Prediction of the Severity of Covid-19 Patients based on Demographics, Comorbidities and Symptoms using Backpropagation Neural

Indah Yanti[1*] Syaiful Anam[1] Zuraidah Fitriah[1] Aurick Yudha Nagara[2]

[1] *Faculty of Mathematics and Natural Science, Brawijaya University*
[2] *Faculty of Medicine, Brawijaya University*
[*]*Corresponding author. Email:* indah_yanti@ub.ac.id

ABSTRACT
Coronavirus is a stranded RNA virus. Coronaviruses belong to the Coronavirdiae family, which infects birds, mammals, and humans, among others. Since the initial report of pneumonia in Wuhan, China in December 2019, SARS-CoV-2 has spread to more than 200 countries and has become a global pandemic. Given that the vast majority of patients who show symptoms related to Covid-19 infection end up negative, there is concern that large numbers of uninfected individuals may come into contact with infected patients in testing centers or emergency departments. Therefore, overcrowded hospitals and clinics can present ideal conditions for virus transmission. Thus, minimizing overcrowding in hospital waiting rooms and clinics is essential to reduce nosocomial spread. The purpose of the study is to provide information in the form of predictions of the severity of Covid-19 patients based on information that can be collected, including demographic data, comorbidities, and complaints suffered by using backpropagation neural networks. From the research that has been done, it can be concluded that the backpropagation artificial neural network using the Gradient Descent optimization method provides better performance than the backpropagation artificial neural network using the Fletcher-Revees optimization method, although the difference in results between the two is not too significant.

Keywords: *COVID-19; demographics; artificial neural network; complaint; comorbid.*

1. INTRODUCTION

Coronavirus is a new virus that infects humans, the coronavirus is commonly called Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2 which was identified in Wuhan, China in December 2019. Since then, the coronavirus has spread around the world and affected more than 180 countries. SARS-CoV-2 has infected humans regardless of age group, ethnicity, or gender spreading through communities at an alarming rate. Clinical symptoms of the coronavirus range from the common cold to more severe illnesses such as bronchitis, pneumonia, acute respiratory distress syndrome (ARDS), multi-organ failure, and even death. It has been believed that for those who have an underlying health condition or comorbidity, then the coronavirus will have a faster and more severe development that often causes death. [1] examine the condition of comorbidity, disease development, and mortality rate in patients of all ages, infected with coronavirus disease. Patients who contract the coronavirus who have accompanying diseases such as hypertension or diabetes mellitus are more likely to cause more severe symptoms of the disease. Furthermore, older patients, especially those aged 65 years and over who have accompanying and infected diseases, have a higher chance of admission to the intensive care unit (ICU) and death from coronavirus disease.

A total of 492 (22.5%) patients in Nigeria had at least one accompanying disease and the most common were hypertension (74.2%) and diabetes (30.3%). The patient mortality rate was 3.3% with a significantly higher proportion of patients with accompanying disease compared to those who did not have the disease. Comorbidities that predict death are hypertension with an Odds Ratio of 2.21, diabetes with an Odds Ratio of 3.69, kidney disease with an Odds Ratio of 12.53, cancer with an Odds Ratio of 14.12, and HIV with an Odds Ratio of 1.77 [2].

In Indonesia, confirmed cases of COVID-19 are 28,233 cases, and this occurs until June 3, 2020. The prevalence of 19 cases of COVID-19 is currently 0.11‰ and transmission is widespread throughout the provinces in Indonesia. The highest mortality rate occurred in old age at 17.68%, while almost one-third of COVID-19 infections were in the 31-45 age group (29.3%). In the number of deaths, 6.84% of them were men. The most common symptoms faced by confirmed patients of COVID-19 are Cough which is 76.2%, fever history by 50.4%, and fever is currently 47.1%. Accompanying diseases experienced by patients with COVID-19 include hypertension, diabetes, and other cardiovascular diseases. The

Y. A. Yusran et al. (eds.), *Proceedings of the 2023 Brawijaya International Conference (BIC 2023)*, Advances in Economics, Business and Management Research 294,
https://doi.org/10.2991/978-94-6463-525-6_41

highest accompanying disease is hypertension at 52.1%, then diabetes at 33.6% and the lowest is other cardiovascular diseases at 20.9% [3].

