Sistemasi: Jurnal Sistem Informasi
Volume 13, Nomor 5, 2024: 1914-1925
ISSN:2302-8149
e-ISSN:2540-9719

# Prediction of Tsunamis in Indonesia Using an Optimized Neural Network with SMOTE

[1]Aisyah Anjani Putri Siregar*, [2]Fauzan Bayu Hera Sudianto, [3]Rani Nooraeni
[1,2,3]Program Studi DIV Statistika, Politeknik Statistika STIS
[1,2,3]Jl. Otto Iskandardinata No.64C, Jakarta, Indonesia 13330
*e-mail: 212111868@stis.ac.id

(***received***: 13 June 2024, ***revised***: 20 August 2024, ***accepted***: 21 September 2024)

**Abstract**

Tsunamis have the potential to have a large impact on the environment, therefore early detection and preparation for tsunami need to be carried out to reduce the impact of casualties and losses incurred. This research aims to predict tsunami events due to large earthquakes in Indonesia as a form of early detection. The optimized neural network method is used in research to classify tsunami events in Indonesia in 2000-2023 for large earthquakes with strength more than 5 magnitudes. The research results show that the neural network structure formed consists of an input layer, a hidden layer, and an output layer. The results of the evaluation of the neural network model with SMOTE obtained an accuracy value of 99.43%, precision of 96.31%, and an F1 score of 97.86%, which means the resulting model is good. Therefore, an optimized neural networks can be applied as a warning system in various regions to detect potential tsunami events in the future.

**Keywords:** earthquake, neural network, smote, tsunami

## 1 Introduction

Natural disasters are natural events that occur due to natural processes that can harm everyone [1]. Natural disasters can change the structure of nature and the environment that have a direct impact on human life [2]. Tsunami is a form of natural disaster that causes many losses. Tsunamis have the potential to wreak havoc when they hit an area and can destroy various assets [3]. The tsunami has great potential to have a major impact on the environment, especially population and community structures and infrastructure in coastal areas [4].

According to Imamura (1937), the term tsunami comes from a Japanese words tu (meaning port) and nami (long wave) [5]. A tsunami is a series of large ocean waves caused by a large and sudden disturbance on the seafloor [6], which then causes changes in sea level. Strike-slip cracks on the seafloor lift the seafloor surface close to horizontal then produce vertical deformations that eventually form waves [7]. Some of the causes of tsunamis include earthquakes, volcanic eruptions, landslides on the seabed, and meteor strikes [8]. Although wave heights are relatively small on the high seas, when tsunamis hit, the waves can reach several meters and can cause loss of life and property damage if they occur on the coast [9]. On the coast, tsunami waves can reach a height of 20 m and spread as far as 5 km depending on the topography and energy of the wave.

Indonesia is geologically and geographically located at the confluence of four major plates, namely Eurasia, Indo Australia, the Philippines, and the Pacific, has a high risk of disasters, especially earthquakes, tsunamis, and volcanic eruptions. The calculation of the 2022 Disaster Risk Index shows that 13 provinces are in the high disaster risk class and 21 provinces are in the medium disaster risk class [10]. In addition, as an archipelagic country with the second longest coastline in the world, many areas in Indonesia surrounded by the sea and more vulnerable to tsunamis, almost along the coast of Indonesia, from the islands of Sumatra, Java, Bali, Nusa Tenggara, Kalimantan, Sulawesi, Maluku and Papua, are not immune from the potential for this natural phenomenon to occur [11].

Indonesian Tsunami Catalog stated several major tsunami events that have occurred include the Palu Dongala tsunami in 2018, which claimed 4340 lives and recorded a total loss of 20.89 trillion rupiah and the Aceh tsunami in 2004 which caused more than 100,000 deaths impacting other countries such as Myanmar, India, Thailand, Sri Lanka, Madagascar, etc [11]. Meanwhile, the tsunami that occurred due to the volcanic earthquake was a tsunami on the coast of Banten and

ISSN:2302-8149
e-ISSN:2540-9719

Lampung with waves reaching a height of 13 m and resulting in 437 casualties, 14,059 injuries, and 33,719 people homeless.

