A Neuro Computing Approach for Designing Optimum, Intelligent Clinical Decision Support System for Dermatology

G.U.Kharat
Fellow IETE
Professor & Dean (Research & Development)
Anuradha Engineering College, Chikhli
(Maharashtra), India.
gukharat@rediffmail.com

Abstract

Neural networks have great potential in areas such as signal and image processing where many hypotheses (or features) are pursued in parallel. Principally human do not resemble statistic computers but pattern recognition systems. Humans can recognize patterns or objects very easily but fail when probabilities have to be assigned to observations. These principal difficulties are not widely known by physicians. Also studies revealed that about 50% of the diagnosis is wrong, Neural networks help in this situation. Artificial Neural Networks (ANNs) are extensively used in diagnostic systems. Benefits of using ANN are that they are not affected by factors such as fatigue, working conditions. The authors has designed an optimal Multilayer Perecptron Neural Network (MLPNN) based decision support system for Dermatology and achieved 100% recognition accuracy in classifying Dermatology diseases.

Keywords:

Multilayer Perceptron (MLP), Support Vector Machine (SVM), Dermatology, classification accuracy, Mean Square Error (MSE), Neural Network (NN).

1 INTRODUCTION

Knowledge based systems are used in medical diagnoses. They have the advantage to give on exploration of a diagnosis. This is very important especially in the domain of medicine where the user wants to have the diagnosis proved. But a main difficulty when dealing with knowledge based system is the acquisition of the domain knowledge. There are several problems with it. It is difficult to transform the explicit and implicit knowledge of the expert's domain, which also partly consists of own experiences, in a form, which is suitable for a knowledge base. The knowledge can also be inconsistent or incomplete. A second problem is that knowledge Based systems are not able to learn from experience or to operate with cases not represented in the knowledge base. Automatic detection and Classification of Dermatology diseases is a challenging problem, which can be effectively handled using Neural Networks.

1.1 Multilayer Perceptron (MLP)

Rosenblatt's perceptron is a pattern-recognition machine that was invented in the 1950s for optical character recognition. The perceptron has multiple inputs fully connected to an output layer with multiple McCulloch-Pitts PEs (Figure1). Each input x_j is multiplied by an adjustable constant w_{ij} (the weight) before being fed to the i^{th} processing element in the output layer, yielding

$$y_i = f(net_i) = f\left(\sum_j w_{ij} xj + b_i\right) \qquad 2\text{-}1$$

Where b_i is the bias for each PE. The number of outputs is normally determined by the number of classes in the data. These PEs add the individual scaled contributions and respond to the entire input space.

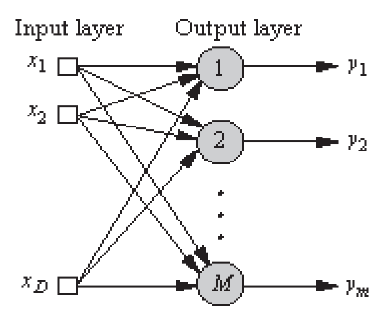

Figure 1 The perceptron with D inputs and M outputs (D-M)

1.2 Support Vector Machines

SVMs have been developed in the reverse order to the development of neural networks (NNs). SVMs evolved from the sound theory to the implementation and experiments, while the NNs followed more heuristic path, from applications and extensive experimentation to the theory. Support vector machines (SVMs) are radically different type of classifier that has attracted a great deal of attention lately due to the novelty of the concepts that they bring to pattern recognition, their strong mathematical foundation, and their excellent results in practical problems. It covered two of the motivating concepts behind SVMs, namely, the idea that transforming the data into a high- dimensional space makes linear discriminant functions practical and the idea of large margin classifiers to train the perceptron.

Here it will couple these two concepts and create the support vector machine.

1. By mapping the input to a sufficiently large feature space, patterns become linearly separable, so a simple perceptron in feature space can do the classification.
2. The RBF places Gaussian kernels over the data and linearly weights their outputs to create the system output i.e. When used as an SVM, the RBF network places a Gaussian in each data sample such that the feature space becomes as large as the number of samples.

