

# Comparing Performance of Different Neural Networks for Early Detection of Cancer from Benign Hyperplasia of Prostate

Mustafa GHADERZADEH<sup>1,\*</sup>, Rebecca FEIN<sup>2</sup>, Arran STANDRING<sup>2</sup>

<sup>1</sup> Department of Health Management and Information Sciences, Tehran University of Medical Sciences.

<sup>2</sup> Applied Health Informatics, Bryan University, Tempe, Arizona 85281, USA.

Emails: ir.medicalinformatics@yahoo.com(\*); rebecca.fein@gmail.com;  
arran.standring@bryanuniversity.edu

\* Author to whom correspondence should be addressed; Tel.:+98-914-980-6771.

Received: 8.5.2013/Accepted: 9.9.2013/ Published online: 23.9.2013

## Abstract

Prostate cancer is one of the most common types of cancer found in men. Presenting a classifier in order classifies between Prostate Cancer (PCa) and benign hyperplasia of prostate (BPH), has been great challenge among computer experts and medical specialists. There are a number of techniques proposed to perform such classification. Neural networks are one of the artificial intelligent techniques that have successful examples when applying to such problems. The increasing demand of Artificial Neural Network applications for predicting the disease shows better performance in the field of medical decision-making. This paper presents a comparison of neural network techniques for classification prostate neoplasia diseases. The classification performance obtained by four different types of neural networks for comparison are Back Propagation Neural Network (BPNN), General Regression Neural Network(GRNN), Probabilistic Neural Network (PNN) and Radial Basis Function Neural Network (RBFNN). Result of these evaluation show that the overall performance of RBFNN can be apply successfully for detecting and diagnosing the cancer from benign hyperplasia of prostate.

**Keywords:** Prostate cancer; benign hyperplasia of prostate; Artificial Neural Network; Back Propagation Neural Network.

## Introduction

Cancer is a wide spread disease responsible for around 13% of all deaths in the world in 2008. In this year, the World Health Organization estimated the number of new cancer cases in the world to be over 7.5 million. Among all types of cancer, prostate cancer is the most frequent in men. Prostate cancer is one of the leading causes of Cancer-related mortality in men and a major health issue. Some studies have estimated that 10% of men in developed countries will suffer from prostate cancer during their lifetime and approximately 12.45% will die from this disease [1, 2]. The American Cancer Society estimates that during 2013 about 238,590 new cases of prostate cancer will be diagnosed in the United States. Moreover, estimation of death in these cases will be 29,720. Prostate cancer is therefore a serious public health concern. If diagnosed early enough prostate cancer is curable or the survivability outcome would be much higher and even at later stages treatment may still be effective. However, prostate cancer has a considerable impact on the

quality of life of adult and elderly men. In general, prostate cancer is a disease that can be diagnosed with prostate biopsy in accordance with the suspicions that arose as a result of PSA test, rectal examination, and transrectal findings [3, 4]. Since the classification of Cancerous from benign tissue by medical application are only based on the intensity variation, classification in most of the cases is difficult. In order to confirm a diagnosis of prostate cancer there is a need for biopsy in addition to transrectal ultrasound and rectal examination. Confirmations of diagnosis are made only after a specialist assesses patients' transrectal ultrasound, rectal examination results, and the amount of prostate-specific antigen (PSA). The PSA level in blood has become one of the most common methods, based off the results of the studies conducted in recent years, for early diagnosis of prostate cancer[4]. But PSA values may not yield conclusive results about existence of prostate cancer because PSA levels can be increased by inflammation of prostate and benign prostate hyperplasia (BPH). Therefore, patients are also given rectal examination, if anomalies are observed at the end of the rectal examination, even if PSA results may seem normal, it is recommended that a prostate biopsy be performed and definitive diagnosis be made. Despite the need for biopsy for conclusive diagnosis, patients with low cancer risk will not have this procedure due to possible complications that may arise, the risk of rectal mucosa being damaged and the high associated costs. Despite the risk of complications and the heavy cost, there is a vital need to develop new techniques in order to classify and detect diseases. In order to solve the classification problems and prediction, many classification techniques have been proposed. Some of the more successful techniques are Artificial Neural Networks (ANN), Support Vector Machines (SVM) and classification trees. There are a number of other techniques that can also be applied to classification problems, for example linear regression, logistic regression, discriminant analysis, genetic algorithms, fuzzy logic, Bayesian networks and k-nearest neighbor techniques[4-6].

