

## Predictions of Laryngeal Cancer Using Neural Network in Kerman Shafa Hospital

Sadegh Nejat-zadeh <sup>1</sup>, Fatemeh Rahimi <sup>2</sup>, Amid Khatibi Bardsiri <sup>\*3</sup> , Elham Vahidian <sup>1</sup>

1- Yasuj University of Medical Sciences, Yasuj, Iran.

2- Ph.D Student in Medical Informatics, Shahid Beheshti University of Medical Sciences, Tehran, Iran.

3- Assistant Professor, Bardsir Branch, Islamic Azad University, Bardsir, Iran.

### ABSTRACT

**Introduction:** One of the challenges facing medical science is the time and correct diagnosis of diseases. Particularly with regard to certain diseases such as the types of cancer, which are the leading causes of death worldwide, their early diagnosis has a significant impact on the control and treatment of this disease. The use of intelligent decision support systems with high precision can be a good way to reduce human error due to fatigue and lack of experience. Therefore, the present study tries to predict the disease by using data mining techniques and taking into account the variables that influence the prediction of laryngeal cancer.

**Material and Methods:** This study is an analytical study. The data from the 249 cases referred to Shafa Hospital in Kerman in 2017 have been obtained. This study is based on the Crisp methodology and in the MATLAB software environment. First, in order to understand the laryngeal cancer, a review of related studies was conducted and interviewed by specialist physicians. Then, according to expert opinion, 24 variables were identified as effective factors in predicting laryngeal cancer. After clearing and preparing data, an artificial neural network model was used to predict the risk of laryngeal cancer. In the following, another model of the combination of the genetic algorithm and the neural network was created. Using genetic algorithm, 9 functional features of prediction of laryngeal cancer were determined from among the 24 selected variables, and artificial neural network was used to predict the risk of laryngeal cancer. Finally, the criteria for accuracy, specificity, and sensitivity were used to evaluate the two models.

**Results:** The genetic algorithm reduced the complexity of the model by reducing the number of features from 24 to 9, but improved the average precision from 80% to 84%. Also, the model made with the characteristics selected by the genetic algorithm, increased the specificity and accuracy criteria by 13% and 8%, respectively.

**Conclusion:** Combining the genetic algorithm with the neural network, in addition to improving the accuracy of prediction of laryngeal cancer, accelerates the diagnosis process, especially at the data collection stage, by reducing the number of characteristics required. Therefore, using this model as a smart decision system is suggested.

*Please cite this article as:*

Nejat-zadeh S, Rahimi F, Khatibi Bardsiri A, Vahidian E. Predictions of Laryngeal Cancer Using Neural Network in Kerman Shafa Hospital. Iran J Med Inform. 2018; 7(1): e4. DOI: [10.24200/ijmi.v7i0.154](https://doi.org/10.24200/ijmi.v7i0.154)

### Article History

Received: 2018-09-03

Accepted: 2018-10-20

Published: 2018-11-01

### Keywords

Laryngeal cancer

Cancer diagnosis

Data mining

Artificial Neural Network

\* Corresponding Author: A Khatibi Bardsiri, Department of Computer and Information Technology, Islamic Azad University, Bardsir Branch, Bardsir, Iran. (Email: khatibi\_amid@yahoo.com)

## INTRODUCTION

Laryngeal cancer is one of the most common types of cancers that usually starts from the laryngeal crater and can gradually spread to the back of the tongue, the various parts of the throat and neck, lungs, and other parts of the body. One of the diagnostic methods for this type of cancer is the use of direct endoscopy, CT scan and biopsy. Since access to the equipment and the cost of diagnosis are not feasible for all, the use of diagnostic tools that meet both low costs and the lack of complex medical technologies seem appropriate.

Cancer research plays a central role in national cancer planning. Given the high cost of cancer treatment at advanced stages of disease, investment in cancer research is economically feasible. In recent years, the use of data mining and artificial intelligence techniques has been used in many areas and the results have shown that in the discussion of medical diagnoses, these techniques can be used with great confidence.

In 2015, Kiu and colleagues used a two-stage fuzzy neural network to build a prostate cancer diagnostic system [1]. In their research, using the classification analysis, the elementary parameters of the membership function were determined and then a combination of a neural network and a particle optimization algorithm were used to examine the relationship between inputs and outputs. The results of the evaluation on the three benchmark functions showed that the proposed two-stage fuzzy neural network is more effective and the proposed model can detect prostate cancer with higher precision. In 2015, Bardway presented a neural network model optimized using genetics to identify malignant tumors of benign tumors in breast cancer [2].