But an SVM is much more than an RBF. To train an RBF network as an SVM, an idea of large margin classifiers is used. Training an RBF for large margins will decouple the capacity of the classifier from the input space and at the same time provides good generalization. SVM extend the Adatron algorithm here in two ways: It's application to kernel-based classifiers such as RBFs, and we modify application to modified training for nonlinearly separable patterns.

2 DERMATOLOGY DATABASE

This database contains 34 attributes, 33 of which are linear Valued and one of them is nominal. The differential diagnosis of erythemato-squamous diseases is a real Problem in dermatology. They all share the clinical features of erythema and scaling, with very little differences. The diseases in these groups are psoriasis, seboreic dermatitis, lichen planus, pityriasis rosea, cronic dermatitis, and pityriasis rubra pilaris. Usually a biopsy is necessary for the diagnosis but unfortunately these diseases share many histopathological features as well. Another difficulty for the differential diagnosis is that a disease may show the features of another disease at the beginning stage and may have the characteristic features at the following stages. Patients were first evaluated clinically with 12 features. Afterwards, skin samples were taken for the evaluation of 22 histopathological features. The values of the histopathological features are determined by an analysis of the samples under a microscope.

In the dataset constructed for this domain, the family history feature has the value 1 if any of these diseases has been observed in the family and 0 otherwise. The age feature simply represents the age of the patient. Every other feature (clinical and histopathological) was given a degree in the range of 0 to 3. Here, 0 indicates that the feature was not present, 3 indicates the largest amount possible, and 1, 2 indicate the relative intermediate values.

Number of Instances: 366

Number of Attributes: 34

3 ATTRIBUTE INFORMATION:

3.1 Clinical Attributes: (take values 0, 1, 2, 3, unless otherwise indicated)

1: erythema

2: scaling

3: definite borders

4: itching

5: koebner phenomenon

6: polygonal papules

7: follicular papules

8: oral mucosal involvement

9: knee and elbow involvement

10: scalp involvement

11: family history, (0 or 1)

34: Age (linear)

3.2 Histopathological Attributes: (take values 0, 1, 2, 3)

12: melanin incontinence

13: eosinophils in the infiltrate

14: PNL infiltrate

15: fibrosis of the papillary dermis

16: exocytosis

17: acanthosis

18: hyperkeratosis

19: parakeratosis

20: clubbing of the rete ridges

21: elongation of the rete ridges

22: thinning of the suprapapillary epidermis

23: spongiform pustule

24: munro microabcess

25: focal hypergranulosis

26: disappearance of the granular layer

27: vacuolisation and damage of basal layer

28: spongiosis

29: saw-tooth appearance of retes

30: follicular horn plug

31: perifollicular parakeratosis

32: inflammatory monoluclear inflitrate

33: band-like infiltrate

3.3 Class Distribution: (Output)

Class code:	Class:	Number of instances:
1	psoriasis	112
2	seboreic dermatitis	61
3	lichen planus	72
4	pityriasis rosea	49

5	cronic dermatitis	52
6	pityriasis rubra pilaris	20

4 COMPUTER SIMULATION EXPERIMENT

The scheme for classification of dermatology diseases is shown in figure 2. Authors have used Multilayer Perceptron (MLP), and Support Vector Machine (SVM) for this pattern recognition problem. Number of input Processing Elements (PEs) must be equal to that of input attributes. Since we have used 34 input attributes, and six types of dermatology diseases, 34 input Processing Elements (PEs) are used in input layer. Six Processing Elements are used in output layer.

