Genetic Algorithm-Based Feature Selection and Optimization of Backpropagation Neural Network Parameters for Classification of Breast Cancer Using MicroRNA Profiles

Amazona Adorada
Department of Informatics
Diponegoro University
Semarang, Indonesia
amazona.ada@gmail.com

Adi Wibowo
Department of Informatics
Diponegoro University
Semarang, Indonesia
bowo.adi@live.undip.ac.id

***Abstract*—Breast cancer is one of the most common types of cancer found in women. Breast cancer mortality increases every year because it has not found an appropriate early detection method. MicroRNA can be used as a potential biomarker, because the profile of the microRNA feature in breast cancer will decrease or increase the value of expression compared to normal conditions. But because of the thousands of types of microRNA that make up breast cancer, a lot of money is needed to detect it entirely. Backpropagation Artificial Neural Network Method has good performance in generalization, so it is suitable to be used as a method for classification with many features. The classification results from the neural network model will be more accurate if the parameters used can be optimized precisely. Genetic algorithms can be used to optimize backpropagation neural network parameters as well as feature selection, because of its global search characteristics. This study aims to compare the performance of backpropagation artificial neural networks optimized parameters as well as feature selection using genetic algorithms (GABPNN_FS) with backpropagation artificial neural networks optimized using genetic algorithms without feature selection (GABPNN). The results showed that the GABPNN had better results with an error value of 0.016115. But GABPNN_FS has a faster average process duration of 53.2689 seconds. The best individual chromosome translation results on GABPNN_FS for breast cancer classification based on microRNA profile are random state = 6098, learning rate = 0.7, number of neuron hidden = 6, and selected features = 707 features that produce accuracy, sensitivity, and specificity ie 97.50 %, 99.00% and 96.00%.**

Keywords— breast cancer, microRNA, neural network, backpropagation, parameter optimization, feature selection, genetic algorithm

I. INTRODUCTION

Cancer is one of the biggest causes of death in the world. Breast cancer is a type of malignant tumor that is most often found in women and its incidence tends to increase [1]. Based on data from the GLOBOCAN (Global Burden of Cancer) project in 2018 there were 626,679 deaths from breast cancer or 15% of female deaths caused by cancer in the world [2]. Breast cancer shows an increase in mortality rates every year because there is no method of early detection and the right prognosis has been found [3]. Usually, if the patient begins to feel pain in the breast and if it is true that it is caused by breast cancer, then it is certain that the cancer has reached an advanced stage [4].

One way that can be taken to make an early diagnosis of breast cancer is by identifying the microRNA profile. MicroRNA is a member of the non-coding RNA family that helps translate the genetic information of DNA into proteins [5]. There are reports that almost all types of cancer undergoing dysregulation of microRNA expression show an important role in carcinogenesis [1]. Profiling with microRNA expression is reported to be used to improve the accuracy of breast cancer diagnosis and therapy because of its ability to classify patients into breast cancer subtypes appropriately [5]. But because there are thousands of types of microRNA that make up breast cancer, it requires a lot of costs if it has to be detected entirely. Therefore it is necessary to choose the best microRNA profile that can be used for breast cancer classification.

Artificial Neural Networks (ANN) is an information processing system designed to mimic the workings of the human brain in solving a problem with the learning process through changes in the weight of the synapses. Data from the past will be studied by ANN so that it has the ability to make decisions about data that have never been studied [6]. Backpropagation is one of the ANN methods that has the ability to solve complex problems because it is trained with supervision learning methods so that it can recognize input patterns with high accuracy [6]. It has been widely known that the classification produced from artificial neural network models will be more accurate if parameters such as the number of hidden layer units and learning rates can be optimized correctly [7].

Feature Selection (FS) is a technique that is often used in the application of pattern recognition. Feature selection can alleviate the problem of overfitting and improve model performance by removing irrelevant, annoying and excessive features from the feature space. Through feature selection can also be obtained more insight and information about data by analyzing the importance of a feature [8]. Because of the complex design and the existence of SF which are also combined, there is an increasing tendency for hybridization of ANN designs with

978-1-7281-4610-2/19/$31.00 ©2019 IEEE

evolutionary algorithms. Genetic Algorithm is one of the most frequently used evolutionary algorithms, which can be used as an optimization algorithm because of its global search characteristics [9].

