Date of publication JUN-28, 2021, date of current version JAN-12, 2021.
www.computingonline.net / computing@computingonline.net

Print ISSN 1727-6209
Online ISSN 2312-5381
DOI 10.47839/ijc.20.2.2165

Genetic Algorithmised Neuro Fuzzy System for Forecasting the Online Journal Visitors

WAYAN FIRDAUS MAHMUDY[1], AJI PRASETYA WIBAWA[2], NADIA ROOSMALITA SARI[3], HAVILUDDIN[4], PURNAWANSYAH[5]

[1] Faculty of Computer Science, Brawijaya University, Malang, East Java, Indonesia (e-mail: wayanfm@ub.ac.id)
[2] State University of Malang, East Java, Indonesia, (e-mail: aji.prasetya.ft@um.ac.id)
[3] Faculty of Economics and Business Sharia, Tulungagung State Islamic Institute, East Java, Indonesia, (e-mail: nadiaroosmalitasari@gmail.com)
[4] Faculty of Computer Science and Information Technology, Universitas Mulawarman, Samarinda, East Kalimantan, Indonesia, (e-mail: haviluddin@unmul.ac.id)
[5] Faculty of Computer Science, Universitas Muslim Indonesia, Makassar, Indonesia, (e-mail: purnawansyah@gmail.com)

Corresponding author: Wayan Firdaus Mahmudy (e-mail: wayanfm@ub.ac.id).

ABSTRACT Artificial Neural Network (ANN) is recognized as one of effective forecasting engines for various business fields. This approach fits well with non-linear data. In fact, it is a black box system with random weighting, which is hard to train. One way to improve its performance is by hybridizing ANN with other methods. In this paper, a hybrid approach, Genetic Algorithm-Neural Fuzzy System (GA-NFS) is proposed to predict the number of unique visitors of an online journal website. The neural network weight is precisely determined using GA. Afterwards, the best weight has been used for testing data and processed using Sugeno Fuzzy Inference System (FIS) for time-series forecasting. Based on experiment, GA-NFS have been produced accuracy with 0.989 of root mean square error (RMSE) that is lower than the RMSE of a common NFS (2,004). This may indicate that the GA based weighting is able to improve the NFS performance on forecasting the number of journal unique visitors.

KEYWORDS Scientific Journal Online; Visitors; Genetic Algorithm (GA); Neural Fuzzy System (NFS); Root Mean Square Error (RMSE).

List of Abbreviations

ARIMA: Auto Regressive Integrated Moving Average
ARMA: Autoregressive Moving Average
ANFIS: Adaptive Neural Fuzzy Inference System
ANN: Artificial Neural Network
BPNN: Backpropagation Neural Network
ETS: Exponential smoothing models from innovation state space
GA: Genetic Algorithm
NARX: Nonlinear Auto-Regressive models with eXogenous inputs
OT: Orthogonal Test
RBF: Radial Basis Function
SARIMA: Seasonal Autoregressive Integrated Moving Average
SVM: Support Vector Machine

I. INTRODUCTION

WITH the growing use of the Internet throughout the world, the number of available websites has also increased. In terms of statistics, the number of visitors and content accessed by users on a site shows the performance of the site. The significance growing number of visitors requires a good resource allocation for the website server. Thus, a forecasting is required to develop an appropriate plan for upgrading the server, power management, and efficient utilization of networking equipment [1]. The forecasting may also be required to estimate a possible revenue from advertising [2].

Several studies related to the forecasting of website traffic were done, most of them used historical data of visitors [2–4]. Various approaches were also applied such as recurrent neural network (RNN) [4], ARIMA [5], ARIMA and LSTM RNN [3], and prophet time series model [2]. Even though there were many studies addressing website traffic forecasting, to the best of our knowledge, there was no study focusing on online scientific journals' visitors.

For online scientific journals, the visitor's traffic may become a benchmark for the success. A large number of visitors may indicate the availability of good collections, play a role in increasing visibility, prestige, and access to scientific papers by readers or visitors. Furthermore, the number of visitors may become a parameter for national journal accreditation.

Forecasting a number of online journal website visitors in the future is needed to review journal management policies and improve the visibility and quality of interface design of journal sites. Thus, an implementation of forecasting methods may become an important part of online journal management.

Numerous forecasting methods using intelligent algorithms such as Backpropagation Neural Network (BPNN), fuzzy logic, Adaptive Neural Fuzzy Inference System (ANFIS), Radial Basis Function (RBF), Support Vector Machine (SVM), and Genetic Algorithm (GA) were developed. The BPPN was implemented for inflation rate prediction [6], pollutant emissions prediction [7], and water simulation [8]. Fuzzy logic based methods were used to address uncertainty in time series data such as wind and wave climate [9], short-term electrical load [10], and long-term electrical load [11].

ANFIS that was developed by incorporating neural networks and fuzzy logic was implemented to address the complexity of atmospheric turbidity forecasting [12], laser cutting roughness prediction [13], and short-term wind and wave conditions prediction [14]. Furthermore, a combination of ANFIS and RBS is developed to predict biodiesel density [15]. SVM that was well known as a powerful method for regression and classification problems was also used for prediction of oil price [16] and photovoltaic power [17]. As optimization tools, GA was used to construct rule-based forecasting [18] and optimize Auto Regressive Integrated Moving Average (ARIMA) model [19] for time series data.