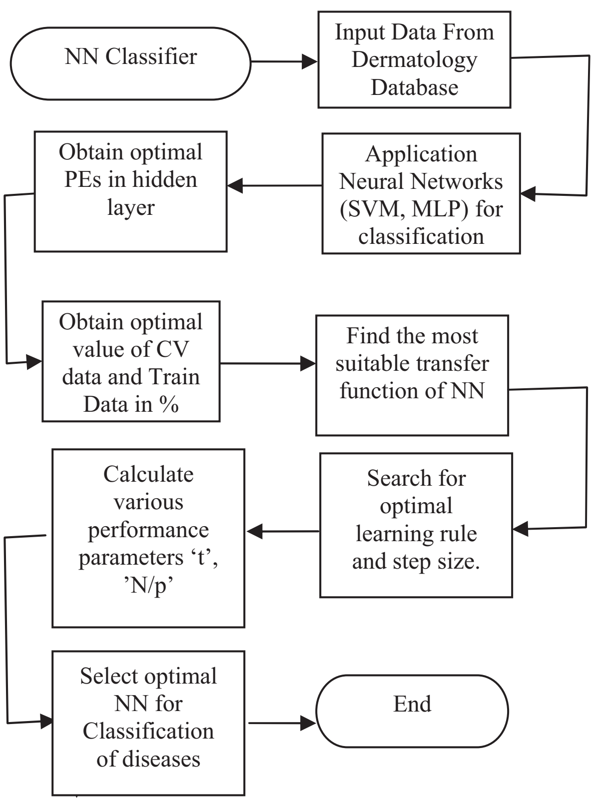

Figure 2: Scheme for Emotion Recognition.

4.1 MLP Neural Network Classifier

For generalization, the randomized data is fed to the network and is trained for different hidden layers. It is observed that MLP with single hidden layer gives better performance. The number of Processing Elements (PEs) in the hidden layer is varied. The network is trained and average minimum Mean Square Error (MSE) on CV data is obtained when 6 PEs are used in hidden layer. Rigorous experimentation is done by varying number of exemplars for training and cross validation data. With 6 PEs in hidden layer, the network is trained three times with different random initialization of connection weights so as to ensure true learning and to avoid any biasing towards choice of specific initial connection weights. Average Minimum MSE on Trained and CV data set along with average classification accuracy is calculated and is shown in figure 3 & 4. It is observed that minimum MSE and maximum classification accuracy is obtained when 20% exemplars are used for cross validation (CV) data and 80% for training.

With 20% CV and 80% train data, various transfer functions are used for training the network and average minimum MSE on training and CV data is measured (Figure5). Average classification accuracy for different transfer function is plotted in figure 6. It is observed that tanh is most suitable transfer function. The MLP neural network is trained with tanh transfer function using various learning rules. Minimum MSE on training and CV data set is measured and is indicated in figure 7. Finally network is tested on training and test dataset (figure 8). It is concluded that momentum is most suitable learning rule for our neural network.

From above experimentation selected parameters for MLP neural network are given below.

MLP NN (34-6-6), Number of examplars = 366

Exemplars for cross validation = 20%

Exemplars for training = 80%

Termination is at 100 epochs without improvement.

Layers of MLP	Transfer Function	Learning Rule	Step size	Momentum
HL#1	Tanh	Momentum	0.1	0.7
Output layer	Tanh	Momentum	0.01	0.7

Time elapsed per epochs per exemplar = 55.67 µSec.

Number of free parameters (P) of MLP= 252

Number of exemplars in training dataset (N) = 193

(N/P) ratio = 1.1626

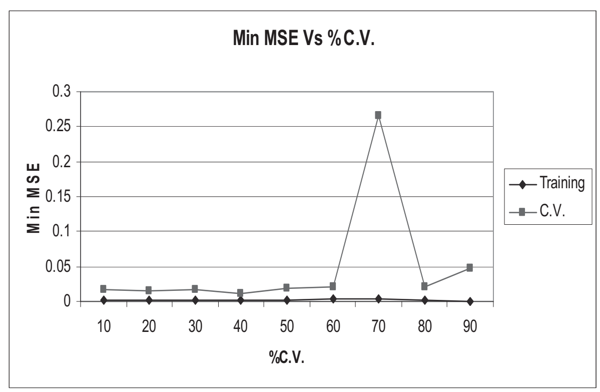

Figure 3: Variation of Av. Min. MSE with % CV.

