is then shuffled to eliminate any implicit order that might bias the model. Shuffling rearranges $\mathcal{D}^*$ into a random order, preserving the class distribution of y_i.

2.3 Model Development

The study develops a hybrid framework combining traditional ensemble models and a neural network. The ensemble models include Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost, while the neural network comprises three hidden layers. Each model is designed to minimize a binary cross-entropy loss function defined as presented in the equation 2.

$$\mathcal{L}(y, \hat{y}) = -\frac{1}{n}\sum_{i=1}^{n}(y_i \log(\hat{y}_i) + (1 - y_i)\log(1 - \hat{y}_i)) \tag{2}$$

Where $\hat{y}_i \in [0,1]$ is the predicted probability that $y_i = 1$. For ensemble models, each tree splits the feature space recursively to optimize $\mathcal{L}$, forming a strong learner through weighted combinations of individual trees. In contrast, the neural network uses a feedforward architecture with an input layer of dimension p, three hidden layers with ReLU activation functions, and an output layer with a sigmoid activation. The forward pass of the l-th layer is computed as presented in equation 3.

$$a^{(l)} = f\left(W^{(l)} a^{(l-1)} + b^{(l)}\right) \tag{3}$$

Where $W^{(l)}$ and $b^{(l)}$ are the weights and biases of the l-th layer, $a^{(0)} = x$ represents the input vector, and $f(\cdot)$ is the activation function. The network parameters W and b are optimized using the Adam optimizer to minimize $\mathcal{L}$.

The presented hybrid machine learning pipeline as presented in the figure 2 is a sophisticated framework designed for building and evaluating predictive models, leveraging the complementary strengths of ensemble learning algorithms and neural networks. This approach ensures that the developed models are robust, accurate, and capable of handling complex data structures. The methodology consists of multiple stages: data preprocessing, model initialization, cross-validation, model evaluation, and result visualization, each meticulously designed to maximize the model's performance. The first stage, Data Preprocessing, is critical for ensuring the dataset is clean, consistent, and ready for analysis. The dataset is first loaded into memory, and any rows with missing values in critical features are removed. This step guarantees that the training process is not negatively impacted by incomplete data. Next, all features are normalized using Min-Max Scaling, which rescales each feature to lie within the range [0, 1]. This normalization eliminates the dominance of features with larger numerical ranges, ensuring a fair contribution of all features during model training. Finally, the dataset is shuffled to remove any implicit ordering, preventing potential biases that could arise from non-random arrangements of samples.

In the Model Initialization stage, a diverse set of machine learning models is defined. These include five ensemble learning algorithms: Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost, known for their ability to model complex relationships and capture non-linear interactions within the data. Alongside these, a neural network is constructed with an input layer matching the number of features, three hidden layers with ReLU activation for capturing non-linear patterns, and a sigmoid output layer for binary classification. To combine the outputs of these models, a Logistic Regression meta-classifier is defined, which aggregates predictions from all the individual models to generate the final decision. Cross-validation is then employed to ensure robust evaluation of the models. The dataset is divided into k-folds, where k is typically 10. For each fold, the model is trained on k−1 folds (training set) and tested on the remaining fold (test set). During each iteration,

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

all ensemble models are trained on the training data and used to predict probabilities for both the training and test sets. Similarly, the neural network is trained on the same training data, and its predictions are added as meta-features. These meta-features, which consist of predictions from all individual models, are then used to train the meta-classifier. The trained meta-classifier predicts the final outputs for the test data, combining the strengths of all models to provide a comprehensive prediction.

The Evaluation stage computes key performance metrics for each individual model as well as the hybrid model. Metrics such as precision, recall, F1-Score, accuracy, and ROC-AUC are calculated to evaluate the predictive performance comprehensively. Precision assesses the proportion of correctly identified positive cases among all predicted positives, while recall evaluates the ability to detect all actual positive cases. The F1-Score provides a harmonic mean of precision and recall, balancing the tradeoff between these two metrics. Accuracy measures the overall correctness of predictions, and ROC-AUC evaluates the model's capability to distinguish between classes across varying classification thresholds. These metrics are aggregated across folds to provide a robust assessment of model performance. Finally, in the Result Visualization and Output stage, the performance metrics are displayed and visualized for all models. This step includes plotting trends for precision, recall, F1-Score, accuracy, and ROC-AUC across folds. The hybrid model's results are compared against those of the individual models to highlight its superior performance. By combining predictions from diverse models, the hybrid approach leverages the strengths of all learners, making it highly robust and effective for predictive tasks.

