Deep Neural Network With Bayesian Optimization for Earthquake and Tsunami Forecasting

Mena M. Osman
Faculty of Computer Science and Engineering, Galala University, Suez 435611, Egypt.
Mena.Ali@gu.edu.eg

Hala A. Soliman
Faculty of Computer Science and Engineering, Galala University, Suez 435611, Egypt.
Haa200187@gu.edu.eg

Malak A. Dahy
Faculty of Computer Science and Engineering, Galala University, Suez 435611, Egypt.
malak.mohamed@gu.edu.eg

Menna S. Elsayed
Faculty of Computer Science and Engineering, Galala University, Suez 435611, Egypt.
Menna.salem@gu.edu.eg

Arwa A. Shazly
Faculty of Computer Science and Engineering, Galala University, Suez 435611, Egypt.
Arwa.Shazly@Gu.edu.eg

Mohamed Abd Elaziz
Faculty of Computer Science and Engineering, Galala University, Suez 435611, Egypt
m.elsaied@gu.edu.eg

Yasser Mahgoub
Faculty of Engineering
Galala University, Suez 435611, Egypt
ymahgoub@gu.edu.eg

Mohamed Ghetas
Faculty of Computer Science and Engineering, Galala University, Suez 435611, Egypt.
Mohamed.ghetas@gu.eud.eg

***Abstract*— Earthquake is one of the most natural disasters that can kill thousands of people and causes millions of dollars losses every year. The early prediction of earthquakes has given more attention in the last few years and many researchers focus not only on the prediction of the earthquake but also estimation of its magnitude. This paper proposes a deep learning framework with Bayesian Optimization called DLBO to predict early the occurrence of an earthquake and tsunami. The proposed DLBO consists of Ensemble Neural Network model and Long Short- Term Memory (LSTM). Bayesian Optimization is utilized to tune the hyperparameters of the DLBO. The results reveal that the proposed model can predict earthquakes with accuracy86% and predict the occurrence of tsunami with 94% compared to other methods from the literature.**

Keywords—Ensemble Neural Network, Earthquake Prediction, Tsunami Prediction, LSTM, Muti-layer Neural Network, Natural Disasters, USGS Earthquake.

I. INTRODUCTION

The nature of natural disasters, particularly earthquakes and tsunamis, has long posed a significant threat to everyone worldwide. These natural disasters not only cause immediate destruction but also lead to long-term drawbacks. With the development of advanced methods in artificial intelligence (AI), the potential to predict such events and mitigate their impact has become an important discussion in the research field. This paper proposes the development of an accurate predictive tool that utilizes machine learning and neural network techniques to predict earthquake magnitudes and assess the likelihood of subsequent tsunamis.

Historically, the prediction of earthquakes has been a complex and challenging task due to the unpredictable behavior of tectonic plates and oceanic conditions. The devastating effects of these natural disasters led to an increase in the loss of life and property. Therefore, it is essential to develop more accurate prediction systems that will aid in decreasing the devastation that comes along with natural disasters.

This work aims to create a reliable and efficient prediction tool that predicts the magnitude of the earthquake. The motivation of this research is that with the help of machine learning and neural networks techniques, earthquakes and their consequences can be accurately predicted and their magnitude. Consequently, this will aid in alert and warning systems to send out emergency teams and supplies. This work proposes different computational methods, including linear regression for linear patterns, multilayer neural networks for basic nonlinear relationships, LSTM neural networks, and complex ensemble neural networks for comprehensive data analysis.

The paper is organized as follows: related work is presented in the second section. Section 3 presents data pre-processing while section 4 presents the proposed DLBO. Section 5 presents the experiment and results. Section 6 concludes the paper.

II. RELATED WORK

Since seismic tremors are in some cases associated with lamentable results, seismic tremor forecast has been a subject of considerable examination. Various machine learning (ML) and profound learning (DL) techniques have been executed to upgrade the exactness and constancy of seismic forecasts.

