Leukemia and colon tumor detection based on microarray data classification using momentum backpropagation and genetic algorithm as a feature selection method

Untari N Wisesty¹, Riris S Warastri² and Shinta Y Puspitasari³
¹,²,³ School of Computing, Telkom University, Bandung, Indonesia
¹ untarinw@telkomuniversity.ac.id

Abstract. Cancer is one of the major causes of morbidity and mortality problems in the worldwide. Therefore, the need of a system that can analyze and identify a person suffering from a cancer by using microarray data derived from the patient's Deoxyribonucleic Acid (DNA). But on microarray data has thousands of attributes, thus making the challenges in data processing. This is often referred to as the curse of dimensionality. Therefore, in this study built a system capable of detecting a patient whether contracted cancer or not. The algorithm used is Genetic Algorithm as feature selection and Momentum Backpropagation Neural Network as a classification method, with data used from the Kent Ridge Bio-medical Dataset. Based on system testing that has been done, the system can detect Leukemia and Colon Tumor with best accuracy equal to 98.33% for colon tumor data and 100% for leukemia data. Genetic Algorithm as feature selection algorithm can improve system accuracy, which is from 64.52% to 98.33% for colon tumor data and 65.28% to 100% for leukemia data, and the use of momentum parameters can accelerate the convergence of the system in the training process of Neural Network.

1. Introduction
Cancer is one of the major causes of morbidity and mortality worldwide, with 14 million new cases and 8.8 million cancer-related deaths by 2015. It is estimated that the number of new cases will increase by about 70% over the next 2 decades [1]. Detection of cancer can be done by obtaining information from the patient's DNA data in the form of microarray data that has thousands of attributes. Data that has many attributes (dimensionality), will take a lot of time and take advantage of excessive computing efforts that are difficult to handle. Therefore, a dimension reduction method is needed to select the optimal attributes so that computational effort can be performed efficiently [2].

Vanitha D.A. [3] use Mutual Information (MI) as feature selection and classification methods, namely ANN with 50.94% accuracy and SVM with 67.74% accuracy on Colon Tumor data. Nurfalah.A [4] construct a scheme using PCA as dimensional reduction and Backpropagation with Conjugate Gradient as microarray data classification, resulting an accuracy of 76.92% for Colon Tumor and 97.14% for Leukemia data. In both studies, it is difficult to determine the number of optimal attributes produced by MI and PCA.

In previous research, microarray data classification process using Evolutionary Artificial Neural Network (EANN) which is a combination of Genetic Algorithm (GA) as feature selection and Artificial Neural Network (ANN) as cancer classification. From the merger is produced very good performance. This is because EANN to get a solution does not require expert knowledge and there is a selection of attributes in the evolutionary process that can avoid the occurrence of ANN network...
structures that are too large which can lead to an increase in the computation process, resulting in lower performance [5]. In addition, other studies related to GA use for feature selection in the classification process i.e. GA and Backpropagation Neural Network (BPNN) combined resulted in an average accuracy of 93.8% [2].

BPNN is one of the learning algorithm of Multi-Layer Neural Network. BPNN provides an efficient computational method for converting weights in a forward network feed with differentiated activation in the function unit. In previous research, BPNN was used to detect breast cancer by producing higher accuracy value than Support Vector Machine (SVM) and Decision Tree, where 99% accuracy was obtained for BPNN, 96.19% for SVM and 92.38% for Decision Tree [6]. Other studies related to BPNN have an accuracy of 93% for colon tumor dataset and 95% in leukemia dataset, but the result of accuracy obtained is still not maximal because it still processes from all attributes, so that research can be continued by applying GA for feature selection process [7]. However, in another study said that BPNN still has a lack of convergence is slow and requires a long training process. This can be overcome by adding momentum to the BPNN used on each iteration, so that it will speed up the training process. Momentum is a parameter used to prevent a system from converging to a local minimum and can increase the convergence speed of the system [8].

