SciencePG
Science Publishing Group

Seismic Detection Model Using Machine Learning to Protect the Public from Landslide and Earthquake Disasters in Kenya

Moses Kung'u Githu[1, *], Edwin Kagereki[2], Serah Munyua[3]

[1]Department of Computer Sciences, St. Pauls University, Nairobi, Kenya

[2]Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Nairobi, Kenya

[3]Department of Environmental Studies, Karatina University, Nyeri, Kenya

Email address:
mosesgithu@gmail.com (M. K. Githu), kagerekie@gmail.com (E. Kagereki), anneserah090@gmail.com (S. Munyua)
[*]Corresponding author

Abstract: Earthquakes and tremors are a common occurrence throughout the world, mostly in China, Japan and Indonesia. In Kenya, we experience a lot of tremors and landslides during the rainy seasons that have extensive negative social, economic, and environmental impacts. These damages include loss of human life, financial loss and destruction of infrastructure. This becomes a lagging factor towards achieving the Vision 2030 and Sustainable Development Goals (SDGs). This study used secondary data, obtained from World Wide Standardized Seismograph Station (WWSSSN) in Kilimambogo. Stochastic artificial neural network was adopted to identify prone areas to the said natural disasters, measure the socioeconomic impacts and build a predictive model for landslides, tremor and earthquakes in Kenya. It was evident that landslides are destructive in nature through observable measurable impacts on people. They increase the social and economic burden on the affected people. 64.76% of the measurable impacts affect human beings directly while the rest affect cattle and crops. Along the Great rift valley, most earthquakes and landslides took place. This is attributed to the active seismic activities. Kenya experiences earthquakes of magnitude m < 4. Our model achieved root mean square of 0.435. Furthermore, we got R^2=0.80 for testing dataset. This implied that 80% of data was trainable by the model. Therefore, the predictive neural network model is efficient and accurate in forecasting, and more importantly is a good fit model.

Keywords: Earthquake, Landslide, Neural Network

1. Introduction

Earthquakes are among the most devastating natural disasters. Tremors or earthquakes of magnitude 4 and below are the most frequent, followed by those of moderate magnitude (4-6) and, those with higher magnitude than 6 are less frequent [3]. They are destructive in terms of human life and financial losses. From 1956-2016, China has recorded the highest frequency of earthquakes. In 2008, the Wenchuan earthquake costed China, an estimate of US 124 billion as direct economic damages. Indirect economic damages were approximated to 40% of the direct economic damages [15]. This is a huge and significant financial loss. Japan is not

spared as well, as it is hit most financially due to earthquakes. It is estimated that Japan incurred losses of about 211 billion USD. This further caused measurable economic impact on the damaged regions, including loss in the country's GDP. The loss was due to Tohoku-Oki Earthquake and Tsunami in 2011 [8]. In 1998-2017, 125 million people were affected by earthquake. In the same period time, at least 747,234 deaths were as a result of earthquakes. Earthquakes lead with the highest fatalities followed by storms [9].

The East African region is characterized by a moderate level of seismicity, mainly controlled by the structural trend of the East African Rift Valley. This makes the area within it, to be prone to landslides, volcanoes, tremors and

earthquakes. In 1928, Kenya was hit by an earthquake, of magnitude 6.9 at Subukia. The earthquake's epicenter lied within the rift valley boundaries [2]. In 2019, a 4.8 magnitude earthquake hit Kenya, Wundanyi as the epicenter. In 2020, Kenya experienced tremors from an earthquake in Tanzania.

Landslide is the movement of mass of earth down the slope. Most of the times, it happens when the soil is loose and there is much rainfall. Between 1998-2017, landslides have affected about 5 million people and at least about 20,000 deaths. Nevertheless, it resulted to 8 billion USD economic losses. Primarily, landslides are associated with mountainous areas however, they can occur in areas of low relief. Landslides are prevalent in any land-based environment with slopes driven by tectonic [4], climatic [11] and/or human activities [14]. Landslides' occurrence is frequent in Kenya, mostly in the rift valley region. Annually, 300 people on average are at risk of being affected on landslides. Analysts estimate that damages caused by landslides is $3 million of GDP annually. Rift valley, Western and Nyanza provinces are greatly affected as they receive heavy rainfall.