Genetic Algorithm (GA) is a search algorithm based on natural selection mechanics and natural genetics [10]. Reference [11] in his research he has compiled four different ways to apply GA to ANN, namely: (1) to optimize the weight on fixed networks; (2) to browse the architectural space; (3) to find optimal learning parameters; (4) a genetic approach that modifies the backpropagation algorithm. Research on artificial neural networks with optimization of genetic algorithms has been carried out by several researchers, including Ahmad, et al. [9] using genetic algorithms to perform parameter optimization and feature selection on three variations of the backpropagation algorithm, namely resilient backpropagation, levenberg marquardt, and gradient descent with momentum with the best accuracy produced by resilient backpropagation which is 98.29%. Another study conducted in [12] used genetic algorithms to determine the relationship between the architecture and the initial weight of ANN in order to obtain an optimal combination of parameters of backpropagation neural networks and produce the highest accuracy of 99.04% in user knowledge data.

Based on this background, a problem can be formulated, namely how to get backpropagation neural network parameters and the best features by feature optimization and selection using genetic algorithms for breast cancer classification based on microRNA profiles.

The objectives to be achieved from this study include: (1) Obtaining the best combination of features and parameters in genetic algorithm based backpropagation neural networks for classification of breast cancer based on microRNA profiles; (2) Comparing the performance of genetic algorithm optimization on backpropagation neural networks with and without feature selection; (3) Knowing the parameters of the best genetic algorithms to produce a combination of the best backpropagation neural network features and parameters.

II. Features Selection and Optimization of Backpropagation Neural Neural Network Parameters Based on Genetic Algorithm

This research was carried out in several stages of the process. These stages start from data collection, data mapping, data normalization, and data sharing. Data sharing is done using the k-fold cross validation which will divide the data into two types of data, namely training data and test data, in the 20% training data will be used as a validation set during the backpropagation training process. The next step is chromosome initialization and translation in genetic algorithms. If feature selection is to be performed, the next step is feature selection using the results of chromosome translation, but if you do not use feature selection, you will enter the backpropagation and genetic surgery training stages. If the stop condition of the genetic algorithm is fulfilled, namely the maximum generation and the fitness value

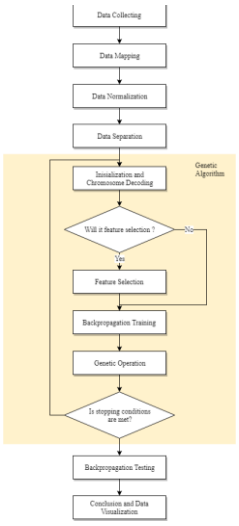

Fig. 1. Outline process of problem solving

conference, the process continues to the backpropagation testing stage, then draws conclusions and visualizes the data, but if the stop condition is not fulfilled the process returns to the chromosome initialization and translation stage. The problem solving outline diagram is shown in Fig 1.

A. Data Collection

The data collected for this study is the value data of quantification of breast cancer microRNA expression. The data was obtained from the National Cancer Institute Genomic Data Commons [13]. The data used in this study amounted to 200 data consisting of 100 data with normal classes and 100 data with breast cancer classes. Each data sample consisted of 1464 profiles of microRNA features that became biomarkers of breast cancer.

B. Data Mapping

MicroRNA is identified in the data mapping stage so that it can be mapped into backpropagation architecture. MicroRNA data consisting of 1464 number of features will be used as input neurons, while class labels will be used as output neurons or targets. The target of breast cancer class data was given a value of 1, while the target of normal class data was 0.

C. Data Normalization

The activation function used in backpropagation must meet several conditions, namely continuous, easily differentiated, and is a function that does not go down. One function that fulfills these three conditions so that it is often used is a binary sigmoid function that has a range of 0 to 1 [14]. The previously mapped data is then normalized using equation (1).

$$s' = 0.8 \times \frac{s - \min(s_k)}{\max(s_k) - \min(s_k)} + 0.1 \qquad (1)$$

D. Data Separation

Stages of data separation are carried out to divide the research data into training data and test data. As per the research conducted on [9] the validation data used is as much as 20% of the training data so that the composition of the data is 70% is training data, 20% is validation data, and 10% is test data.

Distribution of training data and test data using K-Fold Cross Validation. The process of K-Fold Cross Validation in this study uses the value K = 10 so that the dataset will be divided into 10 subsets of the same size, and then iterated 10 times. The K value used is 10 because the use of the best number of folds for the validity test is 10 [15].