One of the JST forecast calculations that can be utilized is the backpropagation calculation. This calculation can be utilized in fathoming issues related to expectation well. In any case, the execution of the backpropagation calculation is affected by the optimization strategies utilized. For the most part, the strategy that's frequently utilized and has ideal execution is the angle plunge strategy. The impediment of the angle plunge strategy is the moderate joining [4]. In expansion to the slope plunge strategy, the strategy that can be utilized in joining issues is Fletcher Reeves. Fletcher Reeves optimization strategy can fathom joining issues many ways better when compared to the slope plunge strategy [5].

This paper proposes a demonstration of foreseeing the seriousness of COVID-19 sufferers utilizing fake neural arrange backpropagation with the Angle Plummet strategy and the Fletcher-Reeves strategy. Prescient models are obtained through tests by combining organized engineering and learning rates to induce the foremost appropriate forecast show.

2. METHOD

The paper proposes a method for predicting the severity of COVID-19 sufferers using backpropagation neural networks with the gradient descent method and the Fletcher-Reeves method. The proposed flowchart method can be seen in Figure 1. Predictions of the severity of COVID-19 sufferers are made based on several factors from demographic data, complaints, and comorbidity from COVID-19 sufferers.

The first step in the research is data processing through data normalization. Data is normalized by transforming it into ranges 0 and 1. In age data, it is assumed that the age of Covid-19 sufferers is in the range of 0 and 100 years. Gender data is normalized by 0 for women and 1 for men. As for the main complaint data, additional complaints, contact history, and comorbidity are normalized by giving a value of 0 if not present and 1 if any.

Optimization is the process of finding the best solution or optimal value of a problem. This optimization method is used to look for maximum values or minimum values. Optimization problems have been applied in everyday life, such as the management of water resources, medicine, agriculture, economics, and others [6], [7], [8], [9].

Backpropagation (BP) algorithms are commonly used to train artificial neural networks (ANN) [10]. Training is usually done by updating weights repeatedly, usually using a negative gradient of the squared error function. The error signal is the difference between the desired output value and the actual output, explaining the tilt of the sigmoid activation function. This error signal is then propagated back to the bottom. Traditionally two parameters, learning rate (LR) and momentum factor (MF) were used to control weight adjustment along the steepest direction of decline and to dampen oscillations. This BP algorithm is popular and widely used for applications. Unfortunately, the convergence rate is relatively slow, especially for networks with more than one hidden layer. The reason is the saturation of the activation function behaviour used for the network layer. Since the output of a unit is in the saturation area, the corresponding decrease gradient requires a very small value, even if the output error is large, leading to very small advances in weight adjustment.

The backpropagation calculation could be a precise strategy for conducting preparation at the ANN layer and is regularly utilized in fathoming complex issues. It has been utilized in numerous applications, such as precipitation expectation [11] and compound work forecast [12]. The backpropagation calculation has three layers, i.e., input, covered up, and yield layers. A covered-up layer consists of m units and a predisposition. The training process of the Backpropagation has three stages, which are the feedforward step from pattern input training, the Backpropagation of associated errors, and weight updating. During the advanced step, each input unit will be counted in the hidden layers to get the output of the pattern. During the training process, the output from the network will be compared with the target, and then the error is calculated. Subsequently, the optimization is carried out so that the factors that distribute the error are obtained and used for updating the weight between the input layer and the output layer [13].

Table 1. Demographic data, major complaints, additional complaints, contact history, and comorbid Covid-19 sufferers.

No	Age	Gender	Main Complaints	Additional Complaints	Contact History	Sum of Comorbid	Degree of Severity
1	54	Woman	3	0	None	3	Critical Symptoms
2	49	Woman	2	1	None	0	Moderate Symptoms
3	64	Man	2	0	None	5	Moderate Symptoms
4	70	Woman	1	0	None	4	Severe Symptoms
5	16	Man	1	4	None	0	Mild Symptoms
6	40	Woman	0	0	None	4	Moderate Symptoms
7	48	Man	3	0	None	0	Moderate Symptoms
8	56	Man	2	4	None	4	Severe Symptoms
9	49	Man	2	1	None	0	Moderate Symptoms
10	55	Man	2	0	None	0	Moderate Symptoms
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
11	38	Woman	0	0	None	0	No Symptoms

