Comparing the Performance of Different Neural Networks for Binary Classification Problems

P. Jeatrakul and K.W. Wong

Abstract— **Classification problem is a decision making task where many researchers have been working on. There are a number of techniques proposed to perform classification. Neural network is one of the artificial intelligent techniques that has many successful examples when applying to this problem. This paper presents a comparison of neural network techniques for binary classification problems. The classification performance obtained by five different types of neural networks for comparison are Back Propagation Neural Network (BPNN), Radial Basis Function Neural Network (RBFNN), General Regression Neural Network (GRNN), Probabilistic Neural Network (PNN), and Complementary Neural Network (CMTNN). The comparison is done based on three benchmark data sets obtained from UCI machine learning repository. The results show that CMTNN typically provide better classification results when comparing to techniques applied to binary classification problems.**

I. INTRODUCTION

Classification is one of the important decision making tasks for many real world problems. Classification will be used when an object needs to be classified into a predefined class or group based on attributes of that object. There are many real world applications that can be categorized as classification problems such as weather forecast, credit risk evaluation, medical diagnosis, bankruptcy prediction, speech recognition, handwritten character recognition, and quality control [1]. Generally, there are two types of classification problems: binary problem and multiclass problem. While a binary problem is a situation in which an outcome of prediction has to be determined with a decision of Yes or No, a multiple classification problem is a condition in which a predicted result is determined as multiple outcomes [2].

In order to solve the classification problems and prediction, many classification techniques have been proposed. Some of the 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 [3, 4]. Moreover, a number of hybrid techniques have also been implemented such as neuro-fuzzy based classification technique [5], recursive partitioning of the majority class (REPMAC) algorithm[6], fuzzy probabilistic neural network [7].

In this paper, we aim to review five types of neural networks for binary classification problems. These are Back Propagation Neural Network (BPNN) [8], Radial Basis Function Neural Network (RBFNN) [9], General Regression Neural Network (GRNN) [10], Probabilistic Neural Network (PNN) [11], and Complementary Neural Network (CMTNN) [12]. The performances of these neural networks based on a range of data sets will be compared.

Each of the networks is selected for different reasons. BPNN is the most commonly used neural network in classification. Therefore, it is chosen as the basis for comparison. RBFNN is chosen for this study because it has several special characteristics. For example, it has a simple architecture, and it performs faster than BPNN. Similar to RBFNN, GRNN and PNN are selected because of their features. While GRNN can reduce the computation complexity, PNN can integrate the characteristics of statistical pattern recognition and BPNN. Furthermore, PNN has the ability to find boundaries between categories of patterns. CMTNN is also chosen because of its particular characteristics. It can integrate the truth and false membership values to deal with the uncertainty of classification while other techniques use only truth membership values.

For the comparison, three binary classification data sets from the UCI machine learning repository [13] are used. These include Pima Indians diabetes, BUPA liver disorders, and Johns Hopkins Ionosphere data set. These data sets are selected because they are benchmark data sets which have been commonly used in the literature. After the accuracies for each network type are evaluated and compared, the performance of each network and directions for further research will be discussed.

II. NEURAL NETWORK TECHNIQUES

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 a neural network is trained from past data, it will be able to generate outputs based on the knowledge extracted form the data [14]. Many research projects have been shown that ANN is a powerful technique for classification [1]. There are several advantages of using ANN for classification. Firstly, ANN can adapt itself to the data without making prior assumption of the

P. Jeatrakul is with the School of Information Technology, Murdoch University, Australia, WA, 6150 (e-mail: p.jeatrakul@murdoch.edu.au)

K.W. Wong is with the School of Information Technology, Murdoch University, Australia, WA, 6150 (e-mail: k.wong@murdoch.edu.au)

978-1-4244-4139-6/09/$25.00 ©2009 IEEE

functions. Secondly, ANN is a universal function approximator. Therefore, ANN is able to approximate any function with arbitrary accuracy. Finally, ANN is a nonlinear model that can be implemented for most complex real world applications [1]. Furthermore, there are many successful real world applications using ANN such as industry, business and science [15]. Example of applications that have implemented using ANN are bankruptcy prediction [16], handwriting recognition [17], fault detection [18], and medical diagnosis [19].

