

Int_2020_J._Phys.__Conf._Ser._1 655_012096_URICSE_Santoso_d kk.pdf *by*

Submission date: 12-Apr-2023 12:28PM (UTC+0700)

Submission ID: 2062296848

File name: Int_2020_J._Phys.__Conf._Ser._1655_012096_URICSE_Santoso_dkk.pdf (1.3M)

Word count: 3977

Character count: 20402

PAPER • OPEN ACCESS

Prediction of Jakarta Composite Index Using Neural Network Model and Genetic Optimization

13 To cite this article: R Santoso *et al* 2020 *J. Phys.: Conf. Ser.* **1655** 012096

View the [article online](#) for updates and enhancements.

You may also like

- Adaptive prediction of respiratory motion
2 motion compensation radiotherapy
Qing Ren, Seiko Nishioka, Hiroki Shirato
et al.

- **19** Multiple model approach to respiratory
motion prediction for real-time IGRT
Devi Putra, Olivier C L Haas, John A Mills
et al.

- Prediction of Indian rainfall during the
14 mer monsoon season on the basis of
links with equatorial Pacific and Indian
Ocean climate indices
Sajani Surendran, Sulochana Gadgil, P A
Francis et al.



*Benefit from connecting
with your community*

ECS Membership = Connection

ECS membership connects you to the electrochemical community:

- Facilitate your research and discovery through ECS meetings which convene scientists from around the world;
- Access professional support through your lifetime career;
- Open up mentorship opportunities across the stages of your career;
- Build relationships that nurture partnership, teamwork—and success!

Join ECS!

Visit electrochem.org/join



Prediction of Jakarta Composite Index Using Neural Network Model and Genetic Optimization

R Santoso¹, B Warsito^{1,2}, and H Yasin¹

¹Department of Statistics, Faculty of Science and Mathematics, Diponegoro University Semarang Indonesia

²Master Program of Information System, School of Postgraduate Studies, Diponegoro University Semarang Indonesia

budiwrst2@gmail.com

Abstract. Researches related to prediction of stock price data have been developing rapidly. Likewise, the modeling techniques used for predictive purposes are also increasing along with advances in the field of computing. This study applied neural network model in predicting the Jakarta Composite Index data as a case of time series. The optimization method used was genetic algorithm. This method is included in one of the heuristic techniques. Unlike standard optimization methods, genetic algorithms do not use gradients as a basis for search techniques. Parameters in the neural network model are obtained from the process of decoding chromosomes from generation to generation. In comparison, the two gradient-based optimization methods were also applied, i.e Conjugate Gradient and Gradient Descent. The results showed the superiority of genetic algorithms compared to other optimization methods in out-sample prediction whereas, the in-sample prediction of gradient-based optimization methods achieve better precision.

1. Introduction

The Jakarta Composite Index (JSX) is an index of all stocks that are traded on the Indonesia Stock Exchange (IDX). In Indonesian the index is called the Indeks Harga Saham Gabungan, abbreviated IHSG. Composite Index (CI) depends on method of combining several variables or indicators to reflect overall assessment. Each method of combining the component indicators results in different values of CI and different rankings from a given dataset [1]. The Jakarta Composite Index is a modified capitalization-weighted index of all stocks listed on the regular board of the Indonesia Stock Exchange. The index was developed with a base index value of 100 as of August 10, 1982. Predicting stock market crashes is a focus of interest for both researchers and practitioners [2]. The prediction of the trends of stocks and index prices is one of the important issues to market participants. Investors have set trading or fiscal strategies based on the trends, and considerable research in various academic fields has been studied to forecast financial markets [3]. Composite Index forecast has important significance for a country in latter regulation and control of the real economy, and the development of regional policy [4]. Many scholars study the fluctuation of stock price. There are many analytical tools available, each with its own strengths and limitations [5]. Among the various research methods they adopt, methods of time series analysis are most widely used. In the various parametric models of time series analysis, the autoregressive integrated moving average (ARIMA) model is the most fitted to forecast the future trend of stock prices [6,7]. However, the severe limitation of these models is the pre-assumed linear form of the associated time series which becomes inadequate in many practical situations [8].