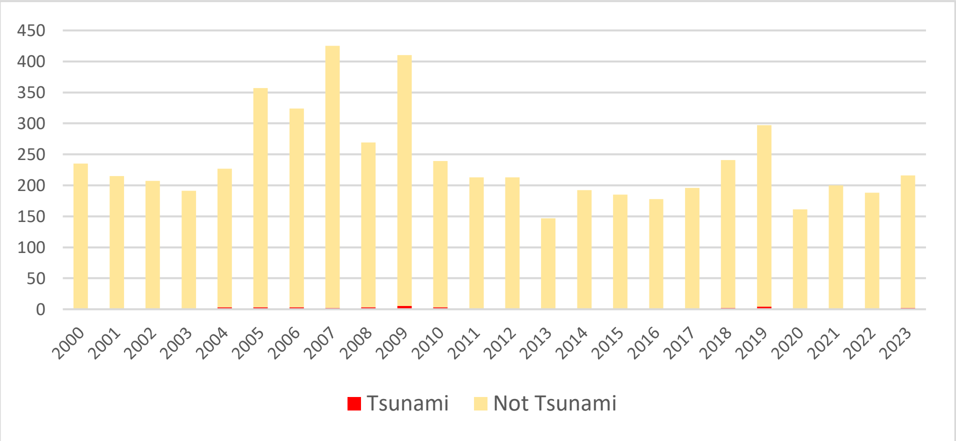


**Figure 1. Indonesia's total large earthquake and tsunami activity 2000-2023**

Based on USGS data, as shown in Figure 1, the consequences of Indonesia's location and geographical position, there are more than 100 major and destructive earthquakes with > 5 mag every year in Indonesia. Meanwhile, in 2022, BMKG recorded 10,792 earthquakes that occurred in Indonesia with 205 earthquakes with a magnitude of more than 5 as a large earthquakes. Where, since 2013, there has been an increase in the trend of earthquake activity in Indonesia [12]. The increase in the frequency of earthquakes raises the potential for new disasters in the form of tsunamis. Tsunamis are secondary hazards posed by megathrust earthquakes which are generally associated with coseismic slip that occurs at shallow depths [13]. Among the various factors that cause tsunamis, most tsunamis occur due to earthquakes [14]. Tsunamis are rare natural disasters that usually occur at low frequency [15]. From the graph on figure 1 it can also be seen that the proportion of tsunamis that occur due to earthquakes is very small, even in certain years, there are no tsunami events caused by earthquakes. However, the high number of earthquakes that occur in Indonesia, especially earthquakes that are large and significant, is a concern for the potential impact of tsunami deaths that occur in Indonesia in the future.

The damage caused by the tsunami in Indonesia can certainly disrupt community activities. Indonesia's territory, which is in the Indonesian Ring of Fire, is prone to earthquakes which can lead to tsunamis. This indicates the need for handling and prevention efforts to minimize the impact of the tsunami disaster from the Government. Early detection and warning of a tsunami is an effective step as a form of preparation in dealing with it. Especially now with advanced devices and technology, early detection and preparation in facing tsunamis will be much better [16]. One form of early detection is by predicting or forecasting tsunamis, so that it can reduce the impact of casualties and losses incurred [17].

There are several conditions that can be used to detect the occurrence of a tsunami. If tsunami is caused by an earthquake, the strength of the earthquake can be a characteristic of whether a potential tsunami will occur or not. Goda (2015) stated that before a tsunami is actually detected, warnings can be given based on earthquake information, for example, the magnitude and location of the hypocenter [18]. Tsunami warnings rely primarily on real-time seismic stations to monitor underwater earthquake activity with earthquake parameters such as time of origin, epicenter, magnitude, focal depth [19].

Based on the background and problem identification, this study was conducted with the aim of modeling tsunami events in Indonesia due to earthquakes with the optimized neural network method as a form of prediction based on variables from historical earthquake characteristics and tsunami event data. In this study, a comparison of modeling results will also be carried out to see the effectiveness of using the resampling method with SMOTE in modeling tsunami events in Indonesia. The use of latitude and longitude as variables in this study is to show the location and distribution of earthquakes that have the potential to cause tsunamis in Indonesia. This study also wants to know the accuracy of using this method in predicting tsunami events in Indonesia.

ISSN:2302-8149
e-ISSN:2540-9719

## 2 Literature Review

Various studies in detecting tsunamis have been conducted, but there are some differences that cause such analyses to continue to be pursued. Konovalov tried to classify earthquakes that caused tsunamis in the northern Pacific Ocean in 1960-2020 using logistic regression with variables of earthquake's depth and magnitude. The results showed that the logit model had a false discovery rate lower than the threshold magnitude criterion used by tsunami warning agencies with an accuracy of about 71%, which shows optimism about tsunami predictions that are quite accurate [20]. In the Indonesian case study, Irfiani (2022) used decision trees and C4.5 algorithms to classify potential tsunamis due to earthquakes. The results showed that the attributes of earthquake strength and location were the highest attributes in the decision tree. Then the application of the C4.5 algorithm to classify the strength of the earthquake and the location of the earthquake showed a very high accuracy value of 99.96% [21]. Then, Amalia and Fuji (2023) also predicts the potential and height of tsunamis in Indonesia with the help of Machine Learning using the Artificial Neural Network (ANN) method. The parameters used are the strength, intensity and location of the earthquake event. The results of this study succeeded in predicting the potential for tsunamis in a country up to 81% with accuracy of 77.45% [22].

Data mining has an important role in the decision-making process as well as to get better long-term results [23]. One method of data mining that can be used to predict is the Neural Network. Although neural networks have disadvantages such as vague algorithmic processes, errors are difficult to detect, requiring high processing power and unpredictable computational time [24]. Neural networks can to work on complex problems with incomplete knowledge [25], have the power to do more than one job at once [26], wherein, data processing is carried out in parallel, so that failures in one nerve element do not affect other elements. In addition, information is stored in the network, so the loss of some information in one place does not hinder the functioning of the network [26]. Regarding these advantages, [27] states that the use of the Neural network method is very suitable for big data because this method is strong against data that contains a lot of noise and can handle uninformative data, which cannot be done by other normal programs [24]

The use of the resampling method is a commonly used approach to overcome data imbalances. The resampling method is used to handle the distribution of target variables/classes. This is because if the data is unbalanced, classification will not work well so re-sampling needs to be done [28]. The imbalance that occurs in the tsunami data will affect and damage modeling, so to overcome this, this study uses the SMOTE resampling method. This is motivated by SMOTE's ability to form new data based on learning patterns from minority data. Many studies have shown that SMOTE's approach to oversampling can improve the accuracy of classifying minority classes [29].