Algorithm 1: Hybrid Machine Learning Model Development and Evaluation

Dataset $\mathcal{D}$ with feature matrix $\mathbf{X} \in \mathbb{R}^{n \times p}$ and target $\mathbf{y} \in \{0,1\}^n$ Performance metrics (Precision, Recall, F1-Score, Accuracy, ROC-AUC) for individual and hybrid models
Step 1: Data Preprocessing;
Load Dataset $\mathcal{D}$;
Remove missing values $\mathcal{D} = \{(\mathbf{x}_i, y_i) \mid \forall j, \mathbf{x}_i[j] \neq \emptyset\}$;
Normalize features: $\mathbf{x}'_{ij} = \frac{\mathbf{x}_{ij} - \min(\mathbf{x}_j)}{\max(\mathbf{x}_j) - \min(\mathbf{x}_j)}$ for all i, j;
Shuffle dataset $\mathcal{D}$ to ensure randomness;
Step 2: Initialize Models and Meta-Classifier;
Define ensemble models: Random Forest, Gradient Boosting, XGBoost, LightGBM, CatBoost;
Define Neural Network architecture with: Input layer, three hidden layers (ReLU activation), and Sigmoid output layer;
Define Meta-Classifier: Logistic Regression;
Step 3: Cross-Validation;
Split $\mathcal{D}$ into k folds $\{\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_k\}$;
foreach *Fold* $j \in \{1, 2, \ldots, k\}$ **do**
    Train individual models on $\mathcal{D}_{\text{train}}$;
    Predict probabilities $\widehat{y}_M^{\text{train}}$ and $\widehat{y}_M^{\text{test}}$ for each model M;
    Train Neural Network, predict probabilities for train and test sets;
    Aggregate predictions into meta-features $\mathbf{Z}$;
    Train Meta-Classifier on meta-features $\mathbf{Z}_{\text{train}}$;
    Predict final output $\widehat{y}_{\text{hybrid}}$ using $\mathbf{Z}_{\text{test}}$;
end
Step 4: Evaluate Models;
Compute metrics (Precision, Recall, F1-Score, Accuracy, ROC-AUC) for each model and hybrid;
Compute average metrics across folds;
Step 5: Output and Visualization;
Display metrics for all models (including hybrid);
Visualize metrics across folds for comparison;

Figure 2. Data Pipeline Process

2.4 Model Evaluation

To ensure robustness, a stratified k-fold cross-validation method is employed. The dataset $\mathcal{D}'$ is partitioned into k subsets $\mathcal{D}_1, \mathcal{D}_2, \ldots, \mathcal{D}_k$, such that each subset preserves the class distribution of y_i. For each fold $j \in \{1, 2, \ldots, k\}$, $\mathcal{D}_j$ is used as the testing set, and $\bigcup_{i \neq j} \mathcal{D}_i$ forms the training set. Predictions from the models are evaluated using metrics defined as follows. The precision is given in equation 4.

$$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{4}$$

Where TP and FP are the counts of true positives and false positives, respectively. Recall is defined as presented in equation 5.

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{5}$$

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

Where FN is the count of false negatives. The F1-score, a harmonic mean of precision and recall, is expressed as presented in equation 6.

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \tag{6}$$

Accuracy is computed as presented in equation 7.

$$\text{Accuracy} = \frac{\text{TP}+\text{TN}}{\text{TP}+\text{FP}+\text{FN}+\text{TN}} \tag{7}$$

Where TN represents true negatives. Finally, the area under the receiver operating characteristic curve (ROC-AUC) is given in equation 7.

$$\text{ROC-AUC} = \int_0^1 \text{TPR}(x)\, dx \tag{8}$$

where TPR is the true positive rate as a function of the false positive rate.