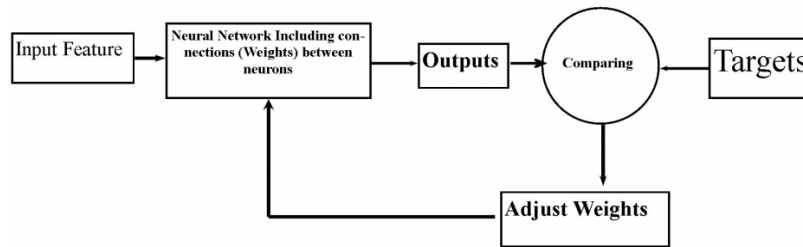
## Material and Method

### *Subjects*

The main purpose of this study is to investigate the applicability and capability of ANN methods, BPNN, GRNN, RBF, and PNN for diagnosing the prostate neoplasia disease based on laboratory and biographic data. Present study used MATLAB Neural Network Toolbox software in order to Classification Prostate Neoplasia Dataset and evaluating performance of different neural network. Artificial Neural Network is a branch of Artificial intelligence that has been accepted as a new technology in computer science. Neural Networks are currently a 'hot' research area in medicine, particularly in the fields of radiology, urology, cardiology, and oncology. It has a huge application scope in many areas, especially Medicine. Neural Networks plays an important role in a classification and detection diseases. Artificial Neural Network (ANN) is a technique based on the neural structure of the brain that mimics the learning capability from experiences. It means that if neural network trained from past data it will be able to generate outputs based on the knowledge extracted from the data. Many research projects have shown that ANN is a powerful technique for classification. There are several advantages of using ANN for classification. ANN is a universal function approximation that can adapt itself to the data without making prior assumption of the Comparing the Performance of Different Neural Networks for Classification Problems. Therefore, ANN is able to approximate any function with arbitrary accuracy. ANN is a nonlinear model that can be implemented for most complex real world applications. Furthermore, there are many successful real world applications using ANN; such as industry, business and science, and especially in Medical diagnosis, detection and classification [7]. This paper presents a comparison of neural network techniques for Prostate neoplasia disorder including Prostate Cancer and Benign Hyperplasia of Prostate (BPH) classification problems. In addition, Selected Neural network were Back Propagation Neural Network (BPNN), Radial, General Regression Neural Network (GRNN), and Probabilistic Neural Network (PNN).

*Back Propagation Neural Network (BPNN)*

Feed forward back-propagation neural network or backpropagation neural network is a simple and effective model of ANN. It contains three layers, which are input, hidden, and output layers. Its structure is multilayer and has a learning process. Figure 1 shows the learning process in a schematic view.

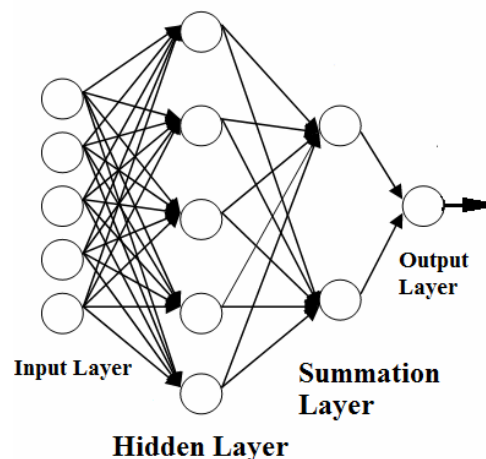


**Figure 1.** Learning cycle in Backpropagation Neural Network.

BPNN is one of the more popular ANN's that has been used for ANN applications. BPNN is a robust neural network that can be applied easily in various problem domains. However, there are also limitations in BPNN. Back propagation employs gradient descent to minimize the squared error between the network output values and desired values for those outputs. These error signals are used to calculate the weight updates, which represent knowledge learned in the networks. The performance of the back propagation algorithm can be improved by adding a momentum term[8, 9].