In this research built a system that can classify and detect a person whether the patient suffered from cancer or not, using GA and Momentum BPNN algorithm in the case of high-dimensional data. Leukemia and colon tumor patients’ data were from Kent Ridge Bio-medical Dataset, where data of colon tumor patients with the number of attributes 2000 [9] and in leukemia patients had the number of attributes of 7129 [10]. With the detection model is expected to help overcome the problems in the medical world.

2. Research Scheme

In this section, we will describe the general system design that will be used to detect Leukemia and Colon Tumor in high-dimensional data. In detecting Leukemia and Colon Tumor and getting good performance, dimensional reduction process is needed to obtain optimal attribute. The method used is a combination of Genetic Algorithm and Momentum Backpropagation (Momentum BPNN). Genetic Algorithm serves to perform dimensional reduction process and Momentum BPNN to perform the process of classification. Data used in this study are leukemia and colon tumor data taken from Kent Ridge Bio-medical Dataset. The design of Leukemia and Colon Tumor detection system that is built can be seen in Figure 1.

![Figure 1. Design of Leukemia and Colon Tumor Detection System using Genetic Algorithm and Momentum BPNN.](image-url)
2.1. Dataset
Data used in this research is data of colon tumor and Leukemia disease. The data were obtained from the Kent Ridge Bio-medical Dataset [9, 10]. The data available on the repository are gene expression data, protein profiling data and genomic sequence data. The data used included into high-dimensional data that having 2000 attributes for colon tumor data and 7129 attributes for Leukemia data, contains 62 samples for colon tumor data and 72 samples for Leukemia data. Colon tumor and Leukemia sample data can be seen in Table 1 and 2.

Table 1. Colon Tumor Data Samples

| No | Attribute 1 | Attribute 2 | Attribute 3 | Attribute 4 | Attribute 5 | ... | Attribute 2000 | Class |
|----|-------------|-------------|-------------|-------------|-------------|-----|----------------|-------|
| 1  | 8589.42     | 5468.24     | 4263.41     | 4064.94     | 1997.89     | ... | 28.70         | Negative |
| 2  | 9164.25     | 6719.53     | 4883.45     | 3718.16     | 2015.22     | ... | 16.77         | Positive  |
| 3  | 3825.71     | 6970.36     | 5369.97     | 4705.56     | 1166.55     | ... | 15.16         | Negative  |
| 4  | 6246.45     | 7823.53     | 5955.84     | 3975.56     | 2002.61     | ... | 16.09         | Positive  |
| 5  | 3230.33     | 3694.45     | 3400.74     | 3463.59     | 2181.42     | ... | 31.81         | Negative  |
| ...| ...         | ...         | ...         | ...         | ...         |     | ...           | ...     |
| 62 | 7472.01     | 3653.93     | 2728.22     | 3494.48     | 2404.67     | ... | 39.63         | Positive  |

Table 2. Leukemia ALL-AML Data Samples

| No | Attribute 1 | Attribute 2 | Attribute 3 | Attribute 4 | Attribute 5 | ... | Attribute 7129 | Class |
|----|-------------|-------------|-------------|-------------|-------------|-----|----------------|-------|
| 1  | -214        | -153        | -58         | 88          | -295        | ... | -37            | ALL   |
| 2  | -139        | -73         | -1          | 283         | -264        | ... | -14            | ALL   |
| 3  | -76         | -49         | -307        | 309         | -376        | ... | -41            | ALL   |
| 4  | -135        | -114        | 265         | 12          | -419        | ... | -91            | ALL   |
| 5  | -106        | -125        | -76         | 168         | -230        | ... | -25            | ALL   |
| ...| ...         | ...         | ...         | ...         | ...         |     | ...            | ...   |
| 72 | -135        | -186        | -70         | 337         | -407        | ... | -10            | AML   |

2.2. Data Preprocessing
At the pre-processing stage will be done some processing steps. The first process is the dataset will be processed using the normalization method. The normalization method is used to convert data containing real numbers into range [0,1]. The second process is to divide the data that has been normalized into two parts, namely testing and training data using K-Fold Cross Validation, where K=10.