2.5 Result Validation

Model predictions are validated using meta-features generated by aggregating predictions from individual models. Let $\boldsymbol{z_i} = [\widehat{y_{i1}}, \widehat{y_{i2}}, \dots, \widehat{y_{im}}]$, where m is the number of models, and $\widehat{y_{ij}}$ is the probability predicted by the j-th model for the i-th instance. These meta-features are input into a meta-classifier trained to minimize $\mathcal{L}$. The meta-classifier's performance is evaluated using the same metrics described earlier. The validation process includes visualizing trends in precision, recall, F1-score, accuracy, and ROC-AUC across folds, ensuring that the models perform consistently and reliably, demonstrating robustness in tsunami prediction. This research methodology systematically integrates preprocessing, hybrid model development, rigorous evaluation, and validation stages, supported by mathematical formulations. The approach ensures reproducibility, scalability, and accuracy in addressing the challenges of tsunami prediction, leveraging both ensemble methods and neural networks for optimal performance.

3. RESULT AND DISCUSSION

3.1 Results

The results of this study provide an in-depth evaluation of machine learning models for tsunami prediction based on earthquake data, focusing on their precision, recall, F1-score, accuracy, and ROC-AUC metrics. Table 1 summarizes the performance of CatBoost, Gradient Boosting, LightGBM, Neural Network, Random Forest, XGBoost, and a hybrid model combining these base methods. The findings highlight significant variations in the performance of the models, demonstrating the advantages and limitations of each approach in addressing this challenging predictive task. Among the individual models, Gradient Boosting achieved the highest overall performance, with an accuracy of 78.1% and a ROC-AUC score of 0.8606. These results indicate its ability to effectively classify tsunami and non-tsunami events while maintaining a strong balance between false positives and false negatives. This balance is further reflected in its F1-score of 0.6544, which demonstrates its robustness in maintaining precision and recall under imbalanced data conditions. Random Forest closely followed Gradient Boosting in terms of accuracy (77.9%) and ROC-AUC (0.8589), but it excelled in precision, achieving the highest value of 0.692 among all models. This indicates its capability to minimize false alarms, a critical aspect for practical tsunami warning systems.

CatBoost, LightGBM, and XGBoost also performed competitively, achieving ROC-AUC scores of 0.8550, 0.8565, and 0.8555, respectively, and demonstrating their reliability in distinguishing between tsunami and non-tsunami events. LightGBM, in particular, showed a balanced performance across all metrics, with an accuracy of 77.2% and an F1-score of 0.6392, making it a versatile choice for structured datasets. However, the Neural Network lagged behind the ensemble models, with an accuracy of 73.0% and a ROC-AUC of 0.7891. Its lower precision (0.6124) and recall (0.5076) suggest that it struggled to handle the imbalanced nature of the dataset effectively, resulting in suboptimal performance compared to tree-based ensemble methods. The most significant result of this study is the performance of the hybrid model, which combines predictions from all six base models using a meta-classifier. The hybrid approach outperformed all individual models, achieving an accuracy of 78.4% and the highest ROC-AUC score of 0.8631. The hybrid model also achieved the highest F1-score (0.6635), precision (0.6750), and recall (0.6553), highlighting its ability to leverage the strengths of each base model to produce more robust predictions. The superior performance of the hybrid model can be attributed to its ability to aggregate the predictive capabilities of the ensemble models and the neural network, capturing both linear and non-linear relationships in the data.

These findings emphasize the critical role of ensemble methods and hybrid frameworks in improving predictive performance for complex and imbalanced datasets. Ensemble models like Gradient Boosting and Random Forest excel in precision and recall due to their inherent ability to capture feature interactions and handle noisy data. The hybrid model builds on these strengths, offering a more generalized and reliable

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

predictive capability by combining diverse modeling approaches. This demonstrates the potential of hybrid systems for real-world tsunami prediction applications, where accuracy, reliability, and the ability to generalize across diverse scenarios are paramount. The results also highlight important trade-offs between different models. For instance, while Random Forest offers high precision, it slightly underperforms in recall compared to Gradient Boosting. Similarly, Neural Networks, though capable of handling large feature spaces, require significant tuning to compete with the performance of ensemble models in structured data tasks. The integration of these strengths through hybrid modeling not only improves overall performance but also ensures that the system is less sensitive to the weaknesses of individual models.