Mainly, researchers stated that intelligent algorithms have better accuracy than traditional algorithms such as Autoregressive Moving Average ARMA [20–22], Auto Regressive Integrated Moving Average (ARIMA) [5, 23], and Seasonal Autoregressive Integrated Moving Average (SARIMA) [24, 25]. However, in some cases, researchers used traditional algorithms to have a good forecasting performance for a specific problem. In other words, each forecasting method has different efficiency due to its characteristics.

Some believe that hybridization methods [26–29] may improve the individual strengths as well as remove the weaknesses of each technique. The hybrid approach combines two forecasting approaches, such as Wavelet-ARMA-NARX [30], ANN-ARIMA [31, 32], ETS-ANN [33], and OT-SVM [34]. Each hybrid technique has its own characteristics and efficiency.

The purpose of this paper is to introduce the Genetic Algorithmised Neuro Fuzzy System (GA-NFS) approach in forecasting visitors on online scientific journal websites. Whilst Neural Fuzzy System (NFS) has been successfully implemented for various forecasting problems [35, 36], there are opportunities to improve its performance by applying optimization method. In this paper, the genetic algorithm that has been proven as a robust optimization technique [37] is used to improve the weight in the neural networks training process.

In this study, our model employs three methods including Neural Network, Fuzzy Logic, and Genetic Algorithm. This strategy provides a hybrid model which produces better results by merging strong points of the methods.

The Neural Network and the fuzzy logic are arranged as a concurrent system. In this system, the methods work continuously together. In the first stage, the data is processed using Neural Network. The Neural Network produces outputs with lower dimension data that are fed to Fuzzy Inference System in the second stage. With lower dimension of input data, a more efficient Fuzzy Inference System could be built with fewer fuzzy rules and also lower computational time.

The main challenge of implementation of the scenario is to ensure that the forecasting result is accurate. To address the challenge, the Genetic Algorithm is employed to adjust weights of the Neural Network. The Genetic Algorithm continuously feeds input data into the Neural Network that produces input data for the Fuzzy Inference System. Output of the Fuzzy Inference System is evaluated by the Genetic Algorithm that then adjusts the Neural Network's weights. By repeating the process, better forecasting results will be achieved each iteration. Thus, the main contribution of this paper is providing an effective mechanism for improving NFS weight using the Genetic Algorithm.

This paper consists of five sections. An introduction highlights the background problem and related research is previously in Section 1. Section 2 presents research methodology, explains the GA-NFS algorithms processes. Section 3 is forecasting using NFS. Section 4 includes results and discussion that explains the findings of forecasting using proposed method. Conclusions are presented at the last section.

II. RESEARCH METHOD

A. STRUCTURE OF GA-NFS

This paper proposes the GA-NFS for unique visitor forecasting. In the first phase, GA is used to optimize the weight in the neural network training process. This process produces the best weight that is used for forecasting with testing data. In the second phase, the previous output is used as an input of Sugeno FIS to get the final estimation.

The baseline, Neuro Fuzzy System (NFS) [38] has two stages in data processing. In the first stage, the data is processed using Neural Network. While in the second stage, the output produced in the first process will be used as input and processed by using Fuzzy Inference System (FIS) method. In this research, fuzzy method used is Sugeno FIS model. This model was chosen because of its efficiency in forecasting. Figure 1 shows the architecture of NFS.

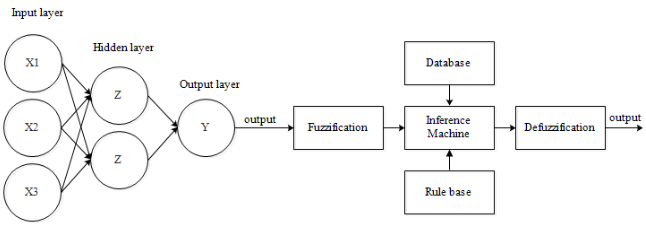

Figure 1. NFS architecture

The proposed method, GA-NFS, optimizes the NFS by producing the best weight for NFS in Figure 1. The weight optimization process using GA is shown in Figure 2.

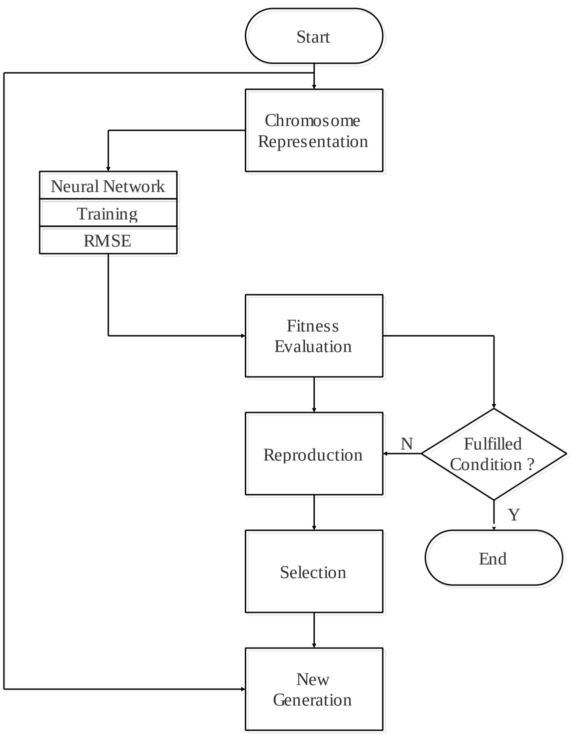

Figure 2. Weight optimization process using GA

The implementation of GA requires the following stages:

1. Building structure of GA by determining chromosome representations and reproduction operators (crossover, mutation, and selection).
2. Testing the best parameters' values of GA that include population size (*PopSize*), generation, crossover rate (*cr*), and mutation rate (*mr*).