Deep learning, specifically, artificial neural network, analysts have grown fond of it. Artificial neural networks are better than classical statistical methods in predicting weather patterns [7]. Recurrent neural network predicts more accurately for earthquakes of magnitude 6^+ [13]. Probabilistic artificial neural network best predicts for magnitude between 4.5 and 6.0 [1]. An artificial neural network is a deterministic algorithm, mainly used for static simulations. Earthquakes, tremors and landslides are complex and dynamic phenomena.

Kavirondo rift and the Great rift valley are characterized by high rates of seismic activities. Furthermore, earthquakes of magnitude 3^+ have been detected [12]. Nevertheless, over the past recent years, Kenya has experienced more occurrences of landslides and tremors. In 2020, West Pokot County suffered from major landslides due to heavy rainfall.

Socioeconomic costs of natural disasters have significantly increased due to consecutive occurrence of natural disasters. In developing countries, there is lack of capability to absorb severe economic impact due to natural disasters and an established effective disaster risk transfer mechanisms [15]. This implies, earthquakes, landslide and tremors inhibit attainment of Vision 2030, better yet, achieving the sustainable development goals and the government's big four agenda.

As a nation, we have the capability to mitigate these natural disasters. The government of Kenya, county governments and other private partnership have taken initiative in combating drought. This can be extended to the other three natural disasters. Therefore, effective and efficient mitigation can be adopted. To accomplish this, research support is necessary, hence the essence of this research.

2. Related Work

Over the past years, there have been successful integrations of data science into seismology. This has motivated most scholars to build models to predict occurrence of the natural disasters. Statistical Seismology aims to bridge the gap between non-statistical models based on physics and non-physical models based on statistics. This scientific domain can then be divided into two categories, with earthquakes as point sources and modeled as non-stationary stochastic point processes, or earthquakes as seismic waves radiating from finite sources.

Zarola and Sil incorporated stochastic techniques to artificial neural networks, so as to estimate earthquakes prevalence in India [16]. They identified 12 regions that are seismic active. Gamma, Lognormal, Weibull and Loglogistic were used in stochastic techniques, estimating time of future earthquakes occurrence. They further used two activation functions; sigmoid and rectified linear unit, trained for a maximum of 25000 training epochs to achieve a sum of squared error between target and estimated outputs of 0.009 and cross validation (k=5).

In verifying landslide susceptibility mapping, [10] used artificial neural network, to which they found it had a prediction accuracy of 93\% for grounds with zero slope. Sigmoid function was the best choice of activation function.

Adeli and Panakkat used probabilistic neural network in predicting earthquake magnitude. Parzen windows classifier and Bayesian approach were used to calculate probabilities. The study concluded that the model was accurate in predicting earthquakes of magnitude between 4 and 6 [1].

3. Methodology

3.1. Data Source

This study used secondary data obtained from World Wide Standardized Seismograph Station in Kilimambogo.

3.2. Artificial Neural Network

Neural network is an interconnected assembly of simple processing units or nodes, the functionality of which is primarily based on the human brain neurons [5]. Figure 1 depicts both a typical neuron and components of a neural network.

This study adopted artificial neural network. Activation function used was Rectified Linear Unit (ReLU). ReLU acts like a linear function, but is actually a nonlinear function, that allows complex relationships in the data to be learned was required. In addition, the function must be more sensitive to the activation sum input and avoid easy saturation. It is defined as:

$$f(x) = max\,[0, x]$$

ReLU is less computational and powerful than sigmoid and tanh functions. It can easily backpropagate the errors and easily activate multi-layer hidden layer. He initialization was

used in assigning values to weights at first before back propagation [6]. This technique will simply multiply random initialization with:

$$\sqrt{\frac{2}{size^{[l-1]}}}$$
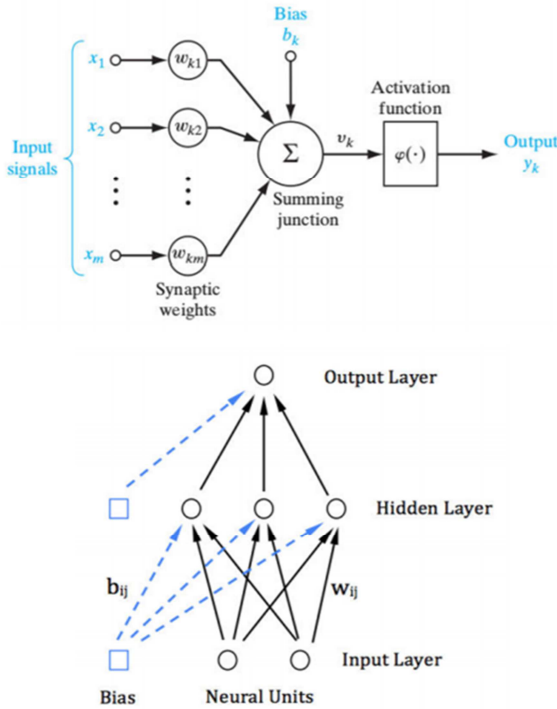
where l-1 is the number of hidden layers.

Figure 1. *Architecture of Neural Network.*

Furthermore, in training the neural network, stochastic gradient descent with backpropagation of errors was applied. Momentum term accelerated the back propagation learning algorithm, which was based on a generalized delta-rule. The momentum term was included to accelerate error convergence during training while maintaining a high learning rate. Weight and bias factors will be adjusted using the below functions;

$$w_{ij}^{k+1} = w_{ij}^{k} + \Delta w_{ij}^{k}$$

$$\Delta w_{ij}^{k} = \eta(\delta_j^k)O_i + \alpha\Delta w_{ij}^{k-1}$$

$$b_j^{k+1} = b_j^{k} + \Delta b_j^{k}$$

$$\Delta b_j^{k} = \eta(\delta_j^k) + \alpha\Delta b_j^{k-1}$$

where η is the learning rate, α is the momentum coefficient, δ is the gradient descent correction term and k is the stand for number of patterns. An iteration represented each pattern presentation. An epoch was presentation of the entire training set.

3.3. Model Assessment

It was obligatory to assess adequacy and appropriateness of the model. For a regression model, root mean square error was used to assess the model. Root Mean Square Error (RMSE) is defined as the measure between predicted and observed values. The RMSE shows us how big of an error we may expect on average from the forecast. RMSE is a good measure of how accurately the model predicts the response, and it is the most important criterion for fit, if the main purpose of the model is prediction.

3.4. Model Diagnostic

Training neural networks is a difficult task that can achieve results that are either much better than expected or far worse and produce solely noise. To avoid over fitting or under fitting, a loss curve was plotted. This curve is plotted in the training stage against epoch. Loss curve provides a brief image of the training process and the direction in which the network learns, indicating either under fitting, over fitting or neither both.

4. Results and Discussions

4.1. Results

For earthquake data, magnitude, depth and age had numerical values. Earthquake data was platykurtic indicating lack of outliers. Moreover, the data was positive skewed. Also, the range of the data was 5.5. This was summarized by Table 1.

Table 1. *Descriptive Statistics for Earthquake data.*

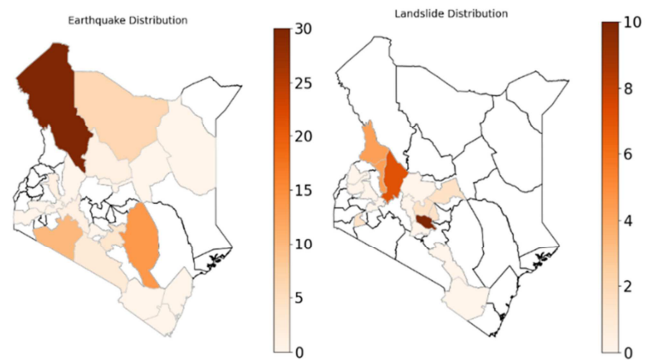| Variable	Min	Max	Mean	Skewness	Kurtosis
Magnitude	1.5	7	3.0789	0.72449	-0.4621

4.2. Geographic Hazard

Figure 2. *Geographic Hazard for Landslide and Earthquake.*

Figure 2, depicts the frequency of occurrence of both landslide and earthquake per county. Turkana and Murang'a are modal counties for earthquake and landslide respectively. The earthquake occurrences are distributed across Kenya unlike landslides, which are concentrated in less than 15 counties. Also, it is evident that, despite Murang'a county, other vulnerable areas prone to landslides are the highlands along rift valley.