Content from this work may be used under the terms of the Creative Commons Attribution 3.0 licence. Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

Published under licence by IOP Publishing Ltd

21

Neural network is an intelligent data mining method. It has been used in many different challenging pattern recognition problems such as stock market prediction [9]. Neural network forms a useful tool in predicting price movement of a particular stock. It can learn the pricing relationship to high precision and be used to make profits with sufficiently large amounts of data. The case of data with low volatility and over a short time period is preferably [10]. Moghaddam et al [11] has investigated the ability of neural network in forecasting the daily NASDAQ stock exchange rate. Devianto et al [12] has also developed a forecast model of the value of the Indonesian composite stock price index using neural network and MARS regression. Several studies related to the use of neural networks for prediction of stock prices have also been developed. One of the main problem in neural network modelling is making algorithm start running and converge to a reasonable solution such as a stationary point [13]. Conventional optimization algorithms rely on stochastic gradient methods that don't scale well to large numbers of cores in a cluster setting [14]. In this research, the using of genetic algorithm for optimizing neural network parameters. Genetic algorithm is a robust adaptive optimization method based on biological principles. It need not search along the contours of the function being optimized and tend not to become trapped in local minima [15]. A combination between neural network modelling and genetic algorithm as optimization method is expected to improve the training process so that more perfect predictions can be obtained. This procedure is applied to the Jakarta Composite Index.

2. Methods

The closing price of daily data of the Jakarta Composite Index was used as case study. It was taken from 05 August 2019 until 03 August 2020. In this experiment, we used about 80% as training data or in-sample prediction and the about 20% remaining as testing data or out-sample prediction. Through the use of training data sets, network architecture is built first. The main goal of prediction process is determining of a set of parameters that can optimize the model so that it is able to produce a predictive value that is as close as possible to the actual. In this case, the predictive model used is neural network. The model based on the algorithm in the work system of these creatures is juxtaposed with the optimization method which is also based on the work system in living things, namely genetic algorithm. It used as the way to obtain the optimal parameters of neural network model.

2.1. Neural Network

Architecture of Neural network for time series prediction is divided into three processing layers. The first is input layer, i.e the past value of the data or often also called lagged time. The number of lags can be determined by using various methods, and for simplicity the partial autocorrelation function was used. The input layer sends the weighted signal to the second layer, which is called the hidden layer. In this layer, a nonlinear function is applied to the incoming signal. The nonlinear function transfers the weighted signal to a small value, usually between 0 and 1. This value is expected to be in accordance with the output which is a continuous data. In this research, logistic sigmoid was used as activation function. The processing is then continued with the weighted sum from the output of the hidden layer to the third layer, i.e output layer. Linear function is applied in the output layer so that the output is in the form of continuous data. This is the predicted value of the model. It is expected that the values match or approach the actual data. To achieve this goal, a number of iterations, or in neural network terminology called epoch, are applied in order to obtain optimal weights that produce the best predictions. In this study, the maximum epoch chosen was 1000. This serves as the criteria for stopping iteration. The architecture of neural network for time series modelling as explained above can be seen in Figure 1.

