Analysis of Sequential Order Incremental Methods in Predicting the Number of Victims Affected by Disasters

Parulian Parulian 1, Medi Hermanto Tinambunan 2, Salomo Ginting 2, M. Khalil Gibran 2, Anjar Wanto 3, La Ode Muharram 4, N Nurawati 5, Gita Widi Bhawika 6

1 Universitas HKBP Nommensen, Medan - Indonesia
2 Universitas Sumatera Utara, Medan - Indonesia
3 STIKOM Tunas Bangsa Pematang Siantar, Medan – Indonesia
4 Universitas Halu Oleo, Kendari – Indonesia
5 Industrial engineering, Universitas 45 Surabaya - Indonesia
6 Institut Teknologi Sepuluh Nopember, Surabaya - Indonesia

* liansirait@uhn.ac.id

Abstract. Disaster is a series of events that threaten and disrupt human life caused by natural factors, non-natural factors and human factors themselves. Therefore, disasters cause casualties, environmental damage, property losses, and psychological impacts. In this study will be discussed about the prediction of the number of victims affected by the disaster, either died, lost, injured, suffered or displaced. Data sources were obtained by the National Disaster Management Agency and the Indonesian Central Statistics Agency. The method used to predict is the Incremental Sequential Order method. This method is one part of the Artificial Neural Network method. With this method, network architecture patterns will be established to predict the number of victims affected by the disaster for years to come. The network architecture models used are 4-5-1, 4-10-1, 4-5-10-1, 4-10-20-1 and 4-15-30-1. Of the five models, the best models will be obtained, namely 4-15-30-1 with an accuracy rate of 80%. With this architectural model, predictions will be made on the number of victims affected by the disaster for years to come.

1. Introducing

Disaster is a series of events that threaten and disrupt human life caused by natural factors, non-natural factors and human factors themselves. Therefore, disasters cause casualties, environmental damage, property losses, and psychological impacts. Nature can be floods, volcanic eruptions, earthquakes, tsunamis, landslides, snow storms, droughts, hail, heat waves, hurricanes, tropical storms, typhoons, tornadoes, wildfires and disease outbreaks. In the 20th century, some of the most common natural disasters were hunger and plague. In the 21st century, more and more natural disasters are climate-related disasters caused by rising earth temperatures (global warming). Global warming is largely followed by floods, droughts, extreme weather and unpredictable seasons. Climate change has the potential to increase poverty and vulnerability in large numbers. At the same time, climate disasters are increasing; more people are affected due to poverty, lack of resources, population growth, movement and placement of people to disadvantaged areas [1].

Indonesia is a country that is very vulnerable to natural disasters such as earthquakes, tsunamis, volcanic eruptions, landslides, floods, and whirlwinds. About 13 percent of the world's volcanoes in the Indonesian archipelago have the potential to cause natural disasters with varying intensity and strength. The earthquake and Indian Ocean tsunami in 2004 took a lot of casualties in the provinces of Aceh (NAD) and North Sumatra, forcing rapid efforts to educate the public to prepare themselves well.
to deal with natural disasters [2]. According to the Indonesian government policy, regional and provincial officials are required to be at the forefront of natural disaster management. While the National Disaster Management Agency and the army can help when needed. However, the policy has not created systematic change at the local level. The regional disaster management agency is planned in all provinces but newly established in 18 regions. Also, the weakness of disaster management in Indonesia is partly due to the lack of resources and skills of local governments that are still dependent on the central government [3]. Therefore the author conducted this research which will produce a prediction of the number of victims of the disaster, with the hope of being able to benefit the central and regional governments as a reference so as early as possible anticipate and reduce the number of victims affected by the disaster in the future. The prediction method that I use is Sequential Order Incremental with learning function.

Sequential Order Incremental Method with learning function is one method of artificial neural networks. Artificial Neural Networks are widely used for solving problems related to estimation, pattern recognition, data analysis, control, and grouping [4]–[7]. This method is part of artificial intelligence, as well as data mining [8]. Several studies have discussed natural disasters and those related to Incremental Sequential Order methods namely: [9] Predict and visualize disaster risks in the Philippines using the Discrete Wavelet Transform (DWT) method, Autoregressive Integrated Moving Average (ARIMA), and Artificial Neural Networks. Of the three methods that are best for predicting disaster victims, damaged houses and damaged property is the ARIMA-ANN model with an accuracy of 53.72%, 36.72%, and 26.36%. But for the whole disaster, the DWT model is better with an accuracy rate of 62.81%, then ARIMA with an accuracy of 39.84% and ANN 33.33%. This study concludes that there is no significant difference between actual data and the predicted results obtained. [10] Conduct research using Incremental Sequence Learning, a simple incremental approach to sequence learning. This study concludes that Additional Sequence Learning significantly accelerates sequence learning and achieves the best level of test performance from regular order learning 20 times faster, reduces test errors by 74%, and generally has better performance.