A. Back propagation neural network (BPNN)

BPNN is a simple and effective model of ANN. It contains three layers which are input, hidden and output layers as shown in Fig. 1 [20]. This network is also known as feedforward back-propagation neural networks.

Fig. 2 shows the learning process of a neural network. In the training phase, the training data is fed into the input layer. It is propagated to both the hidden layer and the output layer. This process is called the forward pass. In this stage, each node in the input layer, hidden layer and output layer calculates and adjusts the appropriate weight between nodes and generate output value of the resulting sum. The actual output values will be compared with the target output values. The error between these outputs will be calculated and propagated back to hidden layer in order to update the weight of each node again. This is called backward pass or learning. The network will iterate over many cycles until the error is acceptable. After the training phase is done, the trained network is ready to use for any new input data. During the testing phase, there is no learning or modifying of the weight matrices. The testing input is fed into the input layer, and the feed forward network will generate results based on its knowledge from trained network.[8]

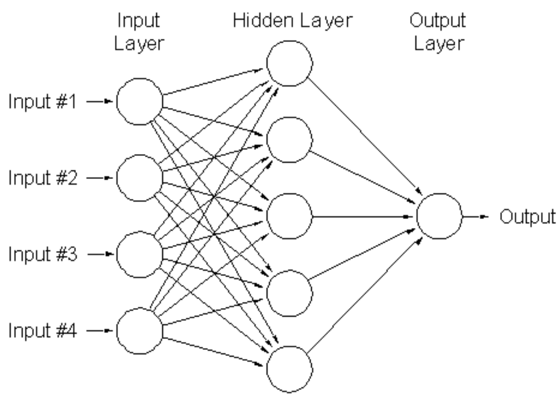

Fig. 1. Feedforward back-propagation network

BPNN is one of the popular ANN that has been used for many ANN applications [21]. BPNN is a robust neural network that can be applied easily in various problem domains [22]. However, there are also limitations in BPNN. This network requires a lot of input and target pairs for training the network. Furthermore, the internal mapping procedures work as a black box that may not be easily understood, and there is no indication that the system can generate all acceptable solutions [20].

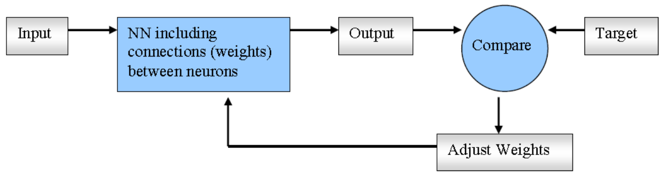

Fig. 2. Learning process of the feedforward back-propagation network [23]

B. 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 Perceptrons (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.

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.[3, 9, 24]

C. General Regression Neural Network (GRNN)

GRNN is a feed-forward neural network for supervised data. It uses nonlinear regression functions for approximation. Similarly to BPNN, GRNN consists of three layers: an input layer, a hidden layer, and an output layer. GRNN uses direct mapping to link the input layer to the hidden layer. While BPNN use the learning rate and momentum for the transfer function, 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. [21, 25]

Due to this process, GRNN can reduce the computational complexity and improve the learning process [26]. GRNN is similar to the probabilistic neural network (PNN) because both of them use non-parametric estimators of probability density functions. While GRNN is suitable to estimate continuous values, PNN is suitable to find boundaries between categories of patterns [10].

D. Probabilistic neural network (PNN)

Probabilistic neural network is a type of radial basis networks [27]. It is related to Bayesian decision rule and Parzen (probability density function estimators), namely the

Bayes-Parzen classification. PNN includes both characteristics of statistical pattern recognition and BPNN. It is applied to several areas including pattern recognition, nonlinear mapping and classification. PNN is a supervised feed-forward neural network. It also consists of three layers with one-pass training algorithm [28]. PNN has the ability to train on sparse data sets. Moreover, it is able to classify data into specific output categories [21]. 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 robust to noise. Furthermore, the training manner of PNN is simple and instantaneous [11].