According to this paper, co-operators and new jumps were introduced, which varied with standard operators. To investigate the superiority of the proposed method, criteria such as accuracy, sensitivity, and specificity were compared with three new algorithms of the neural network optimized with genetics, the classical neural network and the classic post-defragmentation model on the WBCD data from the UCI reference. The results showed that their proposed method works well on breast cancer data [3].

In an article about lung cancer diagnosis using the nearest killing category  $k$  using the genetic algorithm in 2015, an optimization algorithm was proposed to identify the masses in the CT scan in the early stages of lung cancer. As the work of interpreting the typical CNS images of lung cancer is time-consuming and very sensitive, researchers in this paper presented a combination of a genetic algorithm with the closest neighboring  $k$  algorithm to overcome this problem, which categorized with the speed and accuracy of cancer images. Wang published an article entitled "Prediction of the disease using different types of neural network compartments." The main purpose of the paper was to examine the functionality of different categories, including the single category bundles in a collective category and single batches [4].

Also, various evaluation criteria were used to test the performance of these clusters with real life databases. Finally, statistical tests were used to assess the significance of the difference in efficiency between the three groups. The results of the statistical tests showed that a collective class works better than a single model batch. In 2013, Ada conducted a study to diagnose lung cancer using data mining techniques [5].

Lung cancer is formed due to uncontrolled cell growth in the lung tissue, and the timely diagnosis of the disease plays an important role in its treatment. But lung cancer diagnosis includes chest filming with X-ray, CT scan and MRI, which in many parts of the world still cannot use these technologies. In the article, using the

methods used, data from X-ray films from the chest were categorized into two groups: normal and abnormal. Various learning experiments were carried out using feature selection along with the neural network to select the minimum and optimal set of features and compare the results.

Therefore, data mining techniques can be used to extract hidden patterns from a massive range of medical data, and intelligent systems can reduce the flaws and shortcomings caused by the fatigue or lack of exercise of physicians.

Hence, in this paper, an artificial neural network, one of the most important data mining techniques, has been used to detect laryngeal cancer. The purpose of this study was to compare the ability of the neural network algorithm to combine the algorithm derived from the genetic algorithm and the neural network in the diagnosis of laryngeal cancer.

## MATERIAL AND METHODS

### Theoretical Basics

Artificial Neural Networks: A neural network is a "connected" computing system that is organized in a layered manner [6-8]. In general, there are three inputs, hidden, and output layers in a neural network, respectively. Layers consist of a number of neurons that are interconnected internally and each neuron has an activation function. Patterns are provided through network layers. These layers are associated with one or more hidden layers, in which actual processing is performed through a system of weighted connections. The hidden layers are then connected to an external layer, in which the response to the grid problem is specified in the output neurons. Fig. 1 shows the overall structure of a neural network.

One of the important steps in artificial neural network, the learning step is through weight adjustment. In Fig. 1, each line represents a connection between two neurons and shows the path of the flow of information. Each connection has a weighted parameter, which actually controls the signal between two neurons. If the network produces a "good" output, there is no need to adjust the weights. But if the network leads to the production of a "weak" or so-called "error", then the system adapts the system by replacing the weights to improve the sequence of future results.

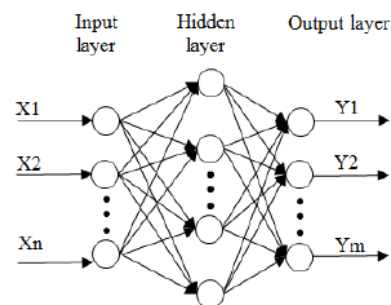


Fig 1: The general structure of a neural network

The most commonly used artificial neural network training technique is the outsourcing algorithm [9, 10]. The successes of the post-transfer algorithm can be expressed in the sense that the network first receives a training sample and uses the weights in the network that were initially randomly assigned. They calculate the output. Then the error is calculated, which is the difference between the calculated result output and the expected value. The error is released inside the network and the weights are reset to

reduce the error, which is the most important part of the weights adjustment. After calculating the prediction error for the first input to the system, weights vary from the last layer to the first layer so that the forecast error decreases. Then the second sample is given to the grid and because the same sample weight may cause the new sample to fail again, the weights are modified in such a way as to produce the least error. This way, after reading the number of samples in the network input, the network converges, which means success in the learning phase. Then the network is ready to be used for the prediction phase.