E. Inisialization and Chromosom Decoding (Genetic Algorithm)

This phase of initialization and translation of chromosomes is the initial stage of optimization of backpropagation artificial neural networks using genetic algorithms. The major stages carried out in it include new population initialization and decoding chromosomes from each individual into features and parameters for the training and testing process in the next stage. In this stage the population size (popsize) must be determined first or the value that states the number of chromosomes that are accommodated in the population. Another value that must be determined first is the length value of each string of chromosomes (stringlen). The length of each chromosome string or the value of stringLen is calculated based on the precision of the solution variable sought [16]. Some parameters of other genetic algorithms must also be determined at the initialization stage, namely crossover probability (cp), mutation probability (mp), and maximum generation. If this stage is not done for the first time, in other words it is an advanced generation, then the set of new solutions used is the result of selection from the previous generation.

This stage starts with generating random binary numbers along stringlen values and as many as popsize or using previous generation selection results if it is the next generation which is also as many as popsize. Binary representations used aim to form the chromosomes that make up the population according to those illustrated in Fig 2. Chromosomes are represented in binary numbers with each bit having a value of 0 or 1. P bits represent random initial weight generators or random initial weight generators allows $2(2^P) - 1$ a combination of different initial weights on the backpropagation network. Q bits represent the value of the learning rate provided that the value obtained will be multiplied by 0.1 with the learning rate range from 0.1 to 0.9. R bits represent the number of hidden neurons by counting the number of bits that have a value of 1, with the maximum number of hidden neurons as much as the value R. The maximum number of hidden neurons used is 9 neurons, this is in accordance with the results of [17] stating that $(8n - 4)/(n - 4)$ with n is the number of neuron inputs is the right number of neuron hidden for the development of artificial neural networks and produces the smallest error compared to the solution from the results of other studies. Then S bits represent the number of features or attributes in the data, with the provisions of values 1 and 0 indicating that the feature

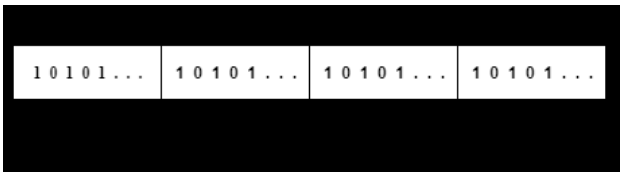

Fig. 2. Chromosome bits representation

is selected or not. Value 1 indicates that the feature in that sequence is selected to proceed to the training and testing process, and the value 0 is the opposite.

F. Backpropagation Training

This backpropagation training process is carried out using a combination of parameters obtained from the chromosome translation process. Some other parameters that will be used must also be determined in advance, namely the epoch maximum and the error target. This training was carried out in each k-fold iteration and each individual in the genetic algorithm population, so that each individual training was carried out as many as K times. In this training process validation data are used to avoid over-training behavior [9].

G. Genetic Operation (Genetic Algorithm)

The genetic operation stage is a stage that contains the advanced process of genetic algorithms after chromosome initialization and translation. This stage of genetic operation includes several processes in genetic algorithms, namely recombination which includes the process of crossover and mutation, selection using a combination of the elitism selection and roullete wheel methods that produce new individual populations for the next stage. The fitness value used is the error value of backpropagation network training (15).

$$Fitness = \frac{1}{Training\ Error} \tag{15}$$

Individuals who have the lowest error value on their network will get the title as the best individual in the population. The following is the genetic operation process carried out in this study:

1) Elitism Selection: A mechanism in genetic algorithms that is used to ensure the survival of the best chromosomes by copying them to the next generation without having to be modified first [9]. The copied chromosome is a ranking based on its fitness value. The number of chromosomes copied is in accordance with the value of the specified elitist number.

2) Roullete Wheel Selection: The Roullete Wheel is used to select individuals based on a probability value to form a swimming pool with individuals as much as the popsize value minus the elitism number. Individuals from theating pool produced will undergo a recombination process to be copied to the population in the new generation.

3) Crossover: The purpose is to exchange the gene pieces between the two parents for all the chromosomes in the swimming pool. The crossover type used is a single point crossover that will cut genes at one point randomly selected. This cut is done on both the chromome parent with the same point location [18]. The results of the crossover process will go into the mutation process.

