Integrated Computer-Aided Engineering 17 (2010) 197–210
DOI 10.3233/ICA-2010-0345
IOS Press

Enhanced probabilistic neural network with local decision circles: A robust classifier

Mehran Ahmadlou[a] and Hojjat Adeli[b,*]
[a]*Amirkabir University of Technology, Tehran, Iran*
[b]*Departments of Biomedical Engineering, Biomedical Informatics, Civil and Environmental Engineering and Geodetic Science, Electrical and Computer Engineering, Neurological Surgery, and Neuroscience, Ohio State University, Columbus, OH, USA*

Abstract. In recent years the Probabilistic Neural Network (PPN) has been used in a large number of applications due to its simplicity and efficiency. PNN assigns the test data to the class with maximum likelihood compared with other classes. Likelihood of the test data to each training data is computed in the pattern layer through a kernel density estimation using a simple Bayesian rule. The kernel is usually a standard probability distribution function such as a Gaussian function. A spread parameter is used as a global parameter which determines the width of the kernel. The Bayesian rule in the pattern layer estimates the conditional probability of each class given an input vector without considering any probable local densities or heterogeneity in the training data. In this paper, an enhanced and generalized PNN (EPNN) is presented using *local decision circles* (LDCs) to overcome the aforementioned shortcoming and improve its robustness to noise in the data. Local decision circles enable EPNN to incorporate local information and non-homogeneity existing in the training population. The circle has a radius which limits the contribution of the local decision. In the conventional PNN the spread parameter can be optimized for maximum classification accuracy. In the proposed EPNN two parameters, the spread parameter and the radius of local decision circles, are optimized to maximize the performance of the model. Accuracy and robustness of EPNN are compared with PNN using three different benchmark classification problems, iris data, diabetic data, and breast cancer data, and five different ratios of training data to testing data: 90:10, 80:20, 70:30, 60:40, and 50:50. EPNN provided the most accurate results consistently for all ratios. Robustness of PNN and EPNN is investigated using different values of signal to noise ratio (SNR). Accuracy of EPNN is consistently higher than accuracy of PNN at different levels of SNR and for all ratios of training data to testing data.

1. Introduction

Since the development of back propagation (BP) algorithm in the 1980's [33,45,67] many attempts have been made to develop more efficient and robust neural networks and training algorithms. Noteworthy among them are the adaptive conjugate gradient neural network algorithm [3,4,18], counter propagation neural networks [14,24,80], recurrent neural networks [60,64], radial basis function neural networks [9, 33,44,45,78], neuro-fuzzy networks [5,6,34,39,71], wavelet neural networks [29,72,98], spiking neural networks [30–32,69], neural dynamics models [7,8, 10,15,16,19,43,62,63,67,79,84], complex-valued neural networks [20,22,27,38,47,51,54,59,66,91], neurogenetic/evolutionary models [26,40], and the probabilistic neural network (PNN) [53,83].

Advantages of PPN are a) simplicity for not requiring a separate training phase for tuning network weights based on some learning rule and b) efficiency [35, 93]. As such in recent years PNN has been used in a large number of applications such as seismic engineering [13], hydrology [86], ocean engineering [50], image processing [65,82], EEG, EKG, and other biomedical data analysis [58,88–90,96], power engineering [28,57, 73,87], and transportation engineering [61].

*Corresponding author: Hojjat Adeli, Departments of Biomedical Engineering, Biomedical Informatics, Civil and Environmental Engineering and Geodetic Science, Electrical and Computer Engineering, Neurological Surgery, and Neuroscience, College of Engineering, 470 Hithcock Hall, 2070 Neil Avenue, Ohio State University, Columbus, OH 43210, USA. E-mail: 1@osu.edu.

ISSN 1069-2509/10/$27.50

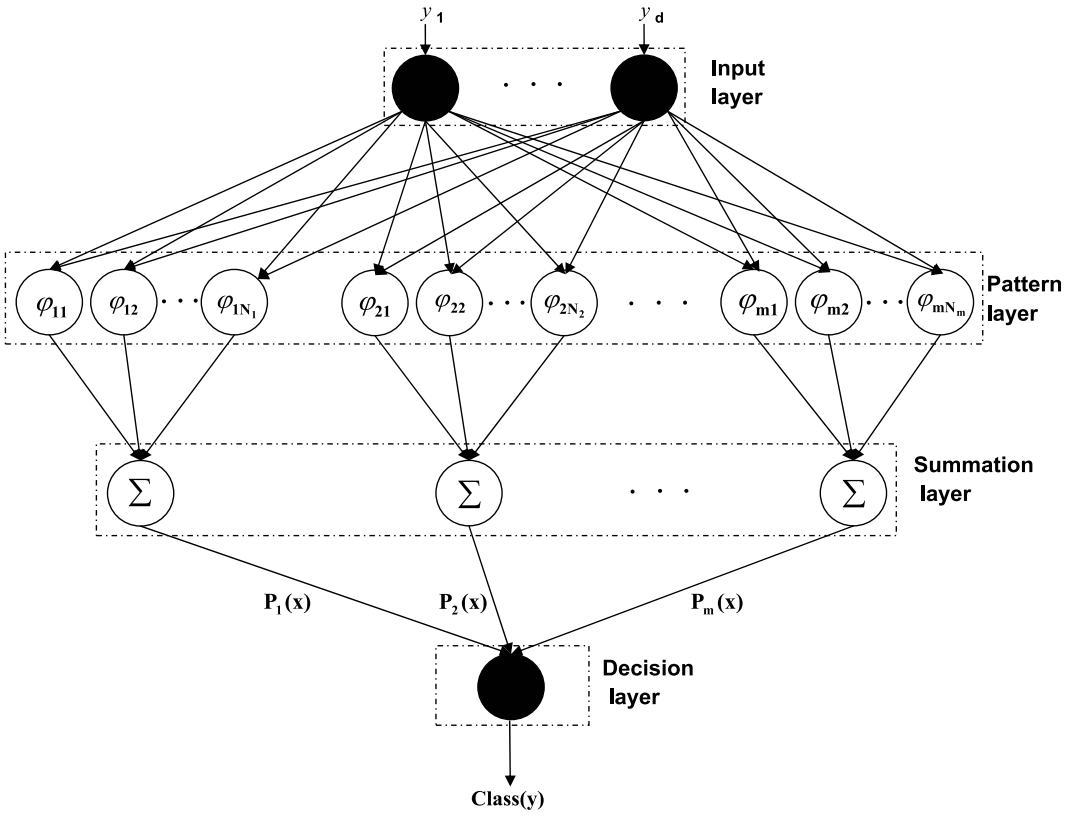

Fig. 1. Architecture of a probabilistic neural network (PNN).

PNN has four layers: an input layer, a pattern layer, a summation layer, and a decision layer (Fig. 1). PNN classifies input data based on a simple Bayesian rule. That is, PNN assigns the test data to the class with maximum likelihood compared with other classes. Likelihood of the test data to each training data is computed in the pattern layer through a kernel density estimation, called Parzen window method, a non-parametric method for estimating the probability density function of a random variable using a simple Bayesian rule. The kernel is usually a standard Probability Distribution Function (PDF). A width or spread parameter is used as a global parameter which determines the width of the kernel. This is the only parameter used to optimize the performance of PNN [55,87]. It is global in the sense that it is the same in all decisions without any assumption on the probable local information in the distribution of the training data. The Bayesian rule in the pattern layer estimates the conditional probability of each class given an input vector without considering any probable local densities or heterogeneity in the training data.

In this paper, an enhanced PNN (EPNN) is presented using *local decision circles* (LDCs) to overcome the aforementioned shortcoming. *Local decision circles* enable EPNN to incorporate local information and inhomogeneity existing in the training population. The *circle* has a radius which limits the contribution of the *local decision*. As such, EPNN uses two Bayesian rules: 1) A global rule which estimates the conditional probability of each class given an input vector of data considering all training data and using the spread parameter. 2) A local rule which estimates the conditional probability of each class, given an input vector of data existing within a *decision circle*, considering only the training data in that *decision circle*. The only parameter of the *decision circle* is its radius. The simultaneous global and local decisions, using the global and local rules, are expected to enhance the accuracy and robustness of the classification. When the radius is zero there is no *local decision* and the proposed EPNN becomes identical to the conventional PNN with only the spread parameter. Consequently, the EPNN can be viewed as a generalized PNN employing local decisions.

Accuracy and robustness of EPNN are compared with PNN using three different benchmark classification problems: iris data, diabetic data, and breast cancer data.

