Performance Improvement of Ann Classifiers using Pso

Ms.Jayshri D.Dhande,
B. D. college of Engg,
Sevagram

D.R.Dandekar,
B. D. college of Engg,
Sevagram

S.L.Badjate
S.B. Jain Institute of
Technology, Management
& Research, Nagpur

ABSTRACT

Data classification has been studied widely in the field of Artificial Intelligence, Machine Learning, data mining, and pattern recognition. Up to the present, the development of classification has made great achievements, and many kinds of classified technology and theory will continue to emerge. The aim of this paper is two fold. First, we present the experimental study of different Artificial Neural Networks classifiers for classification of radar returns from Ionosphere dataset and Bupa liver disorder dataset. Second, we propose a novel classification system based on particle swarm optimization (PSO) to improve the generalization performance of the SVM classifier. For this purpose, we have optimized the kernel parameters of SVM classifier. The experiments were conducted on Jonhs Hopkins Ionosphere dataset and Bupa Liver Disorder dataset. The comparison of different Neural Network classifiers and PSO-SVM is done based on Ionosphere dataset and Bupa Liver Disorder dataset from UCI machine learning repository. The results show that RBFNN typically provide better classification results. When comparing to techniques applied to binary classification problems. Also SVM Classifier with RBF kernel gives best classification accuracy on training set. And PSO-SVM classifier with optimized kernel parameter selection for classification of radar returns from ionosphere dataset and Bupa Liver Disorder gives better accuracy and improves the generalization performance.

Keywords

Artificial Neural Network, BPNN, RBFNN, SVM, and Particle swarm Optimization (PSO).

1.INTRODUCTION

Data classification is based on classification of the target mode. The target model can find a common feature and classify as different categories. 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. Classification has been studied widely in the fields of Artificial Intelligence, Machine Learning [1], Data Mining and Pattern Recognition [2].This also known as the regression issue. There are three methods of data classification, one based on statistical methods, such as Bayesian networks, K-Nearest Neighbor, Support Vector Machine [3], the regression model etc, and the second rule-based methods. Such as decision trees, the association rules and rough set theory etc. and the third based on the connection methods, such as Artificial Neural Networks. In the face of a lot of noisy, clutter, nonlinear data, Artificial Neural Networks (ANN) can not only help make high quality modeling and complete the training in the process of the use of large amounts of data, but also can have a test mode set to assess the performance of ANN. These complementary relationships make NN technology a good prospect. Since ANN access to users of knowledge difficult to understand, the model lacks transparency for its users, like a “black box”[4].The ultimate goal of data classification is to provide a neural network to expositive mechanisms, with the right of matrix to replace the rule , so it can solve the problem of the “black box”. It provides for the practical application of the decision-making based on a higher credibility.

To solve the classification problems, many classification techniques have been proposed and some of the successful techniques are Artificial Neural Networks (ANN), Support vector machine (SVM). Optimal design of classifier is investigated using multilayer perceptron neural network (MLPNN) trained with error back propagation algorithm on this database [6].Using the first 200 instances for training which were carefully split almost 50% positive & 50% negative (equiprobable), MLP NN trained with standard backpropgation algorithm attained an average of about 96% accuracy. Other researcher’s attempts modular ANNs for efficient radar target classification [7]. Radial basis function neural networks are used to classify real life audio radar signals [8]. In [9] paper author describes the aim of this paper in twofold way. First, present a through experimental study to show the superiority of the generalization capability of the SVM approach in automatic classification of ECG beats. Second, propose a novel classification system based on PSO to improve the generalization performance of SVM classifier. For this purpose they have optimized the SVM classifier design by searching for the best value of the parameters that tune its discriminant function and upstream by looking for the best subset of features that feed the classifier. The experiments were conducted on the basis of ECG data from the MIT-BIH arrhythmia database to classify five kinds of abnormal wave forms and normal beats. In particular, they were organized so as to test the sensitivity of the SVM classifier and that of two reference classifier used for comparison i.e. k-nearest neighbor classifier and RBF neural network classifier. The obtained results clearly confirm classifier, and suggest that further substantial improvements in terms of classification accuracy can be achieved by the proposed PSO-SVM classification system. Genetic algorithm (GA) and particle swarm optimization (PSO) techniques have attracted considerable attention among heuristic optimization techniques. The Application of these algorithms for feature selection in retinal image classifications explored [10].