## 3 Research Method

To support this research, data on earthquakes from various sources are used to ensure the accuracy and correctness of the data obtained. Meanwhile, the following is a further explanation of the methods and data used in the research.

### 3.1. Data

In this research we used earthquake catalogs according to their characteristics compiled by the United States Geological System (USGS) as research variables as well as historical data on tsunami events in Indonesia obtained through publications by the Meteorology, Climatology and Geophysics Agency (BMKG) as a label. The observations in this study were 5691 large earthquake events in Indonesia with a magnitude >5.0 in 2000-2023 which were then classified based on tsunami events based of historical data on tsunami events according to the BMKG. It is assumed that there were no tsunami events that were not recorded in BMKG. Using data on earthquakes with magnitudes greater than this indicates large earthquakes that can cause severe damage [12]. The characteristics of earthquakes used in research are spatial characteristics (Longitude and latitude), earthquake magnitude, and earthquake depth. The reason for using the variables such as magnitude, depth, latitude and longitude is based on previous research who said that among the various factors that

 ISSN:2302-8149
e-ISSN:2540-9719

cause tsunamis, most tsunamis occur due to large earthquakes in very shallow areas [14]. The entire definition of characteristic variables in the data is explained in the following Table 1.

**Table 1. Earthquake characteristic definition**

| Characteristic | Definition |
|---|---|
| Latitude | Coordinates of geographic latitude where the earthquake occurred (°), South latitude ((-) or (S)) North latitude (without signs or (N)) |
| Longitude | Geographic longitude coordinates of the place where the earthquake occurred (°) |
| Magnitudo | Magnitude, the strength of an earthquake on the Richter scale |
| Depth | Earthquake depth (Km) |
| Type | Tsunamic Event (Tsunami and Non Tsunami) |

### 3.2. Preprocessing Data

Tsunamis in Indonesia are assumed to be rare events that will cause an imbalance in the proportion of categorical tsunami events (earthquakes that cause tsunami compared to earthquakes that do not cause a tsunami). Therefore, a resampling method is needed as a strategy to overcome imbalances in the data. The use of resampling is aimed at creating a new dataset by changing several samples from the one class to the other. One resampling method that is often used is SMOTE.

SMOTE (Synthetic Minority Oversampling Technique) is a resampling method that is suitable for application in various problems, one of which is over sampling. SMOTE will create a new sample based on observations from the minority class and several of its closest neighbors by interpolation. The new sample calculation formula using the SMOTE method shown in equation (1) [30].

$$O_{new} = O_i + \left(O_{neighbor} - O_i\right) * \alpha \tag{1}$$

Where $O_{new}$ is a newly formed sample, $O_i$ is a sample taken randomly from the minority class, $O_{neighbor}$ is a sample selected based on the closest neighbor with Euclidean distance, and α is a random number in the range 0 to 1.

Then the data was randomly divided into 2 subsets, namely 75% as a training model of tsunami events in Indonesia due to large earthquakes based on the earthquake characteristics and 25% as a testing model for model evaluation.

### 3.3. Neural Network

Artificial neural networks are a method in data mining that can be said to be a powerful data modeling tool because it can capture and represent complex relationships. Neural networks are known to have the ability to model linear and non-linear relationships from available data. [31] describes the Neural Network method as a nonlinear modeling tool mimicking the function of biological neurons with the aim of maximizing network prediction performance by minimizing differences between all network outputs. Just as the human brain has interconnected neurons, neural networks also have interconnected neurons (which called nodes) in various layers of the network [25]. Neural networks allow the use of computational operations to solve complex issues and best to nonlinear resource problems [32].

In general, neural networks have advanced propagation flows called feedforward in calculating prediction values. The function of feedforward propagation is to produce output, so information is needed to flow forward through the network [33]. The input will spread to hidden units in each layer which then produces output. Therefore, an artificial neural network is a network consisting of 3 types of layers, namely input layers as the number of inputs entered, output layers, and hidden layers and neurons that will be determined by trial and error. Feedforward propagation in neural networks is described as a universal approximators network [34]. The following Figure 2 is the illustration of the architecture of the neural network.