4) Mutation: Mutation done in order to change the allele on the chromosome. In this study, bit flip mutation is used which changes the bit value with the opposite value [18]. The selection of genes that have mutations is done randomly ranging from 1 to maximum string.

5) Evaluation: This evaluation phase aims to determine the status of the stop condition of the genetic algorithm used in the system. In this study the stop condition of the genetic algorithm used is the maximum generation and convergence of the average fitness value in the last five generations. If one of the stop conditions is met, the iteration of the genetic algorithm based backpropagation neural network (GABPNN) stops, and the best individual will be taken from the population of the last generation that has the highest fitness value. If the stop condition has not been fulfilled, the iteration will continue and return to the chromosome initialization and translation stage until genetic surgery until the stop condition is fulfilled.

H. Backpropagation Testing

The testing phase is done to test the accuracy of the network resulting from the training process. This testing process is also carried out at each fold like the training process, so that testing is done as many times as using the final weight of each training process. The tested architecture is the best individual with the highest fitness obtained from the final results of the optimization process.

III. DISCUSSIONS

Algorithm testing is done through several experiments using the desired scenarios on a predetermined dataset.

A. Testing Scenario

1) Experiment 1: The tests carried out include three activity scenarios. The first scenario aims to obtain the best combination of genetic algorithm parameters which include the value of crossover probability (cp), mutation probability (mp), and maximum generation. The first scenario testing was carried out by running 10 experiments on each combination of the three parameters which amounted to 27 combinations to get the best variety of possible individuals in each combination with a total of 270 trials. The combination of cp, mp, and maximum generation values that produce the highest average error of the best individual compared to other combinations, will be selected as the best genetic algorithm parameter value.

2) Experiment 2: The second testing scenario aims to get the combination of selected features and the best backpropagation neural network parameters which include random state generators for initial weighting of networks, number of hidden neurons in the hidden layer, and learning rate values. In this scenario you will also get selected features that produce the lowest error. Tests were performed using the best genetic algorithm parameter values in the first scenario 10 times. The best individuals with the lowest errors with the fewest features will be taken as a result of the combination of the best features and parameters. The best individuals obtained will then be tested using test data to determine the value of accuracy, sensitivity, and specificity. In scenario two, it will also compare the performance of the best individual output from scenario one and the two effects of feature selection on performance optimization and classification of genetic algorithm based backpropagation neural networks.

B. Testing Result

In testing the crossover probability and the probability of mutations *used* cp is 0.7, 0.8, and 0.9 and mp is 0.01, 0.05 and 0.08. The maximum generation used is 5, 15 and 30. The population size used is 15 individuals. Elitism number used is 3. The number of bits used on the chromosomes for P, Q, and R are 15, 9, and 9. While other parameters are used, namely maximum epoch = 1000 and target error = 0.001. Each experiment was conducted using K-Fold Cross Validation with a K value of 10. The test was carried out by conducting experiments using a combination of cp, mp, and maximum generation of 27 value combinations. The test results of a combination of crossover probability, mutation probability, and generation maximum are shown in Table 1.

Table 1 shows the value of training error average, fitness value, and process duration of the best individuals in each combination and the average value of the 10 trials carried out in each combination. Based on the data in Table 1, it is known that the testing of the maximum generation combination, crossover probability, and mutation probability that produces the best average value is P-11, which is 40.88704, with an average

TABLE I. AVERAGE TRAINING ERROR VALUE, FITNESS VALUE, AND BEST EXPERIMENT INDIVIDUAL DURATION OF PROCESS 1

ID	Max. Generation	CP	MP	Hasil		
				Error	Fitness	Durasi Proses
P-1	5	0.7	0.01	0.04573	22.80087	55.6005
P-2			0.05	0.042954	24.86359	41.1712
P-3			0.08	0.039216	28.36597	45.4405
P-4		0.8	0.01	0.038711	28.04886	44.8911
P-5			0.05	0.034485	37.43416	47.6774
P-6			0.08	0.051139	27.36477	56.5404
P-7		0.9	0.01	0.044631	23.98797	51.2904
P-8			0.05	0.03304	32.09351	49.4462
P-9			0.08	0.033634	32.46182	53.2809
P-10	15	0.7	0.01	0.036751	31.10727	48.9803
P-11			0.05	0.027466	40.88704	76.8287
P-12			0.08	0.031621	34.9152	70.9653
P-13		0.8	0.01	0.031186	33.44373	60.8208
P-14			0.05	0.031056	37.66119	60.80667
P-15			0.08	0.030808	35.79233	80.3332
P-16		0.9	0.01	0.036034	31.70874	65.4452
P-17			0.05	0.038	30.36333	58.6501
P-18			0.08	0.046064	25.63027	74.2926
P-19	30	0.7	0.01	0.036145	31.40608	70.9404
P-20			0.05	0.037151	30.19952	72.5266
P-21	30	0.7	0.8	0.036823	31.3673	73.3505
P-22		0.8	0.01	0.040043	26.36159	68.3309
P-23			0.05	0.034626	34.18596	76.8052
P-24			0.08	0.042783	27.09911	89.0267
P-25		0.9	0.01	0.04117	28.82216	73.6445
P-26			0.05	0.032005	35.62995	116.7407
P-27			0.08	0.031048	34.7796	86.1315