*General Regression Neural Network (GRNN)*

The General Regression Neural Network (GRNN) is one of the most popular neural networks. GRNN is a feed-forward neural network for supervised data. It uses nonlinear regression functions for approximation. GRNN uses direct mapping to link the input layer to the hidden layer. GRNN employs the smoothing factor as a parameter in learning phase. The single smoothing factor is selected to optimize the transfer function for all nodes. To reduce computational time, GRNN performs one pass training through the network [7, 10, 11]. This network does not require an iterative training procedure that is required in back propagation method. They used four layers: the input layer, pattern layer, summation layer and output layer. Following Figure 2 shows a basic GRNN in schematic view.



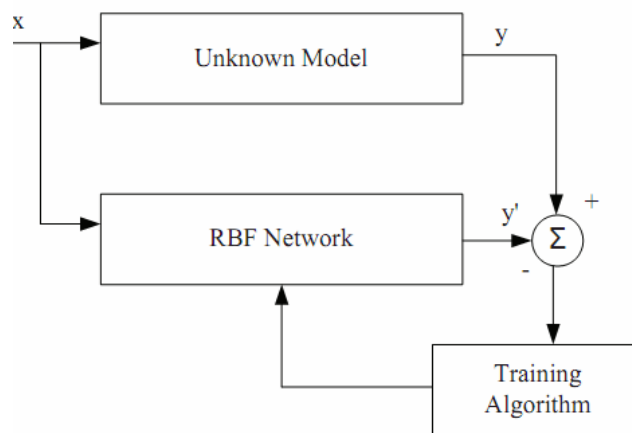
**Figure 2.** Common GRNN structure

*Probabilistic Neural Network (PNN)*

The Probabilistic Neural Network (PNN) was first proposed by Donald Specht in 1990. This is an artificial neural network for nonlinear computing which approaches the Bayes optimal decision boundaries. This is done by estimating the probability density function of the training dataset using the Parzen nonparametric estimator. Bayesian strategies are decision strategies that minimize the expected risk of a classification. The Bayesian decision theory is the basis of many important learning schemes such as the naïve Bayes classifier, Bayesian belief networks, and the EM algorithm. PNN has the ability to train on sparse data sets. Moreover, it is able to classify data into specific output categories [12, 13]. There are a number of advantages of using PNN for classification. For example, the computational time of PNN is faster than BPNN, and it is more robust to noise. Furthermore, the training manner of PNN is simple and instantaneous [7, 14-16].

*Radial basis function neural network (RBFNN)*

Radial Basis Function Neural Network (RBFNN) is a type of multilayer network. It is different from BPNN in its training algorithm. The basic RBFNN structure consists of three layers. These are an input layer, a kernel (hidden) layer, and an output layer. It can be regarded as a special Multi-Layer Perceptron (MLP) because it combines the parametric statistical distribution model and non-parametric linear perceptron algorithm in serial sequence. In the kernel layer, it consists of a set of kernel basis functions called radial basis functions. The output of the RBFNN is a linear combination (weighted sum) of the radial basis function calculated by the kernel units. RBF networks are very popular for function approximation, curve fitting, Time Series prediction, and control problems. Because of a more compact topology than other neural networks and faster learning speed, RBF networks have attracted considerable attention and they have been widely applied in many science and engineering fields [17-20]. A general block diagram of an RBF network is illustrated in Fig.3.