 ISSN:2302-8149
e-ISSN:2540-9719

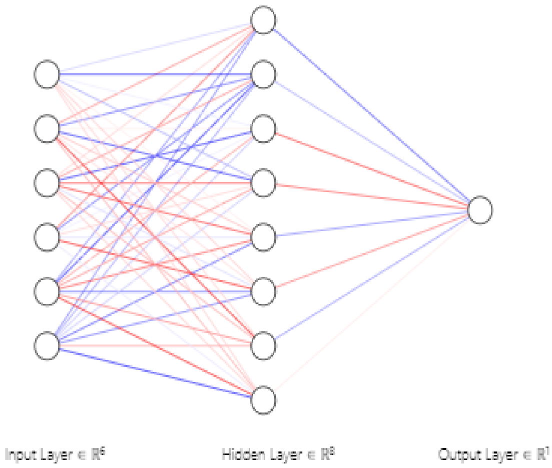


**Figure 2. Architecture neural network**

The input layer will obtain the input value from earthquakes event. When the input layer receives input, the neurons generate output, which serves as input to the network's next layer. Hidden layers handle all additional iterative computations and weight modifications required to create the network, as well as post-design calculations. The weights in the incoming and outgoing hidden layers are modified in accordance with the input and output values. All of the nodes in the hidden layer are fully connected to one another. This means that hidden nodes learn by assessing all input values. The output layer is the layer in which all nodes convey the result value to the user in the form of tsunami and non-tsunami events.

In the process of classifying a neural network into a tsunami event or not, an activation function can be carried out which can determine the results in a neural network based on the input value. In general, the value resulting from an activation function can be formulated as equation (2) [35].

$$y = f\left(\sum_{i=1}^{n} x_i w_{ij} + b_j\right) \quad (2)$$

Where $x_i$ is the input of the neuron, $w_{ij}$ is the weight between the input and hidden layer, and $b_j$ is the polarized value of the neurons that form the hidden layer. Meanwhile y is the output produced for one layer and f is the activation function applied.

There are several types of activation functions, one of which is sigmoid which is commonly used in neural networks. The sigmoid function is suitable for use in binary classification because it can produce an output with a range value of 0 to 1 as the probability of the target class [36]. This activation function provides non-linearity to the model and allows the model to learn more complex relationships between input features as shown on equation (3).

$$y = f(x) = \frac{1}{1+e^{-x}} \quad (3)$$

In the process of evaluating the results of tsunami event predictions using neural networks, a number of evaluation criteria can be used, one of which is the confusion matrix. Confusion Matrix is a table of the number of correct and incorrect predictions for each class which is useful for analyzing how well the classification was carried out [37]. Based on the components in the confusion matrix, several other commonly used evaluation criteria can also be calculated, including the following [38].

Accuracy is a value that shows the proportion of total correct predictions from the total prediction results. The accuracy value of a prediction can be calculated using the following formula as shown on equation (4).

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (4)$$

 ISSN:2302-8149
e-ISSN:2540-9719

Precision is a measure of the accuracy of positive predictions by determining the ratio of true positives compared to all predictions that give positive results (both true positives and false positives) shown on equation (5).

$$Precision = \frac{TP}{TP+FP} \quad (5)$$

Recall also calculates the proportion of positive values that are correctly predicted but compared to observations that are actually positive shown on equation (6).

$$Recall = \frac{TP}{TP+FN} \quad (6)$$

To get the harmonic average between precision and recall, the F1-Score value can be calculated which can be used to consider class imbalance in the data shown on equation (7).

$$F1 - Score = \frac{2*Precission*Recall}{(Precision+Recall)} \quad (7)$$

Sensitivity in evaluating prediction results shows how well the model can identify positive classes correctly, so the sensitivity value can be calculated as shown on equation (8):

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

Specificity is the proportion of correct negative predictions to the total actual negative value so that it can identify the negative class correctly. The following is the formula for calculating the specificity value shown on equation (9).

$$Specificity = \frac{TN}{TN+FP} \quad (9)$$

Where TP is True Positive (the number of correct tsunami predictions), TN is True Negative (the number of correct non-stunami predictions), FP is False Positive (wrong tsunami predictions), and FN is False Negative (wrong non-stunami prediction).

## 4 Results and Analysis

The main focus in this research is on tsunami events that occur as a result of earthquakes. This earthquake was a large destructive earthquake with a magnitude of more than 5 (sizeable destructive earthquake) with variables Focal Depth, magnitude, and earthquake location (latitude and longitude) to predict tsunami potential.

### 4. 1. Characteristic of Large Earthquake

The first step that must be taken in the data preprocessing stage is to look at the variable data structure that has been determined previously. The following are descriptive statistics of the data used in this research.

**Table 2. Summary statistics of large earthquake characteristic variables**

| Summary Statistic | Depth | Mag |
|---|---|---|
| Min | 0,8 | 5 |
| Q1 | 20,9 | 5,1 |
| Q2 | 35 | 5,2 |
| Q3 | 63,2 | 5,5 |
| Max | 663,6 | 9,1 |
| Mean | 63,48 | 5,33 |

ISSN:2302-8149
e-ISSN:2540-9719

Based on Table 2, the average large earthquake that occurred in Indonesia during the last 20 years occurred at a depth of 63.48 KM with an average earthquake magnitude of 5.33. However, it is also known that large earthquakes in Indonesia in 2000-2023 reached a magnitude of 9.