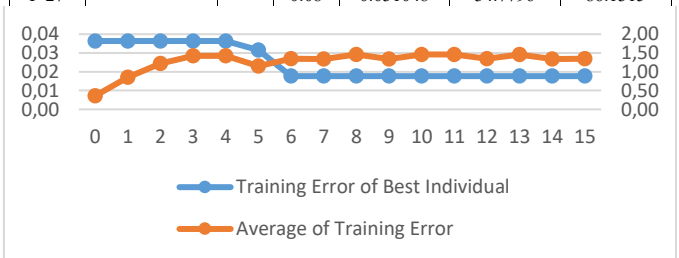

Fig. 3. The best individual training error evolution chart and average P-11 experiment training error

training error value of 0.027466. Then it is determined that the combination of genetic algorithm parameters on P-11 is the combination that produces the best output with a combination of parameters namely maximum generation = 15, cp = 0.7, and mp = 0.05. Table 1 also shows the duration of the optimization process of the best individuals in each combination and the average value of the 10 trials carried out in each combination, where it is known that trials requiring the least processing time average are P-2 with a duration of 41.1712 seconds. One of them can be caused because P-2 has a maximum generation value = 5, so the processing time will tend to be faster. While the average processing time of the selected combination is P-11 is and 76.8287 seconds. In Fig 3 shows the evolution of training errors from the best individuals and averages for each generation for the P-11 experiment.

From Fig 3 we know that the training error of best individual is always decrease, although the average of training error is unstable on each generation.

The testing scenario 2 aims to get the best combination of selected features and parameters of backpropagation neural networks in the feature selection process and optimization of genetic algorithm based backpropagation neural network parameters for breast cancer classification based on microRNA profile (GABPNN_FS) which includes random state generators for initial weighting, number hidden neurons in the hidden layer, and learning rate values. Tests are performed using the best genetic algorithm parameter values in scenario 1.

The experiment was carried out 10 times to obtain various possible results. The value for the parameters used is the maximum generation = 15, cp = 0.7, mp = 0.05, population = 15, maximum epoch = 1000, and target error = 0.001, and the K value for K-Fold Cross Validation is 10. Number of gen bits used on the testing chromosomes of GABPNN_FS for P, Q, R, and S, respectively 15, 9, 9, and 1464. The experimental results in scenario 2 are shown in Table 2.

TABLE II. VALUE OF ERROR TRAINING, VALUE OF FITNESS, AND DURATION OF THE BEST INDIVIDUAL PROCESS 2

ID	Max. Generation	CP	MP	Result		
				Error	Fitness	Process Duration
F-1	5	0.7	0.05	0.056047	17.84201	53.599
F-2				0.052767	18.95126	19.567
F-3				0.035055	28.52632	56.664
F-4				0.033357	29.97899	51.493
F-5				0.046161	21.66326	136.545
F-6				0.031109	32.1453	42.988
F-7				0.024305	41.14433	30.439
F-8				0.028197	35.46419	51.505
F-9				0.037431	26.71586	29.845
F-10				0.039425	25.36474	60.044

TABLE III. RANDOM STATE VALUE, LEARNING RATE, HIDDEN NEURON, AND THE BEST AMOUNT OF INDIVIDUAL FEATURES EXPERIMENT 2

ID	Random State	Learning Rate	Hidden Neuron	Jumlah Fitur
F-1	32695	0.8	8	751
F-2	30981	0.7	3	579
F-3	32767	0.9	9	627
F-4	30503	0.9	8	1294
F-5	30332	0.7	6	1144
F-6	32767	0.7	9	946
F-7	6098	0.7	6	707
F-8	32763	0.9	9	960
F-9	27575	0.8	5	1101
F-10	32767	0.8	9	1384

Table 2 shows the training error value, fitness value, and the best individual process duration of the 10 times the experiment was conducted. Based on the data in Table 2, it is known that the lowest error obtained is 0.024305 on F-7. The fastest process duration was also found in the FS-2 experiment with 19,567 seconds.