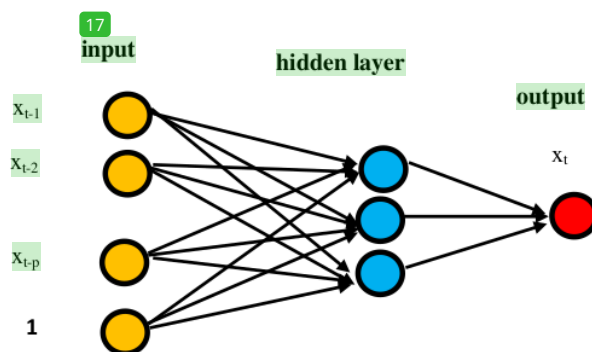


Figure 1. The architecture of neural network modelling for time series

Parameters of neural network model are also often called weights. They can also be divided into certain parts in the form of vectors. The first part is the weight vector from the input to the hidden layer. Models often also add a bias to the input layer, so that weights are also included in the connection between the bias and the hidden layer. The second part indicates the weights of the connection between the hidden layer and the output layer. In addition, there is also a weight between bias and output. The estimated weight of the neural network model is $\{k \times (p + 2)\} + 1$, where k is the number of neurons in the hidden layer and p is lagged time. Unlike standard algorithms on optimizing neural network models that use gradient-based methods, this research uses genetic algorithm which is one of the metaheuristic optimization.

2.2. Genetic Optimization

Genetic algorithm is inspired by the workings of genes in living things. Genetic algorithms originate from theories that state creatures that can adapt better to their environment will have a greater chance of survival and reproduction. Individuals who successfully place themselves in their environment better than others will be stronger and have a greater chance to survive and provide these advantages to the next generation. The concept is translated into an algorithm to find solutions of a problem in a more natural way. The underlying principles of GA are to generate an initial population of chromosomes (search solutions) and then use selection and recombination operators generate a new, more effective population which eventually will have the fittest chromosome (optimal value) among them [16]. The steps of operation of neural network and genetic algorithm hybrid intelligence are as follows.

- 1) The first step is initial population, which begins by representing the solution on the chromosomes. Each chromosomes contains bit strings of randomly generated binary values. The population sizes and chromosome used were 10 and 20, respectively.
- 2) Each parameter in the chromosome which is initialized by binary encoding, is converted into a real value, namely decoding.
- 3) Running neural network model to make prediction of closing price of daily data of Jakarta Stock Exchange Composite Index. The parameters in the model were the real value of decoding at the step 2.
- 4) Fitness evaluation i.e. calculates the fitness value of each chromosome. In this step, the fitness value of each chromosome is prediction accuracy from neural network model.
- 5) Determine the stopping criterion, whether the loop is continued or stopped. The maximum generations as stopping criterion was 1000.
- 6) Elitism is copying and storing a few best chromosomes obtained from the individual evaluation process.
- 7) Linear Fitness Ranking (LFR), which scales fitness values obtained from individual evaluations. LFR is done to avoid the tendency to achieve convergence in local optimal solutions by obtaining new fitness values that have greater variance.

- 8) Selection, which is selecting two chromosomes to make a pair of parent. The method used is the Roulette Wheel Selection.
- 9) Crossover, which is applying an arithmetic crossover operator. It is a linear combination of two chromosomes.
- 10) Mutation, which is to change one of the gene codes on a chromosome into its inverse.
- 11) General replacement, which is replacing the initial population members resulting from initialization with a new population consisting of chromosomes resulting from elitism, crossovers and mutations. The old chromosomes is replaced with two best offspring chromosomes for the next generation.
- 12) Go to Step 2.

The steps can be summarized as in Figure 2.