2. Probabilistic neural network

PNN consists of an input layer with d neurons (d is the dimension of the input data or pattern vector), a pattern layer with N neurons (N is the number of training data set), a summation layer of m neurons (m is the number of classes), and an output or decision layer of one neuron (Fig. 1). Vector $\mathrm{y} = [y_1, \cdots, y_d]$ is a d-dimensional input data vector to be classified into one of the m classes, N_i is number of training data belonging to the i^{th} class, $\mathrm{x_{ij}}$ is the j^{th} data training vector belonging to the i^{th} class, φ_{ij} in the pattern layer is a function of $\mathrm{y} - \mathrm{x_{ij}}$ which represents the likelihood of y being equal to$\mathrm{x_{ij}}$ according to a defined kernel, a standard Gaussian function used as PDF:

$$\varphi_{ij}(\mathrm{y}) = \frac{1}{(2\pi)^{d/2}\sigma^d} \times \exp\left[-\frac{(\mathrm{y}-\mathrm{x_{ij}})^T(\mathrm{y}-\mathrm{x_{ij}})}{2\sigma^2}\right] \quad (1)$$

Where σ is the spread of the Gaussian function or the width parameter which can take a value between 0 and 1. The best value of the spread parameter is the one that yields the most accurate classification result. It is obtained through numerical investigation for any particular classification problem.

The term $P_i(\mathrm{y})$ in the summation layer (Fig. 1) indicates the conditional probability or likelihood of y belonging to the i^{th} class and is obtained through a summation process as follows:

$$P_i(\mathrm{y}) = \frac{1}{N_i}\sum_{j=1}^{N_i}\varphi_{ij}(\mathrm{y}) \quad (2)$$

In the output layer y will be assigned to a class with maximum likelihood, that is,

$$Class(\mathrm{y}) = \arg\max_i\{P_i(\mathrm{y})\} \quad (3)$$

where $Class(\mathrm{y})$ is the class that y belongs to.

3. Enhanced probabilistic neural networkwith local decision circles

Suppose $C(\mathrm{x_{ij}})$ is the class that the data $\mathbf{x}_{\mathrm{ij}}$ belongs to and $S_{r,\mathrm{x_{ij}}}$ is a hyper-sphere with radius r and center $\mathrm{x_{ij}}$. The conditional probability that an embedded data, x, in the hyper-sphere $S_{r,\mathrm{x_{ij}}}$ are from the same class of $\mathrm{x_{ij}}$ (to be classmate of $\mathrm{x_{ij}}$) can be written as:

$$\alpha_{\mathrm{x_{ij}}} = P[\mathrm{x} \in Class(\mathbf{x}_{\mathrm{ij}})|\mathrm{x} \in S_{r,\mathrm{x_{ij}}}] \quad (4)$$

Simply, $\alpha_{\mathrm{x_{ij}}}$ is the ratio of the number of data in the hyper-sphere $S_{r,\mathrm{x_{ij}}}$ that are from the same class of $\mathrm{x_{ij}}$ to the number of all data in the same hyper-sphere. Therefore, $\alpha_{\mathrm{x_{ij}}}$ close to one (or zero) means many (or few) data in the hyper-sphere $S_{r,\mathrm{x_{ij}}}$ (which is around $\mathrm{x_{ij}}$) are from the same class of $\mathrm{x_{ij}}$. Then, the probability of a pattern vector (a test data point) in $S_{r,\mathrm{x_{ij}}}$ belonging to the class of $\mathrm{x_{ij}}$ is expected to be higher than its probability of belonging to other classes with fewer data and vice versa. Quantitatively this can be implemented by changing the width or spread parameter so that when there is a low number of data around $\mathrm{x_{ij}}$ belonging to the same class of $\mathrm{x_{ij}}$ (corresponding to a low value of $\alpha_{\mathrm{x_{ij}}}$), the spread (representing variation around center) of the Gaussian function φ_{ij} becomes small, and vice versa. To achieve this goal, the spread parameter is multiplied by $\alpha_{\mathrm{x_{ij}}}$ as follows:

$$\sigma_{\mathrm{x_{ij}}} = \alpha_{\mathrm{x_{i,j}}} \times \sigma \quad (5)$$

Substituting $\sigma_{\mathrm{x_{ij}}}$ for σ in Eqs (1) and (2), the conditional probability of y belonging to the i^{th} class is computed as follows:

$$P_i(\mathrm{y}) = \frac{1}{N_i}\sum_{j=1}^{N_i} \left\{\frac{1}{(2\pi)^{d/2}\sigma_{\mathrm{x_{ij}}}^d}\exp\left[-\frac{(\mathrm{y}-\mathrm{x_{ij}})^T(\mathrm{y}-\mathrm{x_{ij}})}{2\sigma_{\mathrm{x_{ij}}}^2}\right]\right\} \quad (6)$$

The example two-dimensional data of two classes shown in Fig. 2 are used to illustrate the proposed EPNN. In this figure, black and white dots show data of classes 1 and 2, respectively. $\mathrm{x_{1i}}$ and $\mathrm{x_{1j}}$ are two sample data of class 1 and $S_{r,\mathrm{x_{1i}}}$ and $S_{r,\mathrm{x_{1j}}}$ are the corresponding local decision circles around them with radius r. In this example, the number of data points of class 1 divided by the total number of data in the local decision circle around $\mathrm{x_{1j}}$ (1/4) is much less than that for $\mathrm{x_{1i}}$ (6/6). That is $\alpha_{\mathrm{x_{1,j}}} = 1/4$ and $\alpha_{\mathrm{x_{1,i}}} = 1$. Therefore, the spread of Gaussian kernel function around $\mathrm{x_{1j}}$ is less than that around $\mathrm{x_{1i}}$ in the pattern layer.

Expectation for a pattern vector (a test data point) of class 1to be in $S_{r,\mathrm{x_{1,j}}}$ is substantially lower than in $S_{r,\mathrm{x_{1,i}}}$. Similarly, Bayesian probability leads to the conclusion that the posterior probability for a pattern vector in $S_{r,\mathbf{x}_{1\mathrm{j}}}$ to be from class 1 is substantially less than to be in $S_{r,\mathrm{x_{1i}}}$. According to Eq. (4), $\alpha_{\mathrm{x_{1,j}}}$ and $\alpha_{\mathrm{x_{1,i}}}$ are equal to 1/4 and 6/6, respectively, which are indeed the prior probabilities for a pattern vector (a test data point) in $S_{r,\mathrm{x_{1j}}}$ and $S_{r,\mathrm{x_{1i}}}$ to be from class 1, respectively. In turn, they result in low and high posterior probabilities for the pattern vector to be from

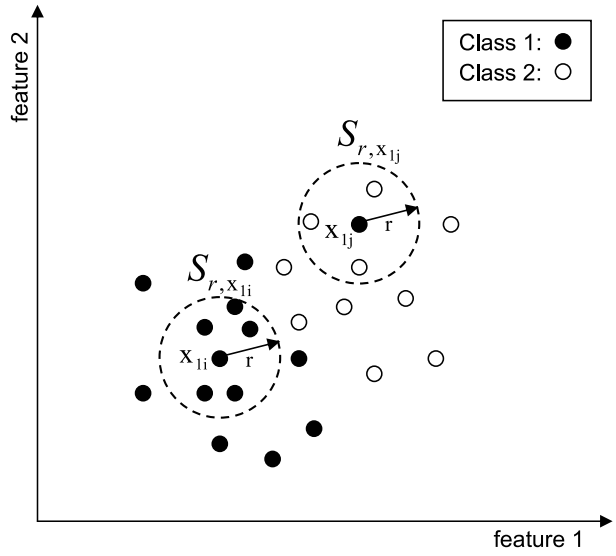

Fig. 2. Example two-dimensional data of two classes to illustrate the proposed EPNN.

class 1, respectively. This correspondence between the expectation and the posterior probability (for a pattern vector not only in the decision circle but in the overall domain of the data) is met by changing the spread of the Gaussian kernel function in $\varphi_{1j}(y)$ and $\varphi_{1i}(\mathbf{y})$ (i.e. σ). Indeed this correspondence is met by narrowing the kernel of $\varphi_{1j}(y)$ through multiplication of σ by $\alpha_{x_{1,j}}[\sigma_{\mathbf{x}_{1,j}} = (1/4{\times}\sigma)]$. Likewise the spread of the kernel of $\varphi_{1i}(y)$ is multiplied to $\alpha_{x_{1,i}}$. But, since $\alpha_{\mathbf{x}_{1,i}}$ is equal to 1, it remains unchanged and $\sigma_{x_{1,i}}$ will be the same as σ. When $r = 0$, the only data in $S_{r,x_{ij}}$ is its central data, x_{ij}, which means (according to Eq. (4)) $\alpha_{x_{ij}} = 1$ and consequently, according to Eq. (5), $\sigma_{x_{ij}} = \sigma$. In this case Eq. (6) is reduced to Eq. (3) which means PNN is a special case of EPNN without local decision circles.