In Table 3 it is known the random state value, learning rate, number of hidden neurons, number of features, and the best number of individual connections from the 10 times the experiment was carried out, which is the result of the chromosome translation phase used to calculate fitness values. Determination of the best output results in scenario 2 is to use the lowest training error value with the least number of selected features. Based on Table 2 and Table 3, it is known that the F-7 has the lowest training error of 0.024305 with the least number of selected features, namely 707 features. The random state value, learning rate, number of hidden neurons, and number of features for FS-7 are 6098, 0.7, 6, and 707. In Fig 4, the evolution of training error values from the best individuals and the average of each generation for the F-experiment is shown. 7.

Fig 4 shows that the training error of best individual is stable and average of training error is slowly increase.

The test uses test data that is run using K-Fold Cross Validation for the best individual in the best experiment, F-7 produces an accuracy value = 97.50%, sensitivity = 99.00%, and specificity = 96.00%.

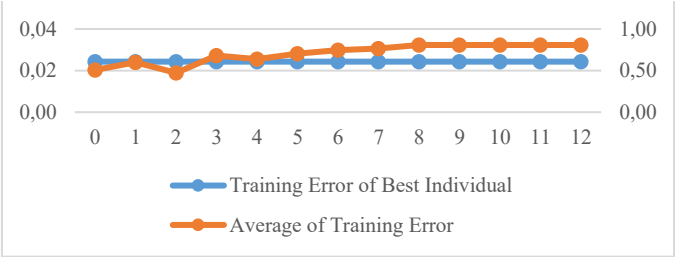

Fig. 4. The best individual training error evolution chart and F-7 experiment training error average

TABLE IV. COMPARISON OF GABPNN PERFORMANCE AND GABPNN_FS PERFORMANCE

	GABPNN		GABPNN_FS	
	Best	Avg.	Best	Avg.
Training Error	0.016115	0.027466	0.024305	0.038385
Random State	3926	18142	6098	28924
Learning Rate	0.8	0.77	0.7	0.79
Hidden Neuron	5	6	6	7
Num. of Features	1464	1464	707	949
Num. of Connections	7331	9969	4255	7011
Duration of Process	59.1009	76.8287	30.439	53.2689

Table 4 compares the best individual results and the average of 10 trials for testing GABPNN in scenario one and testing GABPNN_FS. Based on the data in Table 4 it is known that GABPNN_FS has a higher average training error compared to the GABPNN, which is 0.038385 compared to 0.027466. But it is known that GABPNN_FS takes less time to complete its evolution process with 53.2689 seconds compared to 76.8287 seconds. This can occur due to several reasons, one of which is the smaller GABPNN learning rate value which is 0.77 compared to GABPNN_FS with 0.79. A small learning rate will make the pace of learning slow and requires more iterations, so that even though the maximum number of epochs has been fulfilled, the network has not reached the specified error target. Other reasons can also be caused by fewer features classified in GABPNN_FS.

IV. Conclusion and Future Work

The conclusions that can be drawn from this thesis research are as follows: (1) Combination of selected features and genetic algorithm based backpropagation neural network parameters for breast cancer classification based on the best microRNA profile, namely random state = 6098, learning rate = 0.7, number of hidden neuron = 6, and the number of selected features = 707 features, according to the list of features of the microRNA profile shown in Appendix 4 which produces values of accuracy, sensitivity, and specificity, namely 97.50%, 99.00%, and 96.00%; (2) In this final project, the GABPNN method shows better results than GABPNN_FS with an average training error value of 0.027466 compared to 0.038385. But GABPNN_FS has a faster working duration with an average process duration of 53.2689 seconds compared to 76.8287 seconds; (3) Optimization of backpropagation parameters using genetic algorithms for classification of breast cancer based on microRNA profiles shows the best performance that produces the best individuals by using a combination of genetic algorithm parameters namely population size = 15, crossover probability = 0.7, mutation probability = 0.05, and maximum generation = 15.