2.ARTIFICIAL NEURAL NETWORK CLASSIFICATION TECHNIQUES

2.1 BP Neural Network Classifier

It is shown that from the literature review a BPNN having single layer of neurons could classify a set of points perfectly if they were linearly separable. BPNN having three layers of weights can generated arbitrary decision regions which may be non convex and disjoint. BPNN is based on processing elements, which compute a nonlinear function of the scalar product of the input vector and a weight vector. Its configuration is determined by the number of hidden layers, numbers of the neurons in each of the hidden layers as well as the type of the activation function used for the neurons. Train Levenberg Marquart algorithm is used for determined the connection weights from the samples. The BPNN structure is evaluated on training set and test set. The test data are then used to accesses how well the network has generalized. Figure 1 shows structure of BPNN classifier.

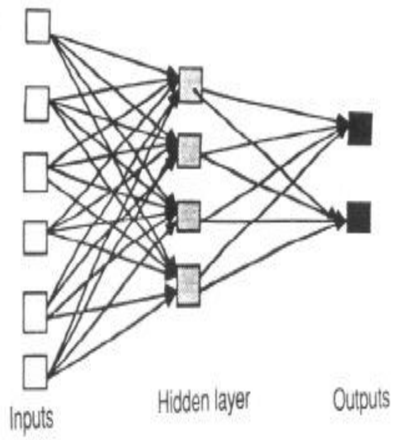

Figure.1 BPNN Classifier

1.2 RBF Neural Network

A radial basis function (RBF) is a real valued function whose value depends only on the distance from the origin, so that $\Phi(x) = \Phi(\ \|x\|)$; or alternatively on the distance from some other point C, called a center. RBF NN is a nearest neighbor classifier. It uses Gaussian transfer function having radial symmetry. The centers coefficient vector W=[W1,W2,W3,…,Wn], f(x) being a real valued vector and x=[x1,x2,x3,..xn] implements the input-output map of the RBFNN. Any arbitrary continuous function can be approximated with an RBFNN if localized Gaussian is placed to cover the space, and the width of each Gaussian is controlled the amplitude of each Gaussian is set. Figure 2 shows RBF NN architecture.

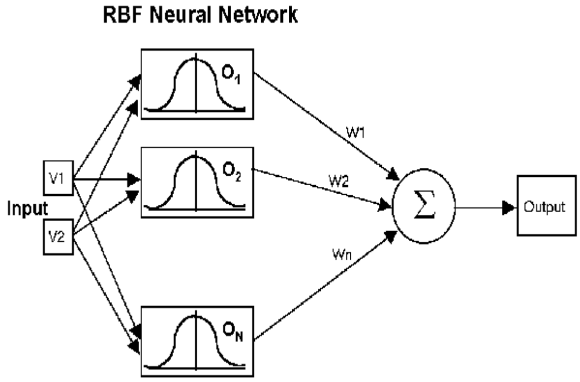

Fig. 2 RBFNN Classifier

1.3 SVM Classifier

Support Vector Machine (SVM) is a relatively new classifier based on strong foundation from the broad area of statistical learning theory. This class of algorithms proposed by Vapnik [11] has shown good performance over real applications and proposed initially for classification problems of two classes. SVM use geometrical properties to exactly calculate the optimal separating hyper plane directly from the training data. They also introduce methods to deal with non linearly separable cases, ie where no separating straight line can be found as well as with cases in which there is noise and /or outliers in the training data, ie some of the training samples may be wrong.