E. Complementary Neural Network (CMTNN)

CMTNN is a technique using a pair of opposite feedforward backpropagation neural networks (Truth neural network and Falsity neural network) for classification problem as shown in Fig 3 [2].

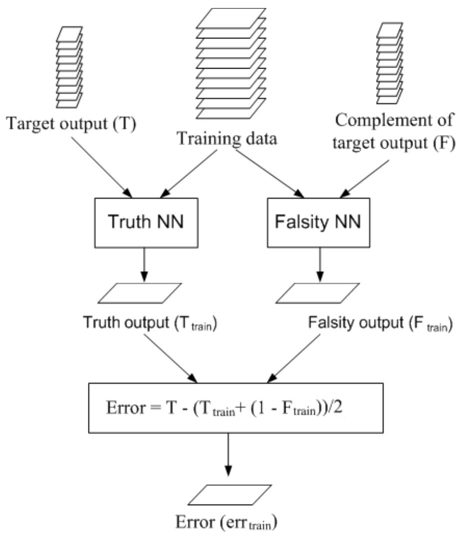

Fig 3. Complementary neural network [2]

This technique has successfully been implemented for both binary and multiclass classification problems. For binary classification, a pair of neural networks is implemented in order to predict degrees of truth and false membership values. Furthermore, the bagging technique is applied to an ensemble of pairs of neural networks in order to improve performances of a single pair of neural networks. The difference between the truth and false membership values can be used to represent uncertainty in the classification [12].

III. DATA SETS

In order to compare the performance of several types of neural network classifiers for binary classification problems, three data sets from benchmarking UCI data sets [13] are employed. These benchmark data sets include PIMA Indians diabetes, BUPA liver disorders, and Johns Hopkins Ionosphere.

- The purpose of PIMA Indians diabetes data set is to predict whether a patient shows signs of diabetes.
- The purpose of BUPA liver disorders data set is to predict whether a male patient shows signs of liver disorders.
- The purpose of Johns Hopkins Ionosphere data set is to predict “Good” or “Bad” radar return from the ionosphere.

The characteristics of these three data sets are shown in Table I.

TABLE I
CHARACTERISTICS OF DATA SETS USED IN THE COMPARISON

	No. of patterns	No. of attributes	No. of patterns in class 1	No. of patterns in class 2
Pima Indians diabetes	768	8	500	268
BUPA liver disorders	345	6	145	200
Johns Hopkins Ionosphere	351	34	225	126

IV. EXPERIMENTAL METHODOLOGY AND RESULTS

In order to compare the performance of neural network techniques, firstly, each data set is split into 80% training set and 20% testing set as shown in Table II. In order to obtain reliable results by using a cross validation method, each data set will be randomly split ten times to form different training and testing data before they are classified by each neural network technique. In the experiment, MATLAB software is used to design and test each neural network type.

TABLE II
NUMBER OF PATTERNS IN THE TRAINING AND TEST DATA SET

	No. of training data	No. of test data	Total
Pima Indians diabetes	614	154	768
BUPA liver disorders	276	69	345
Johns Hopkins Ionosphere	281	70	351

After the test data is classified by each neural network, the average of the ten results of the classification accuracy will be used for comparing the performance of these neural networks.

Table III-V shows the performance results obtained by five neural networks for each data set.

In table III, the classification accuracies obtained by five neural network techniques are compared for Pima Indians diabetes data set. The results show that the average accuracies obtained by each neural network technique are quite compatible. The highest average accuracy is given by RBFNN (76.56%), following by the results of CMTNN (76.49%) and BPNN (76.17%). The results of GRNN and PNN provide the lowest performance with the same accuracies at 75.26%.