B. CHROMOSOME REPRESENTATION

The first step in GA implementation is defining gene or chromosome representation. A good chromosome representation will ease the process of solution finding. In this case, real code chromosomes represent the weights of Neural Networks training process. The number of weights is obtained by counting the amount of inputs and outputs multiplied by the number of neurons on the hidden layer of Neural Networks. The number of neurons in hidden layer will be determined in preliminary computational experiments. For example, there are 10 neurons in the hidden layer, therefore a chromosome will contain a total number of weights equal to 4 (number of inputs) × 10 + 2 (number of outputs) × 10 = 60. An array of real values is used to encode the solutions. By using this real-coded chromosome representation, various genetic operators (crossover and mutation) can be adopted to explore a large feasible search space for Neural Network's weights and increase the possibility to achieve better results. The initial chromosomes values are generated randomly of -1 to 1.

In a previous research, the forecasting errors accuracy was obtained from differences between actual data and estimated results [39]. Hence, the smaller error means a more accurate forecasting result. In this study, we use root mean square error (RMSE) that is the standard deviation of the prediction errors. In the second step fitness functions are calculated based on RMSE. The bigger fitness value indicates more accurate results since the fitness is an inverse of errors.

The third step, crossover, is done by selecting two parents randomly from the population. There are several types of crossover such as ordered based, one-point, two-point, and N-Point crossover. In this paper, one-point method is used. This method selects one random point and swaps the right side of each parent to produce offspring.

Mutation processes are performed in the fourth step by randomly selecting a parent from the population. In this paper, the reciprocal exchange mutation method has been implemented.

The last step, selection process, is an essential chunk of GA because it can produce convergence level. In this paper, the elitism as a selection method has been applied. This method is broadly used in numerous studies related to optimization problems. In principle, the elitism method is performed by gathering all individuals in a population (parents) and offspring in a shelter. Then, the best individual population in this shelter will pass to enter in the next generation. Thus, this method ensures that the best individuals will always get away with it.

C. TESTING THE BEST PARAMETER VALUES

Some tests are done to obtain the best weight during the training data process. The tests performed are population size (*PopSize*) testing, generation, and crossover rate (*cr*) and mutation rate (*mr*) combinations. Determining the best parameter values for GA is a crucial step to ensure that the GA will produce good results in reasonable amount of time. For example, using a higher population size and number of iterations (epoch) will allow the GA to explore a wide search space but a higher computation is required.

Having obtained the best weight, the data is tested until the output is generated at this stage.

D. DATA PROCESSING

Data processing is done in the second phase. At this phase, data is processed by using the Sugeno FIS method. In data processing, the input used in this phase is the output generated in the previous stage. In addition, external factors are also used as input variables at this phase. After data processing using Sugeno FIS is completed, the final output is forecast.

This research uses time-series data of journal visitors in a university website which is recorded in the journal statcounter account. It is daily data in one-year period (*Page Loads* and *Unique Visits*). In this study we used 460 records of data (from August 1, 2016 – November 3, 2017).

Here, the amount of training and testing data is determined based on previous research that is 70:30. As a result, the training and testing data consist of 345 and 115 records respectively. Here, several input variables have been explored. Page loads were divided into three time series variables. They were t-1, t-2, and t-3 as shown in Table 1. Whereas Unique Visits as external factors are shown in Table 2.

Table 1. The example of Time Series Variable

Day	Date	Page Loads	Time Series Variables		
			t-1	t-2	t-3
Monday	1/08/2016	233	247	302	182
Tuesday	2/08/2016	247	302	182	210
…	…	…	…	…	…
Friday	3/11/2017	635	234	166	298

Table 2. The example of External Factor

Day	Date	External Factor
		Unique Visit
Monday	1/08/2016	34
Tuesday	2/08/2016	23
Wednesday	3/08/2016	24
Thursday	4/08/2016	25
Friday	5/08/2016	24
…	…	…
Friday	3/11/2017	68

For both NFS and GA-NFS, this study performs several tests using some parameters such as learning rate, epoch, and neurons. Furthermore, in GA-NFS, the performed tests include the *PopSize*, the number of generations, and the *cr mr* value testing. Genetic Algorithm and Neural Network are stochastic methods and different solutions are obtained in each run, so the methods are run 5 times for each fixed value of the parameters. The average fitness of the runs can be used to measure the stability of the solutions obtained. The computation test is performed to obtain the best parameter values that provide good results in reasonable amount of time. Generally, the higher parameter values such as population size and number of generations will yield more accurate results for forecasting. However, a higher computational time is also required.

The first test is the learning rate test. This test is performed to obtain the value of learning rate (0-1) with a smallest of an average of RMSE value. For the initial test, epoch 2000 and 3 neurons have been used. The test results are as presented in Table 3.

