

Breast Cancer Diagnosis based on Genetic Algorithms and Neural Networks

Khaled M. A. Alalayah
Assistant Professor
IBB University, Science collage,
Computer Department, Yemen
Najran University, Science and Art
Collage, Sharurah, KSA

Siham A. M. Almasani
Assistant
Hodeida University
Faculty of Computer Science and
Engineering, Yemen

Wadeea A. A. Qaid
Assistant Professor
Hodeida University,
Faculty of Computer Science and
Engineering, Yemen

Ibrahim Abdulrab Ahmed
Professor
Najran University, Community College,
Computer Department, KSA

ABSTRACT

This paper proposes to use of hybrid genetic algorithm with artificial neural networks technique for diagnosing the Breast Cancer in order to give more accuracy than using neural network, since the performance of hybrid genetic algorithm with artificial neural networks for diagnosis of the Breast Cancer can measure by the value of fitness function. The number of hidden layers in the neural network has a significant effect on the classification performance and the best diagnoses performance average is attained when the number of layers equal three. The result by the using of GANN technique up to 94%.

Keywords

Classification; Neural network; Genetic Algorithms; Breast Cancer; Performance.

1. INTRODUCTION

Artificial intelligence elements like, artificial neural networks, genetic algorithms, fuzzy logic, machine learning, SVM etc. are tending to emulate the human brain, and consider to be one of the main interests of the researches nowadays.

Classification is one of the most frequent encountered problems in data mining at present. A classification occurs when an object needs to be assigned in predefined classes based on a number of observed attributes related to that object.

The Genetic algorithm and Artificial Neural Network are one of more applicable techniques of the intelligent techniques [7].

GA is an evolutionary optimization approach. It can give set of good solutions and discard bad proposals without affecting the end result negatively. Neural Networks are biologically inspired and mimic the human brain. A neural network consists of neurons which are interconnected with connecting links, where each link has a weight that is multiplied by the signal transmitted in the network [8] At all the number of hidden layers in the neural network has a significant effect on the classification [1]. Hybrid of genetic algorithm with artificial neural networks have been widely applied in many different fields, one of these applications is a classification.

Breast cancer is the fatal disease in women worldwide, which the risk increasing with the age [4], [5]. There are two types of breast lesions malignant and benign. The malignant tumor is in many cases misdiagnosed and the late diagnosis reduces the chances of survival of the patient. Thus, intelligent techniques are required to help the medical practitioners to decrease the instances of misdiagnosed.

Actually, there are various techniques used for the diagnosis of breast cancer,

Our research is based on using GA and ANNs to diagnose Breast Cancer.

In addition, this work proposes a solution based on the concept of Genetic Algorithm and Artificial Neural Networks for classification of the Breast Cancer, due to the the Genetic algorithm works to optimize Neural Networks for diagnosing the breast cancer through adjusting the weights and the bias.

2. MATERIALS AND METHODS

In this research, hybrid technique ANN-GA was devised which, based on input variables to helps predict the type of breast cancers. The proposed technique relied on the available data of 665 cases of patients with breast cancer that were stored in UCI Machine Learning Repository [10]. To implement the neural network and genetic algorithms, the applications and functions were used in Matlab 7.0, and 70% for the data were used for the training phase, whereas the remaining 30% were used for testing phase. Nine clinical variables were considered as the network inputs: clump thickness, uniformity of cell size, uniformity of cell shape, marginal adhesion, single epithelial cell size, bare nuclei, bland chromatin, normal nucleoli, and mitoses.

The existing data were then normalized through the linear method in such a way that numerical values could be expressed in terms of 0-1 binary sets to be used in the ANN-GA technique

2.1 Genetic algorithm

Genetic algorithms have been used in conjunction with neural networks in different ways. For example, this article used Genetic algorithms to set the weights and bias in fixed architectures. The aim of genetic algorithms is to choose the

optimal values for weights and bias for neural networks to give more accurate diagnosis.

The GA maintains a population of n chromosomes with associated the following fitness values.

$$F.F = \frac{1}{n} \sum_{i=1}^n (t - out)^2$$

Where: n-the number of output, t- the target output and out-the actual output.

Parents are selected on the basis of their fitness from the population to produce offspring and to form the new population. There are many methods for selection in genetic algorithm but the best method is Roulette wheel selection. The Roulette wheel selection always chooses the fittest individuals, which have a greater chance of survival than weaker ones. The next step is to generate a new population of solutions from those selected through the operators: crossover and mutation.