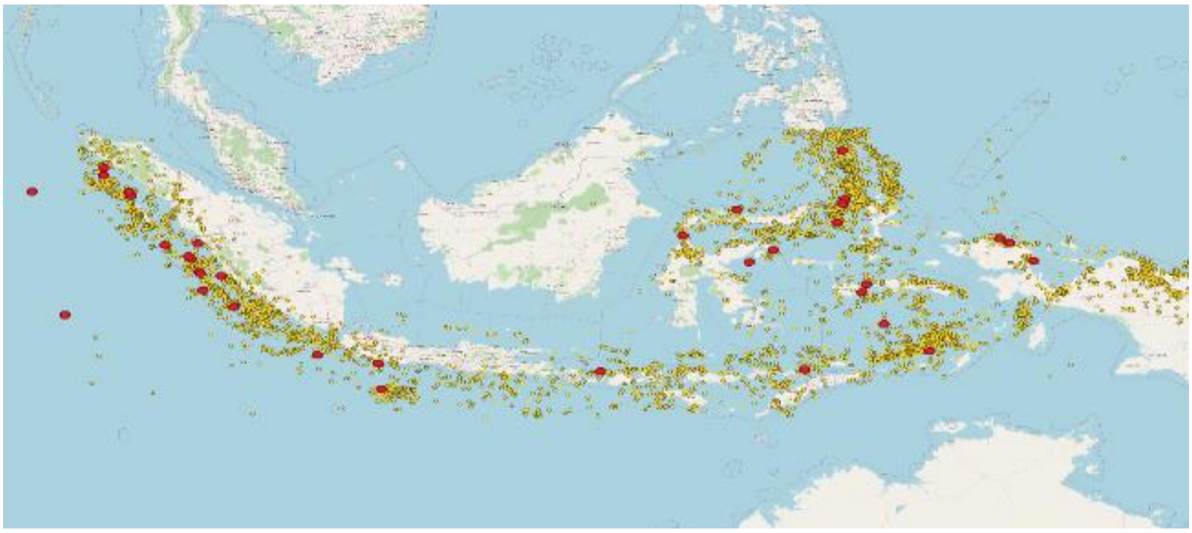

**Figure 3. Tsunami cartilage earthquake spread map**

Then, based on latitude and longitude that shown on Figure 3, can be seen that at the distribution of red dots on Figure 3, the areas with the highest potential for tsunami vulnerability caused by large earthquakes are the West Coast of Sumatra and in the north of Maluku, this is because the two areas are around the subduction zone, because they are located at the meeting of the world's large plates, namely the Eurasian plate, Pacific plate and Australian plate. This causes the intensity of earthquakes in the region to be higher and has a high chance of causing a tsunami [39].

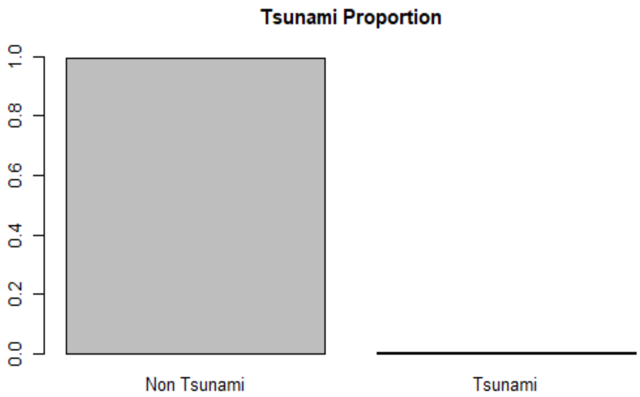


**Figure 4. Tsunamic events proportion**

Then an exploration of the tsunami event in Indonesia was also carried out as in Figure 4. Based on historical data on tsunamis that occurred from 2000 to 2023, only a small portion of tsunamis were caused by large earthquakes which totals 35 observations from 5691 large earthquakes in Indonesia so that causes the proportion of target variables to be unbalanced.

To overcome this imbalance, data resampling is carried out using SMOTE to form new data with a proportion of 735 data on earthquakes that cause tsunamis and 4900 data on earthquakes that do not cause tsunamis. Even though there is still a large difference, the increasing number of tsunami events can still be used to represent the situation of tsunamis in Indonesia without changing the actual conditions much. Then the new data is divided into training data and testing data.

ISSN:2302-8149
e-ISSN:2540-9719

### 4. 2. Tsunami Prediction Modelling With Neural Network

From the training data that has gone through the SMOTE method, a neural network model is produced. A trial and error process was carried out to determine the number of layers and neurons in the hidden part as shown on Table 3.

**Table 3. Comparison of model accuracy with and without SMOTE**

| **Number of Neurons in Hidden Layer** | | **Without SMOTE** | | **With SMOTE** | |
|---|---|---|---|---|---|
| **1st Layer** | **2nd Layer** | **Accuracy** | **Recall** | **Accuracy** | **Recall** |
| 4 | - | 99,371 | 0 | 98,651 | 96,739 |
| 5 | - | 99,371 | 0 | 99,290 | 98,369 |
| 6 | - | 99,371 | 0 | 99,369 | 98,369 |
| **4** | **4** | **99,232** | **0** | **99,432** | **99,456** |
| 5 | 4 | 99,371 | 0 | 99,432 | 98,369 |
| 6 | 4 | 99,022 | 22,222 | 99,503 | 98,913 |
| 6 | 5 | 99,371 | 0 | 98,935 | 97,286 |

Based on Table 3, it can be seen that the hidden layer selected is the hidden layer that has the highest accuracy value, so that the hidden layer used is 2 layer with 4 neurons in each layer. If we look at the accuracy of the model without and using the SMOTE method, it can be seen that without the SMOTE method there is a neural network model with high accuracy but has a recall value of 0. This shows that the model only predicts tsunami events in one category. On the other hand, the neural network model with SMOTE is proven to have high accuracy values along with quite high recall. Therefore, the use of the SMOTE resampling method is proven to be able to improve model evaluation so that predictions of tsunami events can show better results. The following shows the best structure of the optimized neural network model with SMOTE.