2.3. Feature Selection using Genetic Algorithm
The feature selection stage is done after going through the pre-processing and data division. In the feature selection process uses training data as input. Training data will enter at population initialization process by generating a number of chromosomes with a random value of 0 and 1 with a gene length equivalent to the number of attributes entered. Next is the decoding stage that is the selection of attributes from the result of randomly generated population initialization, where the attribute of value 0 is not used and the attribute of value 1 will be used in the next step. The attribute reduction is done by using the feature selection method to generate the optimal attribute. Examples of chromosome representation results can be seen in Figure 2.

![Figure 2. Example of Chromosome Representation Results](image-url)
The next stage is the process of individual evaluation by applying the process of learning and validation using the Momentum BPNN algorithm which will result training accuracy where the training accuracy results are used as the fitness value to enter at the next stages of GA. The highest fitness value obtained will be used as the best solution.

The elitism process is done to store the chromosome with the best fitness value obtained from the previous process, so that the best fitness value or solution ever obtained is not lost [11]. The next stage is every individual through the process of parent selection using Tournament Selection. Tournament Selection method is used because according to the result of literature study which has been done, where the method of Fitness Proportionate Selection (FPS) with Roulette Wheel and Rank Based Selection method still has weakness that is the bottleneck where more computation process is needed to complete the selection process due to the size of the population used [12]. The parent selection process will produce an individual parent who will be recombined on each of the chromosomes. Recombination is done to produce new and more varied individuals, where each child's chromosome has genes from their parents. The recombination process uses 1 point crossover.

New individuals generated from the recombination process will then undergo the mutation process. In the process of mutation begins with the generation of random numbers. If the random number generated is less than the mutation probability then the gene will be replaced by the value of the opposite. Selection of survivors is the process of replacing old chromosomes with new chromosomes. The survivor selection is the generational replacement model. Then check whether the system has reached the maximum generation given or not. If not, then the process will return again in the decoding stage, but if it has reached the maximum value it will produce the most optimal attributes. These attributes will then continue in the process of Leukemia and Colon Tumor detection.

2.4. Training Process of Momentum BPNN

The data used in the training stage of Momentum BPNN is the training data that has been through the feature selection stage. Then enter at the training stage using Momentum BPNN. Where in the first stage is the initialization of the parameters that will be used in the construction of network architecture in Momentum BPNN. Then going into the second stage is a random process of weight and bias that will be used in the training stage. Next will be checked for epoch value, if epoch value is more than maximized epoch value that has been initialized at the beginning then the process is done, but if epoch value less than maximum epoch value specified it will go to next stage [13]. In the Backpropagation algorithm there are two stages, namely feed forward phase and backward phase. In the feed forward phase there are three processes, namely the calculation of the output of the hidden layer and output layer and put it into the activation function in the form of sigmoid function [14], as well as calculating the error value (MSE). With the calculation of the error value then the feed forward stage has been completed.

The next stage is the backward phase that there is a process of calculating the factors of weight and bias changes to generate new weight and bias values based on errors that have been obtained before. In the calculation of weight change factors added momentum parameters to avoid premature convergence, as in equations 1 and 2 [15]. The last stage of calculating the new weights and bias based on the old weights and bias as well as weight change factor. Next will be done looping on the feed forward and backward process in accordance with the epoch value used. The optimal weight and bias with the smallest MSE will be stored for use in the testing phase.

\[
\Delta w_{jk}(t + 1) = \alpha. \delta_k. z_j + mc. \Delta w_{jk}(t) \tag{1}
\]

\[
\Delta v_{ij}(t + 1) = \alpha. \delta_j. z_j + mc. \Delta v_{ij}(t) \tag{2}
\]

Where:
\[
\delta_k : \text{Error information for } k^{th} \text{ output unit}
\]
\[
\delta_j : \text{Error information for hidden unit } Z_j.
\]
\[
z_j : \text{Hidden unit } j^{th}.
\]
\( \alpha \) : Learning rate
\( mc \) : Momentum
\( \Delta w_{jk} \) : The weight change factor of the j-hidden units and k-output units.
\( \Delta v_{ij} \) : The factor changes the weight of the i-input unit to the j-hidden unit.