4.3. Socio-Economic Impacts

Natural disasters are forces to reckon with. That said, they have negative impacts, in terms of socio-economic development. This research focused on damages due to landslide. This was because experienced earthquakes have no measurable negative impacts.
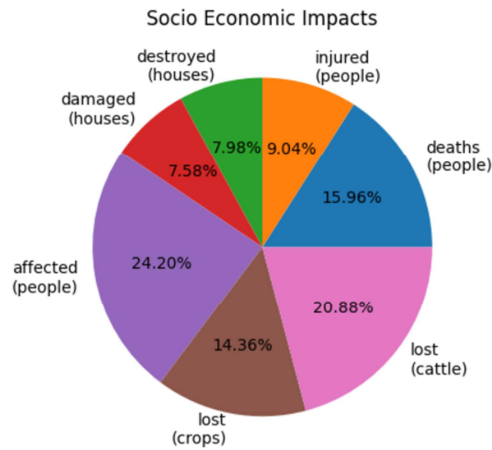

Figure 3. *Measurable Impacts due to Landslides.*

From figure 3, it was evident that landslides significantly impact peoples' lives and their livelihoods. We can see that 64.76% of the measurable impacts affect human beings directly while the rest affect cattle and crops. More people have lost their lives as compare to houses either damaged or destroyed, people being injured, and crops lost due to landslides. Generally, loss and damages of properties have negative economic impact to the affected people. This set back the affected people, more so those with no insurance, or are lower middle economic classes.

4.4. Predictive Model

Artificial neural network was used to build our predictive model. He initialization in conjunction to ReLU was adopted in our algorithm in Tensor Flow-keras. The neural network involved feed forward and back propagation with two hidden layers (first layer has 5 nodes and second layer has 3 nodes). Also, maximum epochs of training were 200.

Input layer consisted of elapsed time between two successive consecutive earthquakes, latitude and longitude. The output layer involved the earthquake's magnitude.

4.5. Model Assessment

It was important to assess the model. This was not exempted in this research. To assess the predictive model, RMSE was adopted. This was because, our predictive model was a regression model. As stated earlier, RMSE is a measure of differences of predicted values and observed values. It was evident that our model produces less deviation, therefore being efficient and accurate in forecasting. The model was a perfect fit as the RMSE has a small value, depicted in Table 2.

Table 2. *Model Assessment Statistic.*

Statistic	Value
RMSE	0.435
R^2	0.80

4.6. Model Diagnostic

Again, the predictive model needed to be examined not to be either under fit or over fit. To do so, loss curve plot was of interest. Loss values were plotted against epochs in the training session. From the loss curve plot, depicted in figure 4, it was evident that our model neither under fit nor over fit.
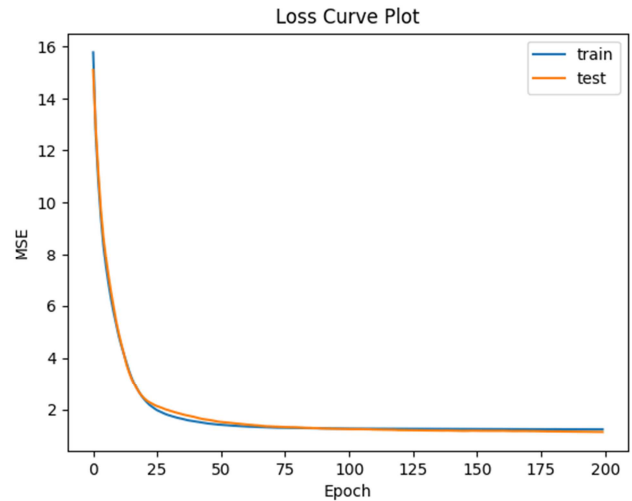

Figure 4. *Loss Curve Plot.*

5. Conclusion and Recommendation

In Kenya, active seismic zone is near and/or within the Great rift valley. This study adopted geographical mapping to establish areas vulnerable to earthquakes and landslides. For earthquake occurrence, Turkana is the modal county. Nevertheless, counties near and/or within the rift valley are vulnerable to earthquakes. The study further found that Kenya experiences earthquakes of magnitude $m \leq 4$. The maximum recorded magnitude was 7, which took place in Subukia, within the rift valley. Also, we can see, apart from Murang'a county, most landslides occur in counties within in the rift valley. Although this does not attribute to the seismic activities solely, as some important factors like soil topography, amount of rainfall were not investigated.