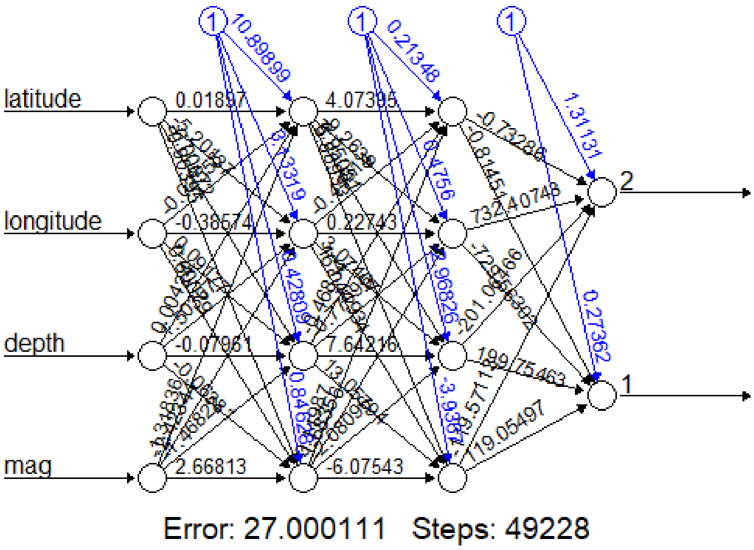


**Figure 8. Optimized neural network structure**

As shown on Figure 8, the best neural network model that is formed consists of four layers, namely the input layer consisting of 4 neurons (latitude, longitude, depth, and mag), one hidden layer consisting of 2 layers with 4 neurons each, and the output layer consisting of 2 neurons (1: tsunami, and 2: no tsunami). Based on the neural network plot, can be seen that each neuron is connected to each other and each connection has a weight (black arrow) with a different value. Where the weight value shows the magnitude of the influence of each neuron.

In each neuron in the hidden layer and output layer there is a bias (blue arrow) which will then be used in the activation function as a constant to produce output values. At this stage, each weight and bias in each neuron will be updated continuously and at each iteration, an evaluation process will be carried out to determine stopping points to produce output in accordance with expectations. Meanwhile, this study uses sigmoid non-linear activation function. And then, if viewed based on the plot, many parameters in the neural network model formed are 4 weight + 4 bias (input-first hidden), 4 weight + 4 bias (first-second hidden), and 4×2 weight + 2 bias ( with steps 177979 and error 24.00.

**Table 4. Confusion matrix**

| | | **Prediction** | |
|---|---|---|---|
| | | Tsunami | No |
| **Observation** | Tsunami | 183 | 1 |
| | No | 7 | 1218 |

From the confusion matrix in Table 4, data that categorized tsunami (true positive) and not (true negative) that were predicted correctly respectively were 183 and 1218 observations, while data categorized as tsunami that was predicted incorrectly (false negative) as well as data categorized as not tsunami (false positive) respectively were 1 and 7 observations.

Furthermore, from four measurements previously, obtained evaluation indicators, including accuracy of 99.43% which showed that 99.43% of the data was predicted correctly, precision of 99.45% showed that 96.45% of data with the tsunami category was predicted correctly, sensitivity of 99,45% stated that the success rate of the model in recovering information was 99.45%, specitivity of 99.43% which showed the capabilities of model in predicting information that are not categorized as tsunamis correctly, and F1 score of 97.86% which means that the balance between precision and sensitivity is 97.86%.

## 5 Conclusion

Based on the results and discussion, it has been obtained that the neural network method can be used to predict potential tsunami events due to large earthquakes. Neural network optimization with SMOTE resampling has been proven to be able to improve the evaluation of the model formed so that it can improve the quality of modeling. The neural network structure formed consists of an input layer consisting of 4 neurons (longitude, latitude, magnitude and depth), a hidden layer consisting of 4 neurons and an output layer with sigmoid activation function. The neural network model formed can be used to predict tsunami events in Indonesia with an accuracy of 99,43% and an f1-score of 97.86 %. It is hoped that the government can apply the use of the optimized neural network method as a warning system in various regions to detect potential future tsunami events. Then for future research, the use of other parameters can be considered to increase the accuracy of predictions and develop methods to shorten data processing time so that it becomes more effective.

## Reference

[1] M. T. I. Khan, S. Anwar, S. A. Sarkodie, M. R. Yaseen, A. M. Nadeem, and Q. Ali, “Natural disasters, resilience-building, and risk: achieving sustainable cities and human settlements,” *Natural Hazards*, vol. 118, no. 1, pp. 611–640, Aug. 2023, doi: 10.1007/s11069-023-06021-x.