The selected values for genetic algorithm can be seen in Table 1. The initial population in GA with a length of the parameters (665), which represent the all weights and bias in the neural network.

Table 1. The genetic algorithm structure specification

Number of Generations	100
Number of Population	20
Crossover Probability	0.8
Mutation Probability	0.01
Selection function	Roulette
Fitness limit	1e-12

The terminating conditions of the generational process in the genetic algorithm as the following [9],[2]:

A solution is found that satisfies minimum criteria when the fitness function less than or equal to 1e-12; Fixed number of generations reached; Allocated budget for computation time reached; The highest ranking solutions fitness is reaching or has reach a plateau such that successive iterations no longer produce better results.

2.2 Artificial Neural Network

Neural networks are able to learn from previous experiences, improve their behavior while being trained. Neural networks have been widely used in various fields such as artificial intelligence, pattern recognition, medical diagnosis, machine learning and so on [3]. The neural network is suitable for solving the classification problem [6].

The major steps of using neural network learning algorithm in breast cancer classification can be summarized as follows: In the beginning give the training samples and the class of sample, the network output of each sample is compared with the actual known class and then the weight of each training sample is adjusted to achieve the purpose of classifying.

Each node in the network corresponds to the output node of a network, while the lines from the bottom into node are regarded as its input. Intermediate layer is called the hidden layer. The output of the hidden layer is considered as the input of output layer.

The use of neural network to classify breast cancer is illustrated in Fig. 1. The network is trained with three hidden layers, the hidden layers neurons number varies from 1 to 18, which the network was trained 40 times.

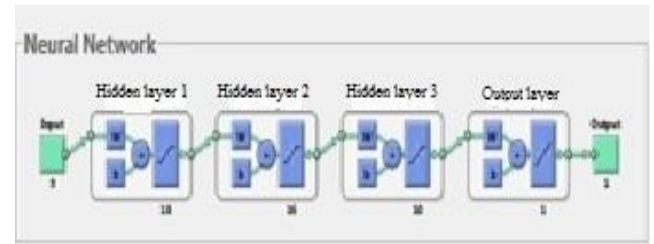


Fig. 1. Structure of neural network to classify breast cancer

Mean Square Error value (MSE) for the 40 training trails: 4.796906490444106E-6.

The main idea related to the hybrid of genetic algorithms and neural network in Fig.2.

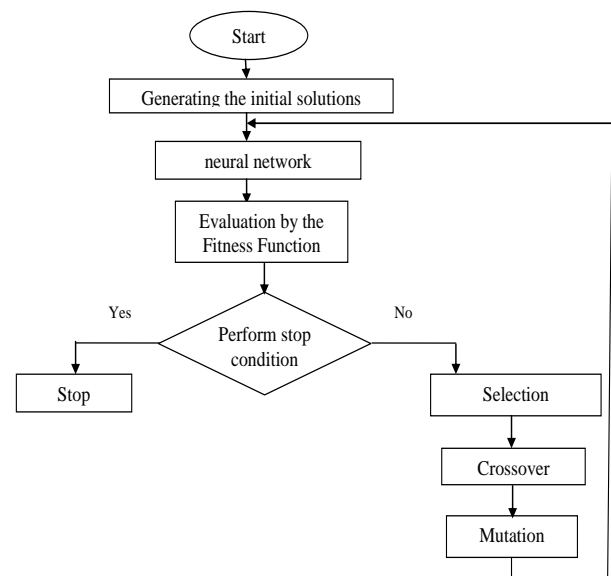


Fig. 2. The structure of GANN techniques

3. THE RESULT OF STUDY

The genetic algorithm runs to optimize the results of neural networks when the number of layers is three to classify the Breast Cancer by adjusting the weights and bias in the neural network. One of the reasons for using the neural networks is its ability to simulate the non-linear functions. But to simulate the effect of the dependent variables on independent variables the genetic algorithm is used.

Fig.3. shows the range of individuals, fitness value and the stopping criteria in each generation and the best individual, fitness value.