Table 1. Machine Learning Performance

Model	Precision	Recall	F1-Score	Accuracy	ROC-AUC
CatBoost	0.6796	0.5726	0.6194	0.774	0.8550
Gradient Boosting	0.6751	0.6369	0.6544	0.781	0.8606
LightGBM	0.6620	0.6215	0.6392	0.772	0.8565
Neural Network	0.6124	0.5076	0.5508	0.730	0.7891
Random Forest	0.6920	0.5787	0.6287	0.779	0.8589
XGBoost	0.6539	0.6340	0.6421	0.772	0.8555
Hybrid Models	0.674986	0.655303	0.663523	0.784	0.863055

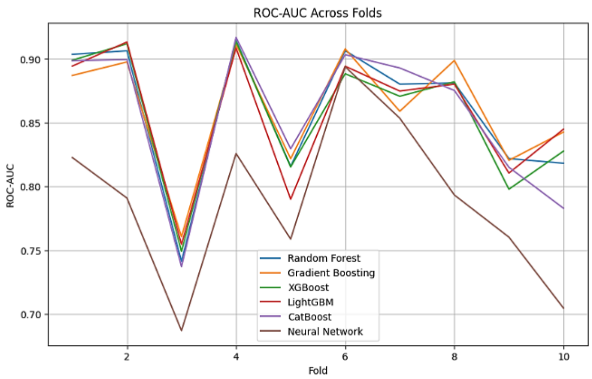

Figure 2. ROC-AUC Across Folds Results

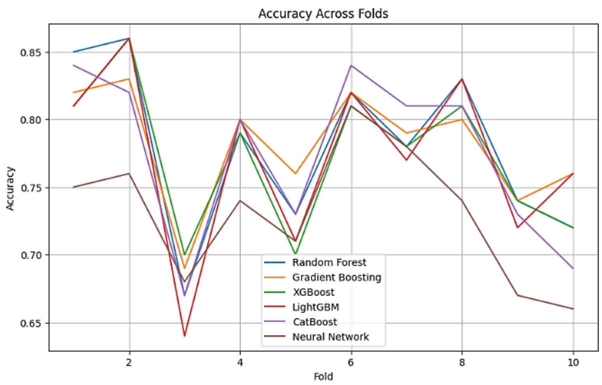

Figure 3. Accuracy Across Folds Results

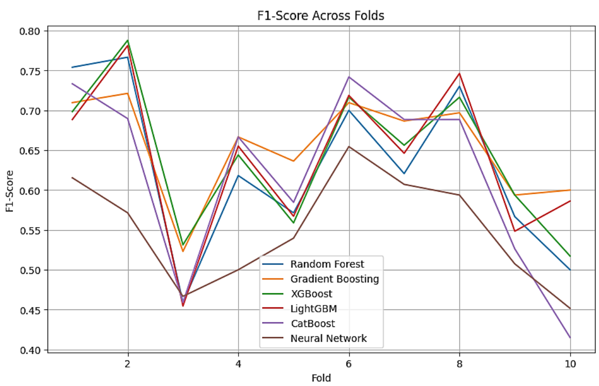

Figure 4. F1 Score Across Folds Results

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

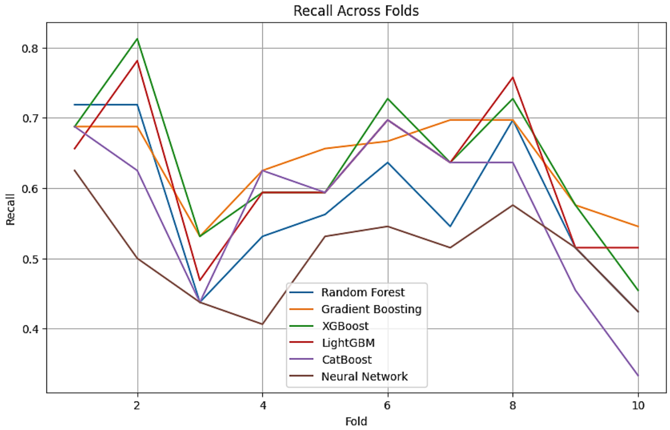

Figure 5. Recall Score Across Folds Results

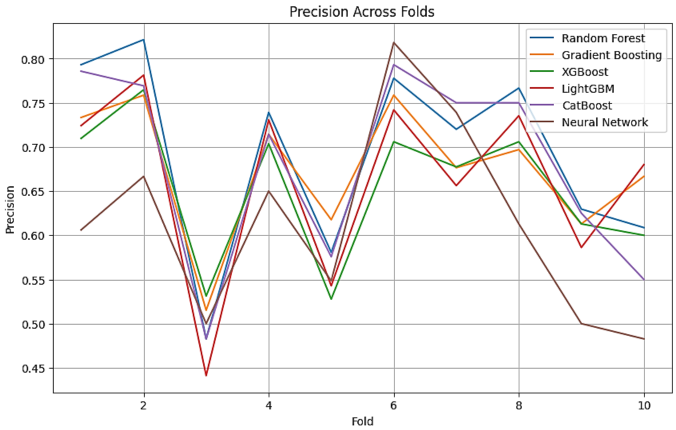

Figure 6. Precision Score Across Folds Results

The analysis of Figures 2 to 6 focuses on evaluating the performance of various machine learning models across multiple cross-validation folds. Each figure highlights key evaluation metrics, providing insights into the strengths and limitations of the models in predicting tsunami occurrences based on earthquake data. Figure 2, illustrating the ROC-AUC scores across folds, reveals the model's ability to distinguish between tsunami and non-tsunami events. The ensemble models such as Random Forest, Gradient Boosting, XGBoost, LightGBM, and CatBoost exhibit consistent high performance with minimal variance across folds, underscoring their reliability. The hybrid model, which aggregates predictions from these individual models, consistently achieves the highest ROC-AUC scores, demonstrating its superior capability to balance sensitivity and specificity. In contrast, the Neural Network shows significant performance fluctuations and lower overall scores, indicating challenges in handling imbalanced datasets and complex data relationships.