Table 3. The results of learning rate testing

Learning rate	Epoch	Experiments					Avg of RMSE
		1	2	3	4	5	
		RMSE					
0.1	2000	0,00276	0,00262	0,00258	0,00266	0,0028	0.002684
0.2	2000	0,00264	0,00267	0,00264	0,00263	0,00264	0.002644
0.3	2000	0,00264	0,00262	0,00262	0,00258	0,00263	0.002618
0.4	2000	0,00263	0,00261	0,00261	0,00261	0,00258	0.002609
0.5	2000	0,00263	0,00262	0,00264	0,00261	0,00262	0.002624
0.6	2000	0,00267	0,00259	0,00261	0,00258	0,00264	0.002617
0.7	**2000**	**0,00261**	**0,00262**	**0,00261**	**0,00262**	**0,00258**	**0.002607**
0.8	2000	0,00264	0,00267	0,00265	0,00263	0,00262	0.002642
0.9	2000	0,00268	0,00268	0,00262	0,00257	0,00268	0.002647
1	2000	0,00262	0,00263	0,00259	0,00269	0,00269	0.002645

Table 3 shows that the best learning rate was 0.7 with an average of RMSE of 0.002607. This value will be used in the next testing, that is, the number of epoch testing, where a number of *epochs* to be tested varies from 5000-100000. The epoch testing results are shown in Table 4.

Table 5 shows that the best number of neurons was 10. Thus, the smallest average of RMSE value of 0.00201 has been obtained. The results of this test acquired the best weight in the training process. This weight can be used in the next stage of NFS testing data. In addition, it can be used to form a chromosome representation in the proposed method, GA-NFS.

Table 4. The result of epoch testing

Learning rate	Epoch	Experiments					Average of RMSE
		1	2	3	4	5	
		RMSE					
0.7	5000	0.003	0.002	0.002	0.002	0.002	0.002
0.7	10000	0.002	0.002	0.002	0.002	0.002	0.002
0.7	30000	0.002	0.002	0.002	0.002	0.002	0.002
0.7	50000	0.002	0.002	0.002	0.002	0.002	0.002
0.7	100000	0.002	0.002	0.002	0.002	0.002	0.002

Table 5. The result of neuron testing

Learning rate	Epoch	Neuron	Experiments					Average of RMSE
			1	2	3	4	5	
			RMSE					
0.7	100000	4	0.004	0.004	0.003	0.003	0.004	0.003
0.7	100000	6	0.003	0.003	0.002	0.002	0.003	0.002
0.7	100000	8	0.002	0.002	0.003	0.002	0.002	0.002
0.7	100000	10	0.002	0.004	0.002	0.002	0.002	0.002

Based on Table 5, 100000 was the best epoch with an average RMSE value of 0.00222. These results can be used in subsequent testing to discover the best number of neurons. The numbers of neurons to be tested were 5-10 neurons.

After the chromosome formation, the number of *PopSize* is tested. In this experiment, the number of chromosomes to produce the best solution is calculated by the number of *Popsize* testing. The number of popSize tested were 10, 20, 30, 40, and 50. The number of initial generations was 150 when the values of *cr* and *mr* of 0.3 and 0.7 were utilized. The *PopSize* test results are shown in Figure 3.

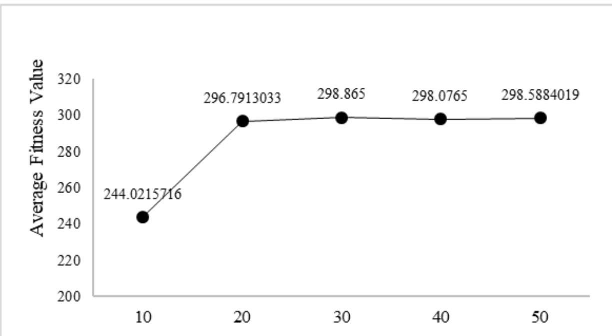

Figure 3. The results of PopSize testing

Based on Figure 3, it can be seen that the greater number of *popSize* may result in the higher value of average fitness. In general, increasing the number of *popSize* may obtain better fitness values because GA has a wider search area. At some point, the fitness value is not increased significantly. A significant increase is not experienced at *popSize* values from 30 to 40 and tends to be stable so that the value of *popSize* 30 is the best *popSize*.

The next testing is the number of generations. This testing aims to determine the best generation by producing the optimum solution to this problem. This test used *popSize* of 30 generated in the previous test.

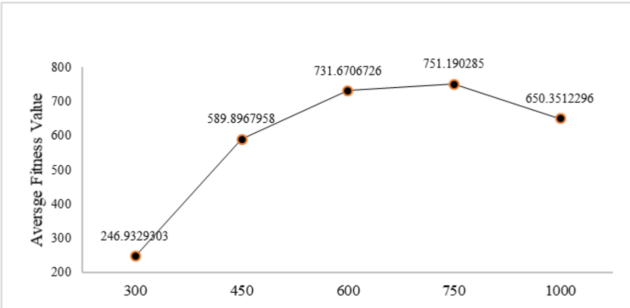

Figure 4. The result of testing generation number

Figure 4 shows that more generations may increase the average fitness value. The number of generations from 300 to 600 has a significant increase, but when it is above 600, the increase in fitness values tends to be stable. Therefore, the best generation is 600.

Furthermore, this study conducted a test to find the best crossover rate (*cr*) and mutation rate (*mr*) value. Appropriate *cr* and *mr* values will help the Genetic Algorithm to balance its exploration and exploitation power. The value of *popSize* was 30 and the number of generations was 600. This value is used because it is considered the number that can produce the most optimal average fitness value after going through various tests. The value used in this test is chosen by considering the *cr* and *mr* values in each test scenario for the total amount equal to 1, because the test should be fair. That is, the number of offspring produced in each test scenario must have the same total.

According to Figure 5, it can be seen that the best *cr* and *mr* values to obtain the optimum solutions are 0.6 and 0.4. A high *cr* value will provide many new chromosomes in the population. However, if the *cr* value is too high, a group of genes that produce the best fitness value do not have the opportunity to stick to each other in chromosomes. That is, these genes are likely to separate and affect the fitness value to be smaller. Otherwise, if the *cr* value is too low it will not produce enough new offsprings.