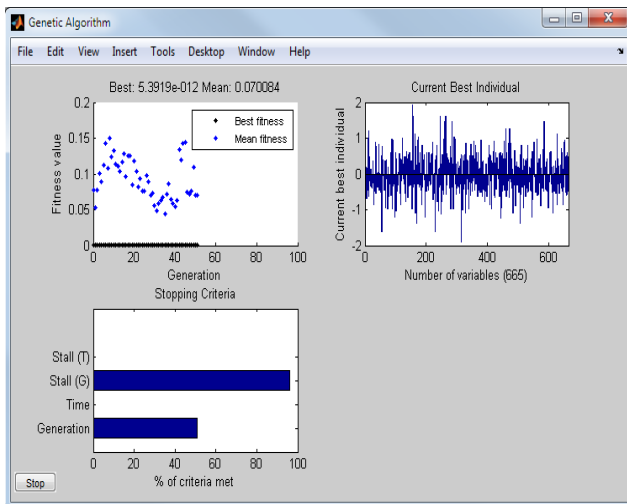


Fig. 3. The result of using the genetic algorithm to optimize the neural networks

The genetic algorithm displays the following results in the Run solver and view results pane:

Objective function value: 5.391906559044106E-12.

Optimization terminated average change in the fitness value less than option TolFun.

The result of objective function value shows that using the genetic algorithm to optimize the neural networks provide us more accurate than only use neural networks to classify the breast cancer. The best results have obtained from hybrid model are shown in Table 2 compared with the result of the usage only neural network to classify the breast cancer by using of mean square error (MSE).

Table. 2. The comparison of between the results of using the hybrid model with neural network to classify the breast cancer

MSE in hybrid model	MSE in neural network
5.4E-12	4.8E-6

The main error in hybrid model (GANN) smaller than MSE in NN.

Table.3 The test Result Technique Performance

Algorithm	TEST DATA	CANCER	MSE	Accuracy
NN	200	182	4.8E-6	0.91
GANN	200	188	5.4E-12	0.94

The accuracy of breast cancer diagnosis using Neural Network based on genetic algorithm is more efficient than using only NN.

4. CONCLUSION

This article discussed the possibility of using genetic algorithm model to optimize the neural networks. The test results of using genetic algorithm to optimize the neural network shows that the error rate is less than the error rate when it is only use neural network to diagnose breast cancer. GA-NN technique showed a proper precision in cancer diagnosis.

The purpose was to achieve a better prediction accuracy by the use of GA-NN hybrid technique than obtained by NN only. Test results for the diagnosis of breast cancer show that the hybrid technique generally gives accurate results, and thus the purpose has been achieved.

5. REFERENCES

- [1] Ahmed Khalid, Noureldien A. Noureldien. Determining the Efficient Structure of Feedforward Neural Network to Classify Breast Cancer Dataset. (IJACSA) International Journal of Advanced Computer Science and Applications, 2013. Vol. XXX, No. XXX, pp:1-4.
- [2] Cook D, Ragsdale C, Major R. Combining a neural network with a genetic algorithm for process parameter optimization. Eng Appl Artif Intell. 2000. Vol.13(4):pp.391-6.
- [3] Hirose Y, Yamashita K, Hijiya S. Back-propagation algorithm which varies the number of hidden units. Neural Networks. 1991. Vol.4(1):pp.61-6.
- [4] Janghel, R.R.; Shukla, A.; Tiwari, R.; Kala, R. (2010). "Breast Cancer Diagnostic Technique using Symbiotic Adaptive Neuro-Evolution (SANE)". Proceedings International conference of Soft Computing and Pattern Recognition (SoCPaR-2010), 7th-10th Dec, ABV-IIITM, Gwalior. pp: 326-329.
- [5] Jemal, A., Siegel, R., Ward, E., Murray, T., Xu, J., Thun, M.J. Cancer Statistics, 2007. CA Cancer J Clin, Vol. 57, 2007. pp: 43-66.
- [6] [K. V. Price, R. M. Storn, and J. A. Lampinen, Differential Evolution—A Practical Approach to Global Optimization. Berlin, Germany: SpringerVerlag, 2005.
- [7] M. McInerney, and A.P. Dhawan., Use of Genetic Algorithms with Backpropagation in Training of Feedforward Neural Networks, Proceeding of IEEE International Conference on Neural Network, 1993, Vol 1, No 3, 28 March-1 April, pg 203-208.
- [8] Netnevitsky, M., Artificial Intelligence: A Guide To Intelligent Techniques, Addison Wesley, Harlow, 2005, 2nd ed.
- [9] S.J. Marshall., & R.F. Harrison, Optimization and Training on Feed-forward Neural Networks by Genetic Algorithms, IEEE 2nd International Conference on Artificial Neural Network, 1991, pg 39-43.
- [10] UCI Machine Learning Repository. [Online]. Available: [https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+\(Original\)](https://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(Original))