In this paper the radius of decision circles r is normalized to interval [0,1] by dividing it by the maximum Euclidean distance of training data pairs. For a given value of the spread factor σ, a maximum threshold for the radius of decision circles, r_T, is expected to exist beyond which no improvement in the performance of PNN is achieved. Values of r larger than r_T can interfere with the global decision of PNN thus reducing its classification accuracy.

The best performance of EPNN, however, is obtained by optimizing σ and r ($0 \leqslant r \leqslant r_T$). In this paper, a bi-level optimization is employed. First, the best value of σ which yields the most accurate results is obtained for PNN (or EPNN with $r = 0$). Then, the optimum value of r is obtained using the best value of σ. The optimization approach used is Particle Swarm Optimization (PSO) [46,68,92,95]. Similar to genetic algorithms [1,2,11,12,21,23,25,34,42,49,52,56,74–77, 81,85,97], PSO is an evolutionary population-based optimization algorithm originally inspired by social behavior of bird flocks. It is known for ease of implementation and relatively high computational speed [41, 95].

4. Results

The proposed EPNN is compared with PNN using three different benchmark classification problems and data sets in the order of increasing dimensionality from the University of California – Irvine: Iris data, diabetic data, and breast cancer data.

4.1. Iris classification problem (*http://archive.ics.uci.edu/ml/datasets/Iris*)

The iris plants database contains 50 instances for 3 classes of irises (setosa, versicolor, and virginica, three different types of iris plant) based on 4 features: sepal length, sepal width, petal length, and petal width (all in cm). One class is linearly separable from the other two but the latter two cannot be distinguished linearly from each other.

4.2. Diabetics Diagnosis Problem (*http://archive.ics.uci.edu/ml/datasets/Pima+Indians+Diabetes*)

This data set known as the Pima Indians diabetes database contains 500 and 268 tested negative and positive for diabetes, respectively, based on 8 features: Number of times pregnant, Plasma glucose concentration 2 hours in an oral glucose tolerance test, diastolic blood pressure (mm Hg), Triceps skin fold thickness (mm), 2-Hour serum insulin, body mass index, diabetes pedigree (heredity) function, and age in years.

4.3. Breast Cancer Classification Problem http://archive.ics.uci.edu/ml/datasets/Breast+Cancer+Wisconsin+(*Original*))

This data set known as Wisconsin breast cancer database (after removing 16 data which contains missing feature values) has been gathered from 444 benign and 239 malignant cancers based on 9 features: Clump Thickness, Uniformity of Cell Size, Uniformity of Cell Shape, Marginal Adhesion, Single Epithelial Cell Size, Bare Nuclei, Bland Chromatin, Normal Nucleoli, and Mitoses [94].

To compare the accuracy of EPNN with the highest accuracy of PNN, first the best value of σ which results in the highest accuracy of PNN is found. To obtain statistically meaningful results, five different ratios of training data to testing data (RTT) were used: 90:10, 80:20, 70:30, 60:40, and 50:50. For each ratio, training and testing data sets were selected randomly and the process was repeated 100 times. This cross-validation method is called repeated random sub-sampling validation. The average value of all accuracies is reported as the accuracy of PNN (i.e., EPNN with $r = 0$) for each σ. Figures 3, 4, and 5 show diagrams of accuracies of PNN in terms of σ for classification of the Iris data (with resolution of 0.01), diabetic data (with resolution of 0.05), and breast cancer data (with resolution of 0.05) for one example RTT of 90:10, respectively. Vertical bars in the diagrams show the range of standard deviations of average accuracies in 100 runs.

According to Figs 3, 4, and 5, the best values of σ are 0.56, 0.35, and 0.45 for classification of the corresponding data. Figures 6, 7, and 8 show the mean accuracies of 100 different runs for EPNN in terms of r using the bi-level optimization approach and the best values of σ obtained for PNN (for the same RTT of 90/10). The value of r_T has been marked in those figures with a dashed vertical line. These figures show accuracy of EPNN is higher than PNN (EPNN with $r = 0$) for the range $0 \leqslant r \leqslant r_T$. Based on bi-level optimization the best accuracies obtained for EPNN are 97.2%, 75.2%, and 96.5% while the best accuracies obtained for PNN are 96.4%, 74.0%, and 96.4% for classification of the iris data, diabetic data, and breast cancer data, respectively. Figure 9 shows a comparison of the accuracies of PNN with the best value of σ (dotted line) and EPNN with the best value of σ for PNN (solid line) for the three classification problems in terms of the ratio of the number of training data to the number of testing data. EPNN provided the most accurate results for all RTTs.

Next, robustness of PNN and EPNN is investigated using different values of Signal to Noise Ratio (SNR) defined as the power of a signal divided by the power of noise in dB scale. Gaussian noise is added to the multi-dimensional data with six different SNR values of 1, 2, 3, 4, 5, and 10 dB. Figures 10, 11, and 12 show the accuracies of PNN with the best value of σ (dotted line) and EPNN (solid line) in terms of RTT (50/50, 60/40, 70/30, 80/20, and 90/10) for the three classification problems and six aforementioned SNR values. Accuracies of PNN and EPNN decrease with decreasing SNR as expected. Accuracy of EPNN is consistently higher than accuracy of PNN at different levels of SNR and for all RTT values, and the difference of their accuracies increases with decreasing SNR (increasing noise levels).

Figure 13 demonstrates why EPNN is more robust to noise than PNN. Figure 13a shows a sample 2-dimensional data of two classes with the same nomenclature as Fig. 2. Figure 13b shows the same data but with added noise. In this figure noise is represented as perturbed data and the small circles indicate the same locations of the circles in Fig. 13a before perturbation. In Fig. 13a the number of data of class 1 divided by the total number of data in the local decision circle is 2/5 (before perturbation) and it remains the same in Fig. 13b after perturbation (note that the center of data in Fig. 13b is different from that in Fig. 13a). That means the alpha parameter of EPNN which plays the role of a robustness factor is not affected by this perturbation or noise. Figure 13b shows adding noise does not change the ratio of the number of data of class 1 to the number of data in a local decision circle which means the same local decisions for both with and without noise conditions for EPNN, while the relative positions of the noisy data are different from those of the original data for PNN. Therefore, the classification of data using EPNN is affected less by noise compared with the conventional PNN.

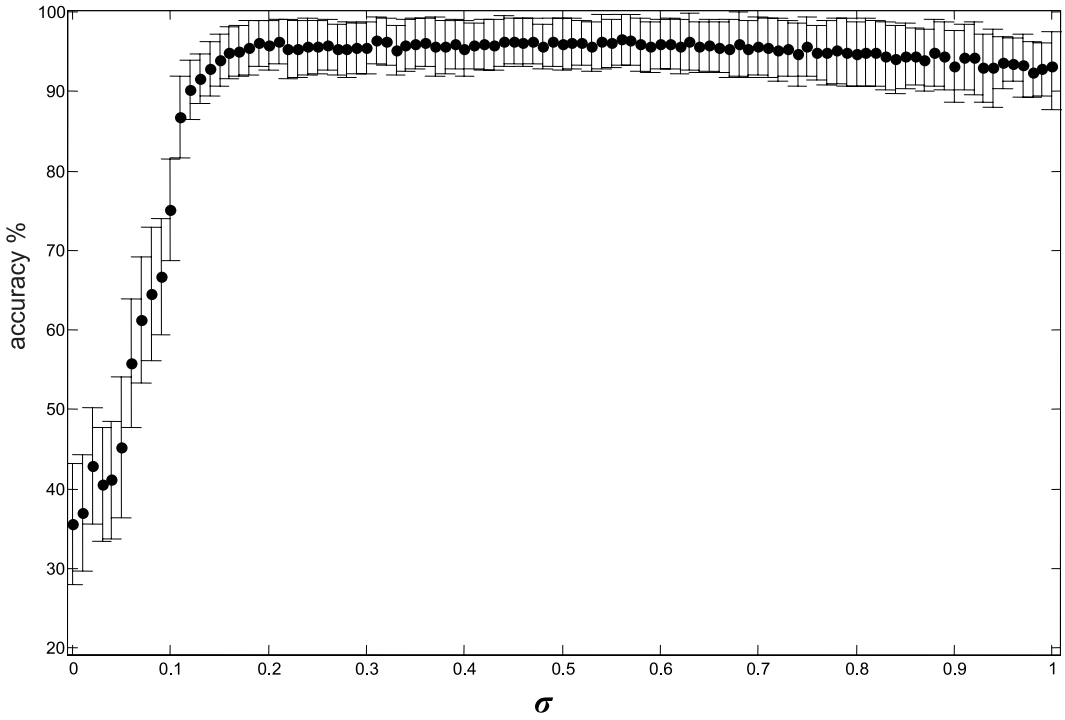

Fig. 3. Accuracy of PNN in terms of σ for the iris data.