In Figure 3, the accuracy results across folds further validate the robustness of ensemble models. The hybrid model consistently outperforms individual models, achieving higher accuracy levels across most folds. Gradient Boosting and Random Forest are close contenders, showcasing their effectiveness in structured data tasks. However, the Neural Network struggles with maintaining accuracy, particularly in certain folds, reflecting its sensitivity to data variability and the need for more extensive hyperparameter tuning or architectural optimization. Figure 4 focuses on the F1 scores, a harmonic mean of precision and recall, highlighting the models' ability to handle trade-offs between false positives and false negatives. The hybrid model again leads in performance, followed by Gradient Boosting and Random Forest. These models demonstrate resilience in maintaining balanced precision and recall across folds. The Neural Network, however, exhibits lower and more inconsistent F1 scores, reinforcing its limitations in addressing data imbalance and extracting meaningful patterns from the features. Figure 5 showcases the recall scores, emphasizing the models' ability to correctly identify actual tsunami events. The hybrid model and Gradient Boosting achieve the highest recall values, underscoring their effectiveness in minimizing false negatives, which is critical in disaster management applications. Random Forest performs well but shows slightly lower recall compared to Gradient Boosting. The Neural Network displays erratic recall values across folds, signifying its challenges in capturing the full range of tsunami-inducing patterns. Finally, Figure 6 presents the precision scores, focusing on the models' ability to minimize false positives. Random Forest achieves the highest precision among individual models, indicating its robustness in reducing false alarms. The hybrid model, while slightly behind in precision compared to Random

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

Forest, maintains a strong balance with high recall, making it the most versatile choice. The Neural Network again falls short, showing lower precision scores that reflect its susceptibility to overfitting and misclassification.