**Genetic Algorithm**

Genetic Algorithm is one of the numerical optimization algorithms inspired by nature and is now used in many sciences to find optimal solutions [11, 12]. Selection, mutation and combination are three main operators of genetic algorithm. Different stages of genetic algorithm in Fig. 2 it has been shown. First, depending on the problem, the variables to be determined are specified. Then these variables are properly encoded and displayed in the form of a chromosome. Based on the target function, a fitness function is defined for chromosomes, and an arbitrary initial population is randomly generated. Then the fitness function is calculated for each primary population chromosome. Next, the parents are selected and the children of the next generation. This process repeats until an appropriate answer or specific repeat count is reached. Finally, the best chromosome seen as the optimal answer is selected.

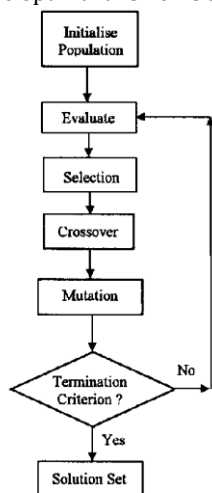


Fig 2: Stages of Genetic Algorithm

**Evaluation criteria**

One of the factors influencing the success of the model is to measure the model with appropriate criteria. In research in the field of data mining, a set of criteria has been used to evaluate the model, which is mentioned below. In this paper, these criteria have been used to evaluate the neural network model for predicting laryngeal cancer.

Sensitivity: This criterion shows the accuracy of the prediction model and is defined as the percentage of classes that are correctly predicted to be error-prone [13].

$$\text{Sensitivity} = TP / (TP + FN) \tag{1}$$

Specificity: This criterion, like the sensitivity criterion, is used to measure the accuracy of the prediction model, which is defined as the percentage of classes that are reasonably predicted to be unpredictable by the following equation [13].

$$\text{Specificity} = TN / (TN + FP) \tag{2}$$

Accuracy is defined as the ratio of the number of correctly predicted classes (including errors and errors) to the total

number of classes. The accuracy criterion is used to measure the overall accuracy of prediction accuracy [13].

$$\text{Accuracy} = (TP + TN) / (TP + TN + FP + FN) \tag{3}$$

Precision shows the number of probable error classes predicted by the model as being prone to error. The best value is 1 and is calculated as follows [14].

$$\text{Precision} = TP / (TP + FP) \tag{4}$$

Recall: The recall time is the number of probable error classes predicted by the model. The best value for it is 1 and is calculated from the following equation [14].

$$\text{Recall} = TP / (TP + FN) \tag{5}$$

**Modeling method**

This study is based on the Crisp methodology and in the MATLAB software environment. In order to construct a prediction model for laryngeal cancer, the required data were collected from patients with suspected laryngeal cancer. For this purpose, 249 patients were recruited from the Shafa hospital in Kerman. These patients are people who have undergone laryngeal sampling at the Shafa Hospital in Kerman, and their eclipse has lasted more than three weeks. Of these, 150 were laryngeal cancer and 99 were healthy. After consultation with relevant specialists, there are 24 significant factors that can be used to predict laryngeal cancer more than other factors. Then, based on the values recorded in the patient records for the selected factors, the final data was obtained for the prediction model. The specifications for this data are given in Table 1.

Table 1: Data specification collected

Characteristics	Description
Number of Samples	249
Number of cancer Samples	150
Number of non-cancer Samples	99
Number of Factors	24
Aim Characteristic	If a person has cancer, it shows 1, else shows 0.
Features	1) age 2) gender 3) Cervical mass 4) Snoring 5) Sore throat and itching or burning 6) Violence and hoarseness 7) Post-nasal drainage 8) Lupus erythematosus 9) Dysphagia 10) Coughing 11) Reflux 12) Edenophagy 13) Ear Pain 14) Whine Breast 15) Overnight transpiration 16) Weight loss 17) Reduce appetite. 18) Liquid aspiration 19) alcohol 20) stool culture 21) Obstructive apnea 22) Oral and dental hygiene weakness 23) history of laryngeal cancer in relatives of degree 24) History of lung cancer, esophagus or oral cancer in relatives

The structure of the neural network used to build the laryngeal cancer prediction model is presented using the collected data. Artificial neural network structure. The structure of the artificial neural network used in Fig. 3 is shown. As shown in this figure, this neural network has three levels of input, hiding, and output. The number of neurons in the input layer equals the number of attributes, it means, 24 neurons. The number of neurons in the output layer is 1. There is no method for precisely determining the number of hidden layer neurons, and it must be determined with the desired value for each problem. The point that should be considered in choosing the number of hidden layer neurons is that its large size increases the complexity of the model and lowers its ability to predict the model. With the many simulations we did, it was found that 5 neurons in the secret layer were a good choice, since the final constructed model had the proper classification accuracy and its complexity was also low.