Basically, the SVM is a linear machine working in the highly dimensional features space formed by the nonlinear mapping of the n-dimensional input vector x into a K-dimensional features space ($k > n$) through the use of a mapping $\Phi(x)$.The following relation gives the equation of hyper plane separating two different classes $Y(x) = W^T \Phi(x)$, where w is the weight vector of the network. The condition $Y(x) > 0$ means one class and $Y(x) < 0$ means the opposite one.

The most distinctive fact about SVM is that the learning task is reduced to quadratic programming by introducing Lagrange multipliers. All operation in learning and testing modes are done in SVM using kernel functions. The kernel is defined as $k(x, xi) = \Phi^T (xi)\ \Phi(x)$.The best known kernels are linear , polynomial, radial basis functions and sigmoid functions. The problem of learning SVM, formulated as the task of separating learning vectors x, into two classes of the destination values either $di=1$

or *di=-1* with maximal separation margin is reduced to the dual maximization problems of the objective function. '*c*' is the regularizing parameter and determines the balance between the maximization of the margin and minimization of the classification error. The solution with respect to Lagrange multiplier gives the optimal weight vector.

In the present study, RBF function is used as kernel and the kernel parameters *γ* and *c*, which provide the best classification. The learning of support vector referred to as the separation of learning vector *xi* in two classes of designation values either *di=1* or *di=-1* with maximal separation margin. Figure 3 shows architecture of Support Vector Machine classifier.

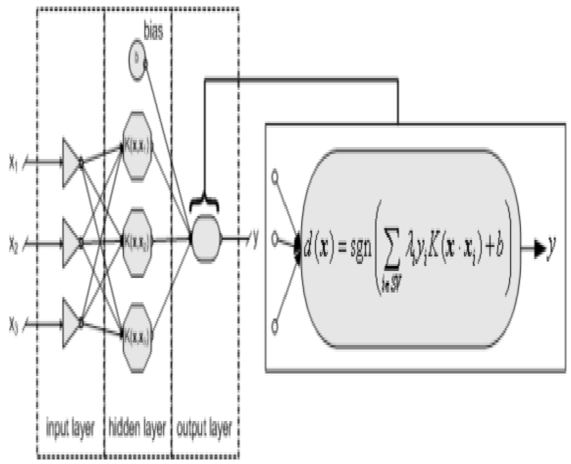

Fig. 3 SVM classifier

3. OPTIMIZATION TECHNIQUES

3.1 Genetic Algorithm

Genetic Algorithm can be viewed as general purpose search method or an optimization method based on biological evolution. Genetic algorithm is similar to the natural evolution, it maintains and evolves a population of chromosomes as potential solution to an optimization problem. A new population of chromosomes is reproduced by applying genetic operators such as mutation and crossover on parent chromosomes in a reproduction step. The new chromosomes are designed fitness values based on an objective function. After that the population undergoes to mate and reproduce. GA has been applied effectively to solve a wide spectrum of optimization problems. In the genetic algorithm, the chromosomes with high fitness have more opportunities to pass to the next generation, participate in genetic manipulation. By this way, the chromosomes are moving towards to the best solution.

A.Particle Swarm Optimization

Particle Swarm Optimization (PSO) is a new swarm intelligent algorithm after GA. Particle swarm optimization is a kind of evolutionary computation techniques motivated by the simulation of social behavior. PSO was developed by Kennedy and Eberhart [12]. It is an iterative optimization similar to genetic algorithm. PSO analogy prey behavior of birds, Such a scenario: a group of bird search food at random .In this area, there is only one food, however all birds don't know where the food is, but know the distance to the food. Then what is optimal strategy to find food? Currently the simplest and most effective method is to search the food from this model to solve this kind problem. Each optimization is to search a bird in space which is called as 'particle'. All particles have fitness value determined by optimization function, every particle also have one velocity to determine direction and distance. Then particles follow the optimization in search space as one random particle (random solution).Particles update themselves by tracking two "extreme" particles. The first is the optimization solution found by particles. This kind of solution is called as Pbest, the other is the optimization found by species. This kind of extremum is called as global extremum. Particles update themselves by equation 1 & 2 to find their own velocity & location.