Source: General Hospital Dr. Saiful Anwar Malang

The goal of the study was to find the best architecture and the right learning rate to predict the severity of COVID-19 sufferers. To date, there has been no precise method that can be used to determine the number of neurons in a hidden layer. So far, the number of neurons in the hidden layer is determined experimentally. The number of neurons in the hidden layer was determined based on previous research. In this study, several architectural models were tried to be used for experiments, namely 16-5-1, 16-20-1, 16-50-1, 16-100-1, and 16-150-1. Architectural models 16-5-1 show that neural networks have 16 neurons in the input layer, 5 neurons in the hidden layer, and one neuron in the output layer. The backpropagation neural network consists of three steps, namely the feedforward step, the backpropagation step, and the weight update step. As with the number of hidden neurons, there is currently no precise method to include the learning rate in neural networks. In this study, the speed of learning was determined through experience. This study used several learning methods, namely 0.001, 0.005, 0.01, 0.1 and 0.2.

Table 2. Normalized dataset

No	X1	X2	X3	X4	…	X28	Y
1	0,54	0	1	1	…	0	4
2	0,49	0	1	1	…	0	2
3	0,64	1	0	1	…	0	2
4	0,7	0	0	1	…	0	3
5	0,16	1	1	0	…	0	1
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
155	0,38	0	0	0	…	0	0

Source: processed data

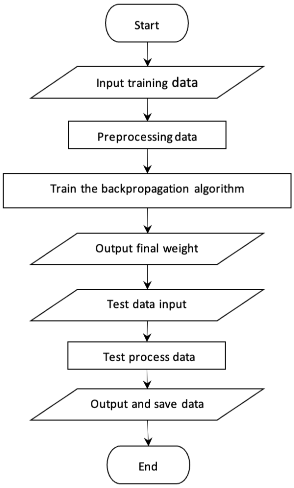

Figure 1 Flowchart of the Backpropagation method.

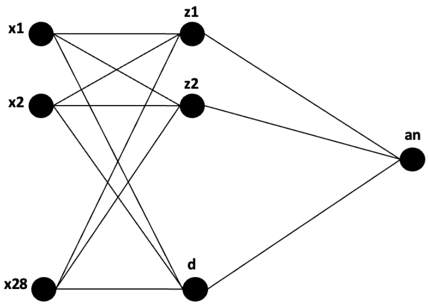

Figure 2 Back Propagation artificial neural network architecture.

3. RESULTS

The results of the program simulation can be seen in Table 3, Table 4, Table 5, and Table 6. To implement the prediction system, the hardware used is a laptop with a 7th Gen Core i7 processor, 2.70 GHz, 2GB of RAM. The programming software is MATLAB.

Input variables in backpropagation artificial neural networks amount to 28 variables consisting of age, gender, main complaints (fever, cough, shortness of breath), additional complaints (weakness, headache, joint pain, throat pain, cold, nausea/ vomiting, diarrhoea, decreased consciousness, anosmia, ageusia), history of contact with Covid-19 sufferers, accompanying disease, diabetes mellitus, heart disease, hypertension, heart failure, coroner heart disease, lung disease, asthma, chronic obstructive pulmonary disease, tuberculosis, kidney disease, and cancer.

Tables 3 and 4 show a performance comparison of the Gradient Descent and Fletcher-Reeves methods for optimization of artificial neural network backpropagation with several different architectures. From these tables, it can be shown that the Gradient Descent method has much better accuracy than the Fletcher-Reeves method. The Fletcher-Reeves method takes longer than the Gradient Descent method for backpropagation neural network architectures with some number of hidden neurons. Tables 3 and 4 are used to show the architectural effects of two backpropagation neural network models. From Tables 3 and 4, the architecture of artificial neural networks or the number of neurons in hidden

Table 3. Performance calculations from the gradient descent method and Fletcher-Reeves using training data for multiple hidden neurons with learning rates of 0.1.