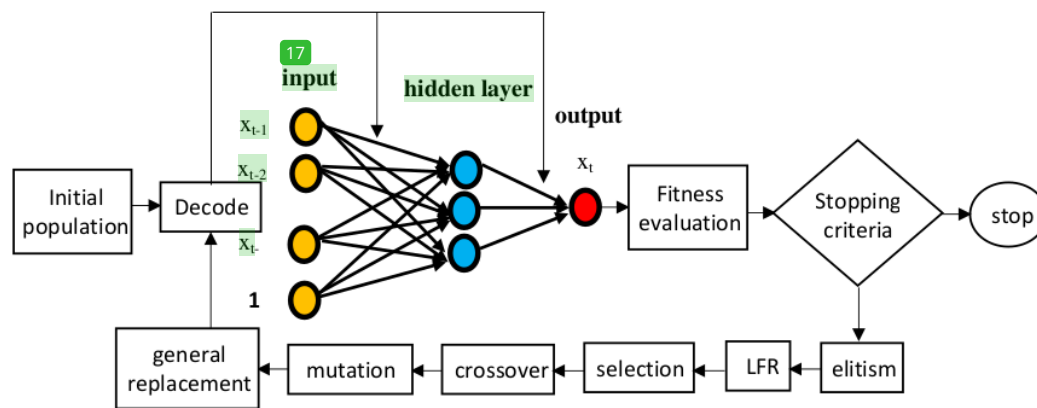


Figure 2. Genetic optimization for optimizing neural network

In each generation, we used accuracy to determine chromosome selection as well as to measure the performance of the prediction model. Fitness values in genetic algorithm were taken as the accuracy values that can be determined by using Mean Square Error (MSE).

3. Results and Discussion

In the first step, we investigated the lags of series as input. Plot of Partial Autocorrelation Function (PACF) was used as a tool for supporting the investigations. The result is that the input of the neural network model were the two variables in lags 1 and 2. For selecting the optimal architecture, the number of hidden units were chosen from 1-4. This is not so different with the number of input. In the proposed procedure, three scenarios are built. In the first scenario, genetic algorithm optimization with 1000 generations is applied. In the second and third scenarios, Conjugate Gradient (CG) and Gradient Descent (GD) algorithms were used for optimizing neural network, respectively. The maximum number of epochs in the two gradient based methods are 1000. In each scenario, the experiments were repeated ten times and the statistics of minimum, maximum, average and variance were obtained. The experimental results are presented in Table 1. Comparison of the three optimization methods can be determined by paying attention to the average and stability of each result which can be seen from the variance.

Table 1. Results of RMSE and MAPE from various scenarios

Accuracy	Optimizer	Hidden units	training				testing			
			min	max	average	var	min	max	average	var
RMSE	NN-GA	1	75.7307	127.2967	90.5514	236.2249	60.8163	144.3706	82.4166	738.6770
		2	78.3395	132.2015	94.5462	353.4831	60.3000	141.8232	82.0060	805.7304
		3	79.2370	107.9231	90.5287	67.0948	61.6970	111.3617	78.3372	206.8492
		4	76.9866	111.6972	85.0733	110.9197	61.1160	125.6663	73.6610	395.2763
	NN-GD	1	92.8120	93.9645	93.4972	0.1816	191.2123	204.0670	199.6590	30.0670
		2	85.1376	123.1075	98.2890	193.5608	66.1208	339.4845	188.2846	9694.0654
		3	92.7456	114.0094	104.6669	55.8777	202.9918	291.3705	244.7518	637.0742
		4	88.8440	1.006498	94.4125	16.0772	64.9216	76.8563	68.2739	19.6029
	NN-CG	1	74.4170	74.5663	74.4754	0.0022	91.1420	96.0560	93.0129	2.0741
		2	73.1392	344.3504	104.4107	7220.315	89.1142	320.3023	122.9736	5024.988
		3	71.2875	130.0188	78.0114	335.1440	81.0130	168.0316	120.4731	684.0028
		4	69.8874	130.1670	79.8548	407.9098	86.2578	205.8264	126.0291	1084.176
MAPE	NN-GA	1	0.9883	1.8517	1.2134	0.0710	0.9170	2.6756	1.3774	0.3204
		2	1.0111	1.7804	1.2658	0.0933	0.9138	2.3428	1.3475	0.2880
		3	1.0151	1.5500	1.1886	0.0221	0.9581	1.9652	1.2908	0.0868
		4	0.9997	1.5542	1.1212	0.0288	0.9227	2.2546	1.1929	0.1675
	NN-GD	1	1.1540	1.1727	1.1610	6.0216	3.6672	3.9306	3.8404	0.0126
		2	1.1200	1.4153	1.2194	0.0115	1.0833	6.5593	3.5187	3.8968
		3	1.1291	1.3427	1.2514	0.0052	3.8462	5.6687	4.6966	0.2720
		4	1.1572	1.2985	1.2333	0.0028	1.0697	1.2871	1.1369	0.0061
	NN-CG	1	0.9630	0.9648	0.9638	3.7361	1.5624	1.6719	1.6022	0.0010
		2	0.9469	5.2718	1.4153	1.8434	1.5426	6.2448	2.1957	2.0869
		3	0.9342	1.5282	1.0031	0.0341	1.3784	2.8419	2.1540	0.2292
		4	0.9123	1.8668	1.0625	0.0978	1.3188	3.9406	2.2388	0.4880