[2] A. K. Das and B. Dutta, “Promoting Data Science in Disaster Risk Reduction: Glimpses from the Global South,” *Journal of Data Science, Informetrics, and Citation Studies*, vol. 2, no. 1, pp. 13–20, Apr. 2023, doi: 10.5530/jcitation.2.1.2.

[3] Triantafyllou, G. A. Papadopoulos, and A. Kijko, “Probabilistic Tsunami Risk Assessment from Incomplete and Uncertain Historical Impact Records: Mediterranean and Connected Seas,” *Pure Appl Geophys*, vol. 180, no. 5, pp. 1785–1809, May 2023, doi: 10.1007/s00024-023-03262-6.

ISSN:2302-8149
e-ISSN:2540-9719

[4] S. Mavroulis, M. Gogou, and E. Lekkas, "Earthquake-Induced Tsunamis in Western Greece (Ionian Sea and Western and Southern Peloponnese): Use of Tsunami Quantities, Impact and ITIS-2012 Intensities for Highlighting Susceptible Areas," *Geosciences (Switzerland)*, vol. 13, no. 2, Feb. 2023, doi: 10.3390/geosciences13020028.

[5] A. Imamura, *TheoreticaI and Applied Seismology. Maruzen*. Tokyo, 1937.

[6] E. Bernard, *Tsunami. In: Natural Disaster Management, a presentation to commemorate the International Decade for Natural Disaster Reduction (IDNR) 1999-2000*. Leicester: J. Ingleton, 1999.

[7] J. A. Álvarez-Gómez, P. Herrero-Barbero, and J. J. Martínez-Díaz, "Seismogenic potential and tsunami threat of the strike-slip Carboneras fault in the western Mediterranean from physics-based earthquake simulations," *Natural Hazards and Earth System Sciences*, vol. 23, no. 6, pp. 2031–2052, Jun. 2023, doi: 10.5194/nhess-23-2031-2023.

[8] J. Selva *et al.*, "Tsunami risk management for crustal earthquakes and non-seismic sources in Italy," *Rivista del Nuovo Cimento*, vol. 44, no. 2. Springer Science and Business Media Deutschland GmbH, pp. 69–144, Feb. 01, 2021. doi: 10.1007/s40766-021-00016-9.

[9] International Tsunami Information Center, "Tsunami Waves," Jan. 2024.

[10] BNPB, "Laporan Kinerja BNPB Tahun 2022," 2022.

[11] Ministry of Energy and Mineral Resources, "Katalog Tsunami Indonesia Tahun 416 – 2021," 2023.

[12] BMKG, "Aktivitas Gempa Bumi di Indonesia Tahun 2022," 2022.

[13] Z. Wang and H. Yue, "Resolving Shallow Moment Releases of Megathrust Earthquakes by Empirical Green's Function-Based Inversions: A New Approach for Tsunami Warning," *J Geophys Res Solid Earth*, vol. 127, no. 7, Jul. 2022, doi: 10.1029/2021JB023765.

[14] S. Lorito, F. Romano, and T. Lay, "Tsunamigenic Major and Great Earthquakes (2004–2013): Source Processes Inverted from Seismic, Geodetic, and Sea-Level Data," in *Complexity in Tsunamis, Volcanoes, and their Hazards*, Springer US, 2022, pp. 247–298. doi: 10.1007/978-1-0716-1705-2_641.

[15] F. Imamura, S. P. Boret, A. Suppasri, and A. Muhari, "Recent occurrences of serious tsunami damage and the future challenges of tsunami disaster risk reduction," *Progress in Disaster Science*, vol. 1. Elsevier Ltd, May 01, 2019. doi: 10.1016/j.pdisas.2019.100009.

[16] I. E. Mulia, N. Ueda, T. Miyoshi, A. R. Gusman, and K. Satake, "Machine learning-based tsunami inundation prediction derived from offshore observations," *Nat Commun*, vol. 13, no. 1, Dec. 2022, doi: 10.1038/s41467-022-33253-5.

[17] M. Kamiya, Y. Igarashi, M. Okada, and T. Baba, "Numerical experiments on tsunami flow depth prediction for clustered areas using regression and machine learning models," *Earth, Planets and Space*, vol. 74, no. 1, Dec. 2022, doi: 10.1186/s40623-022-01680-9.

[18] K. Goda and K. Abilova, "Effects of inaccurate seismic information on tsunami loss Tsunami hazard warning and risk prediction based on inaccurate earthquake source parameters Effects of inaccurate seismic information on tsunami loss," *Nat. Hazards Earth Syst. Sci. Discuss*, vol. 3, pp. 7487–7525, 2015, doi: 10.5194/nhessd-3-7487-2015.

ISSN:2302-8149
e-ISSN:2540-9719

[19] Z. Xu, S. Liang, Abd Rahman, and M.N.B., “Historical earthquakes, tsunamis and real-time earthquake monitoring for tsunami advisory in the South China Sea region,” *Nat Hazards*, vol. 107, pp. 771–793, 2021, doi: https://doi.org/10.1007/s11069-021-04605-z.