2.5. Testing Process of Momentum BPNN
The testing process is used to obtain Leukemia and Colon Tumor detection results in data testing. In the testing stage used only the feed forward process without using the backward stage, so there is no weight update process. At the testing stage we use the best weights and biases that have been obtained in the training process. The classification results are obtained from the output of the neuron at the output layer, where the output value has been converted into 0 or 1. The number 1 states detected cancer, and vice versa. Furthermore, the results obtained will be used in the calculation process of accuracy by using a comparison between the actual data with classification results. In the testing process produces an output of accuracy along with the label of each data.

3. Analysis and Experimental Results
The purpose of system testing to be conducted in this study is to obtain optimal attributes that can detect whether a person is positively infected with cancer or not on microarray data of a high dimension. In the test, things that can affect system performance are structures in network construction on Momentum Backpropagation (Momentum BPNN), as well as parameters of the number of individuals in a population (population size), mutation probability (Pm), and crossover probability (Pc) in GA.

3.1. Influence Analysis of Momentum Parameters on BPNN Algorithm
In the BPNN algorithm has several variations that can be used, one of them is with the addition of parameters momentum (mc) on the process of weight and bias improvement. In this research, the system test using standard BPNN and BPNN using momentum (mc) or commonly known as momentum BPNN, where this test is done to know the influence of the momentum parameter in BPNN. The momentum parameter value (mc) used in momentum BPNN is 0.9, while for other parameter values both in BPNN and in BPNN momentum is the same, as in Table 3. This is done in order to know the difference of standard BPNN with BPNN using momentum.

| Table 3. BPNN Testing Parameters With and Without Momentum Parameters |
|---------------------------------------------------------------|
| **BPNN Parameters** | **GA Parameter** |
| Epoch | Learning Rate | Hidden Layer | Max Generation | Population Size | Pc | Pm |
| 100 | 0.01 | 5 | 20 | 50 | 0.65 | 0.01 |

Figure 3 show the value of MSE generated takes a long time to reach convergent point, where at epoch reach 100-point value of MSE that have not reached convergence point with accuracy value obtained equal to 83.33% for data colon tumor and 91.42% for leukemia data. While in Figure 4, it can be seen that adding momentum to BPNN can accelerate convergence time and increase accuracy value 94.99% for colon tumor and 98.57% data for leukemia data, this is because the momentum parameter can make a significant leap toward MSE value which is produced closer to the minimum position, this is done by taking the weight value on the previous iteration and adding the current weighting value to improve the learning on the BPNN algorithm. In the presence of the momentum parameter, the convergence point can be reached when it is at the epoch value 10.
Thus, it can be concluded that if using momentum on BPNN can accelerate the time to reach the converging point in the training process. So with a faster convergence process will accelerate also in the training process. This proves the results of the analysis ever done that the addition of momentum BPNN can overcome the weakness of the BPNN algorithm is a slow convergence and require a long training process.

3.2. Testing of Number of Hidden Neurons on the Momentum BPNN Algorithm
This test is used to determine the effect of number of hidden neurons to the accuracy produced. The number of hidden neurons used are 5, 10 and 15. The average accuracy obtained from these three values can be seen in Figure 4.

| Number of Hidden Neurons | Accuracy (%) | Colon Tumor | Leukemia |
|--------------------------|--------------|-------------|----------|
| 5                        | 94.62        | 95.71       |
| 10                       | 94.62        | 95.71       |
| 15                       | 94.62        | 95.71       |

In Table 4, it appears that the average accuracy value obtained yields the same result for the different number of hidden layer neurons. In the data colon tumor generated an average accuracy of 94.62% for the number of hidden neurons 5, 10 and 15. While the leukemia data yields an average accuracy of 95.71% for the number of hidden neurons 5, 10 and 15. It can be concluded that in this test the number of hidden neurons does not affect the accuracy of results obtained. This is because the hidden layer is used to build the network architecture that will be used in Momentum BPNN so that the number of neurons in the hidden layer depends on the uniqueness of each dataset pattern used.