**Figure 3.** Block Diagram of a RBF network [21]

In RBF networks, the outputs of the input layer are determined by calculating the distance between the network inputs and hidden layer centers. The second layer is the linear hidden layer and outputs of this layer are weighted forms of the input layer outputs. Each neuron of the hidden layer has a parameter vector called center [17]. RBFNN can overcome some of the limitations of BPNN because it can use a single hidden layer for modeling any nonlinear function. Therefore, it is able to train data faster than BPNN. While RBFNN has simpler architecture, it still maintains its powerful mapping capability. Due to the benefits of these characteristics, RBFNN is an interesting alternative technique for classification problem [5, 22, 23].

## Dataset

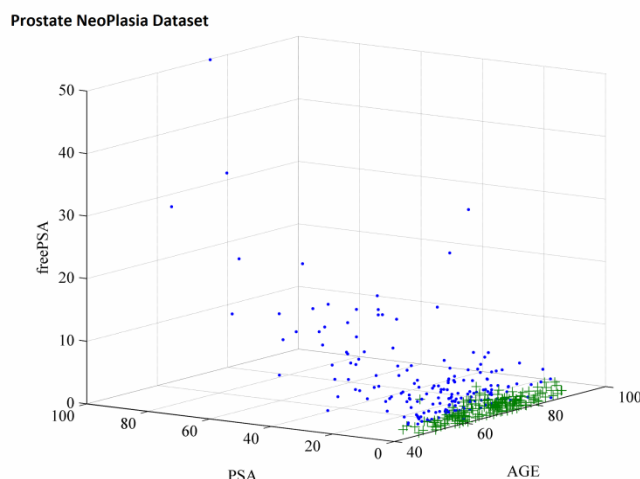
Present study dataset of prostate neoplasia disease is purely real set data. In order to compare the performance of several types of Neural Network we used the database records of the Department of Urology of the Tehran University of Medical Science, Imam Khomeini Hospital. The original database contained 360 records of patients that underwent radical Prostatectomy for prostate cancer between 1 January 2006 and 31 December 2010. The dataset is a two-class problem either positive or negative for prostate cancer and benign hyperplasia of prostate diseases respectively. Laboratory data belonging to 181 cancerous patients and other 179 patient were diagnostic to the BPH, their diseases found by biopsy performed in different urology clinics between 2008 and 2011 (inclusive), and the results of the definitive diagnosis after biopsy indicating whether they had cancer or not were used in the study. The data; included laboratory data like; PSA, freePSA, result of disease, and demographic data like age. Thereafter, the ratio (fPSA/tPSA) was calculated and the database was formed. This dataset contained 181 patients having prostate cancer, which can be interpreted as “1” and the remaining patients not having prostate cancer and can be interpreted as “0”. MATLAB Neural Network Toolbox software version 2012 was used in order to devise an ANN and conduct analyses. The descriptive statistics of preoperative parameters are given in Table 1:

**Table 1.** Pro Disease Database description of attributes\*

Num	Attribute description	Values of Attribute	Mean	Standard deviation
1	AGE	45-91	69	9.89
2	TOTAL PSA	0.1-100	13	17.15
3	FREE PSA	0.07-49.9	2.84	4.74
4	RATIO	0.018-2.9	0.2549	0.2

\*N = 360 OBSERVATIONS INCLUDING 181 MALIGNANT AND 179 BENIGN.

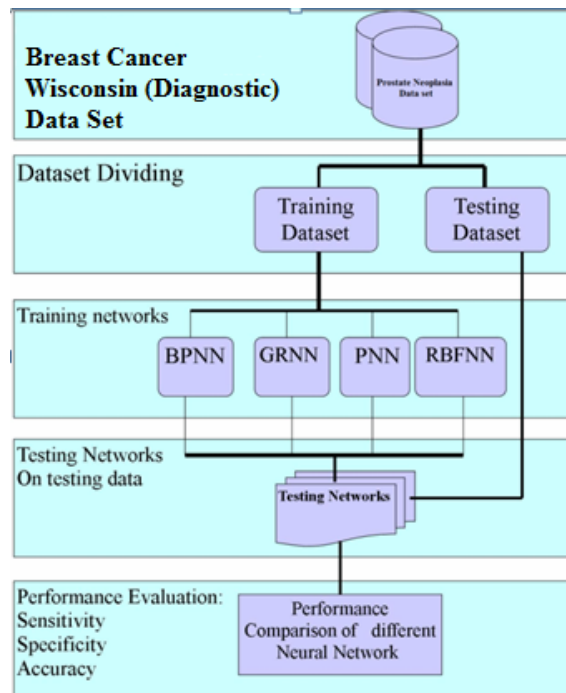
In order to observe the distribution prostate neoplasia data set including prostate cancer and benign hyperplasia of prostate, the data has presented in Fig.4. It represents the distribution of raw cancer and benign hyperplasia of prostate dataset according to the first three features (Age, PSA, and freePSA attributes).