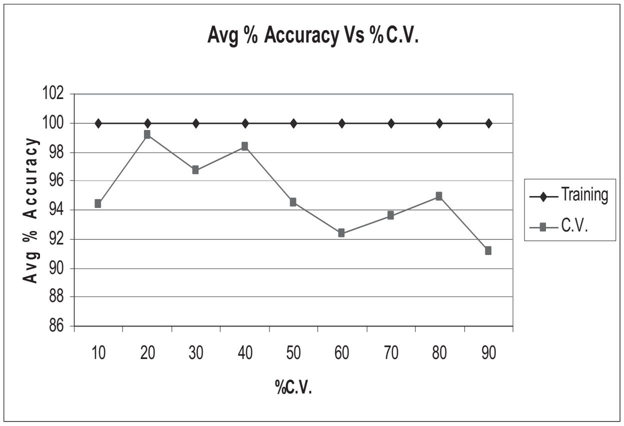

Figure 4: Variation of Av. Accuracy with % CV.

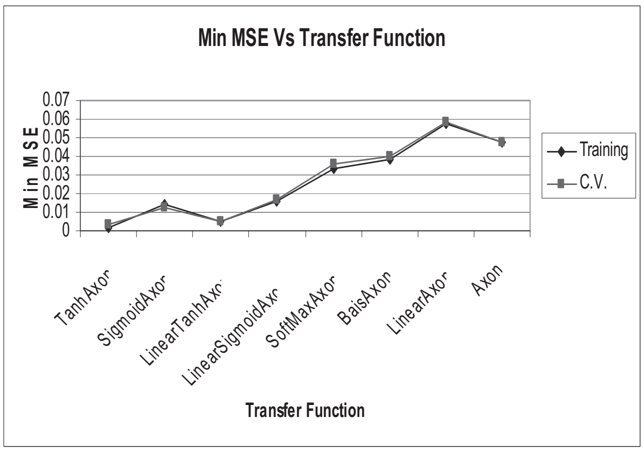

Figure 5: Variation of Transfer Fun. with % CV.