3.3. Testing Momentum and Learning Rate Parameters on Momentum BPNN Algorithm
This test is used to determine the effect of learning rate and momentum parameter values on the resulting accuracy. The value of momentum used is 0.9, 0.5 and 0.1, while the value of learning rate used is 0.01, 0.1 and 0.2. Accuracy value obtained from the combination of momentum and learning rate in colon tumor data can be seen in Table 5 and leukemia data in Table 6.

From the Table 5 and Table 6, can be seen that the colon tumor data get the highest accuracy value is by using a combination of momentum parameter value (mc) 0.9 and the value of learning rate parameter (lr) 0.2 with an accuracy of 98.33%. While the leukemia data obtained results have similarities with the data Colon Tumor is the highest accuracy value obtained on the combination of
momentum parameter value (mc) 0.9 and the value of learning rate (lr) 0.2 with 100% accuracy value obtained.

### Table 5. Testing Result of Momentum and Learning Rate Parameters In Colon Tumor Data to System Accuracy

| Momentum | Learning Rate | Accuracy (%) |
|----------|---------------|--------------|
| 0.9      | 0.01          | 94.99%       |
|          | 0.1           | 96.66%       |
|          | 0.2           | 98.33%       |
| 0.5      | 0.01          | 96.66%       |
|          | 0.1           | 94.99%       |
|          | 0.2           | 96.66%       |
| 0.1      | 0.01          | 83.33%       |
|          | 0.1           | 93.32%       |
|          | 0.2           | 96.66%       |

### Table 6. Testing Result of Momentum and Learning Rate Parameters In Leukimia Data to System Accuracy

| Momentum | Learning Rate | Accuracy (%) |
|----------|---------------|--------------|
| 0.9      | 0.01          | 98.57%       |
|          | 0.1           | 97.14%       |
|          | 0.2           | 100%         |
| 0.5      | 0.01          | 95.71%       |
|          | 0.1           | 92.85%       |
|          | 0.2           | 98.57%       |
| 0.1      | 0.01          | 91.42%       |
|          | 0.1           | 94.28%       |
|          | 0.2           | 92.85%       |

From the results of both data accuracy can be concluded that the greater the value of momentum (mc) and learning rate (lr) used can achieve optimum accuracy. This is because if the greater the value of momentum used it will accelerate the network weight changes to achieve the minimum error or reach the convergent point, so that the faster in achieving the minimum error and convergence will affect the value of accuracy obtained. The larger momentum parameter will avoid the convergence at local minimum, so we can get the optimum weight and bias value. With the acceleration of the training process, the accuracy of the value obtained will also be higher. However, if the value of learning rate is too large, it will cause the learning process of Momentum BP not get optimum convergence. This makes the system unable to produce the expected accuracy value.

3.4. Testing of Genetic Algorithm Parameters

Based on the results of the test, the combination of GA parameters used are maximum generation, population size, crossover probability (Pc), and mutation probability (Pm) by using pre-processing process first then will be entered into the system with the data division of training and testing using 10-Fold Cross Validation. The resulting output is a classification class and accuracy of the class that is detected correctly.

### Table 7. Result of Population Size Experiment

| Population Size | Accuracy (%) |
|-----------------|--------------|
|                 | Colon Tumor  | Leukimia    |
| 50              | 91.78        | 97.36       |
| 100             | 94.53        | 97.83       |
| 200             | 94.87        | 98.12       |

In Table 7, it can be seen that in colon tumor data using population size at GA that is 50 produce accuracy 91.78% and with population size 100 produce accuracy 94.53%, while using population size 200 obtained highest accuracy is 94.53%. In leukemia data obtained the test results with 97.26% accuracy for population size 50 and 97.83% for population size 100. While the highest accuracy obtained from the test results using population size 200. So it can be concluded that the greater the size of the population the greater the accuracy obtained. This is because when the size of the population gets bigger, then the number of samples taken as a candidate solution will also be more and more. So that will increase the probability of getting the best solution among the large solution space.
The second parameter used in GA is the crossover probability (Pc) used to determine what percentage of the probability of doing crossover in order to obtain a good individual. The effect of crossover probability can be seen in Table 8.