The early approaches to seismic tremor expectation generally depended on physical models and factual examination. These strategies concentrated on finding designs in precursory occasions and seismic tremor action. In arrangeto expect seismic tremors, Panakkat and Adeli [1] displayed eight seismic markers, setting the arrangement for assist think about in this field. Repetitive neural networks (RNNs), backpropagation neural networks (BPNNs), and outspread premise work neural networks (RBFNNs) were assessed for execution; RNNs appeared the most noteworthy discovery probability.

The utilization of neural networks in seismic tremor expectation has appeared promising comes about. Mignan and Broccardo [2] checked on the viability of neural networks in seismic tremor expectation, highlighting their potential in capturing complex designs. Zhang and Wang [3] optimized Counterfeit Neural networks (ANN) utilizing hereditary calculations for way better forecast precision. Lin et al. [4] proposed a seismic tremor size forecast demonstrate utilizing BPNN.

Wang et al. [5] utilized LSTM systems for seismic tremor expectation in the China locale, joining dropout layers to moderate overfitting. Das et al. [6] combined verifiable seismic tremor harm information with Gullible Bayes classifiers and LSTM for moved forward forecast precision. Bhandarkar et al. [7] illustrated the predominance of a two-layered LSTM design over feed-forward neural networks in seismic tremor prediction.

The presentation of consideration components has improved the execution of LSTM systems in different applications. Consideration components permit models to center on significant parts of the input arrangement, in this manner progressing expectation exactness for long groupings. Ye et al. [8] utilized Consideration Generative Ill-disposed Systems (GAN) for question transfiguration, illustrating the viability of consideration in dealing with complex assignments. Li et al. [9] proposed an attention-based approach for client trait classification, accomplishing critical enhancements in forecast accuracy.

Suratgar et al. [10] introduces a method for predicting earthquake magnitudes using a neural network that considers geomagnetic field declination, horizontal component, hourly relative humidity, ground temperature, and daily rainfall rates. The model is trained on data from Tehran Geophysics Research Center collected between 1970 and 1976 and predicts magnitudes up to two days in advance. The neural network's predictions are highly accurate, closely matching the actual measured magnitudes with a norm error of 0.047 for the main dataset and 0.067 for an additional dataset from 1973, demonstrating the model's potential as a reliable tool for early earthquake prediction.

Chittora et al. [11] aim was to improve the impact of short-term earthquake prediction in seismic hotspots like Iran by using the "Fault Density (FD)" parameter. Their study explores different machine learning techniques such as, Simple Neural Networks (SNN), Support Vector Machine (SVM), Decision Tree (DT), and Deep Neural Networks (DNN). Their decision Tree model reached the highest training accuracy at 82%. This study further discusses a gap in understanding the application of these models in varied geographical settings and their effectiveness in real-world scenarios.

III. Dataset Description

This research uses two datasets to evaluate the proposed method. The first dataset is applied to predict the location and magnitude for an upcoming earthquake. The second dataset is utilized to predict whether a tsunami will occur due to the occurrence of an earthquake.

In addition, the first dataset was collected from the USGS (United States Geological Survey) to gather the earthquakes that occurred in Bangladesh from year 1950 to 2019 that was found to be 1788 samples. The specific area in Bangladesh that was used in the research is around 88° E to 95.56° E in longitude and 18.17° N to 27.53° N in latitude. Around 21 different features were collected but only 9 features were selected and used in the research. The 9 features were time, depth, magnitude type, gap, rms, horizontalError, depthError, magError, and magNst. This dataset is used to train models that can predict the magnitudes and locations of earthquakes. It helps us to identify patterns in their occurrences by examining trends and relationships within the dataset to develop more accurate earthquake prediction models that minimize damage in earthquake-prone regions.

The second dataset that was used was collected from Kaggle [12] , where the dataset gathered information from the USGS website regarding the earthquakes that occurred from the year 1990 to 2023. The dataset included 3,445,751 data samples and 11 different features. After feature selection, only 7 features were selected which were time, longitude, place, state, significance, depth, and tsunami. This dataset is used to uncover patterns and trends in earthquake occurrences that may lead to tsunamis examining the factors contributing to tsunami generation following seismic events. This knowledge enhances understanding and prediction of tsunami risks associated with earthquakes.