2. Metodologi

2.1. Data Collection

Data is taken from Information and Disaster Data in Indonesia, BNPB and the Indonesian Central Statistics Agency.

| No | Victim | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 |
|----|--------|------|------|------|------|------|------|
| 1  | Die    | 691  | 399  | 721  | 769  | 320  | 369  |
| 2  | Lost   | 262  | 114  | 128  | 229  | 125  | 109  |
| 3  | Injured| 1097 | 1528 | 3612 | 2348 | 463  | 457  |
| 4  | Suffer | 11687| 826753| 3386955| 1967167| 1043677| 2139124|
| 5  | Evacuate| 21259| 142238| 511431| 865720| 184258| 282038|

Source: Source: Indonesia's Disaster Information and Data, BNPB

2.2. Stages of Research

The stages of this study are as follows:

![Figure 1. Stages of Research](image-url)
Explanation from Figure 2: it can be explained that collecting data in a study is the first thing to do. The second stage was carried out a literature study to complement the basic knowledge and theories used in this study. The third stage identifies the problem to process the conversion phase of the data obtained in accordance with the specified weight. The fourth stage preprocesses with the aim to facilitate understanding of the contents of the record. The fifth stage determines the pattern and determination of the network architecture model that is adapted to the research problem faced. The sixth stage examines the results of data processing using the Matlab application. The seventh stage predicts that is to see comparisons of several architectural models used in the study in order to obtain the best architectural models and the most accurate level of accuracy. The eighth stage evaluates the end to find out whether the results of data processing are as desired.

2.3. Data Normalization
The formula used for the normalization of the initial data is as follows [11]–[14]:

$$x' = \frac{0.8(x - a)}{b - a} + 0.1 \quad (1)$$

Explanations:
- $x'$: Data transformation, $x$: Data to be normalized, $a$: The lowest value data, $b$: Data with the highest value.

Data on the Number of Death, Missing, and Injured Victims Affected by Disaster that will be predicted, are first divided into 2 (table 1). First is the training data section, namely 2011-2014 with 2015 as the target. Second is the data testing section, namely 2012-2015 with 2016 as the target. Normalization of training and testing can be seen in table 2 and table 3 below:

| Table 2. Normalization of Data Training |
|----------------------------------------|
| **D a t a** | **2011** | **2012** | **2013** | **2014** | **T a r g e t** |
| 1 | 0.100136 | 0.100067 | 0.100143 | 0.100155 | 0.100049 |
| 2 | 0.100035 | 0.100000 | 0.100003 | 0.100027 | 0.100003 |
| 3 | 0.100232 | 0.100334 | 0.100826 | 0.100528 | 0.100082 |
| 4 | 0.102734 | 0.295259 | 0.900000 | 0.564634 | 0.346498 |
| 5 | 0.104995 | 0.133571 | 0.220777 | 0.304463 | 0.143496 |

| Table 3. Normalization of Data Testing |
|----------------------------------------|
| **D a t a** | **2012** | **2013** | **2014** | **2015** | **T a r g e t** |
| 1 | 0.100069 | 0.100145 | 0.100156 | 0.100500 | 0.100061 |
| 2 | 0.100001 | 0.100004 | 0.100028 | 0.100004 | 0.100000 |
| 3 | 0.100335 | 0.100827 | 0.100529 | 0.100084 | 0.100082 |
| 4 | 0.295260 | 0.900000 | 0.564635 | 0.346499 | 0.605252 |
| 5 | 0.133572 | 0.220778 | 0.304464 | 0.143497 | 0.166594 |