Results of Table 1 are focussed on the two measurements of accuracy. The RMSE provides hints that genetic algorithm and conjugate gradient give similar accuracy in the in-sample prediction. However, when combined with the out-sample prediction the genetic algorithm appears to be the best, with a thin difference. In the results of gradient descent optimization, the out-sample predictions are often unstable. In two and four hidden units, the results obtained are relatively close to the best of genetic algorithm results. However, in one and three hidden units the results are very much different. Architecture with the four neurons at hidden layer and genetic optimization is recommended for being chosen as optimal results. This result is similar with Göçken et al. [17], genetic algorithm given a better result than regular optimization for out-sample prediction when using RMSE criteria. Slightly different results occur at the in-sample prediction, because regular method is better.

The focus of the discussion continues on the accuracy of using MAPE. Overall, the results of in-sample predictions obtained on accuracy using MAPE are in the range of 1-2 percent. In the in-sample prediction, genetic algorithm and gradient descent given similar results in both range and stability. Conjugate gradient also given rather similar results, except the two hidden units. Significantly different results were obtained in the out-sample prediction. Genetic optimization seems to give better and more stable results compared to the two gradient-based methods. Once again, architecture with the four neurons at hidden layer and genetic optimization is recommended with high confidence. A similar

result can be found at [18], where the genetic optimization give better results than the algorithms based on the error gradient comp¹¹tion. To see the results in more detail, a visual representation is presented. Plots of the actual vs in-sample prediction by using training data and the actual vs out-sample prediction by using testing data can be seen in Figure 3.

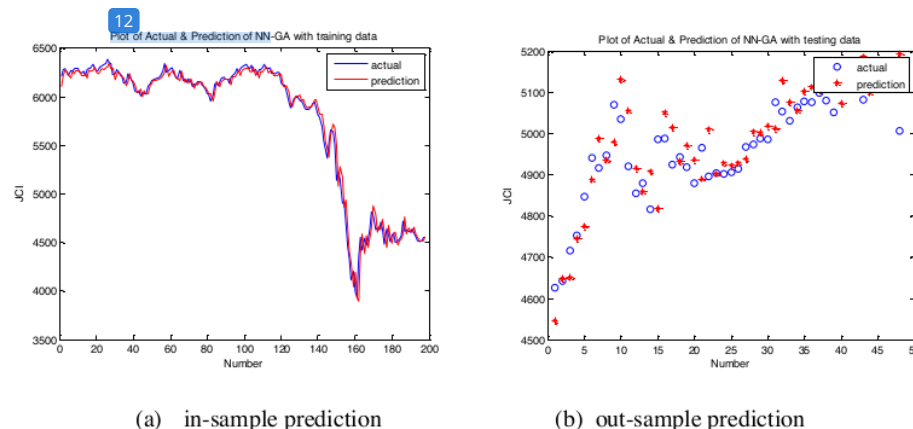


Figure 3. Actual vs prediction of in-sample and out-sample data from NN-GA model

³In the two plots, the blue line showed the actual whereas, the red line was the in-sample and out-sample predictions. In the training data, we can see that the predicted model succesfull of approaching the actual although the pattern go down rapidly in a few periods and extremly up again also in a short periods. The steep decline was caused by the condition of the co-19 pandemic that occurred in Indonesia and in most of the world. During that period the Indonesian government implemented a large-scale social restrictions policy. This condition has an impact on the value of the stock price index. In the following period, policy easing was implemented which impacted on increasing socio-economic activities and again increasing the index of stock prices. The proposed procedure has also given a high precision in the out-sample prediction. The rise and fall of data patterns can be approached well by the model. A few points are in a position somewhat far from the actual data however, in general, the model successfully predicts the testing data well.