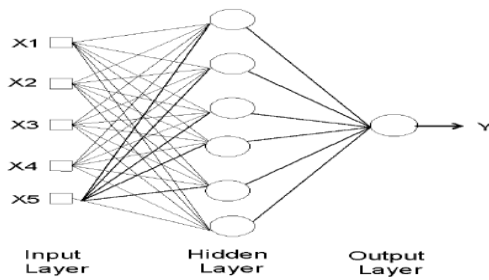


Fig 3: Artificial Neural Network structure used

**Artificial Neural Network Training**

To train the neural network, the Langbjarmark algorithm [15, 16] has been used with a number of 1000 replications. As a function of activation of neurons, a tangent Sigmoid has also been used (Fig. 4). Table 2 shows the parameters of the neural network used.

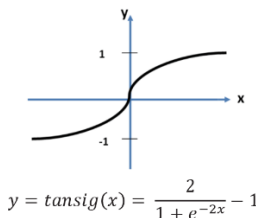


Fig 4: Activation function of neurons (tangent Sigmoid)

Table 2: Artificial Neural Network Parameters

Parameters	Value
Hidden Layer number	1
Number of neurons	24
Number of neurons in hidden layer	5
Number of neurons in output layer	1
Active function of neurons	tangent Sigmoid
Network training function	Langbjarmark

**Reduced feature with genetic algorithm**

One of the important points in modeling with the neural network and more generally of each method is the identification of effective features among all the features [17, 18]. Data from patients suspected cases of laryngeal cancer are 24 characteristics for each patient. Certainly, all of the features may not play a role in predicting laryngeal cancer. For this reason, it is necessary to identify the features that are effective in making the model using the feature selection methods. If the number of

attributes is equal to n, then the total number of modes that can be used to select the attribute is 2<sup>n</sup>. Naturally, it's not possible to check all possible states to get the best set of attributes because the number of states is exponential. Hence, the genetic algorithm is used for purposeful search in the collection of possible states of features.

To select the attribute using the genetic algorithm, chromosomes are considered to be 24-bit strings. If the bit value of it equals to 1, it means the selection and the zero value means no selection of the "i". The evaluation is considered to be accurate, after implementing the genetic algorithm of 50 generations, the best chromosome in the last generation. , The best set of features to build the model. This chromosome is shown in Fig. 5. The parameters of the genetic algorithm used in Table 3 are presented.

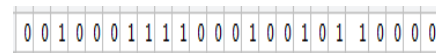


Fig 5: The best chromosome obtained with genetic algorithm

Table 3: Parameters of the genetic algorithm used to select the attribute

Parameters	Value
Type of chromosomes	Bit
The length of each chromosome	24
Number of generations	50
Number of chromosomes	100
Selection Operator	tournament
Associate operator	Scattered
Mutation operator	monotone
Coordinate rate	0.8
Mutation rate	0.01

**RESULTS**

A 5-layer validation method was used to evaluate the performance of the neural network in predicting patients with laryngeal cancer. For this purpose, data was first randomly divided into 5 parts, each time using 4 parts for making the model and one for evaluating the model using accuracy, sensitivity, specificity, accuracy and recall criteria. Finally, the minimum, average, and maximum values of each metric have been reported. Table 4 shows the results of simulation of the neural network model with all the characteristics.

Table 4: Simulation results of the model made using all the features

	Precision	Sensitivity	Specificity	Accuracy	Recall
Minimum	0.67	0.71	0.58	0.69	0.71
Mean	0.80	0.84	0.75	0.83	0.84
Maximum	0.90	0.90	0.93	0.97	0.90