[20] A. Konovalov, G. Samsonov, and A. Stepnov, “A logit-based binary classifier of tsunamigenic earthquakes for the Northwestern Pacific Ocean,” *Pure Appl Geophys*, vol. 5, pp. 1623–1637, 2022, doi: https://doi.org/10.1007/s00024-022-03194-7.

[21] E. Irfiani *et al.*, “Classification of Tsunami Potential Based on Earthquakes in Indonesia Using the C4.5 Algorithm,” *Jurnal Mantik*, vol. 6, no. 3, pp. 3730–3736, 2022, doi: https://doi.org/10.35335/mantik.v6i3.3342.

[22] Amalia Listiani and Fuji Lestari, “Tsunami Potential Prediction with Artificial Neural Network,” *Int J Sci Res Sci Eng Technol*, pp. 231–236, Jan. 2023, doi: 10.32628/ijsrset2310130.

[23] R. Rastogi and M. Bansal, “Diabetes prediction model using data mining techniques,” *Measurement: Sensors*, vol. 25, Feb. 2023, doi: 10.1016/j.measen.2022.100605.

[24] H. Kukreja, N. Bharath, C. S. Siddesh, and S. Kuldeep, “AN INTRODUCTION TO ARTIFICIAL NEURAL NETWORK Article in International Journal Of Advance Research And Innovative Ideas In Education,” *IJARIIE*, vol. 1, no. 5, pp. 27–30, 2016, [Online]. Available: https://www.researchgate.net/publication/319903816

[25] M. J. Mhapsekar and M. R. Sharma, “Characteristics and applications of artificial neural network,” *J Emerg Technol Innov Res*, vol. 10, no. 4, pp. 44–49, 2023, [Online]. Available: www.jetir.orgh44

[26] M. M. Mijwel, “Artificial Neural Networks Advantages and Disadvantages,” *Mesopotamian Journal of Big Data*, pp. 29–31, Aug. 2021, doi: 10.58496/mjbd/2021/006.

[27] X. Shu and Y. Ye, “Knowledge Discovery: Methods from data mining and machine learning,” *Soc Sci Res*, vol. 110, Feb. 2023, doi: 10.1016/j.ssresearch.2022.102817.

[28] R. Ghorbani and R. Ghousi, “Comparing Different Resampling Methods in Predicting Students’ Performance Using Machine Learning Techniques,” *IEEE Access*, vol. 8, pp. 67899–67911, 2020, doi: 10.1109/ACCESS.2020.2986809.

[29] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, “SMOTE: Synthetic minority over-sampling technique,” *Journal of Artificial Intelligence Research*, vol. 16, pp. 321–357, 2002, doi: 10.1613/jair.953.

[30] A. J. Mohammed, “Improving Classification Performance for a Novel Imbalanced Medical Dataset using SMOTE Method,” *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 3, pp. 3161–3172, Jun. 2020, doi: 10.30534/ijatcse/2020/104932020.

[31] A. M. Sarhan, “Data Mining in Internet of Things Systems: A Literature Review,” *Journal of Engineering Research (ERJ)*, vol. 6, no. 5, p. 252, 2022, doi: 10.21608/erjeng.2023.186378.1143.

[32] A. K. Singha, S. Zubair, A. Malibari, N. Pathak, S. Urooj, and N. Sharma, “Design of ANN Based Non-Linear Network Using Interconnection of Parallel Processor,” *Computer Systems Science and Engineering*, vol. 46, no. 3, pp. 3491–3508, 2023, doi: 10.32604/csse.2023.029165.

*ISSN:2302-8149*
*e-ISSN:2540-9719*

[33] T. Chen, I. Goodfellow, and J. Shlens, "Net2Net: Accelerating Learning via Knowledge Transfer," Nov. 2015, [Online]. Available: http://arxiv.org/abs/1511.05641

[34] K. Hornik, M. Stinchcombe, and H. White, "Dermition 2.1."

[35] A. Lazcano, P. J. Herrera, and M. Monge, "A Combined Model Based on Recurrent Neural Networks and Graph Convolutional Networks for Financial Time Series Forecasting," *Mathematics*, vol. 11, no. 1, Jan. 2023, doi: 10.3390/math11010224.

[36] S. W. Sidehabi, A. Suyuti, I. S. Areni, and I. Nurtanio, "Classification on Passion Fruit's Ripeness using K-Means Clustering and Artificial Neural Network," in *International Conference on Information and Communications Technology (ICOIACT)*, 2018, pp. 304–309. doi: DOI:10.1109/ICOIACT.2018.8350728.

[37] A. Yaqin, A. D. Laksito, and S. Fatonah, "Evaluation of Backpropagation Neural Network Models for Early Prediction of Student's Graduation in XYZ University," *International Journal on Advanced Science Engineering Information Technology (IJASEIT)*, vol. 11, no. 2, 2021, doi: 10.18517/ijaseit.11.2.11152.

[38] N. Agarwal, I. Arora, H. Saini, and U. Sharma, "A Novel Approach for Earthquake Prediction Using Random Forest and Neural Networks," *EAI Endorsed Transactions on Energy Web*, vol. 10, 2023, doi: 10.4108/EW.4329.

[39] BNPB, "Tsunami Puluhan Meter Melanda Maluku 346 Tahun Lalu," 2020.