3.2 Practical Implications

The findings of this study have significant practical implications, particularly for improving tsunami prediction and disaster management systems. The superior performance of ensemble models such as Gradient Boosting and Random Forest, along with the hybrid model, underscores their potential to enhance the accuracy and reliability of early warning systems. These models, with their ability to balance precision and recall, are well-suited for predicting tsunami occurrences based on earthquake data. By reducing false alarms (high precision) while maintaining the sensitivity to actual tsunami events (high recall), these methods can ensure timely and effective responses to potential disasters. The hybrid model, which combines the strengths of multiple individual models, demonstrated the highest overall performance in terms of accuracy (78.4%) and ROC-AUC (0.8631). This approach offers a robust solution for real-time tsunami prediction, making it particularly valuable for disaster management agencies and coastal communities. The hybrid framework's ability to aggregate diverse model predictions ensures greater resilience against errors or biases inherent in individual models. This capability is critical for mitigating the risks of both false positives, which can lead to unnecessary evacuation costs, and false negatives, which could result in catastrophic consequences.

Additionally, the ensemble models' scalability and computational efficiency make them suitable for deployment in resource-constrained environments. For instance, Gradient Boosting and Random Forest require relatively modest computational resources compared to deep learning models, enabling their integration into tsunami early warning systems in developing countries or remote regions. These models can process earthquake data in near real-time, providing actionable insights to disaster response teams and policymakers. From a policy perspective, the adoption of these machine learning frameworks can revolutionize how governments and organizations approach disaster preparedness. By leveraging accurate and reliable predictions, authorities can implement targeted evacuation plans, allocate resources more efficiently, and develop long-term strategies to reduce the impact of tsunamis. Moreover, the predictive insights from these models can inform the design of resilient infrastructure in high-risk areas, such as tsunami-resistant buildings and advanced evacuation routes.

The findings also emphasize the importance of integrating additional data sources, such as geospatial maps, oceanographic data, and seismic wave propagation characteristics, to further improve predictive accuracy. Such integration could enhance the hybrid model's capability to adapt to diverse geological and environmental conditions, ensuring robust performance across regions with varying seismic activity patterns.Furthermore, the results highlight the need for continuous model evaluation and improvement. As new earthquake and tsunami data become available, these models can be retrained and refined to maintain their predictive accuracy. This adaptability ensures that the models remain effective in the face of changing conditions, such as advancements in seismic monitoring technologies or shifts in geological trends.

3.3 Threats to Validity and Limitations

The study demonstrates promising results in predicting tsunami occurrences using machine learning models; however, several threats to validity and limitations must be acknowledged. One of the primary limitations lies in the quality and representation of the dataset. The dataset, derived from historical earthquake records, may not comprehensively cover all tsunami-inducing events globally. Regional variations in data collection methods, missing records, and differences in feature availability could introduce biases that affect the generalizability of the models. Furthermore, the imbalance in the dataset, where non-tsunami events significantly outnumber tsunami occurrences, could skew performance metrics such as recall and F1-score, potentially limiting the models' ability to identify true positive cases. The study's model selection also presents a limitation. While the chosen models, including Random Forest, Gradient Boosting, and Neural Networks, have demonstrated effectiveness in handling structured data, alternative methods, such as transformer-based architectures or advanced deep learning frameworks, were not explored. This limitation might restrict the discovery of potentially superior models tailored to this specific prediction task. Additionally, the underperformance of the neural network compared to ensemble models suggests that its architecture may not have been optimal for the dataset, warranting further tuning and experimentation to improve its efficacy. Another limitation involves feature engineering and interpretability. The study relies on predefined features such as magnitude, latitude, longitude, and depth without extensive exploration of additional or derived features that could enhance model performance. For instance, incorporating geospatial patterns, seismic wave propagation data, or temporal dynamics might yield better predictive outcomes. Moreover, while ensemble models like Gradient Boosting and Random Forest exhibit strong performance, their lack of interpretability makes it challenging to derive actionable insights for disaster management from their predictions, which could hinder real-world implementation.