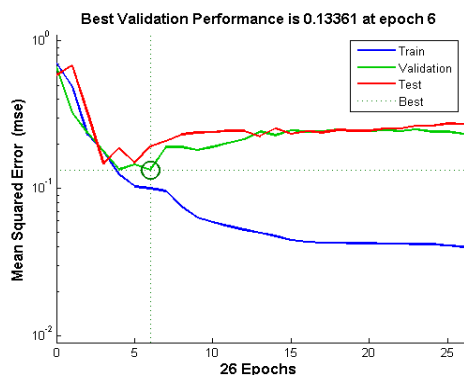


Fig 6: Artificial Neural Network Training Charts with Total Characteristics

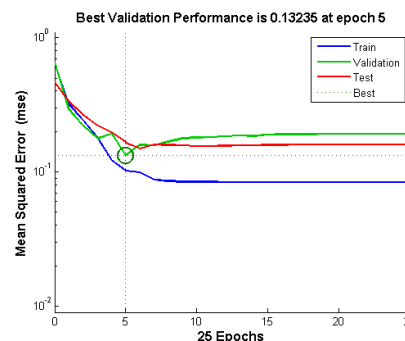


Fig 7: Artificial Neural Network Training Charts with Characteristics Selected by Genetic Algorithm

Fig. 6 illustrates the training diagram of an artificial neural network with all features. As it is known, the neural network has achieved its best parameters after 6 replications. In fact, the training of the neural network proceeds to a point where the network error decreases on the validation data set. If we continue to train the network afterwards, although its error decreases on the training data, it will not perform well at the time of testing due to the reduced ability to generalize.

In Table 5, the simulation results of the constructed model have been reported with the characteristics selected by the genetic algorithm. As Table 5 shows, the accuracy of the neural network constructed with 9 characteristics selected by genetic algorithm (cervical mass, postnatal discharge, dysphagia, appetite loss, alcohol consumption, wheezing, cough, lupus erythematosus, stool culture) Average is 84%. In other words, if we use a neural network constructed with the characteristics determined by the genetic algorithm to predict patients with laryngeal cancer, on average, with a probability of 84%, the prediction of the neural network would be correct. Hence, the use of the neural network as a smart decision maker will be of great help to the specialist and will be a kind of supportive judgment of the physician. In addition to the average degree of classification accuracy, the minimum, maximum and average values of other criteria have also been remarkably improved in comparison to the neural network with all the features in some cases.

Table 5. Simulation results of the model made with the characteristics selected by the genetic algorithm

	Precision	Sensitivity	Specificity	Accuracy	Recall
Minimum	0.76	0.72	0.80	0.84	0.72
Mean	0.84	0.82	0.88	0.91	0.82
Maximum	0.94	0.93	0.95	0.96	0.93

Fig. 7 illustrates the training of an artificial neural network with characteristics selected by the genetic algorithm. As it is known, the neural network has achieved its best parameters after 5 repetitions.

## DISCUSSION

One of the most important criteria in evaluating each categorized model is the degree of precision that the model provides in the categorization of samples. The neural network model with a total of 24 features has an average precision of 80%. To reduce the complexity and improve the power of the classification of the neural network model, we also investigated the feature selection with genetic algorithm. As expected, the genetic algorithm reduced the complexity of the model by reducing the number of features from 24 features to 9, but improved the average precision from 80% to 84%. Also, the model made with the characteristics selected by the genetic algorithm has increased the specificity and accuracy criteria by 13% and 8%, respectively. As it is known, the trained neural network has a better performance than the trained neural network with all its features in terms of criteria of accuracy, specificity and precision.

The accuracy criterion means that if the neural network identifies an individual with laryngeal cancer, how much is the correct prediction? The accuracy of the trained neural network with the total characteristics is 83%, while this amount in the trained neural network has increased by 91% with the selected features by the genetic algorithm, which is a significant improvement. Also, the level of sensitivity indicates that if a person does not have laryngeal cancer, what is the probability of a nerve lining for her laryngeal cancer? In the case of the trained neural network with all the features, this value is equal to 75%, which is about 88% higher for the trained neural network with selectable features by genetic algorithm, which is still a significant increase. The sensitivity criterion specifies that if a person really has a laryngeal cancer, how likely does the nervous system predict a positive outcome to him? In this case, the performance of the trained neural network has dropped by 2% with selective features. Finally, the precision of the neural network also indicates that all the results that the neural network predicts are likely to be reliable, which is about 84% higher for the trained neural network with selectable features by the genetic algorithm. Although the use of all features is also appropriate in making the categorization model, but reducing features from 24 to 9, in addition to enhancing the improvement of the categorized model, it will require fewer inputs for new ones. In other words, for the new patients it is enough to collect only 9 characteristics in place of the 24 features. The advantage of diminishing features is that cost-saving features are cost-effective, and cost-saving, and time-consuming.

## CONCLUSION

One of the challenges facing medical science is the timely and correct diagnosis of diseases. Particularly for certain diseases, such as cancers, can properly affect the course of the disease if diagnosed properly in the early stages. In diagnosing for various reasons such as fatigue, brain drain, failure of test devices, etc., the result of the diagnosis may be mistaken or even costly in terms of time and money. Hence, the use of data mining in diagnostic decision making is of particular importance.