4. Conclusion

⁷Utilization of genetic algorithm in neural network modelling for financial time series was developed. The proposed procedure has been appied in the closing price of daily data of Jakarta Stock Exchange Composite Index. This has also been compared with several gradient based methods. The superiority of genetic optimization indicated its performance. Unfortunately, the time needed for getting the optimal solutions cause a serious problem from the using of this algorithm. Development of the hybrid procedure as a combination with gradient methods or other metaheuristic algorithms been the interesting future works.

Acknowledgement

This research was funded by the Directorate of Research and Community Service, Deputy for Strengthening Research and Development, Ministry of Research and Technology / National Agency for Research and Innovation with contract number 257-21 / UN7.6.1 / PP / 2020.

References

- [1] Chakrabartty S N 2017 Composite index: Methods and properties *Journal of Applied Quantitative Methods* **12**(2) 25-33

- [2] Lleo S and Ziemba W T 2019 A tale of two indexes: predicting equity market downturns in China Available at SSRN 2698422
- [3] Pyo S, Lee J, Cha M and Jang H 2017 Predictability of machine learning techniques to forecast the trends of market index prices: Hypothesis testing for the Korean stock markets *PloS one* **12(11)** e0188107
- [4] Li C, Yang B and Li M 2017 Forecasting Analysis of Shanghai Stock Index Based on ARIMA Model In *MATEC Web of Conferences* **100** 02029 EDP Sciences
- [5] Beard E, Marsden J, Brown J, Tombor I, Stapleton J, Michie S and West R 2019 Understanding and using time series analyses in addiction research *Addiction* **114(10)** 1866-1884
- [6] Haijian W and Qianqian L 2017 Empirical Study on Shanghai Composite Index Forecast Based on ARIMA Model *Journal of World Economic Research* **6(6)** 71-74
- [7] Ariyo A A, Adewumi A O and Ayo C K 2014 Stock price prediction using the ARIMA model In 2014 UKSim-AMSS 16th International Conference on Computer Modelling and Simulation 106-112
- [8] Adhikari R and Agrawal R K 2013 An introductory study on time series modeling and forecasting *arXiv preprint arXiv:1302.6613*
- [9] Naeini M P, Taremi H and Hashemi H B 2010 Stock market value prediction using neural networks In 2010 international conference on computer information systems and industrial management applications (CISIM) 132-136
- [10] Jabin S 2014 Stock market prediction using feed-forward artificial neural network *International Journal of Computer Applications* **99(9)** 4-8
- [11] Moghaddam A H, Moghaddam M H and Esfandiyari M 2016 Stock market index prediction using artificial neural network *Journal of Economics, Finance and Administrative Science* **21(41)** 89-93
- [12] Devianto D, Permathasari P, Yollanda M and Ahmad A W 2020 The Model of Artificial Neural Network and Nonparametric MARS Regression for Indonesian Composite Index IOP Conf. Series: Materials Science and Engineering **846** (2020) 012007 IOP Publishing doi:10.1088/1757-899X/846/1/012007
- [13] Sun R 2019 Optimization for deep learning: theory and algorithms *arXiv preprint arXiv:1912.08957*
- [14] Taylor G, Burmeister R, Xu Z, Singh B, Patel A and Goldstein T 2016 Training neural networks without gradients: A scalable admm approach *International conference on machine learning* 2722-2731
- [15] Whitley D, Starkweather T and Bogart C 1990 Genetic algorithms and neural networks: Optimizing connections and connectivity *Parallel computing* **14(3)** 347-361
- [16] Inthachot M, Boonjing V and Intakosum S 2016 Artificial neural network and genetic algorithm hybrid intelligence for predicting thai stock price index trend *Computational intelligence and neuroscience*
- [17] Göçken M, Özçalıcı M, Boru A and Dosdoğru A T 2016 Integrating metaheuristics and artificial neural networks for improved stock price prediction *Expert Systems with Applications* **44** 320-331
- [18] Halloua H, Elhassnaoui A, Saifi A, Elamiri A, Obbadi A, Errami Y and Sahnoun S 2017 Neural networks and genetic algorithms for the evaluation of coatings thicknesses in thermal barriers by infrared thermography data *Procedia Structural Integrity* **5** 997-1004