Suggestions that can be given for the development of further research are as follows: (1) Comparing the performance of optimization and feature selection using genetic algorithms with other optimization algorithms and feature selection; (2) Dig deeper information about each selected feature and find out whether these features can be re-selected so that they can obtain fewer features with high accuracy results; (3) Comparing the Backpropagation method with other classification methods to get the best accuracy and features; (4) Comparing single-point crossover and bit flip mutation methods with other crossover and mutation methods.

References

[1] S. L. Anwar, S.J. Haryon, T Aryandono. & S. M. Haryana, 2018. *MicroRNA : Biogenesis, Fungsi, dan Perannya dalam Proses Karsinogenesis dan Penatalaksanaan Kanker.* Yogyakarta: Gadjah Mada University Press.

[2] IARC, 2018. *Cancer Today IARC.* [Online]. Available at: https://gco.iarc.fr/today/online-analysis-pie. [Accessed 6 December 2018].

[3] Zhu, J. et al., 2014. Different MiRNA Expression Profiles Between Human Breast Cancer Tumors and Serum. *Frontiers in Genetics,* 5(149).

[4] Saputro, A. A. S. & Kusumawati, Y., 2016. Pengaruh Pelatihan Sadari Terhadap Pegetahuan, Sikap dan Cara Deteksi Dini Kanker Payudara Pada Siswi SMK Dwija Dharma Boyolali. *Skripsi Thesis UMS.*

[5] Rosenfeld, N. et al., 2008. MicroRNAs Accurately Identify Cancer Tissue Origin. *Nature Biotechnology,* Volume 26, pp. 462-469.

[6] Sukmawati, N. E. et al., 2016. Backpropagation dan Aplikasinya. In: *Ilmu Komputer Studi Kasus dan Aplikasi.* s.l.:Undip Press, pp. 135-146.

[7] Zamani, A. M., Amaliah, B. & Munif, A., 2012. Implementasi Algoritma Genetika pada Struktur Backpropagation Neural Network untuk Klasifikasi Kanker Payudara. *Jurnal Teknik ITS,* 1(A), pp. 222-227.

[8] Yan, K. & Zhang, D., 2015. Feature Selection and Analysis on Correlated Gas Sensor Data with Recursive Feature Elimination. *Sensors and Actuators,* B(212), pp. 353-363.

[9] Ahmad, F. et al., 2015. A GA-based Feature Selection and Parameter Optimizaion of an ANN in Diagnosing Breast Cancer. *Pattern Anal Applic,* Issue 18, pp. 861-870.

[10] Goldberg, D. E., 1989. *Genetic Algorithms: in Search, Optimization, and Machine Learning.* Canada: Addison-Wesley Publishing Company.

[11] Castillo, P. A. et al., 2000. G-Prop: Global Optimization of Multilayer Perceptrons Using GAs. *Naurocomputing,* 35(1-4), pp. 149-163.

[12] Suhendra, C. D. & Wardoyo, R., 2015. Penentuan Arsitektur Jaringan Syaraf Tiruan Backpropagation (Bobot Awal dan Bias Awal) Menggunakan Algoritma Genetika. *IJCCS,* 9(1), pp. 77-88.

[13] NCI, 2017. *National Cancer Institute Genomic Data Commons - The Next Generation Cancer Knowledge Network.* [Online]. Available at: https://gdc.cancer.gov/. [Accessed 9 11 2018].

[14] Siang, J. J., 2009. *Jaringan Syaraf Tiruan dan Pemrogramannya Menggunakan Matlab.* 1st ed. Yogyakarta: Andi Offset.

[15] Pitria, P., 2014. *Analisis Sentimen Pengguna Twitter pada Akun Resmi Samsung Indonesia dengan Menggunakan Naive Bayes.* Bandung: Teknik Informatika Universitas Komputer Indonesia.

[16] Mahmudy, W. F., 2013. *Algoritma Evolusi.* Malang.

[17] Madhiarasan, M. & Deepa, S. N., 2015. A Novel Criterion to Select Hidden Neuron Numbers in Improved Back Propagation Networks for Wind Speed Forecasting. *Applied Intelligence,* 44(4), p. 878–893.

[18] Lin, S. M. et al., 2017. Crossover and Mutation Operators of Genetic Algorithms. *International Journal of Machine Learing and Computing,* 7(1), pp. 1-12.