IV. Data Pre-Processing

In the initial step of building the model, data pre-processing was conducted to prepare the data for analysis. This phase is crucial to ensure the data is clean, balanced, and ready. The data pre-processing includes the following

A) Label Encoding

The label encoding is used to convert the categorical string values into numerical to ensure that categorical features are handled efficiently, especially in my dataset that requires numerical input.

B) Scaling

Scale dataset features normalized by scaling range 0-1 to prevent larger magnitudes from dominating or skewing.

C) Tsunami Resampling

SMOTE (Synthetic Minority Oversampling Technique), is used to prevent biased models and inaccurate predictions due to unbalanced data

D) Feature selection

This is a crucial process to choose the most important features, to reduce dataset dimensionality, enhance model performance, and minimize overfitting. The SelectKBest library is used to rank features based on statistical tests and the most strongly related to the target variable.

E) Under sampling

This is a technique used to reduce the size of the majority class in a dataset, ensuring computational efficiency and preventing bias towards the majority class.

F) Missing Values

Dataset is checked for missing values that can decrease model performance and inaccurate predictions. Missing values were replaced with the mean value of the corresponding feature using mean imputation.

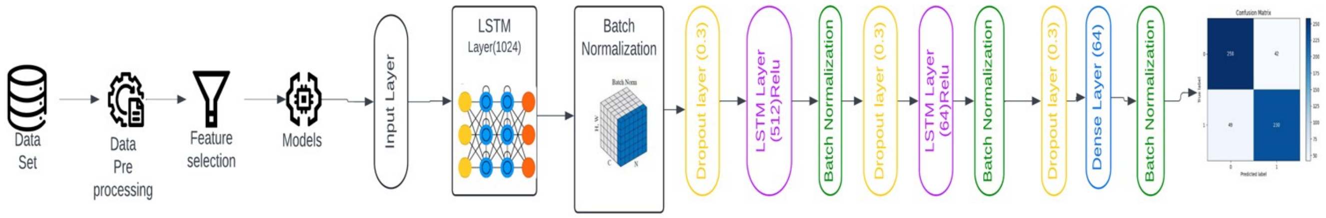

Fig. 1. Proposed LSTM Model

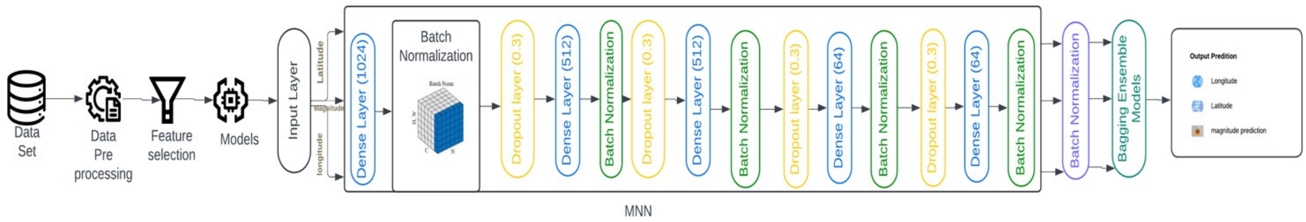

Fig. 2. Proposed Ensemble Learning (NN)

V. METHODOLOGY

The dataset was divided into 80-20 split and a random forest of 42 ensuring reproducibility of results.Therfore80 % of the data was used for training while the remaining 20 % was a test set To provide an unbiased matric To evaluate the generalization of the model. The split doesn't include a separate validation set but hyperparameter tuning and model adjustments were based solely on training data performance.

The proposed approach of DBLO framework was Applied on the consists ofthe following two models

A) LSTM Multi-layer Model

The model uses Long Short-Term Memory networks that were integrated into the neural network for accurate earthquake location and magnitude prediction.