Both landslides and earthquakes are destructive in nature. This study focused on the socio-economic impacts due to landslides. It is evident that 64.76% of the measurable impacts, directly affect the people in those environs. Loss of human lives is higher than damages of properties, which is terrifying. It is evident that landslides have significant negative impacts on people's lives and their livelihoods.

This study used neural network to build a predictive. This is because neural networks outperform traditional statistical approaches. The predictive model yields a small value root mean square error of 0.435 and R^2 of 0.80. This implies the predictive model is efficient and accurate in forecasting, and

more importantly it is not poor fit.

For further work, it is in the best interest to focus on the measurable risks due to landslides and earthquakes, and their mitigation. Also, it will be great for insurance covers be provided to people who live in landslides prone areas. The government could utilize this model to plan for disaster management and evacuation to reduce the negative significant impacts of socio-economic impacts in the prone areas.

References

[1] Adeli, H. and Panakkat, A. (2009). A probabilistic neural network for earthquake magnitude prediction. Neural networks, 22 (7): 1018–1024.

[2] Ambraseys, N. (1991). Earthquake hazard in the kenya rift: the subukia earthquake 1928. Geophysical Journal International, 105 (1): 253–269.

[3] Asim, K. M., Idris, A., Iqbal, T., and Martínez-Álvarez, F. (2018). Earthquake prediction model using support vector regressor and hybrid neural networks. PloS one, 13 (7): e0199004.

[4] Bennett, G. L., Miller, S. R., Roering, J. J., and Schmidt, D. A. (2016). Landslides, threshold slopes, and the survival of relict terrain in the wake of the mendocino triple junction. Geology, 44 (5): 363–366.

[5] Gurney, K. (1997). An introduction to neural networks. CRC press.

[6] He, K., Zhang, X., Ren, S., and Sun, J. (2015). Delving deep into rectifiers: Surpassing human-level performance on imagenet classification. In Proceedings of the IEEE international conference on computer vision, pages 1026–1034.

[7] Jafri, Y. Z., Sami, M., Waseem, A., Murtaza, G., and Akbar, S. (2012). Stochastic approaches for time series forecasting of rate of dust fall: A case study of northwest of Balochistan, Pakistan. International Journal of Physical Sciences, 7 (4): 676–686.

[8] Kajitani, Y., Chang, S. E., and Tatano, H. (2013). Economic impacts of the 2011 tohoku-oki earthquake and tsunami. Earthquake Spectra, 29 (1_suppl): 457–478.

[9] Koubi, V. (2018). Climate change and disasters: Sustainable development impacts, trends, interlinkages with migration and conflicts, and likely scenarios.

[10] Lee, S. and Evangelista, D. (2006). Earthquake-induced landslide-susceptibility mapping using an artificial neural network. Natural Hazards and Earth System Sciences, 6 (5): 687–695.

[11] Moreiras, S. M. (2005). Climatic effect of enso associated with landslide occurrence in the central andes, Mendoza province, Argentina. Landslides, 2 (1): 53–59.

[12] Mulwa, J. K., Kimata, F., Suzuki, S., and Kuria, Z. N. (2014). The seismicity in kenya (east africa) for the period 1906–2010: A review. Journal of African Earth Sciences, 89: 72–78.

[13] Panakkat, A. and Adeli, H. (2007). Neural network models for earthquake magnitude prediction using multiple seismicity indicators. International journal of neural systems, 17 (01): 13–33.

[14] Petley, D. N., Hearn, G. J., Hart, A., Rosser, N. J., Dunning, S. A., Oven, K., and Mitchell, W. A. (2007). Trends in landslide occurrence in Nepal. Natural hazards, 43 (1): 23–44.

[15] Wu, J., Li, N., Hallegatte, S., Shi, P., Hu, A., and Liu, X. (2012). Regional indirect economic impact evaluation of the 2008 Wenchuan earthquake. Environmental Earth Sciences, 65 (1): 161–172.

[16] Zarola, A. and Sil, A. (2017). Artificial neural networks (ann) and stochastic techniques to estimate earthquake occurrences in northeast region of India. Annals of Geophysics, 60 (6): 0658.