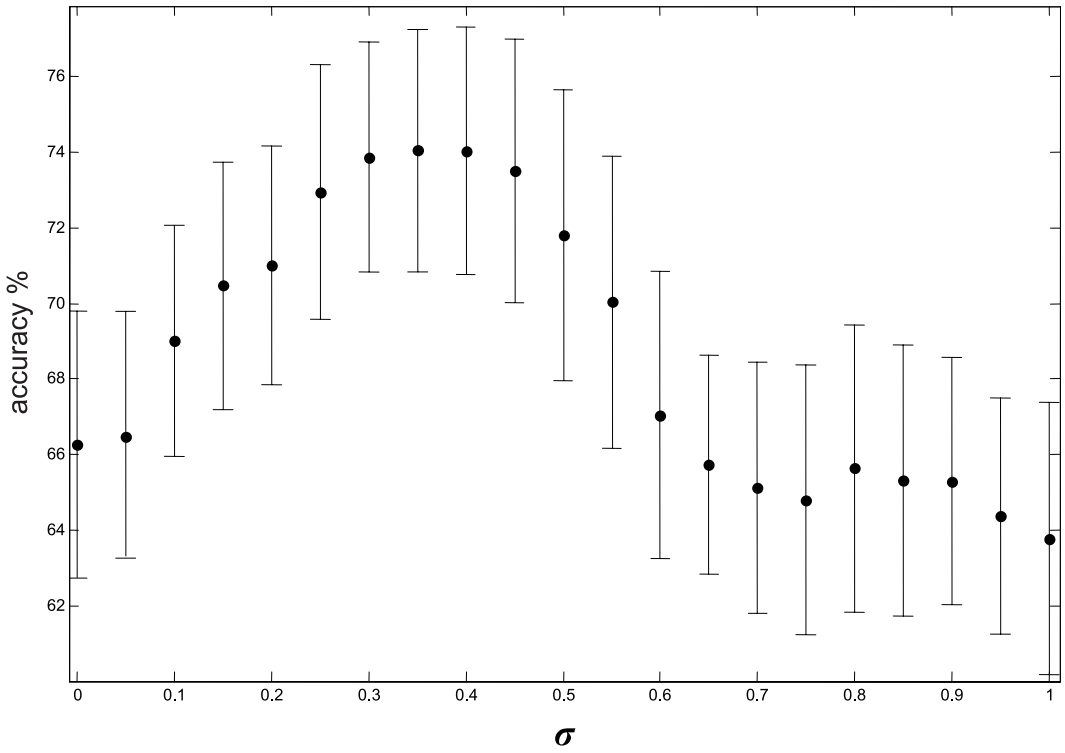

Fig. 4. Accuracy of PNN in terms of σ for diabetic data.

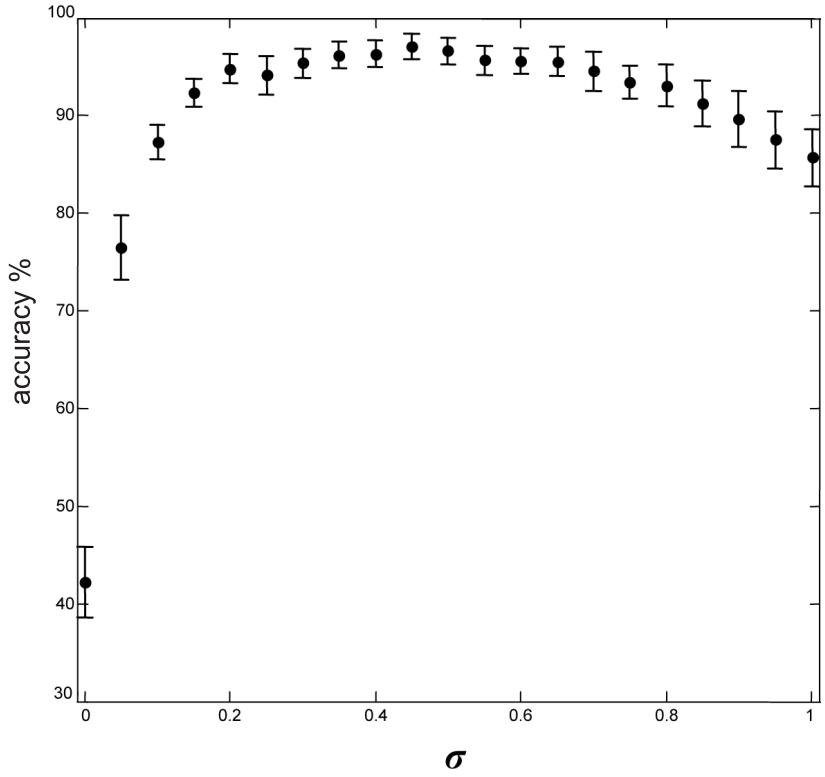

Fig. 5. Accuracy of PNN in terms of σ for breast cancer data.

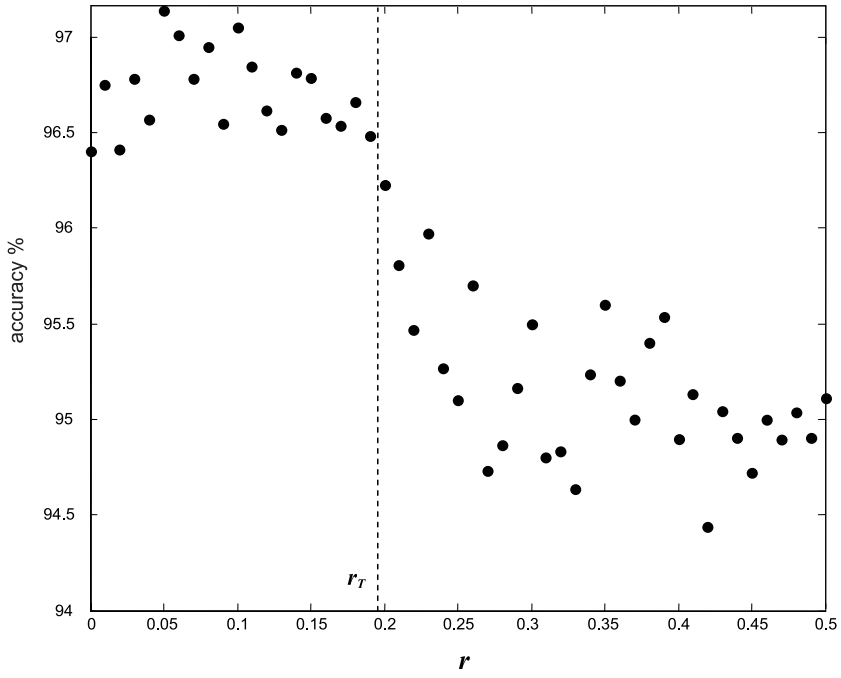

Fig. 6. Accuracy of EPNN in terms of r for the iris data using the best value of σ for PNN.

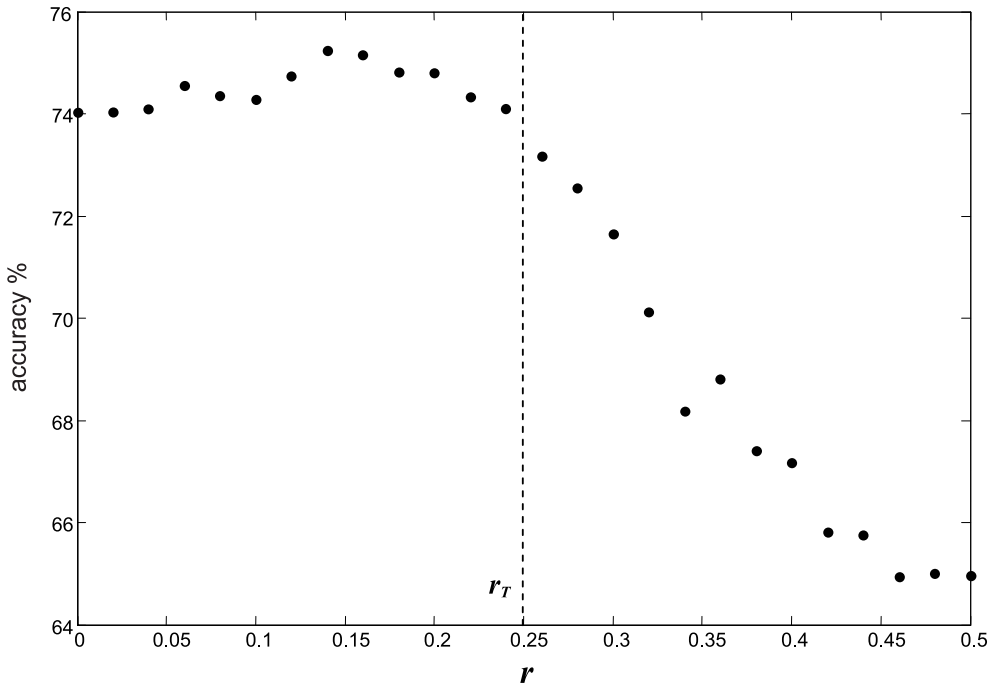

Fig. 7. Accuracy of EPNN in terms of r for the diabetic data using the best value of σ for PNN.