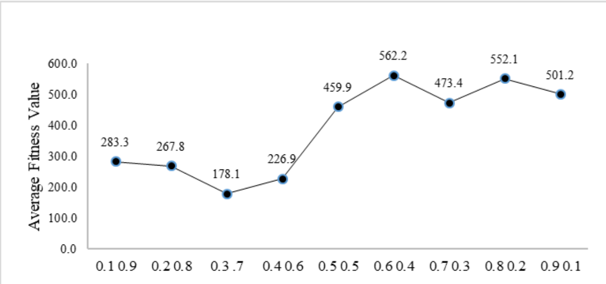

Figure 5. The result of testing cr and mr value

III. FORECASTING USING NFS

The best weights generated in the data training process have been used in the data testing stage. Table 6 shows the results of the data testing process.

The output shown in Table 6 will be used as input to be reprocessed using Sugeno FIS to produce final forecasting. This stage is the final phase of data processing using NFS. Based on Table 7, it can be seen that the accuracy produced by the NFS method is 2.005 of RMSE.

Table 6. The results of data testing process

Periods	Time series variable			Actual	Output (NN)
	t-1	t-2	t-3		
12/07/2017	291	420	221	411	410.2313659
13/07/2017	420	221	320	291	293.2863146
14/07/2017	221	320	364	420	420.1420272
15/07/2017	320	364	317	221	222.2054394
16/07/2017	364	317	272	320	321.3531680
17/07/2017	317	272	440	364	365.8488516
18/07/2017	272	440	206	317	317.2957780
19/07/2017	440	206	98	272	273.8643448
20/072017	206	98	325	440	440.3970469
…	…	…	…	…	
3/11/2017	561	517	298	635	635.2557478

Table 7. Forecasting results using NFS

Periods	Actual	Output (NN)	Ext. Factor Unique Visits	NFS Forecasting	RMSE
12/07/2017	411	410.231	104	410.30484	0.483240
13/07/2017	291	293.286	103	293.93357	8.605833
14/07/2017	420	420.142	98	420.10029	0.010058
15/07/2017	221	222.205	59	222.72397	2.972090
16/07/2017	320	321.353	50	321.27723	1.631330
17/07/2017	364	365.849	92	366.01275	4.051183
18/07/2017	317	317.296	104	317.83452	0.696424
19/07/2017	272	273.864	88	274.44368	5.971572
20/072017	440	440.397	116	440.45201	0.204318
…	…	…	…	…	…
3/11/2017	635	635.256	234	635.634720	0.4028695
RMSE					462.2646

In order to know the performance of the system, this research uses Root Mean Square Error (RMSE) analysis technique. This technique is used to determine the accuracy of the resulting system due to its certainty [40]. The smaller the RMSE value is, the more accurate are the forecast models.

IV. FORECASTING USING GA-NFS

This section discusses the implementation of same test on GA-NFS architecture. Based on the results of these tests, the best weight was obtained to be used in testing data. The results of data testing process with the best weight of genetic algorithmised-NN (GA-NN) are shown in Table 8.

Table 8. The results of testing data process (GA-NN)

Periods	Time series variable			Actual	Output (GA-NN)
	t-1	t-2	t-3		
12/07/2017	291	420	221	411	411.914860
13/07/2017	420	221	320	291	292.311801
14/07/2017	221	320	364	420	422.496457
15/07/2017	320	364	317	221	223.615624
16/07/2017	364	317	272	320	323.029207
17/07/2017	317	272	440	364	364.421372
18/07/2017	272	440	206	317	318.674001
19/07/2017	440	206	98	272	271.210135
20/072017	206	98	325	440	440.036648
…	…	…	…	…	…
3/11/2017	561	517	298	635	638.783311

In the next step, this research will perform data processing using Sugeno FIS by using the output data in Table 8. The data is used as input in the next stage of final forecasting on GA-NFS method. Based on Table 9, the forecasting result using GA-NFS shows the RMSE value of 0.989.

Table 9. Forecasting results using GA-NFS

Periods	Actual	Output GA-NN	Ext. Factor Unique Visits	GA-NFS Forecasting	Error
12/07/2017	411	411.915	104	411.563	0.31697
13/07/2017	291	292.312	103	291.958	0.91776
14/07/2017	420	422.496	98	421.132	1.28142
15/07/2017	221	223.616	59	221.174	0.03028
16/07/2017	320	323.029	50	320.569	0.32376
17/07/2017	364	364.421	92	364.045	0.00202
18/07/2017	317	318.674	104	318.322	1.74768
12/07/2017	272	271.210	88	270.826	1.37828
…	…	…	…	…	…
3/11/2017	635	638.783	234	636.691	2.85948
					112.6661
RMSE					**0.989801**

Fig. 6 shows the comparison of forecasting results between the NFS and GA-NFS methods. Based on the comparison of the forecasting results shown in Figure 6 it can be seen that the GA-NFS forecasting tends to approach the actual data compared to the green curve. It is also supported by the value of accuracy produced by both methods. The accuracy generated by GA-NFS (RMSE = 0.98) is higher than that of NFS (RMSE = 2.004). This shows that the performance of the proposed system using (GA-NFS) is better than its baseline (NFS), yet with higher computation complexity.