3. Results and Discussion
There are 5 architectural models used in this study, namely: 4-5-1 (4 are input layers, 5 are hidden layer neurons, and 1 is output), 4-10-1 (4 are input layers, 10 are hidden layers neurons, and 1 is output), 4-5-10-1 (4 is the input layer, 5 is the first hidden layer neuron, 10 is the second hidden layer neuron, and 1 is output), 4-10-20-1 (4 is the input layer, 10 is the first hidden layer neuron, 20 is the second hidden layer neuron, and 1 is the output) and 4-15-30-1 (4 is the input layer, 15 is the first hidden layer neuron, 30 is the hidden layer neuron second, and 1 is output). Four which is the input layer taken from 2011-2014 and 2012-2015, the determination of hidden layers (no standard rules), but the authors use 5, 10, 5-10, 10-20, and 15-30. While output 1 is the target or output of the network.
The error rate is used 0.001-0.03. While training in Sequential Order Incremental method parameters with learning functions (trains) that are used in general can be seen in the following Figure:

![Figure 2. Parameters Used](image)

Explanation from Figure 2: That after the training data is entered into Matlab, the next step is to create a new network (Example using a 4-15-30-1 network model). Input data is not included in the formula when building a new network, because the input has been entered and processed in advance (P and T) which is normalization (Table 2). The transfer function used is 'tansig' and 'logsig'. While the method used is Sequential Order Incremental with learning functions (trains). Show GUI training (showWindow) = true, Generate command line output (showCommandLine) = false, Epochs between display (show) = 25, maximum epoch number to train (Epoch) = 20000, maximum time to practice in seconds (time) = inf, Performance target (target) = 0, Maximum validation failure (max fail) = 6. Program code >> net = train (net, P, T) is a data simulation stage, where this process will bring up training results in the form of error values and training charts, while >> [a, Pf, Af, e, Perf] = sim (net, P, [], [], T) are used to see the results released by the network.

|   | Training | Testing |
|---|---|---|
|   | Data | Target | Output | Error | SSE | Target | Output | Error | SSE | Results |
| 1 | 0.10005 | 0.12930 | -0.02925 | 0.0008556410 | 0.10006 | 0.12780 | -0.02774 | 0.0007694291 | 1 |
| 2 | 0.10000 | 0.12730 | -0.02730 | 0.0007451481 | 0.10000 | 0.12650 | -0.02650 | 0.0007022500 | 1 |
| 3 | 0.10008 | 0.13080 | -0.03072 | 0.0009435687 | 0.10008 | 0.13280 | -0.03272 | 0.0010704544 | 1 |
| 4 | 0.34650 | 0.38470 | -0.03820 | 0.0014593732 | 0.60525 | 0.16690 | 0.43835 | 0.1921528085 | 0 |
| 5 | 0.14350 | 0.11180 | 0.03170 | 0.0010046582 | 0.16659 | 0.16690 | -0.00031 | 0.0000000937 | 1 |
|   | Total | 0.0050083892 | Total | 0.1946950358 | 80% |

|   | Training | Testing |
|---|---|---|
| MSE | 0.0010016778 | MSE | 0.0389390072 |

Explanation:

1 = True
0 = False
Error = Target-Output
SSE = Error^2
Accuracy = correct results on ((Data / 5) * 100), yields 80%
MSE = Total SSE / 5 (number of datas)

In table 5 it can be seen that the best model that will be used to predict the number of victims affected by disasters in Indonesia is 4-15-30-1 with epoch 20000 iterations, time 2 minutes 55 seconds, MSE 0.0285523266 and truth level 80%.

Table 5. Architectural Results Sequential Order Incremental with learning function Method

| Sequential order incremental | Architecture | Epoch | Time | MSE         | Accuracy |
|------------------------------|--------------|-------|------|-------------|----------|
|                              | 4-5-1        | 20000 | 02.32 | 0.0285785631 | 60%      |
|                              | 4-10-1       | 20000 | 02.31 | 0.0379603865 | 80%      |
|                              | 4-5-10-1     | 20000 | 02.59 | 0.0826960215 | 60%      |
|                              | 4-10-20-1    | 20000 | 03.03 | 0.0411551165 | 80%      |
|                              | 4-15-30-1    | 20000 | 02.55 | 0.0285523266 | 80%      |

Figure 3. Graph of Accuracy and MSE Level of Sequential Order Incremental Method

Explanations Figure 3: the best model architectural model of the five architecture comparisons is 4-15-30-1 with 80% accuracy and MSE 0.0285523266.