**Figure 4.** Distribution of row NPD including prostate cancer ‘.’ And benign hyperplasia of prostate ‘+’, according to three feature

## Experiment Results and Discussion

Medical diagnosis by neural network is the black box approach. A network is chosen and trained with examples of all classes. After successful training, the system is able to diagnose the unknown cases and to make predictions. In this experiment work, we applied four neural networks in order to classify prostate neoplasia diseases. MATLAB 2012a is used for simulations, results, and the procedure used for the classification is shown in Fig.5 After the accuracies for each network type are evaluated and compared the performance of each network and directions for further research will be discussed.



**Figure 5.** Black Box of training, testing and performance comparing of different Neural Network

The basic idea behind medical tests is to calculate the probability of patients being sick on the basis of the patient's test results. Receiver Operating Characteristics (ROC) analysis is an established method of measuring diagnostic performance for the analysis of medical test performance. The ROC curve is a good measure when the performance of different classifiers needs to be compared. ROC analysis is a standard approach used to determine the sensitivity and specificity of the diagnosis. Sensitivity is also known as the ability to distinguish the sick from the true ill; and specificity is the ability to distinguish the healthy from the true healthy. Sensitivity and specificity are the basic expressions for the diagnostic test interpretation of the ROC analysis. In the following equations Sensitivity, specificity and accuracy are expressed [7, 20, 24]:

$$\text{Sensitivity} = \frac{\text{true positive (TP)}}{\text{true positive (TP)} + \text{false negative (FN)}} \quad (1)$$

$$\text{Specificity} = \frac{\text{true negative (TN)}}{\text{true negative (TN)} + \text{false positive (FP)}} \quad (2)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (3)$$

### Simulation Result

After testing each neural network, the mean square error (MSE), specificity, and sensitivity of the classification will be used for comparing the performance of these neural networks. In this study, a BPNN with Gradient descent with momentum algorithm that composed of an input layer, a hidden layer, and an output layer was used. After the ANN structure was configured, in order to determine the BPNN structure that yielded the best result, the neurons in the hidden layer were changed from 1 to 100 neurons. The back propagation algorithm was used in the training procedure. Different transfer functions (Purelin, Tansig, Logsig, etc...) were used and tried on the neurons in the hidden and output layers and Tansig-Sigmoid (Tansig) was selected as the transfer function that yielded the best result. The learning rate and momentum coefficient were changed among 0-1.

After Training and Testing BPNN, the best classification was yielded by the ANN structure that the learning rate and momentum value were 0.01 and 0.55. As can be seen in Fig.6, the best classification was yielded by the BPNN structure that has 12 neurons in its hidden layer. Time of Training of each neural network were different, but BPNN is an iterative network, these iterations or cycles are called Epochs. The network structure with a hidden layer of 12 neurons that was thus obtained was tested with different epoch numbers. The best epoch number was determined as 1000. At the end of these procedures, the network structure that yielded the best classification is given in Table 2.