v= *v*+*c1**rand ()*(pbest-present) + *c2**rand()*(gbest-present) (1)

Present= present + *v* (2)

Where v is the particles velocity, present is the particles position currently rand () is random number among (0, 1).*c1*, *c2* as learning factor. Commonly *c1*=*c2*=2.The velocity in any dimension is limited in maximum velocity exceeds vmax, and then the velocity is vmax.

Data sets and Experiment Design

For the experiments, data sets were obtained from the University Of California at Irvine (UCI) machine learning repository and are Ionosphere dataset and Bupa liver disorder dataset. The detailed information is shown in Table 1.

Table 1: Information about the Experimental datasets

Dataset	Attributes	Size	Training set	Testing set
Ionosphere	34	351	200	151
Bupa liver Disorder	06	345	200	145

The back propagation neural network classifier is designed with trainlm function, one hidden layer with 12 hidden neurons, input layer and output layer for both datasets. RBFNN classifier designed with spread constant 4 of Gaussian function. SVM classifier designed with number of support vector (*nvs*), rbf kernel having spread (*g*=1), penalty constant(*c*=1).The proposed PSO based SVM classifier design with C_1, C_2 =1.3, number of particles (*np*=5), Dim size equal to 4 and *C* and gamma equal to 1and 3 respectively. The performance of classifier is measure in term of its classification accuracy for training and testing set. Also observe the generalization ability of designed classifier for unseen data.

EXPERIMENTAL RESULTS

In this section, experimental results regarding the evaluation process of the developed classifiers are presented .In order to compare the performance of neural network techniques .Firstly dataset is splits into training set and testing set. In the experiment, MATLAB is used to design and test all classifiers. Table 2 shows the performance results obtained by different classifiers and proposed PSO-SVM classifier on both dataset.

Table 2: Experimental performance result

Designed Classifiers	Ionosphere Dataset		Bupa Liver Disorder	
	Training set Accuracy (%)	Testing set Accuracy (%)	Training set Accuracy (%)	Testing set Accuracy (%)
BPNN	84	81.75	73	63
RBFNN	100	82.78	100	75
SVM	100	68.21	100	78
PSO-SVM	100	82.78	100	84

CONCLUSION

We have implemented BPNN, RBFNN, and SVM and proposed PSO based SVM classifiers for classification. We explored various other implementations such as SVM with different kernels functions and empirically selected this particular network since it achieved the best performance. This paper presents the comparison of three neural network classifier BPNN, RRFNN, and PSO-SVM for classification of ionosphere dataset and Bupa liver disorder. Each neural network techniques selected for this comparison has different structures and different advantages and disadvantages. While RBFNN and SVM have simpler architectures and they can train data faster than BPNN. For parameter optimization of classifiers the particle swarm optimization techniques is used. In terms of performance comparison based on the classification accuracy as shown in figure 4., we found that the results achieved by RBFNN and PSO-SVM are higher than other techniques with best generalization for ionosphere dataset but not for Bupa liver disorder dataset.