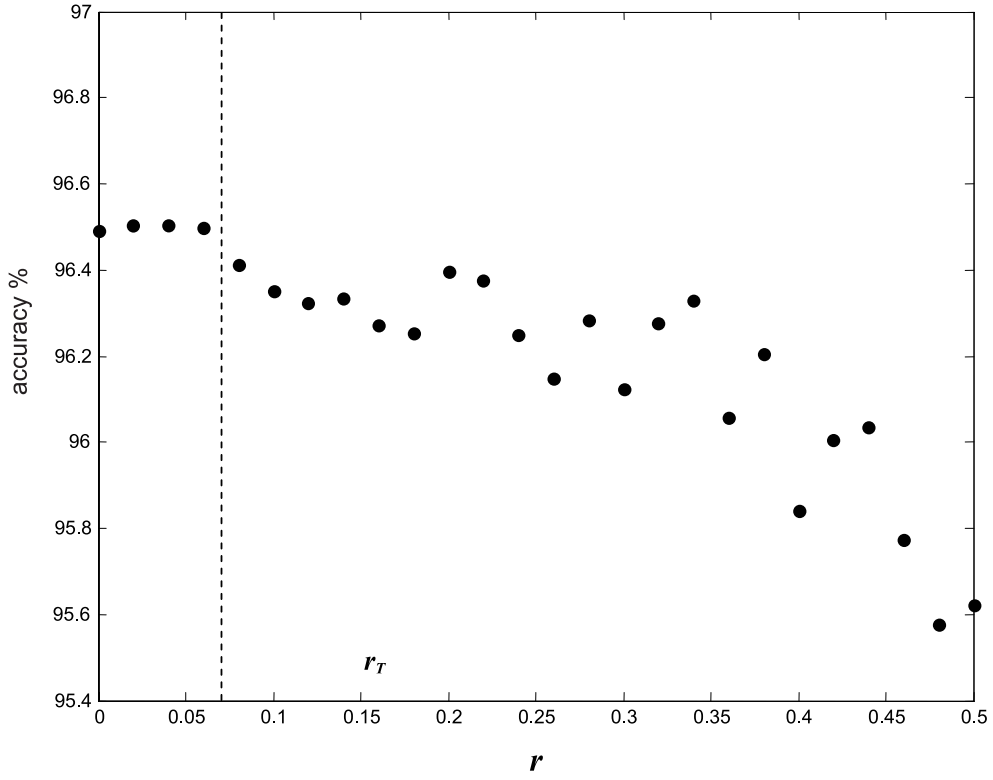

Fig. 8. Accuracy of EPNN in terms of r for the breast cancer data using the best value of σ for PNN.

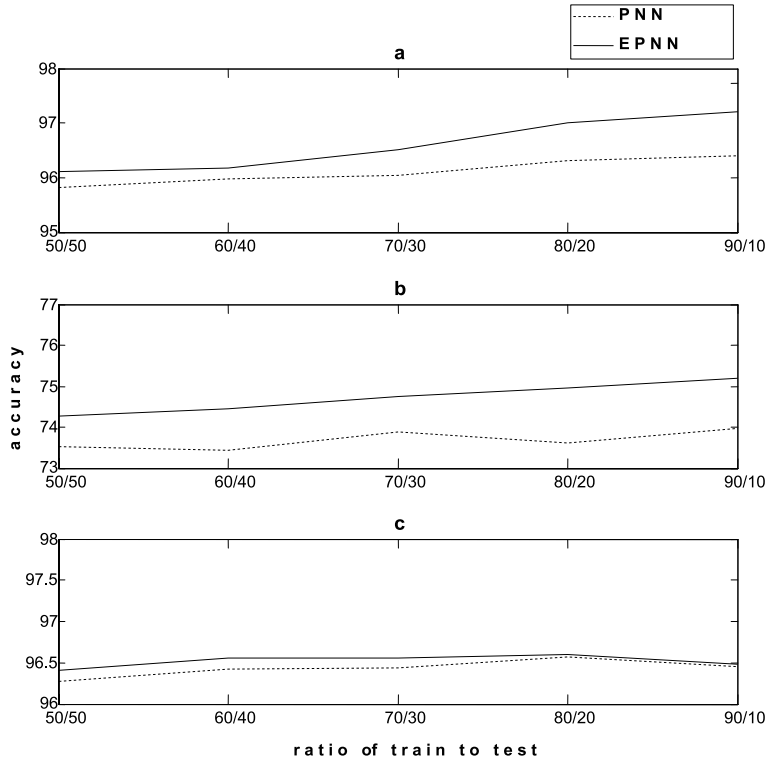

Fig. 9. Comparison of accuracies of PNN with the best value of σ (dotted line) and EPNN with the best value of σ for PNN (solid line) in classification of a) Iris data, b) diabetic data, and c) breast cancer data in terms of the ratio of the number of training data to the number of testing data.

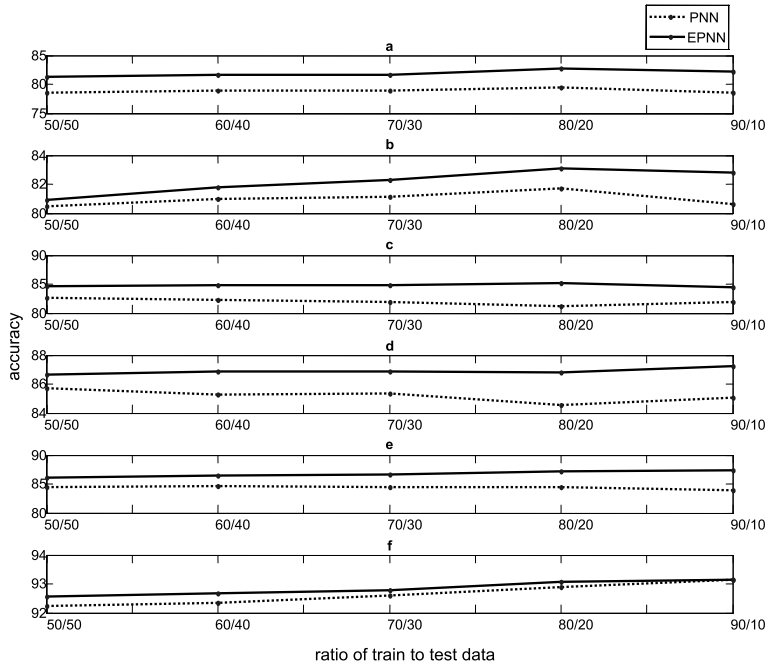

Fig. 10. Accuracies of PNN with the best value of σ (dotted line) and EPNN with the best value of σ for PNN (solid line) in terms of RTT for the iris classification problem using six different SNR values of 1(a), 2(b), 3(c), 4(d), 5(e), and 10(f) dB.

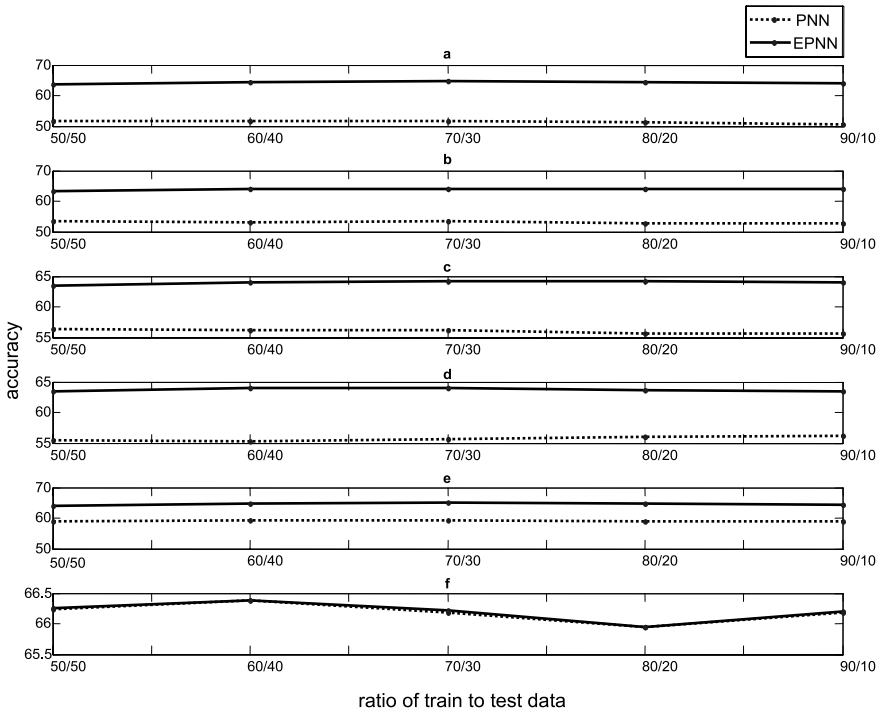

Fig. 11. Accuracies of PNN with the best value of σ (dotted line) and EPNN with the best value of σ for PNN (solid line) in terms of RTT for the diabetic classification problem using six different SNR values of 1(a), 2(b), 3(c), 4(d), 5(e), and 10(f) dB.