**Table 2.** Parameters and properties used in BPNN

Parameters	Properties
Number of neurons in the input layer	4
Number of the hidden layers	1
Number of neurons in the hidden layer	12
Number of neurons in the output layer	2
Learning rate (a)	0.01
Coefficient of momentum (b)	0.55
Learning algorithm	Gradient descent with momentum (Traingdm)
Transfer function	Tangent-Sigmoid(Tansig)

For GRNN and RBF applications, the optimum spread values were found by trial-and-error and used for training and the classification of test data. For GRNN and RBF, spread value of 2.5 and 1.5 was used respectively. The performance of a PNN is largely influenced by the spread parameter. In this paper, this parameter was found and selected by test trial and error. Spread values were changed among 1-3.

In Table 3, the classification accuracies obtained by four neural network techniques are compared for Prostate Neoplasia Dataset (PND). The results show that the accuracies obtained by each neural network technique are quite compatible. The highest accuracy is given by RBFNN (90%), following by the results of BPNN (87%) and PNN (83.33%). The results of GRNN and provide the lowest performance with the accuracy at 81.67%.

**Table 3.** Classification accuracies obtained by five techniques of neural networks for prostate neoplasia dataset

Algorithm	Sensitivity (%)	Specificity (%)	Accuracy (%)	Error (MSE)
<b>BPNN</b>	88.61	90.34	87	0.088
<b>GRNN</b>	95.65	73	81.67	0.1079
<b>PRNN</b>	71.88	96.43	83.33	0.167
<b>RBFNN</b>	90.62	89.28	90	0.0913

In Table3, the Network performance in classification accuracies for prostate neoplasia disorder data set are compared. It can conclude from table 3 that the best result for this data set is classified

by RBFNN (90%). The following figure shows the RBFNN training performance with error close to '0' as shown in Figure 7.

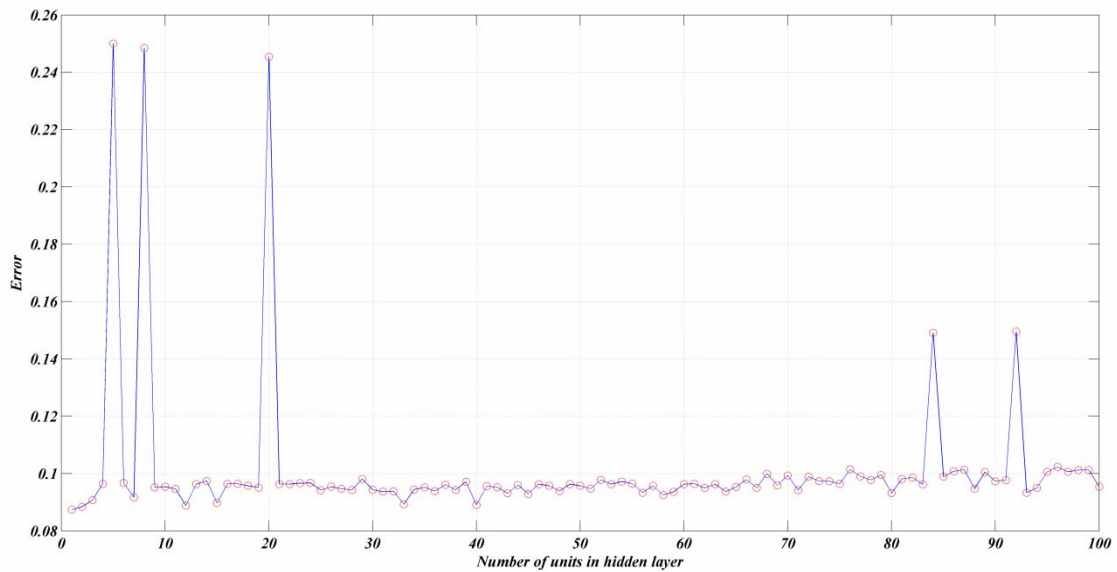


Figure 6. The relationship between the number of neurons in the hidden layer and the ANN performance.