TABLE III
CLASSIFICATION ACCURACIES (%) OBTAINED BY FIVE TECHNIQUES OF NEURAL NETWORKS FOR PIMA INDIANS DIABETES DATA SET

Random data No.	BPNN	GRNN	RBFNN	PNN	CMTNN
1	77.27	74.68	79.22	74.68	77.92
2	76.62	79.87	79.22	79.87	76.62
3	70.13	70.13	74.03	70.13	72.08
4	85.71	81.82	79.22	81.82	83.77
5	75.97	75.97	77.27	75.97	75.32
6	70.78	70.13	72.08	70.13	72.08
7	75.32	72.73	76.62	72.73	75.97
8	79.22	78.57	77.27	78.57	79.22
9	74.68	74.68	76.62	74.68	75.32
10	75.97	74.03	74.03	74.03	76.62
Average (%)	76.17	75.26	76.56	75.26	76.49

TABLE IV
CLASSIFICATION ACCURACIES OBTAINED BY FIVE TECHNIQUES OF NEURAL NETWORKS FOR BUPA LIVER DISORDERS DATA SET

Random data No.	BPNN	GRNN	RBFNN	PNN	CMTNN
1	75.36	71.01	69.57	71.01	75.36
2	62.31	69.57	65.22	69.57	66.67
3	71.01	65.23	66.67	65.23	69.57
4	75.36	63.77	66.67	63.77	72.46
5	79.71	68.12	71.01	68.12	78.26
6	72.46	68.12	71.01	68.12	75.36
7	63.77	57.97	66.67	57.97	69.57
8	66.67	53.62	69.57	53.62	69.57
9	72.46	69.57	72.46	69.57	73.91
10	60.86	53.62	56.52	53.62	52.17
Average (%)	70.00	64.06	67.54	64.06	70.29

TABLE V
CLASSIFICATION ACCURACIES (%) OBTAINED BY FIVE TECHNIQUES OF NEURAL NETWORKS FOR JOHNS HOPKINS IONOSPHERE DATA SET

Random data No.	BPNN	GRNN	RBFNN	PNN	CMTNN
1	87.14	90.00	95.71	82.86	90.00
2	94.29	94.29	95.71	85.71	95.71
3	87.14	91.43	90.00	85.71	92.86
4	94.29	95.71	82.86	90.00	95.71
5	94.29	94.29	91.43	88.57	95.71
6	82.86	94.29	91.43	85.71	88.57
7	87.14	91.43	88.57	77.14	90.00
8	88.57	94.29	85.71	91.43	94.29
9	94.29	92.86	88.57	84.29	95.71
10	92.86	92.86	91.43	84.29	95.71
Average (%)	90.29	93.15	90.14	85.57	93.43

In table IV, the classification accuracies for BUPA liver disorder data set are compared. The accuracies of the GRNN, RBFNN and PNN techniques achieved are less than 70%. The performance of the BPNN and CMMN techniques are better than those three techniques. The best result for this data set is classified by CMTNN (70.29%), following by BPNN at 70%.

In table V, the classification accuracies obtained by five neural network techniques are compared for John Hopkins Ionosphere data set. The results show that four techniques can provide the average classification accuracies higher than 90%. These are BPNN (90.29%), GRNN (93.15%), RBFNN (90.14%) and CMTNN (93.43%), whereas PNN provides the lowest average accuracies at 85.57%. The CMTNN again provides the best result in this data set.

V. DISCUSSION AND CONCLUSION

This paper presents the comparison of five neural network techniques for binary classification on three benchmarking UCI data sets. Each neural network technique selected for this comparison has different structures and different advantages and disadvantages. While RBFNN, GRNN and PNN have simpler architectures and they can train data faster than BPNN, BPNN is a robust model and it can provide competent results in various problems. In addition, CMTNN has special features based on BPNN. It uses a pair of opposite BPNN to deal with uncertainty.

In terms of performance comparison based on the classification accuracy as shown in Figure 4, we found that generally the results achieved by CMTNN are higher than other techniques. In additional, BPNN gives the good performance results in every test case.