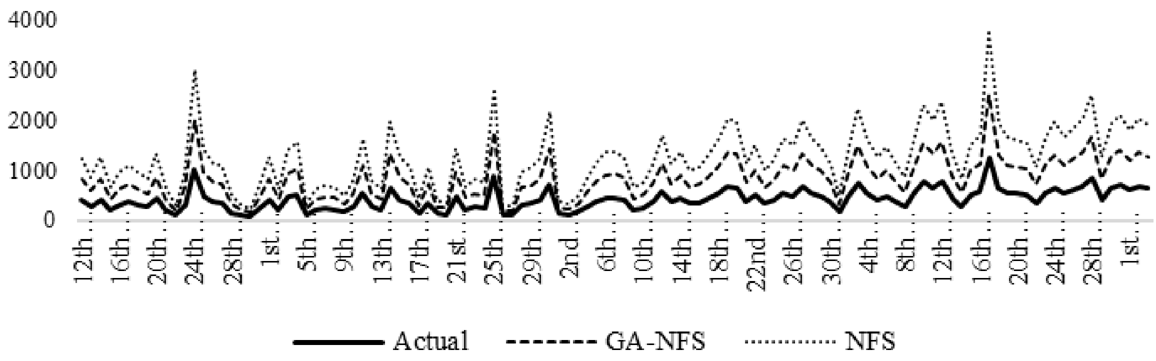

Figure 6. The comparison of forecasting results between GA-NFS and NFS

V. CONCLUSION

This paper presents the performance comparison of intelligent algorithms, namely Neuro Fuzzy System (NFS) and Genetic Algorithm (GA), in learning time series data. Based on the forecasting results, it can be concluded that GA can be used to determine the optimum weight for NFS based forecasting.

The best algorithm parameters used to obtain the optimal solution are 30 for popSize, 600 for number of generations, 0.6 for *cr* and 0.4 for *mr*. These parameters produce the best forecasting with an accuracy of 0.989. In other words, GA-NFS outperformed NFS as comparison method due to weighting optimization in each neuron.

A further research in a long short-term memory (LSTM) should be conducted. The research may discover the efficiency of GA-NFS while the data availability is limited.

References

[1] M. F. Iqbal, M. Zahid, D. Habib, and L. K. John, "Efficient prediction of network traffic for real-time applications," *Journal of Computer Networks and Communications*, vol. 2019, p. 4067135, 2019, https://doi.org/10.1155/2019/4067135.

[2] A. Subashini, S. K, S. Saranya, and U. Harsha, "Forecasting website traffic using prophet time series model," *International Research Journal of Multidisciplinary Technovation*, vol. 1, no. 1 SE-Articles, pp. 56–63, 2019, https://doi.org/10.34256/irjmt1917.

[3] T. Shelatkar, S. Tondale, S. Yadav, and S. Ahir, "Web traffic time series forecasting using ARIMA and LSTM RNN," *ITM Web of Conferences*, vol. 32, pp. 3017, 2020, https://doi.org/10.1051/itmconf/20203203017.

[4] N. Petluri and E. Al-Masri, "Web traffic prediction of Wikipedia pages," *Proceedings of the 2018 IEEE International Conference on Big Data (Big Data)*, 2018, pp. 5427–5429, https://doi.org/10.1109/BigData.2018.8622207.

[5] P. Purnawansyah, H. Haviluddin, R. Alfred, and A. F. O. Gaffar, "Network traffic time series performance analysis using statistical methods," *Knowledge Engineering and Data Science*, vol. 1, no. 1, pp. 1–7, 2017, https://doi.org/10.17977/um018v1i12018p1-7.

[6] P. Purnawansyah, H. Haviluddin, H. J. Setyadi, K. Wong, R. Alfred, and A. P. Wibawa, "An inflation rate prediction based on backpropagation neural network algorithm," *International Journal of Artificial Intelligence Research*, vol. 3, no. 2, pp. 92–98, 2019, https://doi.org/10.29099/ijair.v3i2.112.

[7] X. Feng, T. M. Fu, H. Cao, H. Tian, Q. Fan, and X. Chen, "Neural network predictions of pollutant emissions from open burning of crop residues: Application to air quality forecasts in southern China," *Atmospheric Environment*, vol. 204, pp. 22–31, 2019, ttps://doi.org/10.1016/j.atmosenv.2019.02.002.

[8] Z. Zhang, X. Wang, and Y. Ou, "Water simulation method based on BPNN response and analytic geometry," *Procedia Environmental Sciences*, vol. 2, no. 5, pp. 446–453, 2010, https://doi.org/10.1016/j.proenv.2010.10.049.

[9] C. N. Stefanakos and E. Vanem, "Nonstationary fuzzy forecasting of wind and wave climate in very long-term scales," *Journal of Ocean Engineering and Science*, vol. 3, no. 2, pp. 144–155, 2018, https://doi.org/10.1016/j.joes.2018.04.001.

[10] G. C. Silva, J. L. R. Silva, A. C. Lisboa, D. A. G. Vieira, and R. R. Saldanha, "Advanced fuzzy time series applied to short term load forecasting," *Proceedings of the 2017 IEEE Latin American Conference on Computational Intelligence*, 2017, pp. 1–6, https://doi.org/10.1109/LA-CCI.2017.8285726.

[11] D. Ali, M. Yohanna, M. I. Puwu, and B. M. Garkida, "Long-term load forecast modelling using a fuzzy logic approach," *Pacific Science Review A: Natural Science and Engineering*, vol. 18, no. 2, pp. 123–127, 2016, https://doi.org/10.1016/j.psra.2016.09.011.