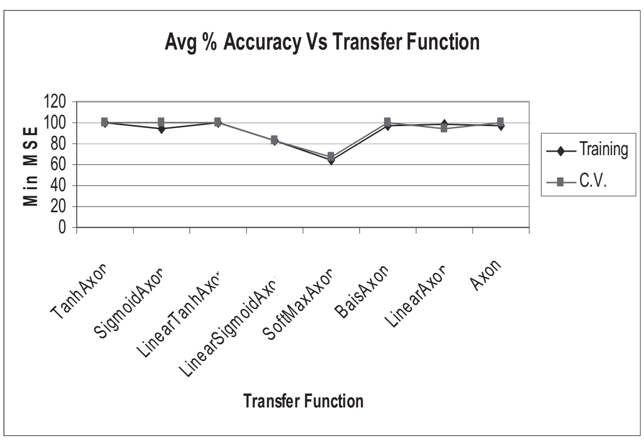

Figure 6: Variation of Av. % Accuracy with Transfer Function

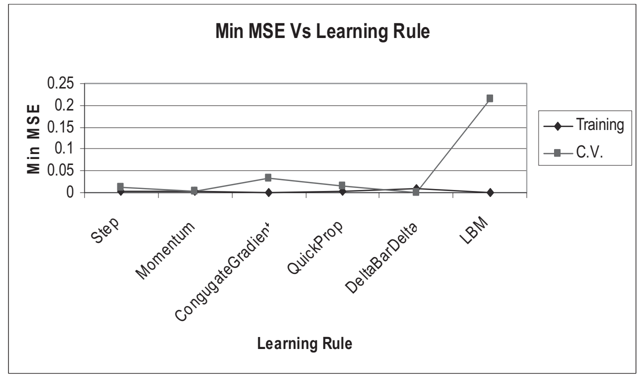

Figure 7: Variation of Av. Min. MSE with Learning Rules

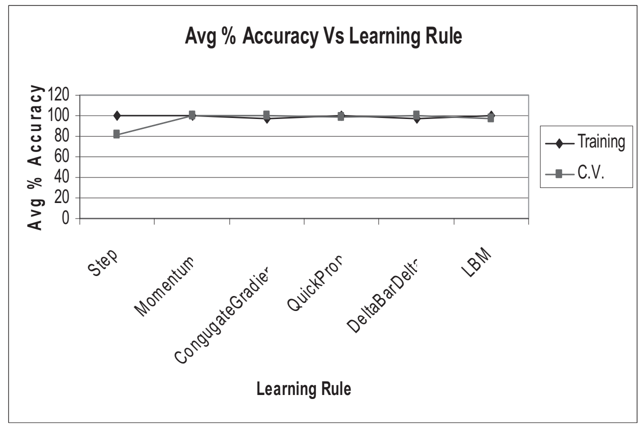

Figure 8: Variation of Av. % Accuracy with Learning Rules

5 SVM NEURAL NETWORK CLASSIFIER

The randomized data is fed to the SVM network. The network is trained three times by varying the number of exemplars for training and CV data. The average minimum MSE for train and CV data and percentage average classification accuracy is calculated and is shown in figure 9 & 10. The optimal results are obtained when 40% data is used for cross validation.

With 40% CV and 60% train data the step size for training the SVM is varied from 0.01 to 1.0 and each time network is trained and tested on training and CV dataset. The graph of average minimum MSE and % average classification accuracy is plotted against step size in figure 11 & 12 respectively. It is observed that optimal results are obtained for 0.9 step size. Optimally designed SVM is

Learning control	=	Supervised
Weight update	=	Batch
Step size	=	0.9
Number of epochs	=	366
Number of runs	=	03
Termination of training	=	100 epochs without improvement.

Time elapsed per epochs per exemplar = 733.97 µSec.