Architecture	MSE		Computing Time	
	Gradient Descent	Fletcher-Reeves	Gradient Descent	Fletcher-Reeves
16-5-1	**0,55**	0,74	1,21	**0,54**
16-20-1	**0,71**	0,79	1,30	**0,55**
16-50-1	**0,75**	0,81	2,26	**0,84**
16-100-1	**0,79**	0,81	3,26	**1,22**
16-150-1	**0,82**	0,83	3,73	**1,78**

Table 4. Performance calculations from the gradient descent method and Fletcher-Reeves using test data for multiple hidden neurons with learning rates of 0.1.

Architecture	MSE		Computing Time	
	Gradient Descent	Fletcher-Reeves	Gradient Descent	Fletcher-Reeves
16-5-1	**0,55**	0,74	1,21	**0,54**
16-20-1	**0,71**	0,79	1,30	**0,55**
16-50-1	**0,75**	0,81	2,26	**0,84**
16-100-1	**0,79**	0,81	3,26	**1,22**
16-150-1	**0,82**	0,83	3,73	**1,78**

layers also affects the performance of artificial neural networks backpropagation. Tables 3 and 4 show that if the number of hidden neurons in the backpropagation neural network algorithm using the Gradient Descent method and the Fletcher-Reeves method increases then means square error (MSE) resulting from backpropagation neural networks using both methods also increase in test data. However, in general, the performance of backpropagation neural networks using the Gradient Descent method is better than the Fletcher-Reeves method's backpropagation neural network. The results of the experiment showed that the number of hidden neurons as many as 5 neurons gave the best results. So that the next number of neurons to be used in determining the learning rate is 5 neurons. Table 3 also shows the computing time of both methods. The backpropagation neural network algorithm with the Gradient Descent method has a faster computational time than the backpropagation neural network algorithm with the Fletcher-Reeves method. In both methods, an increase in the number of neurons influenced the increase in computing time required for training.

Table 5 and 6 presented an evaluation of the performance of the backpropagation neural network with the Gradient Descent method and the Fletcher-Reeves method to predict the severity of COVID-19 sufferers with different learning rates. The table shows that the best learning rate is 0.1. The best MSE produced by a backpropagation neural network with the Gradient Descent method for training data is 0.76, and for test data is 0.63. The best MSE produced by a backpropagation neural network with the Fletcher-Reeves method for training data was 0.79, and for test data of 0.64.

Table 5. Performance calculations from gradient descent and Fletcher-Reeves methods using training data for multiple learning rates with 5 hidden neurons.

Learning rate	MSE		Computing Time	
	Gradient Descent	Fletcher-Reeves	Gradient Descent	Fletcher-Reeves
0,001	**0,77**	0,79	2,11	**1,15**
0,005	**0,76**	0,81	2,12	**0,92**
0,01	**0,76**	0,81	2,07	**0,92**
0,1	**0,76**	0,80	2,15	**0,91**
0,2	**0,76**	0,81	2,11	**0,91**

Gradient Descent method has a faster computational time than the Backpropagation neural network algorithm with the Fletcher-Reeves method. In both methods an increase in the number of neurons resulted in an increase in the computing time required for training. Learning rate of 0.1 provides the best results. Overall, the backpropagation neural system with gradient descent optimization method provides better results compared to backpropagation neural networks with Fletcher-Reeves optimization method to predict the severity of COVID-19 sufferers.

Table 6. Performance calculations from gradient descent and Fletcher-Reeves methods using test data for multiple learning rates with 5 hidden neurons.

Learning rate	MSE	
	Gradient Descent	Fletcher-Reeves
0,001	0,65	**0,64**
0,005	**0,64**	0,67
0,01	**0,63**	0,65
0,1	**0,63**	0,64
0,2	**0,63**	0,65

4. DISCUSSION

From the results of experiments and discussions the performance of artificial neural networks backpropagation depends on the number of neurons in hidden layers (hidden neurons) and optimization algorithms for learning. When the number of hidden neurons is too much, the method's generalization ability decreases. From the number of neurons in the selected hidden layer, the appropriate learning rate can be selected. The simulation results showed that the algorithm of backpropagation neural networks with Gradient Descent method has a faster computational time than the Backpropagation neural network algorithm with the Fletcher-Reeves method. In both methods an increase in the number of neurons resulted in an increase in the computing time required for training. Learning rate of 0.1 provides the best results. Overall, the backpropagation neural system with gradient descent optimization method provides better results compared to backpropagation neural networks with Fletcher-Reeves optimization method to predict the severity of COVID-19 sufferers.