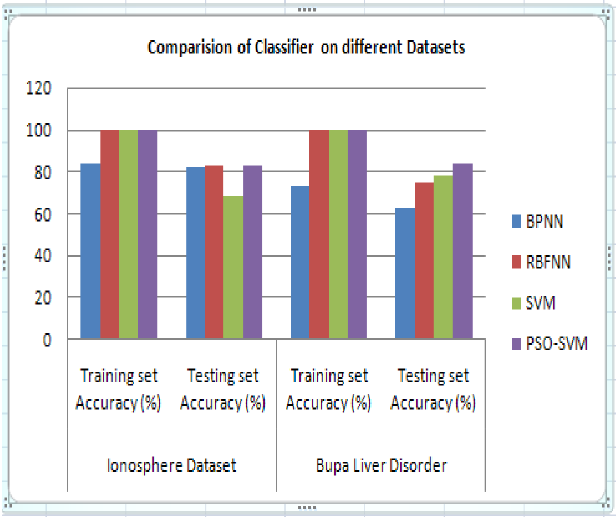

Figure 4.The comparison of classification Accuracies on Ionosphere dataset and Bupa Liver Disorder dataset. From results, it can be concluded that RBFNN is suitable for given tasks. The results of the experiments show that RBFNN can provide good results for both set. The proposed PSO–SVM Classifier is optimal classifier for classification.

REFERENCES

[1] Tom M. Mitchell, “Machine Learning” McGraw-Hill.1997.

[2]Richard O. Duda, Peter E. Hart, Divid G. Stork, “pattern Classification”, second Edition, Wiley Interscience, October, 2000.

[3] Zhang Xuegogn,”Introduction to Statistical Learning Theory and Support vector machins”ACTA, Automatica Sinica, vol. 26, N0.1, Jan, 2000.

[4] J.M.Benitez, J.L.Csstro, I.Requena,”Are Artificial Neural Networks Black boxes?” IEEE transactions on neural networks, vol.8, no.5, September 1997.

[5] S Vince. ‘Applied Physics Laboratory ‘.Johns Hopkins University, Johns Hopkins Road, Laurel. MD 20723, John Hokins University Ionosphere database.1989.

[6] V G Sigillito ,S P Wing ,L V Hutton and K.B Baker.’Classification of Radar Returns from the Ionosphere using Neural Networks’ Jonhs Hopkins APL Technical Digest ,vol 10,1989,pp 262-266.

[7] J Wenli, Z Huoji, LQizhong and Z Yiyu ‘Efficient Radar Target Classificafication using Modular Neural Networks’. Proceedings of International Conference on Radar .CIE2001,15-18 October 2001,p 1031.

[8] T M Conaghy, Leung, E Bosse and V Varadan.”Classification of Audio Radar signals using Radial Basis Function Neural Networks’.IEEE Transactions on

Instrumentation and Measurements, vol 5, no 6 December 2003, pp. 1771.

[9] Farid Melgani ;"Classification of Electrocardiogram signals with supports vectors machines and particle swarm optimization" 2008 IEEE.

[10] J. Anitha, C. Kezi selva vijila and D.jude Hemanth; "Comparative Analysis of Genetic algorithm & particle swarm optimization techniques for SOFM based Abnormal Retinal Image classification" International journal of recent trends in Engineering, vol 2,No. 3,November 2009.

[11] Vapnik V.N. The Nature of Statistical Learning Theory, Springer, New York.1995.

[12] Kenned Y J, Eberhart R."Particle Swarm Optimization ".IEEE International Conference on Neural Networks.Perth: IEEE Neural Network society, 1995, pp.1942.

[13] Ayat N.E., Cheriet M.,Suen C.Y "Automatic Model selection for the optimization of svm kernels".Pattern Recognition, 2005,pp1733-1745

[14] Jayshri Dhande, D R Dandekar"Particle Swarm intelligence for Neural Network Classifier" International conference proceeding on Emerging Trends in Soft computing and ICT(SCICT-2011),G.G Vishwavidyalaya,Bilaspur(C.G)

[15] Jayshri Dhande, D R Dandekar "Design of Optimal Neural Networks and SVM Classifier for Classification of Radar Returns from Ionosphere" IFRSA'S International Journal of Computing, Vol.1Issue3, July, 2011

[16] Jayshri Dhande, D R Dandekar" PSO based SVM as an Optimal Classifier for Classification of Radar returns from Ionosphere": International journal on Emerging Technologies (IJET), (ISSN NO. 0975-8364), Dec 2011