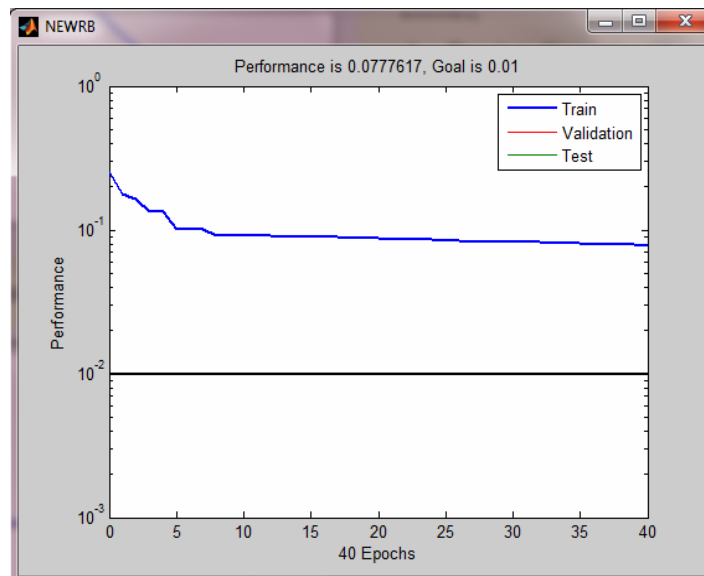


Figure 7. RBFNN training performance.

## Discussion and Conclusion

This paper presents the comparison of four neural network techniques for classification on local prostate neoplasia data sets. Each neural network technique selected for this comparison has different structures and different advantages and disadvantages. The RBFNN showed very promising results as shown in classification PND samples. BPNN exhibits show learning for big datasets, and this may be due to the iterative learning used by BPNN. Nevertheless, if topology and



neurons were used in a logical manner for BPNN training the NN can attain good learning performance, but that does not ensure accurate and good generalization. BPNN network makes a real-valued prediction between 0 and 1. BPNN is a robust model and it can provide competent results in various problems. BPNN gets trapped on the local minima and it is almost impossible to get global minima. It is still very effective and unlike most other neural networks, which produce good results for some datasets and bad for others, BPNN performs fairly well on most of the medical datasets. GRNN and PNN have simpler architectures, they can train data faster than BPNN model and it can provide competent results in various scenarios. The RBFNN with two layers and spread 7.5 is the best model for diagnosis of prostate cancer from benign hyperplasia of prostate disease. Its accuracy is 90% to diagnosis the prostate neoplasia diseases. It correctly classified the 54 new instances from 60 instances. Neither of other networks under consideration is superior in terms of classification accuracy to RBFNN. On the other hand, GRNN as a non-iterative neural network is more accurate than others, and it is parallel and takes less time during the training and testing phase. Thus, finally we come to conclusion that Radial Basis Function Neural Network (RBFNN) and multilayer perceptron trained with back propagation are best algorithms for prostate cancer and diagnosis. Future medical application contains such classification algorithms, and this system can help doctors in making accurate decisions about prostate cancer and benign hyperplasia of prostate. By using such systems, doctors can remove unnecessary biopsy and reduces cost. In addition, this system can speed up diagnostic time.