AUTHORS' CONTRIBUTIONS

The title "AUTHORS' CONTRIBUTIONS" should be in all caps.

ACKNOWLEDGMENTS

Thank you to LPPM Universitas Brawijaya and Prof. Akhmad Sabarudin for the advice and the discussion.

REFERENCES

[1] A. Sanyaolu, C. Okorie, A. Marinkovic, R. Patidar, K. Younis, P. Desai, Z. Hosein, I. Padda, J. Mangat and M. Altaf, Comorbidity and its impact on patients with COVID-19. SN comprehensive clinical medicine, 2020, 2, pp.1069-1076. DOI: https://doi.org/10.1007/s42399-020-00363-4

[2] A. Osibogun, M. Balogun, A. Abayomi, J. Idris, Y. Kuyinu, O. Odukoya, Akinroye, Outcomes of COVID-19 patients with comorbidities in southwest Nigeria, PloS one, 2021 March, 1 6(3). DOI: https://doi.org/10.1371/journal.pone.0248281

[3] D.R. Karyono, A.L. Wicaksana, Current prevalence, characteristics, and comorbidities of patients with COVID-19 in Indonesia, J Community Empowerment Health. 2020 Jul, 3(2):77-84. DOI: https://doi.org/10.22146/jcoemph.v3i2.60

[4] F.D. Marleny, Comparison of Conjugate gradient and Gradient Descent Algorithms on MLPNN for Maternal Knowledge Level. Proceedings of the National Conference on Systems & Informatics, 2017, pp. 507–512.

[5] T.M. Bafitlhile, Z. Li, Q. Li, Comparison of Levenberg Marquardt and Conjugate Gradient Descent Optimization Methods for Simulation of Streamflow Using Artificial Neural Network, Adv in Eco and Envi Research, 2018, vol. 3, pp. 217–237.

[6] D.V. Morankar, K.S. Raju and D. N. Kumar, Integrated Sustainable Irrigation Planning with Multiobjective Fuzzy Optimization Approach, Water Resource Management, vol. 27, pp. 3981–4004, 2013. DOI 10.1007/s11269-013-0391-3

[7] P. Schulthess, V. Rottschäfer, J.W.T. Yates, and P.H. van der Graaf, Optimization of Cancer Treatment in the Frequency Domain, The AAPS Journal, 2019, vol. 21, no. 106 p.m. 1-9. DOI: 10.1208/s12248-019-0372-4

[8] S.D. Kumara, S. Esakkirajan, C. Vimalraj, and BK Veena, Whale Optimization-Based Disease Prediction Method Design Employing Neural Networks in Tomatoes, Today's Material: Proceedings, 2020. DOI: https://doi.org/10.1016/j.matpr.2020.08.450

[9] S. Zhu, X. Hu, K. Huang, and Y. Yuan, Optimization of Product Category Allocation in Multiple Warehouses to Minimize the Separation of Online Supermarket Customer Orders, European Journal of Operational Research, 2020. DOI: https://doi.org/10.1016/j.ejor.2020.08.024

[10] D.E. Rumelhart, J.L. McClelland, Parallel Distributed Processing: Explorations in the Microstructure of Cognition, 1986, Vol. I, MIT Press, Cambridge, MA.

[11] Y.H. Zweiri, J.F. Whidborne, L.D. Seneviratne. A three-term backpropagation algorithm. Neurocomputing, 2003, 50, 305-318. DOI: https://doi.org/10.1016/S0925-2312(02)00569-6

[12] Ampazis N, Perantonis SJ, Taylor JG. Dynamics of multilayer networks in the vicinity of temporary minima. Neural Networks, 1999, 12, 43–58. DOI: https://doi.org/10.1016/S0893-6080(98)00103-8

[13] S. Anam, M.H.A.A Maulana, N. Hidayat, I. Yanti, Z. Fitriah, and D.M. Mahanani, Predicting the number of COVID-19 sufferers in Malang city using the backpropagation neural network with the Fletcher–Reeves method. Applied Computational Intelligence and Soft Computing, 2021, pp.1-9. DOI: https://doi.org/10.1155/2021/6658552