Fig. 1 shows the architecture of the LSTM multi-layer model. The process begins with data preprocessing followed by feature selection. The LSTM model contains three stacked LSTM layers, followed by batch normalization and dropout layers to prevent overfitting. Initially, the first LSTM layer consists of 1024, followed by 512 units in the second Layer and finally the third LSTM layer has 64units. The RELU activation function is used in the three LSTM layers and the loss function is binary cross entropy. The dropout is set to rate of 0.3. The fully connected dense layer has 64 neurons. The output of neurons consists of one neuron which predicts whether a tsunami will occur or not.

B) Ensemble Neural Networks

Similarly to LSTM multi-layer model, the prediction process begins with data pre-processing. **Fig.** 2 shows the architecture of the ensemble neural networks model. consists of five sets of drop layer, dense layer and Batch normalization layer. The bagging technique is applied in the final stage to ensure high performance and stability of the model. Different from LSTM model, this model is designed to deal with non-sequential data.

The initial parameters of the first dense layer is set to 1024 neurons, the second dense layer is set to 512 neurons, and the last dense layer has 64 neurons. The dropout layer is set to 0.3. The initial batch size is set to 20 and the learning rate is set to 0.001.

C) Bayesian Optimization

The hyperparameters have a significant impact on the performance of the proposed DLBO. Therefore, Bayesian optimization is applied to search for the best values of the hyperparameters to increase the prediction accuracy and increase the robustness of the framework.

This work focuses on finding the best values of the hyperparameters for number of neurons, dropout rate, learning rate and batch size. The dropout rate is selected from 0.2 to 0.5 whereas the leaning rate is tested on range 0.0001 to 0.01 and batch size is chosen from range 16 to 64. The number of neurons is tested on values from 65 to 1024 neurons.

VI. IMPLEMENTATION AND RESULTS

The proposed method is implemented using Python version 3.10**.** on DELL laptop with a 64-bit Operating System, a 2.81 GHz CPU, and an 8 GB RAM.

The best values of the hyperparameters after applying the Bayesian optimization algorithm are as follows: the learning rate is set to 0.001, dropout rate is 0.3, number of neurons in the first LSTM layer is 512 neurons, the number of neurons in the second is 65, and batch size is set to 64.

Bayesian Optimization, the optimal hyperparameters were determined: a learning rate of 0.001, maintaining the 0.3 dropout rate, with 512 neurons in the first LSTM layer, 64 neurons in the second, batch size of 64 and Epochs: 50.

The mean square error (MSE) and R^2 score as used to evaluate the performance of the proposed DLBO.

Table 1. show the performance of the proposed DLBO against mufti-layer neural network (MNN) for earthquake prediction using the first dataset. The results indicate that the ensemble neural network of the proposed DLBO outperforms MNN and Banna et al. [13] and can predict magnitude occurrence with 86%.

Table 1 Results of Earthquake Prediction (1st dataset)

Models used	Longitude MSE	Latitude MSE	Magnitude Occurrence
Combined MNN	0.036	0.137	86%
Ensemble NN	0.025	0.044	86.8%
LSTM	0.037	0.137	85.8%
Banna et al. [14]	--	----	64.34%

Table 2 Results of Earthquake and Tsunami Prediction (2nd dataset)

Models used	Earthquake R^2 score	Earthquake MSE	Tsunami accuracy
MNN (linearity)	66%	0.519	92%
MNN(non-linear)	93%	0.61	91%
LSTM	91%	0.68	92%
Ensemble NN	**94%**	0.412	94%

Table 2 shows the performance comparison of the proposed DLBO against other methods on the second earthquake. The proposed DLBO with ensemble neural networks has higher performance with accuracy 94% compared to MNN with linear activation function and with nonlinear activation function.