The external validity of the study's findings is also a concern. The results are based on a specific dataset covering a defined time period, and the models' performance may not generalize well to other regions or datasets collected under different conditions. Factors such as changes in seismic monitoring technologies or geological variations could significantly impact the predictive capabilities of the models. This limitation underscores the

Journal of Computer System and Informatics (JoSYC)
ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

need for further testing on diverse datasets to evaluate the robustness of the models in various real-world scenarios. Computational complexity is another limitation, particularly for ensemble models such as Gradient Boosting and CatBoost, which require significant computational resources. This constraint could restrict their deployment in resource-constrained environments or applications requiring rapid, real-time predictions. The reliance on computationally intensive methods may limit the scalability of these models in practical settings, especially for large-scale disaster prediction systems. Lastly, ethical and practical concerns must be considered when applying these models to tsunami prediction. Predicting tsunami occurrences carries significant ethical implications, as false positives could lead to unnecessary evacuation efforts and false negatives could result in catastrophic consequences. Although the models achieve acceptable performance metrics, the possibility of failure in critical scenarios remains. Reliance on automated predictions without human oversight could pose risks to disaster management processes, highlighting the need for a balance between automation and expert evaluation.

4. CONCLUSION

This study demonstrates the potential of machine learning models, particularly ensemble methods and hybrid frameworks, to enhance tsunami prediction systems based on earthquake data. Among the evaluated models, Gradient Boosting emerged as the most effective standalone method, achieving the highest accuracy and ROC-AUC, followed closely by Random Forest and CatBoost, which also demonstrated strong performance in precision and recall. However, the hybrid model, which integrates the predictions of all individual models through a meta-classifier, outperformed them all. It achieved the highest overall accuracy (78.4%), ROC-AUC (0.8631), and F1-score (0.6635), highlighting its ability to leverage the strengths of diverse modeling approaches for robust and reliable predictions. The findings underscore the critical role of ensemble methods and hybrid frameworks in addressing the challenges posed by complex and imbalanced datasets. By combining the strengths of individual models, the hybrid approach minimizes the risk of overfitting, reduces the impact of false positives and false negatives, and enhances the generalizability of predictions. This makes it a particularly valuable tool for real-time tsunami prediction systems, where timely and accurate forecasts are essential for disaster mitigation and response. The study also highlights the limitations of individual models, such as the Neural Network's difficulty in handling imbalanced datasets, and emphasizes the need for continued research into hybrid techniques and additional features, such as geospatial or seismic propagation data, to further improve predictive performance. Practical implications include the potential for deploying these models in resource-constrained environments and integrating them into disaster preparedness frameworks to optimize resource allocation, evacuation planning, and long-term risk reduction strategies.

REFERENCES

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

[2] J. Rosselló, S. Becken, and M. Santana-Gallego, “The effects of natural disasters on international tourism: A global analysis,” *Tour. Manag.*, vol. 79, p. 104080, 2020.

[3] R. K. Chadha, “Living safely with earthquakes in Asia,” in *Extreme Natural Events: Sustainable Solutions for Developing Countries*, Springer, 2022, pp. 361–379.

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

[5] F. Qin, C. Huang, and Z. Lin, “Big data and artificial intelligence-driven natural disaster prediction and prevention: Technological advances and prospects,” *Geogr. Res. Bull.*, vol. 3, pp. 381–398, 2024.

[6] K. Dissanayaka, N. Tanaka, and T. L. C. Vinodh, “Integration of Eco-DRR and hybrid defense system on mitigation of natural disasters (Tsunami and Coastal Flooding): a review,” *Nat. hazards*, vol. 110, no. 1, pp. 1–28, 2022.

[7] J. Zhao *et al.*, “Deep-learning seismology,” *Rev. Geophys.*, vol. 58, no. 3, p. 103876, 2022.

[8] M. S. Abdalzaher, M. Krichen, D. Yiltas-Kaplan, I. Ben Dhaou, and W. Y. H. Adoni, “Early detection of earthquakes using iot and cloud infrastructure: A survey,” *Sustainability*, vol. 15, no. 15, p. 11713, 2023.