| Victim | Preliminary data | Predictions results |
|--------|------------------|---------------------|
|        | 2011 2012 2013 2014 2015 2016 | 2017 2018 2019 2020 |
| 1 Die   | 691 399 721 769 320 369 | 426 595 662 679 |
| 2 Lost  | 262 114 128 229 125 109 | 961 595 662 696 |
| 3 Injured | 1097 1528 3612 2348 463 457 | 1981 1270 999 747 |
| 4 Suffering | 11687 826753 3386955 1967167 1043677 2139124 | 1351164 539202 136324 9293 |
| 5 Evacuate | 21259 142238 511431 865720 184258 282038 | 178182 71508 18571 1832 |

4. Conclusion
The conclusion of this study is:
a. The 4-15-30-1 architectural model using Sequential Order Incremental Method produces 80% accuracy.
b. There are three architectural models that have the same level of accuracy of 80%. However, the smaller 4-15-30-1 architectural models are Mean Square Error (MSE) compared to the 4-10-1 and 4-10-20-1 architectural models.
References

[1] O. B. Paper, “Climate Alarm Disasters increase as climate change bites,” *Oxfam Briefing Paper*, no. November, pp. 1–28, 2007.

[2] U. Matsuura and D.-G. Koichiro, *Natural Disaster Preparedness and Education for Sustainable Development*. 2007.

[3] E. A. Forum, “Natural disasters in Indonesia: Strengthening disaster preparedness,” *Natural disasters in Indonesia: Strengthening disaster preparedness*, pp. 1–3, 2010.

[4] A. Wanto, M. Zarlis, Sawaluddin, and D. Hartama, “Analysis of Artificial Neural Network Backpropagation Using Conjugate Gradient Fletcher Reeves in the Predicting Process,” *Journal of Physics: Conference Series*, vol. 930, no. 1, pp. 1–7, 2017.

[5] A. Wanto, A. P. Windarto, D. Hartama, and I. Parlina, “Use of Binary Sigmoid Function And Linear Identity In Artificial Neural Networks For Forecasting Population Density,” *International Journal Of Information System & Technology*, vol. 1, no. 1, pp. 43–54, 2017.

[6] N. Nasution, A. Zamsuri, L. Lisnawita, and A. Wanto, “Polak-Ribiere updates analysis with binary and linear function in determining coffee exports in Indonesia,” *IOP Conference Series: Materials Science and Engineering*, vol. 420, no. 12089, pp. 1–9, 2018.

[7] B. Febriadi, Z. Zamzami, Y. Yunefri, and A. Wanto, “Bipolar function in backpropagation algorithm in predicting Indonesia’s coal exports by major destination countries,” *IOP Conference Series: Materials Science and Engineering*, vol. 420, no. 12089, pp. 1–9, 2018.

[8] S. Sudirman, A. P. Windarto, and A. Wanto, “Data Mining Tools | RapidMiner: K-Means Method on Clustering of Rice Crops by Province as Efforts to Stabilize Food Crops In Indonesia,” *IOP Conference Series: Materials Science and Engineering*, vol. 420, no. 12089, pp. 1–8, 2018.

[9] G. L. V Alquisola, D. J. A. Coronel, B. M. F. Reolope, and J. N. A. Roque, “Prediction and Visualization of the Disaster Risks in the Philippines Using Discrete Wavelet Transform (DWT), Autoregressive Integrated Moving Average (ARIMA), and Artificial Neural Network (ANN),” *2018 3rd International Conference on Computer and Communication Systems (ICCCCS)*, pp. 146–149, 2018.

[10] E. D. De Jong, “Incremental sequence learning,” *arXiv:1611.03068v2 [cs.LG]* 1 Dec 2016, no. 1991, pp. 1–18, 2016.

[11] M. Fauzan *et al.*, “Epoch Analysis and Accuracy 3 ANN Algorithm Using Consumer Price Index Data in Indonesia,” 2018, pp. 1–7.

[12] A. Wanto *et al.*, “Analysis of Standard Gradient Descent with GD Momentum And Adaptive LR for SPR Prediction,” 2018, pp. 1–9.

[13] A. Wanto *et al.*, “Levenberg-Marquardt Algorithm Combined with Bipolar Sigmoid Function to Measure Open Unemployment Rate in Indonesia,” 2018, pp. 1–7.

[14] S. P. Siregar and A. Wanto, “Analysis of Artificial Neural Network Accuracy Using Backpropagation Algorithm In Predicting Process (Forecasting),” *International Journal Of Information System & Technology*, vol. 1, no. 1, pp. 34–42, 2017.