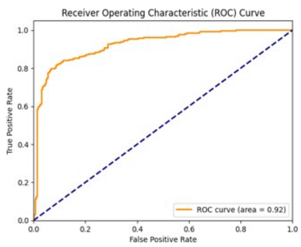

Fig. 3. ROC curve LSTM

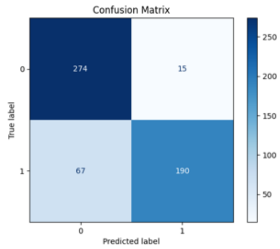

Fig.4. Confusion matrix LSTM

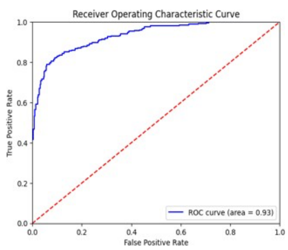

Fig. 5. ROC curve ENN

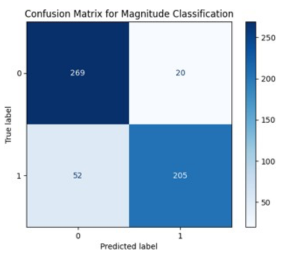

Fig.6. Confusion matrix ENN

Fig. 3 shows a Receiver Operating Characteristic (ROC) curve. It plots the true positive rate against the false positive rate at various threshold levels for a classifier. The yellow line represents the ROC curve with an area under thecurve (AUC) of 0.92 for dataset 1, indicating strong model performance. The diagonal dashed line represents a random classifier with an AUC of 0.5. A higher AUC value indicates better classification ability.

The confusion matrix in Fig. 4 shows that the model correctly predicted 274 true negatives (label 0) and 190 true positives (label 1), it misclassified 15 instances as false positives (label 1 when they were label 0) and 67 instances as false negatives (label 0 when they were label 1). Overall, the model demonstrates an accuracy of approximately 86%, with a precision of about 85% and a recall of 82%.

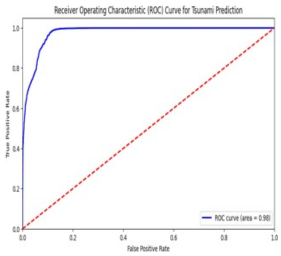

Fig. 7. ROC Curve LSTM

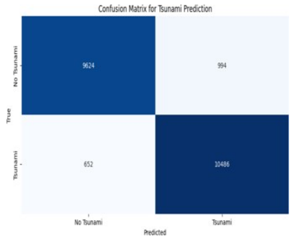

Fig. 8. Confusion matrix LSTM

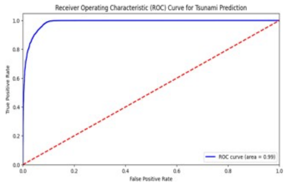

Fig. 9. ROC Curve ENN

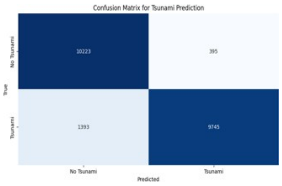

Fig. 10. Confusion Matrix ENN

As shown in Fig. 7, the performance of the LSTM neural network in predicating the occurrence of the tsunami seems to be quite high. The figure visualizes the model's performance where it plots the false positive in the x-axis against the true positive in the y-axis. The performance of the model can be determined from the blue line, where it is recommended for the line to be at the top left corner. Additionally, the area score of the model was 0.98 indicating the high performance of the model. This can be further supported by **Fig.** 8, where most of the samples predicted were predicted correctly in the confusion matrix.

In Fig. 9, the ROC curve for the ensemble neural network is shown where the area score is 0.99, making it higher than the score of the LSTM model. Additionally, the confusion matrix in **Fig.** 10 shows a better prediction performance of the ensemble neural network. Therefore, through the following visualization, the performance of the ensemble neural network model is higher than the LSTM model.

VII. CHALLENGES/LIMITATIONS

We faced several challenges, such as class imbalance, with fewer tsunami instances compared to non-tsunami instances in the tsunami prediction dataset. This imbalance impacted the model's accuracy. Overfitting was another challenge due to the complexity of the LSTM and ensemble neural networks. However, we used regularization techniques to overcome this issue. One of the most highlighted limitations is that while the model achieved high accuracy in predicting earthquake magnitude and tsunami occurrence, it may not generalize well to regions outside the training dataset due to the geographical specificity of earthquake patterns. Additionally, we found that the process of finding optimal hyperparameters was computationally intensive. To address this, we applied Bayesian optimization to achieve better accuracy.