In this paper, the neural network, an important data mining technique, is used to detect laryngeal cancer. Laryngeal cancer is one of the most common types of cancers that usually starts from the laryngeal crater and can gradually spread to the back of the tongue, the various parts of the throat and neck, lungs, and other parts of the body. One of the diagnostic methods for this type of cancer is the use of direct endoscopy, CT scan and biopsy. Since access to this equipment and the cost of diagnosis are not feasible for all, the use of diagnostic tools that meet both low costs and the lack of complex medical technologies seem appropriate.

The simulation results showed that the use of artificial neural networks in diagnosing laryngeal cancer as a smart decision making system helped a specialist physician and is a good solution to this problem because of the lack of advanced technologies and cost reduction. Using evolutionary algorithms to regulate neural network weights, using fuzzy neural networks and investigating neural networks with more than one layer of hidden layers, including the ones we will look at in the future.

## REFERENCES

1. Kuo R-J, Huang M-H, Cheng W-C, Lin C-C, Wu Y-H. Application of a two-stage fuzzy neural network to a prostate cancer prognosis system. *Artif Intell Med*. 2015; 63(2): 119-33. PMID: 25576196 DOI: 10.1016/j.artmed.2014.12.008 [PubMed]
2. Bhardwaj A, Tiwari A. Breast cancer diagnosis using genetically optimized neural network model. *Expert Systems with Applications*. 2015; 42(10): 4611-20.
3. Bhuvaneshwari P, Therese AB. Detection of cancer in lung with k-nn classification using genetic algorithm. *Procedia Materials Science*. 2015; 10: 433-40.
4. Weng C-H, Huang TC-K, Han R-P. Disease prediction with different types of neural network classifiers. *Telematics and Informatics*. 2016; 33(2): 277-92.
5. Ada, Kaur R. A study of detection of lung cancer using data mining classification techniques. *International Journal of Advanced Research in Computer Science and Software Engineering*. 2013; 3(3): 131-4.
6. Schumacher M, Roßner R, Vach W. Neural networks and logistic regression: Part I. *Computational Statistics & Data Analysis*. 1996; 21(6): 661-82.
7. Wang J, Tsapakis I, Zhong C. A space-time delay neural network model for travel time prediction. *Engineering Applications of Artificial Intelligence*. 2016; 52: 145-60.
8. Schiavo ML, Prinari B, Gronski JA, Serio AV. An artificial neural network approach for modeling the ward atmosphere in a medical unit. *Mathematics and Computers in Simulation*. 2015; 116: 44-58.
9. Riedmiller M. Advanced supervised learning in multi-layer perceptrons-from backpropagation to adaptive learning algorithms. *Computer Standards and Interfaces*. 1994; 16(3): 265-78.
10. Bhattacharyya S, Bhattacharjee S, Mondal NK. A quantum backpropagation multilayer perceptron (QBMLP) for predicting iron adsorption capacity of calcareous soil from aqueous solution. *Applied Soft Computing*. 2015; 27: 299-312.
11. Goldberg DE. *Genetic algorithms in search, optimization and machine learning*. Reading, Addison-Wesley Longman Publishing Co.; Boston, USA. 1989.
12. Coley DA. *An introduction to genetic algorithms for scientists and engineers*. World Scientific Publishing Company; Singapore. 1999.
13. Van Stralen KJ, Stel VS, Reitsma JB, Dekker FW, Zoccali C, Jager KJ. Diagnostic methods I: sensitivity, specificity, and other measures of accuracy. *Kidney Int*. 2009; 75(12): 1257-63. PMID: 19340091 DOI: 10.1038/ki.2009.92 [PubMed]
14. Guazzo M. Retrieval performance and information theory. *Information Processing & Management*. 1977; 13(3): 155-65.
15. Levenberg K. A method for the solution of certain non-linear problems in least squares. *Quart Appl Math*. 1944; 2(2): 164-8.
16. Marquardt DW. An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics*. 1963; 11(2): 431-41.
17. Cong Y, Wang S, Fan B, Yang Y, Yu H. UDSFS: Unsupervised deep sparse feature selection. *Neurocomputing*. 2016; 196: 150-8.
18. Zhang X, Mei C, Chen D, Li J. Feature selection in mixed data. *Journal of Pattern Recognition*. 2016; 56: 1-15.