Number of free parameters (P) of SVM = 252

Number of exemplars in training dataset (N) = 193

(N/P) ratio = 1.1626

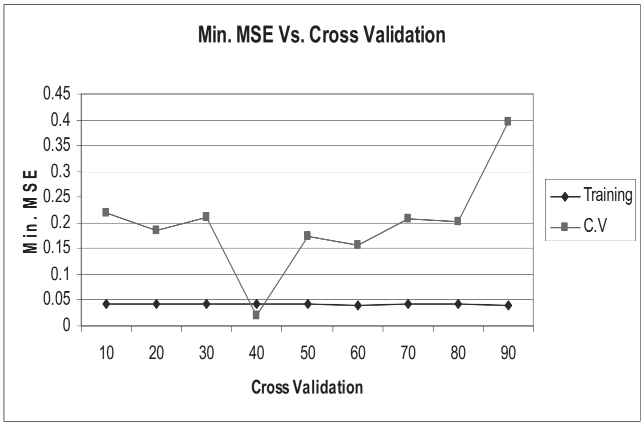

Figure 9: Variation of Av. Min. MSE with CV

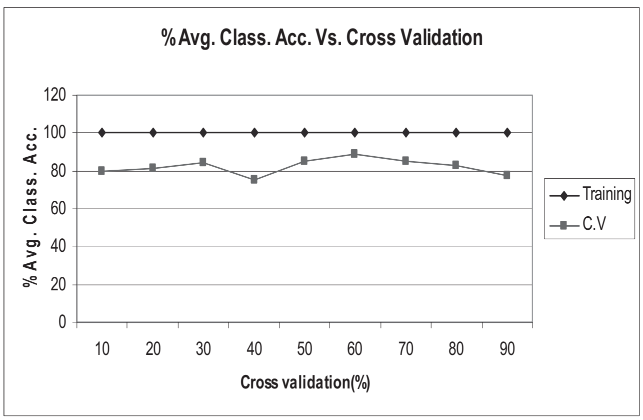

Figure 10: Variation of Av. % Accuracy with CV

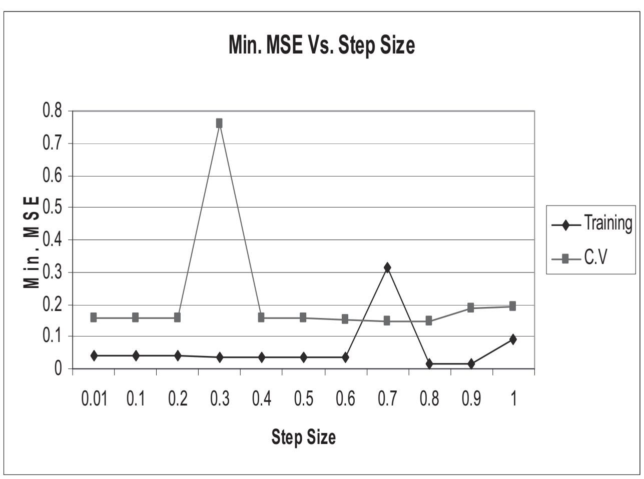

Figure 11: Variation of Av. Min. MSE with Step Size

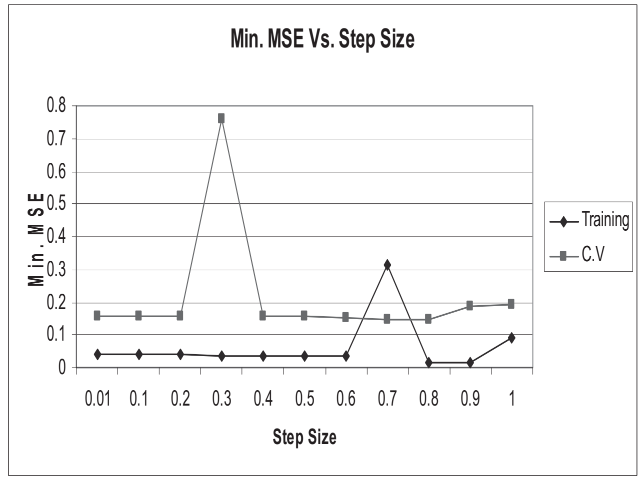

Figure 10: Variation of Av. % Accuracy with Step Size

6 RESULT & CONCLUSION

In this paper, the author evaluated the performance of the two Neural Networks MLP and SVM for classification of dermatology diseases. The results are shown in Table 1.

Table 1 show classification accuracy on training and testing data set using MLP and SVM. The accuracy of recognition is 100% on both test dataset and training dataset. The time elapsed per examplar per epochs per run for MLP NN is quite less than that of SVM. Hence Multi Layer Perceptron Neural Network is recommended for the classification of dermatology diseases.

Table 1: Comparison of Performance Parameters for MLP & SVM

	Min. MSE		%Avg. Class. Acc.		N/P Ratio	Time Elapsed 't' (Microsecond.)
	Training	C.V	Training	C.V		
MLP	0.0010	0.0077	100	100	1.16	55.67
SVM	0.0100	0.0257	100	100	1.16	733.97

7 REFERENCES

[1] Herbin, M. Venot et.al., Color Quantitation for Image Processing in Dermatology, IEEE Transaction on Medical Images Vol 9, Issue 3, PP 262-269 ISSN: 0278-0062

[2] C. Grana, Color Calibration for a Dermatological Video Camera System, IEEE Transaction on Medical Images Vol 19-7, July 2000 pp 722-730

[3] Multicenter Post-Infarction Research Group. Mechanism of death and prevalence of myocardial ischemic symptoms in

the terminal event after acute myocardial infarction. Am JCardiol 1988; 61: 8-15.

[4] MIT-BIH Data Reposiratories

[5] Artificial Neural Networks in Medicine http://www.emsl.pnl.gov:2080/docs/cie/techbrief/NN.techbrief.html

[6] Pattern Recognition and Machine Learning by Christopher M. Bishop