| Crossover Probability | Colon Tumor | Leukimia |
|-----------------------|-------------|----------|
| 0.65                  | 91.80       | 97.39    |
| 0.9                   | 95.65       | 98.15    |

In Table 8, it can be seen that in colon tumor data using crossover probability value (Pc) 0.65 resulted 91.80% accuracy and 95.65% accuracy using crossover probability value (Pc) 0.9. Leukemia data obtained 97.39% accuracy by using probability value of crossover (Pc) 0.65 and get 98.15% accuracy using crossover probability value (Pc) 0.9. So the test results from both data, it can be concluded that the higher the probability value of crossover (Pc) used then the higher the accuracy produced. This is because the greater probability of crossover (Pc) will cause greater chromosomal chances of crossover, thus increasing the chances of getting the best individual that will have an effect on the accuracy obtained.

| Mutation Probability | Colon Tumor | Leukimia |
|----------------------|-------------|----------|
| 0.65                 | 91.80       | 97.39    |
| 0.9                  | 95.65       | 98.15    |

In Table 9, it can be seen that in colon tumor data using mutation probability value (Pm) 0.01 yield accuracy of 93.92% and accuracy of 97.53% if using mutation probability (Pm) of 0.3. While on leukemia data, on graph of test result can be seen that by using mutation probability value (Pm) 0.01 yield accuracy 93.54% and 98.01% if using mutation probability value (Pm) 0.3. Can be concluded from the test results on both data is generated the highest accuracy if the value of mutation probability (Pm) is used greater. Where mutation probability (Pm) is used to control how many chromosomes will be mutated. So the greater the probability value of mutation (Pm) the more chromosomes will undergo the mutation process. This causes more new chromosomes to be generated from the mutation process and used in subsequent generations, so that the individual generated will also be more varied and can prevent the system from being trapped in the local optimum.

3.5. Testing of Feature Selection Effect on System Performance
Based on the results of tests that have been done on colon tumor data with the number of attributes 2000 and leukemia data with the number of attributes 7129 done classification with Momentum BPNN using GA and Momentum BPNN without using GA, where GA is used as feature selection process for optimal attribute selection. The result of comparison on detection of cancer patient by using Momentum BPNN with GA and without GA can be seen in Table 10.

| Algorithm                     | Colon Tumor | Leukimia |
|-------------------------------|-------------|----------|
|                              | Number of Attributes | Accuracy (%) | Number of Attributes | Accuracy (%) |
| Momentum BPNN                | 2000        | 64.516%  | 7,129        | 65.277%       |
| Momentum BPNN and GA         | 968         | 98.33%   | 3526         | 100%          |
Based on Table 10, accuracy obtained from testing of detection process on cancer patient using Momentum BPNN without using GA that is in colon tumor data have accurate value 64.516% with attribute count used is 2,000 attributes, while leukemia data has accurate value 65.277% with attributes used as many as 7,129 attributes. Both accuracy values are still low, this is because all attributes that exist in both data are still processed as a whole, so the attribute is less informative also used in the process of classification.

While in Momentum BPNN by using GA get accuracy value equal to 98.33% with optimal attribute that used is 968 for colon tumor data, while for leukemia data yield 100% accuracy value with optimal attribute count is 3,526 attributes. In the test results it can be seen that Leukemia and Colon Tumor detection using Momentum BPNN and GA has a higher accuracy value compared to Momentum BPNN without using GA, this is because the attributes processed in the classification process has been reduced so that only the optimal attributes used in the classification process.