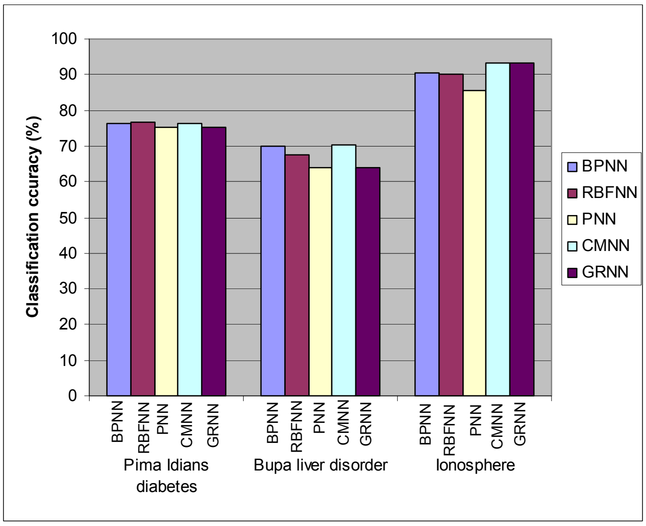

Fig 4. The comparison of classification accuracies in each data set

Furthermore, we notice that GRNN performs well in Ionosphere data set, which includes a large number of input attributes, 34 attributes. This is because GRNN can work well in noisy environment if it is trained with enough data [25]. In addition, RBFNN show the best performance in Pima Indians diabetes data set, which composes of a large number of training data. GRNN and PNN are consistently providing lower accuracy in the two data sets, Pima Indians diabetes data and BUPA disorders data, where they compose of small number of input attributes, eight and six attributes respectively. Moreover, all neural network types perform well with compatible results in the large training data set, Pima Indians diabetes data.

As can be seen from the results, it can be concluded that CMTNN is suitable for the binary classification problems. The results of the experiment show that CMTNN can provide good results in most cases. This is because of their unique features of using the true and false networks. CMTNN can deal with the uncertainty better than other techniques by using a pair of complementing neural network.

In future studies, the relationships of the number of attributes, the size of training set and the type of neural network can be studied intensively. The computational time of each neural network type can be evaluated and compared as well.

REFERENCES

[1] G. P. Zhang, "Neural networks for classification: a survey," *Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on,* vol. 30, pp. 451-462, 2000.

[2] P. Kraipeerapun, "Neural network classification based on quantification of uncertainty," Murdoch University, 2008.

[3] S. B. Kotsiantis, "Supervised Machine Learning: A Review of Classification Techniques," *Informatica,* vol. 31, pp. 249-268, 2007.

[4] K. Kayaer and T. Yildirim, "Medical diagnosis on pima diabetes using general regression neural networks," in *artificial neural networks and neural information processing (ICANN/ICONIP)* Istanbul, Turkey, 2003.

[5] Y. Sun, F. Karray, and S. Al-sharhan, "Hybrid soft computing techniques for heterogeneous data classification," in *Fuzzy Systems, 2002. FUZZ-IEEE'02. Proceedings of the 2002 IEEE International Conference on*, 2002, pp. 1511-1516.

[6] H. Ahumada, G. L. Grinblat, L. C. Uzal, P. M. Granitto, and A. Ceccatto, "REPMAC: A new hybrid approach to highly imbalanced classification problems," in *Hybrid Intelligent Systems, 2008. HIS '08. Eighth International Conference on*, 2008, pp. 386-391.

[7] Y. Huang and C. Tian, "Research on credit risk assessment model of commercial banks based on fuzzy probabilistic neural network," in *Risk Management & Engineering Management, 2008. ICRMEM '08. International Conference on*, 2008, pp. 482-486.

[8] J. Nazari and O. K. Ersoy, "Implementation of back-propagation neural networks with Matlab," School of electrical engineering, Prudue university1992.

[9] P. Venkatesan and S. Anitha, "Application of a radial basis function neural network for diagnosis of diabetes mellitus," *Current Science,* vol. 91, pp. 1195-1999, 2006.

[10] D. F. Specht, "A general regression neural network," *Neural Networks, IEEE Transactions on,* vol. 2, pp. 568-576, 1991.

[11] S. G. Wu, F. S. Bao, E. Y. Xu, Y.-X. Wang, Y.-F. Chang, and Q.-L. Xiang, "A leaf recognition algorithm for plant classification using probabilistic neural network," in *Signal Processing and Information Technology, 2007 IEEE International Symposium on*, 2007, pp. 11-16.