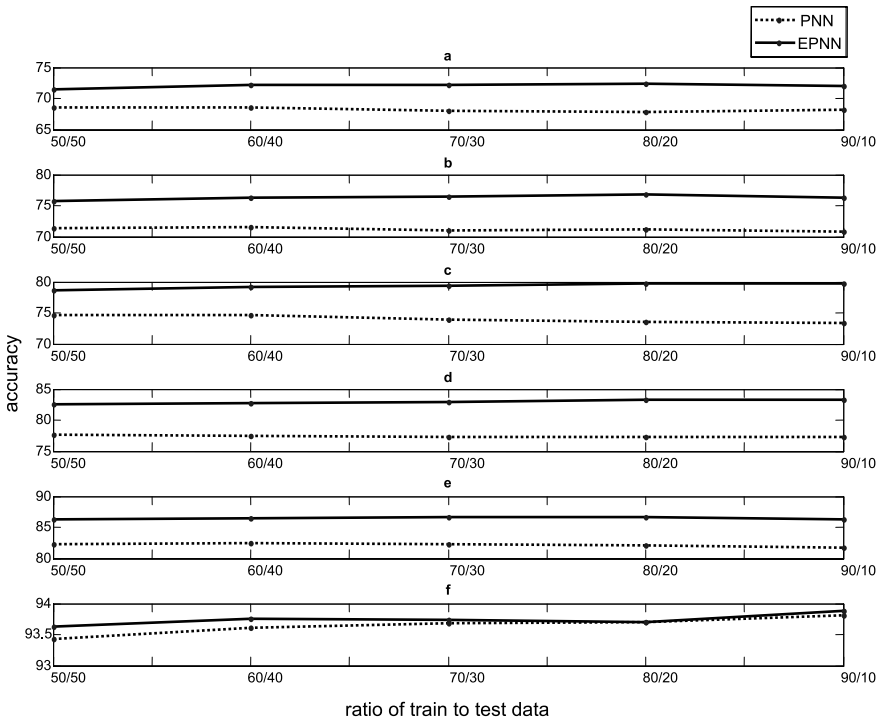

Fig. 12. Accuracies of PNN with the best value of σ (dotted line) and EPNN with the best value of σ for PNN (solid line) in terms of RTT for the breast cancer classification problem using six different SNR values of 1(a), 2(b), 3(c), 4(d), 5(e), and 10(f) dB.

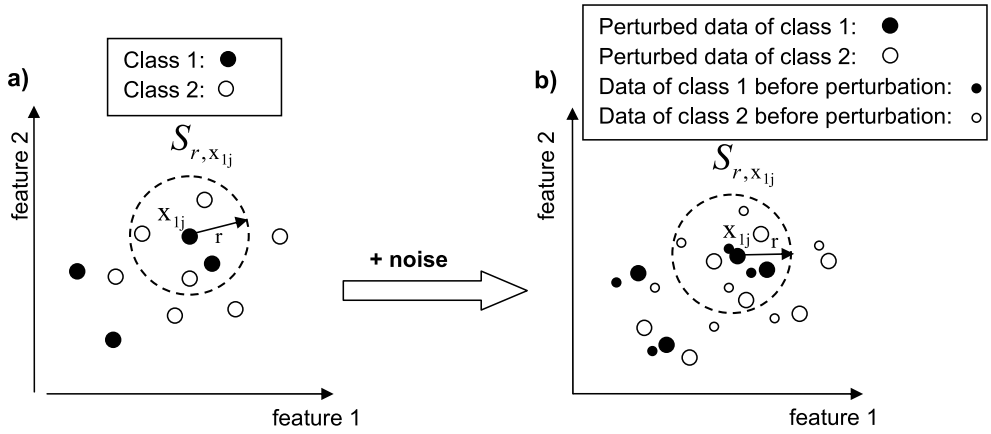

Fig. 13. Illustration of a sample two dimensional data of two classes in two conditions: a) without noise and b) with noise. Number of data of class 1 divided by the number of data in the local decision circle is 2/5 in Fig. a before perturbation and remains the same in Fig. b after perturbation.

5. Conclusion

An enhanced PNN was presented using local decision circles. The performance of EPNN was evaluated using three different benchmark data sets. It was shown that the accuracy of all three classification problems was improved by the proposed method somewhat. But, the major advantage of EPNN is its robustness when a bi-level optimization is performed where first the best value of the spread parameter is obtained followed by the optimum value for the radius of decision circles.

References

[1] H. Adeli and N.-T. Cheng, Augmented Lagrangian Genetic Algorithm for Structural Optimization, *Journal of Aerospace Engineering*, ASCE, **7**(1) (1994), 104–118.

[2] H. Adeli and N.-T. Cheng, Concurrent Genetic Algorithms for Optimization of Large Structures, *Journal of Aerospace Engineering*, ASCE, **7**(3) (1994), 276–296.

[3] H. Adeli and S.L. Hung, A Concurrent Adaptive Conjugate Gradient Learning Algorithm on MIMD Machines, *Journal of Supercomputer Applications*, MIT Press, **7**(2) (1993), 155–166.

[4] H. Adeli and S.L. Hung, An Adaptive Conjugate Gradient Learning Algorithm for Effective Training of Multilayer Neural Networks, *Applied Mathematics and Computation* **62**(1) (1994), 81–102.

[5] H. Adeli and X. Jiang, Neuro-Fuzzy Logic Model for Freeway Work Zone Capacity Estimation, *Journal of Transportation Engineering* **129**(5) (2003), 484–493.

[6] H. Adeli and X. Jiang, Dynamic Fuzzy Wavelet Neural Network Model for Structural System Identification, *Journal of Structural Engineering* **132**(1) (2006), 102–111.

[7] H. Adeli and A. Karim, Neural Network Model for Optimization of Cold-Formed Steel Beams, *Journal of Structural Engineering*, ASCE, **123**(11) (1997), 1535–1543.

[8] H. Adeli and A. Karim, Scheduling/Cost Optimization and Neural Dynamics Model for Construction, *Journal of Construction Management and Engineering*, ASCE, **123**(4) (1997), 450–458.

[9] H. Adeli and A. Karim, Fuzzy-Wavelet RBFNN Model for Freeway Incident Detection, *Journal of Transportation Engineering* **126**(6) (2000), 464–471.

[10] H. Adeli and H. Kim, Cost Optimization of Composite Floors Using the Neural Dynamics Model, *Communications in Numerical Methods in Engineering* **17**(11) (2001), 771–787.

[11] H. Adeli and S. Kumar, Distributed Genetic Algorithms for Structural Optimization, *Journal of Aerospace Engineering* **8**(3) (1995), 156–163.

[12] H. Adeli and S. Kumar, Concurrent Structural Optimization on a Massively Parallel Supercomputer, *Journal of Structural Engineering*, ASCE, **121**(11) (1995), 1588–1597.

[13] H. Adeli and A. Panakkat, A probabilistic neural network for earthquake magnitude prediction, *Neural Networks* **22**(7) (2009), 1018–1024.

[14] H. Adeli and H.S. Park, Counter Propagation Neural Network in Structural Engineering, *Journal of Structural Engineering*, ASCE, **121**(8) (1995), 1205–1212.

[15] H. Adeli and H.S. Park, A Neural Dynamics Model for Structural Optimization-Theory, *Computers and Structures* **57**(3) (1995), 383–390.

[16] H. Adeli and H.S. Park, Optimization of Space Structures by Neural Dynamics, *Neural Networks* **8**(5) (1995), 769–781.

[17] H. Adeli and H.S. Park, Fully Automated Design of Super-highrise Building Structure by a Hybrid AI Model on a Massively Parallel Machine, *AI Magazine* **17**(3) (1996), 87–93.