## References

1. Llobet R, Pérez-Cortés JC, Toselli AH, Juan A. Computer-aided detection of prostate cancer. *Int J Med Inf.* 2007;76(7):547-56.
2. Ferlay J, Shin HR, Bray F, Forman D, Mathers C, Parkin DM. Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer.* 2010;127(12):2893-917.
3. Ella Hassanien A, Al-Qaheri H, El-Dahshan E-SA. Prostate boundary detection in ultrasound images using biologically-inspired spiking neural network. *Applied Soft Computing.* 2011;11(2):2035-41.
4. Saritas I, Ozkan IA, Sert IU. Prognosis of prostate cancer by artificial neural networks. *Expert Systems with Applications.* 2010;37(9):6646-50.
5. Kotsiantis S, Zaharakis I, Pintelas P. Supervised machine learning: A review of classification techniques. *Frontiers in Artificial Intelligence and Applications.* 2007;160:3.
6. Kayaer K, Yıldırım T, editors. Medical diagnosis on Pima Indian diabetes using general regression neural networks. *Proceedings of the international conference on artificial neural networks and neural information processing (ICANN/ICONIP);* 2003.
7. Jeatrakul P, Wong K, editors. Comparing the performance of different neural networks for binary classification problems. *Natural Language Processing, 2009 SNLP'09 Eighth International Symposium on;* 2009: IEEE.
8. Nazzal JM, El-Emary IM, Najim SA. Multilayer perceptron neural network (MLPs) for analyzing the properties of Jordan oil shale. *World Applied Sciences Journal.* 2008;5(5):546-52.
9. Zhu Y, Williams S, Zwiggelaar R. Computer technology in detection and staging of prostate carcinoma: a review. *Med Image Anal.* 2006;10(2):178.
10. Fung CC, Iyer V, Brown W, Wong KW, editors. Comparing the performance of different neural networks architectures for the prediction of mineral prospectivity. *Machine Learning and Cybernetics, 2005 Proceedings of 2005 International Conference on;* 2005: IEEE.
11. Frost F, Karri V, editors. Performance comparison of BP and GRNN models of the neural network paradigm using a practical industrial application. *Neural Information Processing, 1999 Proceedings ICONIP'99 6th International Conference on;* 1999: IEEE.
12. Specht DF. Probabilistic neural networks and the polynomial adaline as complementary techniques for classification. *Neural Networks, IEEE Transactions on.* 1990;1(1):111-21.

13. Oliveira E, Ciarelli PM, Souza A, Badue C, editors. Using a probabilistic neural network for a large multi-label problem. *Neural Networks, 2008 SBRN'08 10th Brazilian Symposium on*; 2008: IEEE.
14. Wu SG, Bao FS, Xu EY, Wang Y-X, Chang Y-F, Xiang Q-L, editors. A leaf recognition algorithm for plant classification using probabilistic neural network. *Signal Processing and Information Technology, 2007 IEEE International Symposium on*; 2007: IEEE.
15. Gorunescu F, Gorunescu M, El-Darzi E, Ene M, Gorunescu S, editors. Statistical comparison of a probabilistic neural network approach in hepatic cancer diagnosis. *Computer as a Tool, 2005 EUROCON 2005 The International Conference on*; 2005: IEEE.
16. Berrar DP, Downes CS, Dubitzky W, editors. Multiclass cancer classification using gene expression profiling and probabilistic neural networks. *Proceedings of the Pacific Symposium on Biocomputing*; 2002.
17. Kurban T, Beşdok E. A comparison of RBF neural network training algorithms for inertial sensor based terrain classification. *Sensors*. 2009;9(8):6312-29.
18. Devaraj D, Yegnanarayana B, Ramar K. Radial basis function networks for fast contingency ranking. *International journal of electrical power & energy systems*. 2002;24(5):387-93.
19. Fu X, Wang L. Data dimensionality reduction with application to simplifying RBF network structure and improving classification performance. *Systems, Man, and Cybernetics, Part B: Cybernetics, IEEE Transactions on*. 2003;33(3):399-409.
20. Kumar K, Abhishek B. *Artificial Neural Networks for Diagnosis of Kidney Stones Disease*. 2012.
21. Delican Y, Ozyilmaz L, Yildirim T, editors. Evolutionary algorithms based RBF neural networks for Parkinson's disease diagnosis. *Electrical and Electronics Engineering (ELECO), 2011 7th International Conference on*; 2011: IEEE.
22. Venkatesan P, Anitha S. Application of a radial basis function neural network for diagnosis of diabetes mellitus. *Curr Sci*. 2006;91(9):1195-9.
23. LUO J-c, ZHOU C-h, LEUNG Y, editors. A knowledge-integrated RBF network for remote sensing classification. Paper presented at the 22nd Asian Conference on Remote Sensing; 2001.
24. Raza A, Khalid MU, Zaidi A. Comparison between Back-propagation and General Regression Neural Networks for Underwater Mine detection.