[12] P. Kraipeerapun, C. C. Fung, and S. Nakkrasae, "Porosity prediction using bagging of complementary neural networks," in *Advances in Neural Networks – ISNN 2009*, 2009, pp. 175-184.

[13] A. Asuncion and D. J. Newman, "UCI Machine Learning Repository," University of California, Irvine, School of Information and Computer Sciences, 2007.

[14] P. J. G. Lisboa, A. Vellido, and B. Edisbury, *Business applications of neural networks: The state of the art of real world applications*: World scientific, 2000.

[15] B. Widrow, D. E. Rumelhard, and M. A. Lehr, "Neural networks: Applications in industry, business and science," *Commun. ACM,* vol. 37, pp. 93-105, 1994.

[16] C. Charalambous, A. Charitou, and F. Kaourou, "Comparative analysis of artificial neural network models: application in bankruptcy prediction," in *Neural Networks, 1999. IJCNN '99. International Joint Conference on*, 1999, pp. 3888-3893 vol.6.

[17] L. Gang, B. Verma, and S. Kulkami, "Experimental analysis of neural network based feature extractors for cursive handwriting recognition," in *Neural Networks, 2002. IJCNN '02. Proceedings of the 2002 International Joint Conference on*, 2002, pp. 2837-2842.

[18] D. R. Hush, C. T. Abdallah, G. L. Heileman, and D. Docampo, "Neural networks in fault detection: a case study," in *American Control Conference, 1997. Proceedings of the 1997*, 1997, pp. 918-921 vol.2.

[19] Z.-H. Zhou and Y. Jiang, "Medical diagnosis with C4.5 rule preceded by artificial neural network ensemble," *Information Technology in Biomedicine, IEEE Transactions on,* vol. 7, pp. 37-42, 2003.

[20] D. Anderson and G. McNeill, "Artificial neural neworks technology: A DACS state-of-the-art report," Kaman Sciences Corporation20 August 1992.

[21] C. C. Fung, V. Iyer, W. Brown, and K. W. Wong, "Comparing the performance of different neural networks architectures for the prediction of mineral prospectivity," in *Machine Learning and Cybernetics, 2005. Proceedings of 2005 International Conference on*, 2005, pp. 394-398.

[22] S. C. Chen, S. W. Lin, T. Y. Tseng, and H. C. Lin, "Optimization of back-propagation network using simulated annealing approach," in *Systems, Man and Cybernetics, 2006. SMC '06. IEEE international conference on*, 2006, pp. 2819-2824.

[23] R. Ball and P. Tissot, "Demonstration of artificial neural network in Matlab," Division of Nearhsore research, Texas A&M university, 2006.

[24] J.-c. Luo, C.-h. Zhou, and Y. Leung, "A knowledge-integrated RBF network for remote sensing classification," in *The 22nd Asian Conference on Remote Sensing*, 2001.

[25] F. Frost and V. Karri, "Performance comparison of BP and GRNN models of the neural network paradigm using a practical industrial application," in *Neural Information Processing, 1999. Proceedings. ICONIP '99. 6th International Conference on*, 1999, pp. 1069-1074 vol.3.

[26] Y. Song and Y. Ren, "A Predictive model of nonlinear system based on generalized regression neural network," in *Neural Networks and Brain, 2005. ICNN&B '05. International Conference on*, 2005, pp. 2009-2012.

[27] S.-Y. Ooi, A. B. J. Teoh, and T.-S. Ong, "Compatibility of biometric strengthening with probabilistic neural network," in *Biometrics and Security Technologies, 2008. ISBAST 2008. International Symposium on*, 2008, pp. 1-6.

[28] F. Gorunescu, M. Gorunescu, E. El-Darzi, and S. Gorunescu, "An evolutionary computational approach to probabilistic neural network with application to hepatic cancer diagnosis," in *Computer-Based Medical Systems, 2005. Proceedings. 18th IEEE Symposium on*, 2005, pp. 461-466.