[18] H. Adeli and A. Samant, An Adaptive Conjugate Gradient Neural Network – Wavelet Model for Traffic Incident Detection, *Computer-Aided Civil and Infrastructure Engineering* **15**(4) (2000), 251–260.

[19] F. Ahmadkhanlou and H. Adeli, Optimum Cost Design of Reinforced Concrete Slabs Using Neural Dynamics Model, *Engineering Applications of Artificial Intelligence* **18**(1) (2005), 65–72.

[20] S. Buchholz and N.L. Bihan, Polarized Signal Classification by Complex and Quaternionic Multilayer Perceptrons, *International Journal of Neural Systems* **18**(2) (2008), 75–85.

[21] L. Carro-Calvo, S. Salcedo-Sanz, R. Gil-Pita, A. Portilla-

Figueras and M. Rosa-Zurera, An Evolutive Multiclass Algorithm for Automatic Classification of High Range Resolution Radar Targets, *Integrated Computer-Aided Engineering* **16**(1) (2009), 51–60.
[22] V.S. Chakravarthy, N. Gupte, S. Yogesh and A. Salhotra, Chaotic Synchronization using a Network of Neural Oscillators, *International Journal of Neural Systems* **18**(2) (2008), 157–164.
[23] T.M. Cheng and R.Z. Yan, Integrating Messy Genetic Algorithms and Simulation to Optimize Resource Utilization, *Computer-Aided Civil and Infrastructure Engineering* **24**(6) (2009), 401–415.
[24] A. Dharia and H. Adeli, Neural Network Model for Rapid Forecasting of Freeway Link Travel Time, *Engineering Applications of Artificial Intelligence* **16**(7–8) (2003), 607–613.
[25] L. Dridi, M. Parizeau, A. Mailhot and J.P. Villeneuve, Using Evolutionary Optimisation Techniques for Scheduling Water Pipe Renewal Considering a Short Planning Horizon, *Computer-Aided Civil and Infrastructure Engineering* **23**(8) (2008), 625–635.
[26] H.M. Elragal, Improving Neural Networks Prediction accuracy Using Particle Swarm Optimization Combiner, *International Journal of Neural Systems* **19**(5) (2009), 387–393.
[27] S. Fiori, Learning by Criterion Optimization on a Unitary Unimodular Matrix Group, *International Journal of Neural Systems* **18**(2) (2008), 87–103.
[28] D. Gerbec, S. Gasperic, I. Smon and F. Gubina, Allocation of the Load Profiles to Consumers Using Probabilistic Neural Networks, *IEEE Transactions on Power Systems* **20**(2) (2005), 548–555.
[29] S. Ghosh-Dastidar and H. Adeli, Wavelet-Clustering-Neural Network Model for Freeway Incident Detection, *Computer-Aided Civil and Infrastructure Engineering* **18**(5) (2003), 325–338.
[30] S. Ghosh-Dastidar and H. Adeli, Improved Spiking Neural Networks for EEG Classification and Epilepsy and Seizure Detection, *Integrated Computer-Aided Engineering* **14**(3) (2007), 187–212.
[31] S. Ghosh-Dastidar and H. Adeli, Spiking Neural Networks, *International Journal of Neural Systems* **19**(4) (2009), 295–308.
[32] S. Ghosh-Dastidar and H. Adeli, A New Supervised Learning Algorithm for Multiple Spiking Neural Networks with Application in Epilepsy and Seizure Detection, *Neural Networks* **22**(10) (2009), 1419–1431.
[33] S. Ghosh-Dastidar, H. Adeli and N. Dadmehr, Principal Component Analysis-Enhanced Cosine Radial Basis Function Neural Network for Robust Epilepsy and Seizure Detection, *IEEE Transactions on Biomedical Engineering* **55**(2) (2008), 512–518.
[34] A.M.A. Haidar, A. Mohamed, M. Al-Dabbagh, A. Aini Hussain and M. Masoum, An Intelligent Load Shedding Scheme Using Neural networks & Neuro-Fuzzy, *International Journal of Neural Systems* **19**(6) (2009), 473–479.
[35] T. Hoya, On the Capability of Accommodating New Classeswithin Probabilistic Neural Networks, *IEEE Transactions on Neural Networks* **14**(2) (2003), 450–453.
[36] S.L. Hung and H. Adeli, Parallel Backpropagation Learning Algorithms on Cray Y-MP8/864 Supercomputer, *Neurocomputing* **5**(6) (1993), 287–302.
[37] S.L. Hung and H. Adeli, A Parallel Genetic/Neural Network Learning Algorithm for MIMD Shared Memory Machines, *IEEE Transactions on Neural Networks* **5**(6) (1994), 900–909.
[38] T. Isokawa, H. Nishimura, N. Kamiura and N. Matsui, Associative memory in quaternionic hopfield neural network, *International Journal of Neural Systems* **18**(2) (2008), 135–145.
[39] X. Jiang and H. Adeli, Dynamic Fuzzy Wavelet Neuroemulator for Nonlinear Control of Irregular Highrise Building Structures, *International Journal for Numerical Methods in Engineering* **74**(7) (2008), 1045–1066.
[40] X. Jiang and H. Adeli, Neuro-Genetic Algorithm for Nonlinear Active Control of Highrise Buildings, *International Journal for Numerical Methods in Engineering* **75**(8) (2008), 770–786.
[41] Y. Jiang, T. Hu, C. Huang and X. Wu, An improved particle swarm optimization algorithm, *Applied Mathematics and Computation* **193**(1) (2007), 231–239.
[42] M.W. Kang, P. Schonfeld and N. Yang, Prescreening and Repairing in a Genetic Algorithm for Highway Alignment Optimization, *Computer-Aided Civil and Infrastructure Engineering* **24**(2) (2009), 109–119.
[43] A. Karim and H. Adeli, CONSCOM: An OO Construction Scheduling and Change Management System, *Journal of Construction Engineering and Management* **125**(5) (1999), 368–376.
[44] A. Karim and H. Adeli, Comparison of the Fuzzy – Wavelet RBFNN Freeway Incident Detection Model with the California Algorithm, *Journal of Transportation Engineering* **128**(1) (2002), 21–30.
[45] A. Karim and H. Adeli, Radial Basis Function Neural Network for Work Zone Capacity and Queue Estimation, *Journal of Transportation Engineering* **129**(5) (2003), 494–503.
[46] A. Kaveh and S. Shojaee, Optimal Design of Scissor-Link Foldable Structures Using Ant Colony Optimization Algorithm, *Computer-Aided Civil and Infrastructure Engineering* **22**(1) (2007), 56–64.
[47] S. Kawata and A. Hirose, Frequency-Multiplexing Ability of Complex-Valued Hebbian Learning in Logic Gates, *International Journal of Neural Systems* **18**(2) (2008), 173–184.
[48] A. Khashman, Blood Cell Identification Using a Simple Neural Network, *International Journal of Neural Systems* **18**(5) (2008), 453–458.
[49] H. Kim and H. Adeli, Discrete Cost Optimization of Composite Floors using a Floating Point Genetic Algorithm, *Engineering Optimization* **33**(4) (2001), 485–501.
[50] D. Kim, D.H. Kim and S. Chang, Application of probabilistic neural network to design breakwater armor blocks, *Ocean Engineering* **35**(3–4) (2008), 294–300.
[51] M. Kobayashi, Pseudo-relaxation learning algorithm for complex valued associative memory, *International Journal of Neural Systems* **18**(2) (2008), 147–156.
[52] H. Lee, E. Kim and M. Park, A genetic feature weighting scheme for pattern recognition, *Integrated Computer-Aided Engineering* **14**(2) (2007), 161–171.
[53] E. Lopez-Rubio and J.M. Ortiz-de-Lazcano-Lobato, Dynamic Competitive Probabilistic Principal Components Analysis, *International Journal of Neural Systems* **19**(2) (2009), 91–103.
[54] D.P. Mandic, P. Vayanos, M. Chen and S.L. Goh, Online Detection of the Modality of Complex Valued Real World Signals, *International Journal of Neural Systems* **18**(2) (2008), 67–74.
[55] K.Z. Mao, K.C. Tan and W. Ser, Probabilistic Neural-Network Structure Determination for Pattern Classification, *IEEE Transactions on Neural Networks* **11**(4) (2000), 1009–1016.
[56] S. Mathakari, P. Gardoni, P. Agarwal, A. Raich and T. Haukaas, Reliability-based Optimal Design of Electrical Transmission Towers Using Multi-objective Genetic Algorithms, *Computer-Aided Civil and Infrastructure Engineering*