ORIGINALITY REPORT

23%
SIMILARITY INDEX

15%
INTERNET SOURCES

21%
PUBLICATIONS

6%
STUDENT PAPERS

PRIMARY SOURCES

1 Montri Inthachot, Veera Boonjing, Sarun Intakosum. "Artificial Neural Network and Genetic Algorithm Hybrid Intelligence for Predicting Thai Stock Price Index Trend", Computational Intelligence and Neuroscience, 2016 **3%**
Publication

2 Jiejie Dai, Hui Song, Xiaoqi Wan, Yingjie Yan, Gehao Sheng, Xiuchen Jiang. "A Prediction Method for Power Transformer State Parameters Based on Grid Long Short-Term Memory Network", Journal of Physics: Conference Series, 2019 **3%**
Publication

3 www.scik.org **1%**
Internet Source

4 article.sciencepublishinggroup.com **1%**
Internet Source

5 www.readkong.com **1%**
Internet Source

6	www.science.gov Internet Source	1 %
7	Yasin Hasbi, Warsito Budi, Santoso Rukun. "Feed Forward Neural Network Modeling for Rainfall Prediction", E3S Web of Conferences, 2018 Publication	1 %
8	Submitted to KDI School of Public Policy and Management Student Paper	1 %
9	www.jaqm.ro Internet Source	1 %
10	www.indonesia-investments.com Internet Source	1 %
11	B Warsito, H Yasin, A Prahutama. "Particle swarm optimization versus gradient based methods in optimizing neural network", Journal of Physics: Conference Series, 2019 Publication	1 %
12	Budi Warsito, Hasbi Yasin, Rukun Santoso. "Incremental Contribution Method to Determine the Size of Neural Network Optimized by Genetic Algorithm", Asian Journal of Scientific Research, 2020 Publication	1 %

13	R Dipinto, R Santoso, A Prahutama. "The feed forward neural network with genetic algorithm for daily stock prediction", Journal of Physics: Conference Series, 2019 Publication	1 %
14	Thanh Le, Kyung-Ja Ha, Deg-Hyo Bae. "Projected response of global runoff to El Niño-Southern oscillation", Environmental Research Letters, 2021 Publication	1 %
15	Submitted to University of Limerick Student Paper	1 %
16	pdfs.semanticscholar.org Internet Source	1 %
17	ndl.ethernet.edu.et Internet Source	1 %
18	www.semanticscholar.org Internet Source	1 %
19	Ying Huang, Xiaoyan Huang, Huasheng Zhao, Jian Jin. "The influence of environmental factors on the intensity of typhoon Mirinae in the South China Sea and its forecast analysis", IOP Conference Series: Earth and Environmental Science, 2021 Publication	1 %

20

D Whitley, T Starkweather, C Bogart. "Genetic algorithms and neural networks: optimizing connections and connectivity", Parallel Computing, 1990

Publication

1 %

21

www.meltdata.com

Internet Source

1 %

Exclude quotes On

Exclude matches < 1 %

Exclude bibliography On