[9] M. S. Abdalzaher, M. Krichen, and F. Falcone, “Emerging technologies and supporting tools for earthquake disaster management: A perspective, challenges, and future directions,” *Prog. Disaster Sci.*, vol. 23, p. 100347, 2024.

[10] P. S. Chaudhary, M. R. Khurana, and M. Ayalasomayajula, “Real-World Applications of Data Analytics, Big Data, and Machine Learning,” in *Data Analytics and Machine Learning: Navigating the Big Data Landscape*, Springer, 2024, pp. 237–263.

[11] T. T. Khoei and A. Singh, “Data reduction in big data: a survey of methods, challenges and future directions,” *Int. J. Data Sci. Anal.*, pp. 1–40, 2024.

[12] A. Tursunalieva, D. L. J. Alexander, R. Dunne, J. Li, L. Riera, and Y. Zhao, “Making Sense of Machine Learning: A Review of Interpretation Techniques and Their Applications,” *Appl. Sci.*, vol. 14, no. 2, p. 496,

ISSN 2714-8912 (media online), ISSN 2714-7150 (media cetak)
Volume 6, No. 1, November 2024, Page 274-283
https://ejurnal.seminar-id.com/index.php/josyc
DOI 10.47065/josyc.v6i1.6291

2024.

[13] P. Kourehpaz and C. Molina Hutt, "Machine learning for enhanced regional seismic risk assessments," *J. Struct. Eng.*, vol. 148, no. 9, p. 4022126, 2022.

[14] A. K. F. Kouassi *et al.*, "Identification of karst cavities from 2D seismic wave impedance images based on Gradient-Boosting Decision Trees Algorithms (GBDT): Case of ordovician fracture-vuggy carbonate reservoir, Tahe Oilfield, Tarim Basin, China," *Energies*, vol. 16, no. 2, p. 643, 2023.

[15] E. Yavuz, M. C. Iban, and E. Arpaz, "Identifying the source types of the seismic events using discriminant functions and tree-based machine learning algorithms at Soma Region, Turkey," *Environ. Earth Sci.*, vol. 82, no. 11, p. 265, 2023.

[16] C. E. Yavas, L. Chen, C. Kadlec, and Y. Ji, "Predictive modeling of earthquakes in los angeles with machine learning and neural networks," *IEEE Access*, 2024.

[17] C. E. Yavas, L. Chen, C. Kadlec, and Y. Ji, "Improving earthquake prediction accuracy in Los Angeles with machine learning," *Sci. Rep.*, vol. 14, no. 1, p. 24440, 2024.

[18] L. Dong, H. Shu, Z. Tang, and X. Yan, "Microseismic event waveform classification using CNN-based transfer learning models," *Int. J. Min. Sci. Technol.*, vol. 33, no. 10, pp. 1203–1216, 2023.

[19] S. Yang, J. Hu, H. Zhang, and G. Liu, "Simultaneous earthquake detection on multiple stations via a convolutional neural network," *Seismol. Soc. Am.*, vol. 92, no. 1, pp. 246–260, 2021.

[20] 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.

[21] V. Özbey, S. Ergintav, and E. Tar\i, "GNSS Time Series Analysis with Machine Learning Algorithms: A Case Study for Anatolia," *Remote Sens.*, vol. 16, no. 17, p. 3309, 2024.

[22] S. Aymaz, "Boosting medical diagnostics with a novel gradient-based sample selection method," *Comput. Biol. Med.*, vol. 182, p. 109165, 2024.

[23] V. Linardos, M. Drakaki, P. Tzionas, and Y. L. Karnavas, "Machine learning in disaster management: recent developments in methods and applications," *Mach. Learn. Knowl. Extr.*, vol. 4, no. 2, 2022.

[24] C. Reis, M. Lopes, M. A. Baptista, and S. Clain, "Towards an integrated framework for the risk assessment of coastal structures exposed to earthquake and tsunami hazards," *Resilient Cities Struct.*, vol. 1, no. 2, pp. 57–75, 2022.

[25] Warcoder, "Earthquake Dataset," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/warcoder/earthquake-dataset. [Accessed: Nov. 2, 2024].