[12] J. Nou, R. Chauvin, A. Traoré, S. Thil, and S. Grieu, "Atmospheric turbidity forecasting using side-by-side ANFIS," *Energy Procedia*, vol. 49, pp. 2387–2397, 2014, https://doi.org/10.1016/j.egypro.2014.03.253.

[13] Y. L. Zhang and J. H. Lei, "Prediction of laser cutting roughness in intelligent manufacturing mode based on ANFIS," *Procedia Engineering*, vol. 174, pp. 82–89, 2017, https://doi.org/10.1016/j.proeng.2017.01.152.

[14] M. Wu, C. Stefanakos, Z. Gao, and S. Haver, "Prediction of short-term wind and wave conditions for marine operations using a multi-step-ahead decomposition-ANFIS model and quantification of its uncertainty," *Ocean Engineering*, vol. 188, p. 106300, 2019, https://doi.org/10.1016/j.oceaneng.2019.106300.

[15] Y. Kassem and H. Çamur, "Prediction of biodiesel density for extended ranges of temperature and pressure using adaptive neuro-fuzzy inference system (ANFIS) and radial basis function (RBF)," *Procedia Computer Science*, vol. 120, no. 2017, pp. 311–316, 2017, https://doi.org/10.1016/j.procs.2017.11.244.

[16] L. T. Zhao and G. R. Zeng, "Analysis of timeliness of oil price news information based on SVM," *Energy Procedia*, vol. 158, pp. 4123–4128, 2019, dhttps://doi.org/10.1016/j.egypro.2019.01.821.

[17] M. Malvoni, M. G. De Giorgi, and P. M. Congedo, "Data on support

vector machines (SVM) model to forecast photovoltaic power," *Data in Brief*, vol. 9, pp. 13–16, 2016, https://doi.org/10.1016/j.dib.2016.08.024.

[18] L. Hulianytskyi and A. Pavlenko, "Development and analysis of genetic algorithm for time series forecasting problem," *International Journal*, vol. 4, no. 1, pp. 13–29, 2015.

[19] Y. K. Al-Douri, H. Al-Chalabi, and J. Lundberg, "Time series forecasting using a two-level multi-objective genetic algorithm: A case study of maintenance cost data for tunnel fans," *Proceedings of the ADVCOMP 2018 : The Twelfth International Conference on Advanced Engineering Computing and Applications in Sciences*, 2018, pp. 4–9, https://doi.org/10.3390/a11080123.

[20] N. H. J. Al-Saadi, "Forecasting traffic accidents according to (types of roads and their causatives) in Iraq using ARMA of low ordered combinations models," *International Journal of Engineering & Technology*, vol. 8, no. 1.5, pp. 9–15, 2019.

[21] E. Akdeniz, E. Egrioglu, E. Bas, and U. Yolcu, "An ARMA Type Pi-Sigma Artificial Neural Network for nonlinear time series forecasting," *Journal of Artificial Intelligence and Soft Computing Research*, vol. 8, no. 2, pp. 121–132, 2018, https://doi.org/10.1515/jaiscr-2018-0009.

[22] W. Wang, "Improved short term load forecasting of power system based on ARMA model," *Advances in Economics, Business and Management Research*, vol. 30, pp. 12–19, 2016, https://doi.org/10.2991/iconfem-16.2016.2.

[23] C. Liu, S. C. H. Hoi, P. Zhao, and J. Sun, "Online ARIMA algorithms for time series prediction," *Proceedings of the 30th AAAI Conference on Artificial Intelligence, AAAI 2016*, 2016, pp. 1867–1873.

[24] M. Milenković, L. Švadlenka, V. Melichar, N. Bojović, and Z. Avramović, "SARIMA modelling approach for railway passenger flow forecasting," *Transport*, vol. 33, no. 5, pp. 1113–1120, 2018, https://doi.org/10.3846/16484142.2016.1139623.

[25] R. Rossetti, "Forecasting the sales of console games for the Italian market," *Ekonometria*, vol. 23, no. 3, pp. 76–88, 2019, https://doi.org/10.15611/eada.2019.3.07.

[26] M. Shafaei, J. Adamowski, A. Fakheri-Fard, Y. Dinpashoh, and K. Adamowski, "A wavelet-SARIMA-ANN hybrid model for precipitation forecasting," *Journal of Water and Land Development*, vol. 28, no. 1, pp. 27–36, 2016, https://doi.org/10.1515/jwld-2016-0003.

[27] H. Ince and T. B. Trafalis, "A hybrid forecasting model for stock market prediction," *Economic Computation and Economic Cybernetics Studies and Research*, vol. 51, no. 3, pp. 263–280, 2017.

[28] M. Siddique, D. Panda, S. DAs, and S. K. Mohapatra, "A hybrid forecasting model for stock value prediction using soft computing skill," *International Journal of Computer Sciences and Engineering*, vol. 117, no. 19, pp. 357–363, 2017.

[29] C. Wu, P. Luo, Y. Li, L. Wang, and K. Chen, "Stock price forecasting: Hybrid model of artificial intelligent methods," *Engineering Economics*, vol. 26, no. 1, pp. 40–48, 2015, https://doi.org/10.5755/j01.ee.26.1.3836.

[30] H. Nazaripouya, B. Wang, Y. Wang, P. Chu, H. R. Pota, and R. Gadh, "Univariate time series prediction of solar power using a hybrid wavelet-ARMA-NARX prediction method," *Proceedings of the 2016 IEEE/PES Transmission and Distribution Conference and Exposition*, 2016, pp. 1–5, https://doi.org/10.1109/TDC.2016.7519959.