It can be concluded that the use of Momentum BPNN algorithm using GA has a higher accuracy value than without using GA. Increased accuracy is due to GA can be used for dimensional reduction process through the feature selection stage to select the optimal attribute without reducing the information from the data. So with dimension reduction with GA as feature selection which only chooses the optimal attribute which will be used in the classification process, then the attribute is processed only the selected attribute is not the whole of the attribute. Whereas, if without using GA will cause the overall attributes contained in the data used in the classification process, so the presence of attributes that do not provide information to be processed in the process of classification. This may affect the accuracy and performance results of the system being built because not all attributes have the information used for the classification process.

4. Conclusion
Based on the system testing that has been done, it can be concluded that the system built has been very good in detecting colon tumor and leukemia disease with the best accuracy of 98.33% for colon tumor data and 100% for leukemia data. The feature selection process using Genetic Algorithm has an effect on improving system accuracy, from 64.52% to 98.33% for colon tumor data and 65.28% to 100% for leukemia data. This is due to the feature selection process so as to produce the optimal attributes that will be used in the classification process. The existence of momentum parameters at BPNN can accelerate the time to reach the convergent minimum point in the training process. In BPNN test results without using momentum parameters convergent point were achieved at epoch 90 with an accuracy of 83.33% for colon tumor data and 91.42% for leukemia data, whereas when the momentum parameters were added to BPNN, convergent points could be achieved at epoch 10, with accuracy obtained for 98.33% in colon tumor data and 100% in leukemia data.

Acknowledgments
The authors would like to thank Telkom University for financially supporting in this research.

References
[1] World Health Organization 2016 Available Online: http://www.who.int/mediacentre/factsheets/fs297/en/. [Accessed 22 October 2016]
[2] Khobragade V P and Vinayababu D A 2012 A Classification of Microarray Gene Expression Data Using Hybrid Soft Computing Approach International Journal of Computer Science Issues vol 9 no 6 pp 246-255
[3] Vanitha C D A Devaraj D and Venkatesulu M 2015 Gene Expression Data Classification Using Support Vector Machine and Mutual Information-based Gene Selection Procedia Computer Science 47 pp 13-21
[4] Nurfalah A Adiwijaya and Suryani A A 2016 Cancer Detection Based on Microarray Data Classification Using PCA And Modified Back Propagation Far East Journal of
Electronics and Communications 16(2) pp 269-281

[5] Kim K J and Cho S B 2004 Prediction of Colon Cancer Using An Evolutionary Neural Network Neurocomputing pp 361-379

[6] Pawar P S and Patil D 2012 Breast Cancer Detection Using Backpropagation Neural Network with Comparison Between Different Neuron IEEE International Conference on Parallel, Distributed and Grid Computing pp 170-173

[7] Satpathy S and Mahapatra P 2014 Microarray Classification Using Intelligent Techniques International Journal of Scientific & Engineering Research vol 5 no 7 pp 1663-1667

[8] Yu C C and Liu B D 2002 A Backpropagation Algorithm With Adaptive Learning Rate and Momentum Coefficient IEEE pp 1218-1223

[9] Kent Ridge Bio-medical Dataset Available Online: http://datam.i2r.a-star.edu.sg/datasets/krbd/ColonTumor/ColonTumor.html [Accessed 22 October 2016]

[10] Kent Ridge Bio-medical Dataset Available Online: http://datam.i2r.a-star.edu.sg/datasets/krbd/Leukemia/ALLAML.html [Accessed 22 October 2016]

[11] Suyanto 2008 Evolutionary Computation Komputasi Berbasis "Evolusi" dan "Genetika" Bandung: Informatika

[12] Suyanto 2008 Soft Computing "Membangun Mesin Ber-IQ Tinggi" Bandung: Informatika

[13] Noertjahyana A and Yulia Studi Analisa Pelatihan Jaringan Syaraf Tiruan Dengan dan Tanpa Algoritma Genetika

[14] Andrian Y and Putra P H 2014 Analisis Penambahan Momentum Pada Proses Prediksi Curah Hujan Kota Medan Menggunakan Metode Backpropagation Neural Network Seminar Nasional Informatika pp 165-172

[15] Hermawan A 2006 Jaringan Saraf Tiruan, Teori dan Aplikasi Yogyakarta: Andi Offset