22(4) (2007), 282–292.
[57] S. Mishra, C.N. Bhende and B.K. Panigrahi, Detection and Classification of Power Quality Disturbances Using S-Transform and Probabilistic Neural Network, *IEEE Transactions on Power Delivery* **23**(1) (2008), 280–287.
[58] A.M.S. Muniz, H. Liu, K.E. Lyons, R. Pahwa, W. Liu, F.F. Nobre and J. Nadal, Comparison among probabilistic neural network, support vector machine and logistic regression for evaluating the effect of subthalamic stimulation in Parkinson disease on ground reaction force during gait, *Journal of Biomechanics* **43**(4) (2010), 720–726.
[59] T. Nitta, The Uniqueness Theorem for Complex-Valued Neural Networks with Threshold Parameters and the Redundancy of the Parameters, *International Journal of Neural Systems* **18**(2) (2008), 123–134.
[60] A. Panakkat and H. Adeli, Recurrent Neural Network for Approximate Earthquake Time and Location Prediction Using Multiple Seismicity Indicators, *Computer-Aided Civil and Infrastructure Engineering* **24**(4) (2009), 280–292.
[61] A. Pande and M. Abdel-Aty, A Computing Approach Using Probabilistic Neural Networks for Instantaneous Appraisal of Rear-End Crash Risk, *Computer-Aided Civil and Infrastructure Engineering* **23**(7) (2008), 549–559.
[62] H.S. Park and H. Adeli, A Neural Dynamics Model for Structural Optimization-Application to Plastic Design of Structures, *Computers and Structures* **57**(3) (1995), 391–399.
[63] H.S. Park and H. Adeli, Distributed Neural Dynamics Algorithms for Optimization of Large Steel Structures, *Journal of Structural Engineering*, ASCE, **123**(7) (1997), 880–888.
[64] G. Puscasu and B. Codres, Nonlinear System Identification Based on Internal Recurrent Neural Networks, *International Journal of Neural Systems* **19**(2) (2009), 115–125.
[65] J.J. Quan, X.B. Wen and X.Q. Xu, Multiscale probabilistic neural network method for SAR image segmentation, *Applied Mathematics and Computation* **205**(2) (2008), 578–583.
[66] V.S.H. Rao and G.R. Murthy, Global Dynamics of a Class of Complex Valued Neural Networks, *International Journal of Neural Systems* **18**(2) (2008), 165–171.
[67] G.G. Rigatos and S.G. Tzafestas, *Neurodynamics and Attractors in Quantum Associative Memories* **14**(3) (2007), 224–242.
[68] A. Rodriguez and J. Reggia, A Distributed Learning Algorithm for Particle Systems, *Integrated Computer-Aided Engineering* **16**(1) (2009), 1–20.
[69] J.L. Rossello, V. Canals, A. Morro and J. Verd, Chaos-based Mixed Signal Implementation of Spiking Neurons, *International Journal of Neural Systems* **19**(6) (2009), 465–471.
[70] D.E. Rumelhart, D.E. Hinton and R.J. Williams, Learning internal representation by error propagation, in: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, D.E. Rumelhart, J.L. McClelland and the PDP Research Group, eds, MIT Press, Cambridge, MA, 1986, pp. 318–362.
[71] C. Sabourin, K. Madani and O. Bruneau, Autonomous Biped Gait Pattern based on Fuzzy-CMAC Neural Networks, *Integrated Computer-Aided Engineering* **14**(2) (2007), 173–186.
[72] A. Samant and H. Adeli, Enhancing Neural Network Incident Detection Algorithms using Wavelets, *Computer-Aided Civil and Infrastructure Engineering* **16**(4) (2001), 239–245.
[73] S.R. Samantaray, B.K. Panigrahi and P.K. Dash, High impedance fault detection in power distributionnetworks using time–frequency transform andprobabilistic neural network, *IET Generation, Transmission & Distribution* **2**(2) (2008), 261–270.
[74] K. Sarma and H. Adeli, Fuzzy Genetic Algorithm for Optimization of Steel Structures, *Journal of Structural Engineering*, ASCE, **126**(5) (2000), 596–604.
[75] K. Sarma and H. Adeli, Fuzzy Discrete Multicriteria Cost Optimization of Steel Structures, *Journal of Structural Engineering*, ASCE, **126**(11) (2000), 1339–1347.
[76] K.C. Sarma and H. Adeli, Bi-Level Parallel Genetic Algorithms for Optimization of Large Steel Structures, *Computer-Aided Civil and Infrastructure Engineering* **16**(5) (2001), 295–304.
[77] K.C. Sarma and H. Adeli, Life-Cycle Cost Optimization of Steel Structures, *International Journal for Numerical Methods in Engineering* **55**(12) (2002), 1451–1462.
[78] R. Savitha, S. Suresh and N. Sundararajan, A Fully Complex-valued Radial Basis Function Network and its Learning Algorithm, *International Journal of Neural Systems* **19**(4) (2009), 253–267.
[79] A.B. Senouci and H. Adeli, Resource Scheduling using Neural Dynamics Model of Adeli and Park, *Journal of Construction Engineering and Management*, ASCE, **127**(1) (2001), 28–34.
[80] G. Sirca and H. Adeli, Neural Network Model for Uplift Load Capacity of Metal Roof Panels, *Journal of Structural Engineering* **127**(11) (2001), 1276–1285.
[81] J.F. Smith, H. ThanhVu and T.H. Nguyen, Autonomous and cooperative robotic behavior based on fuzzy logic and genetic programming, *Integrated Computer-Aided Engineering* **14**(2) (2007), 141–159.
[82] T. Song, M.M. Jamshidi, R.R. Lee and M. Huang, A Modified Probabilistic Neural Network for Partial Volume Segmentation in Brain MR Image, *IEEE Transactions on Neural Networks* **18**(5) (2007), 1424–1432.
[83] D.F. Specht, Probabilistic neural networks, *Neural Networks* **3**(1) (1990), 109–118.
[84] A.R. Tashakori and H. Adeli, Optimum Design of Cold-Formed Steel Space Structures Using Neural Dynamic Model, *Journal of Constructional Steel Research* **58**(12) (1990), 1545–1566.
[85] F. Teklu, A. Sumalee and D. Watling, A Genetic Algorithm for Optimizing Traffic Control Signals Considering Routing, *Computer-Aided Civil and Infrastructure Engineering* **22**(1) (2007), 31–43.
[86] Torfs, P. and Wójcik, R. (2001), Local probabilistic neural networks in hydrology, *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* **26**(1), 9–14.
[87] M. Tripathy, R.P. Maheshwari and H.K. Verma, Power Transformer Differential Protection Based on Optimal Probabilistic Neural Network, *IEEE Transactions on Power Delivery* **25**(1) (2010), 102–112.
[88] E.D. Ubeyli, Probabilistic neural networks employing Lyapunov exponents for analysis of Doppler ultrasound signals, *Computers in Biology and Medicine* **38**(1) (2008), 82–89.
[89] E.D. Ubeyli, Implementing eigenvector methods/probabilistic neural networks for analysis of EEG signals, *Neural Networks* **21**(9) (2008), 1410–1417.
[90] E.D. Ubeyli, Lyapunov exponents/probabilistic neural networks for analysis of EEG signals, *Expert Systems with Applications* **37**(2) (2010), 985–992.
[91] D. Vigliano, R. Parisi and A. Uncini, Flexible Nonlinear blind signal separation in the complex domain, *International Journal of Neural Systems* **18**(2) (2008), 105–122.
[92] B.J. Vitins and K.W. Axhausen, Optimization of Large Transport Networks Using the Ant Colony Heuristic, *Computer-Aided Civil and Infrastructure Engineering* **24**(1) (2009), 1–14.

[93] P. Wasserman, *Advanced Methods in Neural Networks*, Van Nostrand Reinhold, New York, NY, 1993.
[94] W.H. Wolberg and O.L. Mangasarian, Multisurface method of pattern separation for medical diagnosis applied to breast cytology, *Proceedings of the National Academy of Sciences of the United States of America* **87**(23) (1990), 9193–9196.
[95] Z. Yang, B. Yu and C. Cheng, A Parallel Ant Colony Algorithm for Bus Network Optimization, *Computer-Aided Civil and Infrastructure Engineering* **22**(1) (2007), 44–55.
[96] S.N. Yu and Y.H. Chen, Electrocardiogram beat classification based on wavelet transformation and probabilistic neural network, *Pattern Recognition Letters* **28**(10) (2007), 1142–1150.
[97] J.A. Zeferino, A.P. Antunes and M.C. Cunha, An Efficient Simulated Annealing Algorithm for Regional Wastewater Systems Planning, *Computer-Aided Civil and Infrastructure Engineering* **24**(5) (2009), 359–370.
[98] W. Zou, Z. Chi and K.C. Lo, Improvement of Image Classification Using Wavelet Coefficients with Structured-based Neural Network, *International Journal of Neural Systems* **18**(3) (2008), 195–205.