[31] R. M. K. T. Ratnayaka, D. M. K. N. Seneviratne, W. Jianguo, and H. I. Arumawadu, "A hybrid statistical approach for stock market forecasting based on artificial neural network and ARIMA time series models," *Proceedings of the 2015 International Conference on Behavioral, Economic and Socio-Cultural Computing, BESC 2015*, 2015, pp. 54–60, dhttps://doi.org/10.1109/BESC.2015.7365958.

[32] I. Khandelwal, R. Adhikari, and G. Verma, "Time series forecasting using hybrid arima and ann models based on DWT decomposition," *Procedia Computer Science*, vol. 48, pp. 173–179, 2015, dhttps://doi.org/10.1016/j.procs.2015.04.167.

[33] S. Panigrahi and H. S. Behera, "A hybrid ETS–ANN model for time series forecasting," *Engineering Applications of Artificial Intelligence*, vol. 66, no. November 2019, pp. 49–59, 2017, https://doi.org/10.1016/j.engappai.2017.07.007.

[34] Y. Liu, Y. Sun, D. Infield, Y. Zhao, S. Han, and J. Yan, "A hybrid forecasting method for wind power ramp based on Orthogonal Test and Support Vector Machine (OT-SVM)," *IEEE Transactions on Sustainable Energy*, vol. 8, no. 2, pp. 451–457, 2017, https://doi.org/10.1109/TSTE.2016.2604852.

[35] N. R. Sari, A. P. Wibawa, and W. F. Mahmudy, "Comparison of ANFIS and NFS on inflation rate forecasting," *Proceedings of the 2017 5th International Conference on Electrical, Electronics and Information Engineering: Smart Innovations for Bridging Future Technologies, ICEEIE 2017*, 2017, pp. 123-130, https://doi.org/10.1109/ICEEIE.2017.8328775.

[36] D. Marcek, "Forecasting of financial data: a novel fuzzy logic neural network based on error-correction concept and statistics," *Complex & Intelligent Systems*, vol. 4, no. 2, pp. 95–104, 2018, https://doi.org/10.1007/s40747-017-0056-6.

[37] L. W. Yan, Z. Y. Sun, and K. Y. Mao, "Robust Optimization Based on an Improved Genetic Algorithm," *Advanced Materials Research*, vol. 655–657, pp. 955–958, 2013, https://doi.org/10.4028/www.scientific.net/AMR.655-657.955.

[38] N. R. Sari, W. F. Mahmudy, and A. P. Wibawa, "The effectiveness of hybrid backpropagation Neural Network model and TSK Fuzzy Inference System for inflation forecasting," *Journal of Telecommunication, Electronic and Computer Engineering*, vol. 9, no. 2, pp. 111–117, 2017.

[39] C. Tofallis, "A better measure of relative prediction accuracy for model selection and model estimation," *Journal of the Operational Research Society*, vol. 66, no. 8, pp. 1352–1362, 2015, https://doi.org/10.1057/jors.2014.103.

[40] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature," *Geoscientific Model Development*, vol. 7, no. 3, pp. 1247–1250, 2014, https://doi.org/10.5194/gmd-7-1247-2014.

WAYAN FIRDAUS MAHMUDY *is Senior Lecturer in Faculty of Computer Science, Universitas Brawijaya, Indonesia. He obtained bachelor degree in mathematics from Universitas Brawijaya, master degree in information technology from Institut Teknologi Sepuluh November (ITS), Indonesia, and completed his Ph.D. in Manufacturing Engineering at University of South Australia. His research interests include optimisation of combinatorial problems and machine learning.*

AJI PRASETYA WIBAWA *is a head of Electrical Engineering Department in Universitas Negeri Malang (UM), Indonesia. He has PhD of Electrical and Information Engineering from University of South Australia. He leads a research group, focused on socio-language computing. Theory and Application. His research interests are AI, data mining, machine translation, and social informatics.*

NADIA ROOSMALITA SARI *is a lecturer in the Faculty of Islamic Economic and Business of IAIN Tulungagung since 2019. Nadia completed her Master Computer Science University of Brawijaya and her undergraduate at Universitas Negeri Malang. Her research interests are in the area of artificial intelligence, ranging from theory about fuzzy, neural network, and genetic algorithm.*

HAVILUDDIN *is a lecturer at Faculty of Computer Science and Information Technology (CSIT), Mulawarman University, Indonesia, since 2002. He completed his Ph.D. in Computer Science from Universiti Malaysia Sabah, Malaysia in 2016. He is the coordinator of publication and intellectual property rights of Research Institute and Community Service of Mulawarman University. He is a member of the Institute of Electrical and Electronic Engineers (IEEE), Association for Scientific Computing Electronics and Engineering (ASCEE), Institute of Advanced Engineering and Science (IAES), and Indonesian Computer, Electronics, Instrumentation Support Society (IndoCEISS), and Association of Computing and Informatics Institutions Indonesia (APTIKOM) societies. His research interest is in artificial intelligent area.*

PURNAWANSYAH *was born in Makassar, South Sulawesi, Indonesia. He completed a Master at Gadjah Mada University, Yogyakarta in 2005 in the field of Computer Science. He is also a Lecturer in the Faculty of Computer Science, Universitas Muslim Indonesia, Makassar, South Sulawesi, Indonesia since 2002. He is a member of the Institute of Electrical and Electronic Engineers (IEEE), Head of Sulawesi Selatan Region for Association of Higher Education Institutions in computer science and information technology (APTIKOM). His research interest is in artificial intelligent area.*