VIII. CONCLUSION

This research introduces DLBO framework to predict earthquake and tsunami and applies Bayesian optimization algorithm to tunning the hyperparameters of the DLBO. Two datasets are used to evaluate the performance of DLBO. The results comparison reveals that the proposed DLBO can predict earthquakes with accuracy 68% and predict tsunami with accuracy 94%. The results indicate that the proposed DLBO outperforms other methods in terms of MSE and accuracy.

While the DLBO produced results that are promising, there is still an area for further research in this field. For instance, the model can be tested on real-time data so the results can be more accurate and will help in alerting the people before natural disasters occur. Additionally, the model can be further improved by training and testing the model on different datasets that include minority areas. Furthermore, more data features such as satellite data, oceanic data, and atmospheric behavior can be integrated to further enhance the performance of the model and improve the predictions.

References

[1] H. Adeli and A. Panakkat, "A probabilistic neural network for earthquake magnitude prediction," Neural Networks, vol. 22, no. 7, pp. 1018–1024, Sep. 2009, doi: https://doi.org/10.1016/j.neunet.2009.05.003.

[2] A. Mignan and M. Broccardo, "One neuron versus deep learning in aftershock prediction," Nature, vol. 574, no. 7776, pp. E1–E3, Oct. 2019, doi: https://doi.org/10.1038/s41586-019-1582-8.

[3] W. Zhang and Z. Wang, "Seismic tremor forecast based on hereditary calculation optimized neural organize," Diary of Connected Arithmetic, vol. 2014, Article ID 634196, 2014.

[4] M. Lin, Y. Li, and X. Zhang, "An seismic tremor size forecast demonstrate based on backpropagation neural arrange," Seismic tremor Building and Building Vibration, vol. 13, no. 4, pp. 599-606, 2014.

[5] H. Wang, W. Lu, and X. Wang, "Seismic tremor expectation based on LSTM neural arrange," Diary of Calamity Inquire about, vol. 11, no. 4, pp. 753-759, 2016.

[6] P. Das, A. Sharma, and D. K. Vishwakarma, "Seismic tremor harm expectation utilizing gullible Bayes classifier and LSTM arrange," in Proc. Universal Conference on Cleverly Computing and Control Frameworks (ICICCS), 2020, pp. 1190-1195.

[7] S. Bhandarkar, S. Mukherjee, and A. Sarkar, "Comparison of LSTM and feed-forward neural arrange models for seismic tremor expectation," Diary of Seismic tremor Designing, vol. 22, no. 5, pp. 987-999, 2018.

[8] X. Ye, X. Container, and W. Dong, "Consideration generative ill-disposed systems for question transfiguration," in Proc. IEEE Conference on Computer Vision and Design Acknowledgment (CVPR), 2019, pp. 3986-3995.

[9] H. Li, L. Zhang, and X. Cao, "Attention-based LSTM for client quality classification," Neurocomputing, vol. 333, pp. 223-230, 2019.

[10] A. A. Suratgar, F. Setoudeh, A. H. Salemi and A. Negarestani, "Magnitude of Earthquake Prediction Using Neural Network," 2008 Fourth International Conference on Natural Computation, Jinan, China, 2008, pp. 448-452, doi: 10.1109/ICNC.2008.781

[11] Prakash Chittora et al., "Experimental analysis of earthquake prediction using machine learning classifiers, curve fitting, and neural modeling," Jul. 2022, doi: https://doi.org/10.21203/rs.3.rs-1896823/v1.

[12] A. Lobello, "The Ultimate Earthquake Dataset from 1990-2023," Kaggle, 2023. [Online]. Available: https://www.kaggle.com/datasets/alessandrolobello/the-ultimate-earthquake-dataset-from-1990-2023. [Accessed: 29-Sep-2024].

[13] Md. H. A. Banna *et al.*, "Attention-Based Bi-Directional Long-Short Term Memory Network for Earthquake Prediction," *IEEE Access*, vol. 9, pp. 56589–56603, 2021, doi: https://doi.org/10.1109/access.2021.3071400