Intelligent Diagnosis of Heart Diseases using Neural Network Approach

RANJANA RAUT

Reader

P. G. Department of Applied Electronics

SGB Amravati University, Amravati,

(M.S.) India 444602

DR. S. V. DUDUL

Professor & Head,

P. G. Department of Applied Electronics

SGB Amravati University, Amravati,

(M.S.) India 444602

ABSTRACT

Experiments with the Switzerland Heart Disease database have concentrated on attempting to distinguish presence and absence. The classifiers based on various neural networks, namely, MLP, PCA, Jordan, GFF, Modular, RBF, SOFM, SVM NNs and conventional statistical techniques such as DA and CART are optimally designed, thoroughly examined and performance measures are compared in this study. With chosen optimal parameters of MLP NN, when it is trained and tested over cross validation (unseen data sets), the average (and best respectively) classification of 98±2.83 % (and 100%), 96.67±4.56% overall accuracy, sensitivity 96±5.48, specificity 100% are achieved which shows consistent performance than other NN and statistical models. The results obtained in this work show the potentiality of the MLP NN approach for heart diseases classification.

Keywords: Heart disease; MLP neural network; Error back propagation algorithm; Performance

1. INTRODUCTION

A major challenge, facing healthcare organizations (hospitals, medical centers) is the provision of quality services at affordable costs. Quality service implies diagnosing patients correctly and administering treatments that are effective [Bonow et. al, 2006]. Integration of clinical decision support with computer-based patient records could reduce medical errors, enhance patient safety, decrease unwanted practice variation, and improve patient outcome. Global burden of disease estimates for 2001 by World Bank Country Groups shows severity statistics indicated in year 2001 is 25.2 % for India and from literature survey now it has increased to 46% [Mathers et al, 2004]. In spite of the rapid development of pathological research and clinical technologies, more than 60,000 people die suddenly each year in India due to arrhythmias and heart diseases.

The aim of the present study is to identify the combination of clinical and a laboratory noninvasive variable, easy to obtain in most patients, that best predicts the occurrence of heart diseases. Taking cardiologist's as gold standard it is aimed to minimize the difference by means of machine learning tools. From exhaustive and careful experimentations, it is observed that proposed NN classifiers ensures true estimation of the complex decision boundaries, remarkable discriminating ability and does outperform the statistical discriminant analysis and classification tree rule based predictions [Principe, 1999].

1.1 Types and Existing methods

Clinical decisions are often made based on doctors' intuitions and heuristics experience rather than on the knowledge rich data hidden in the database. This practice leads to unwanted biases, errors and excessive medical costs which affects the quality of service provided to patients [Yu et al, 2007]. In face of uncertainty of heart disease symptoms even experienced cardiologists need complimentary assistance from intelligent decision system to arrive at precise diagnosis of cardiac disease [Itchhaporia et al., 1996]. A number of techniques have been used for identification of heart diseases including waveform analysis, time frequency analysis, complexity measures, Neuro Fuzzy RBF NN [Sandhu S., 1989] and a total least square based Prony modeling algorithm [Panayiota et al., 2004]. But it has been observed that classification accuracies were not good (only upto 79 %) with these techniques and still enough scope in improving by choosing appropriate NN model. In the early stages, angina tends to be asymptomatic and can be detected only through screening.

Modern digital computers outperform humans in the domain of numeric computation and related symbol manipulation [Hansen and Salamon, 1990]. The echocardiogram, ECG information and other patient data are gathered in real time and sent to the intelligent DSS system. This stored data can then be processed to detect the various complexes and then detect specified heart diseases.

1.2 Benchmark Data Set

Data is obtained from UCI (University of California, Irvine C. A.) center for machine learning and intelligent systems [Murphy, 2004]. This database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. Namely Age in years, Sex, Chest pain type (typical, atypical, non-anginal, asymptomatic), Resting blood pressure, Serum cholesterol in mg/dl, Fasting blood sugar > 120 mg/dl, Resting electrocardiographic results, Maximum heart rate achieved, Exercise induced angina, ST depression induced by exercise relative to rest, The slope of the peak exercise ST segment, Number of major vessels, and 14th feature is output based on Angiographic diagnosis of heart disease.

The "goal" field refers to the presence of heart disease in the patient. Missing data is first preprocessed by estimation through interpolation method. This machine learning benchmark dataset consists of 123 heart recordings from different patients (08 normal, 115 abnormal) and includes about 0.63% missing attribute values, so the prediction accuracy of any model built using it cannot be perfect. It is reasonably difficult data with incomplete and ambiguous and only 6.50% normal samples. Class distribution of this data set is very unfair.

1.2 Complexity of the Decision Boundaries

Figure 1 shows a typical feature plot of features, maximum heart rate versus resting blood pressure. Although, the plots show that these features are reliable indicators for the classification of heart disease type, the features clusters are not linearly separable. It is evident that the clusters formed are complex and hence classification of heart diseases angina types) is more challenging. To solve it efficiently, NNs are used as classifiers [Haykin, 2007]. These features together are significantly sensitive for discrimination of heart function as normal or abnormal.
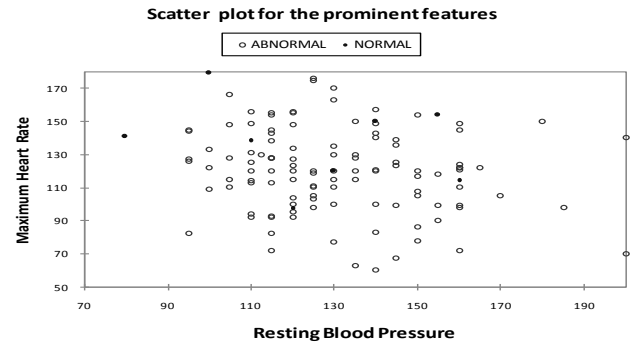

Figure 1 Scatter plot for the features maximum heart rate and resting blood pressure

2. DESIGN OF INTELLIGENT SYSTEM

The neural network design mainly consists of defining the topology (i.e. the arrangement of PEs, connections, and patterns into the neural network) and the architecture (i.e. the selection of the number of PEs for each layer necessary for the specific application of the topology) of the network [Bose, 2001]. For generalization the randomized data is fed to the networks and similar methodical experimentation work is done by choosing different NN models and configurations, varying data partitioning. The learning and generalization ability of the estimated NN based classifier is assessed on the basis of certain performance measures such as average and overall classification accuracy, sensitivity, specificity, area under ROC curve, training time, network complexity and MSE [Tokan, 2006]. From the comparison of performance measures outcome it is obvious that 13-09-02 MLP architecture has the largest, 13-111-02 SVM, 13-16-02 Jordan, 13-11-02 PCA, 13-08-02 Modular have moderate and, 13-07-02 GFF, 13-10-02 SOFM, 13-25-02 RBF and conventional statistical DA have the smallest performance measures. From comparative analysis of the result obtained on various Neural Network models, it is observed that MLP (13-09-02) with tangent hyperbolic activation function at hidden and output and momentum learning rule performed elegantly providing the highest performance measures.

2.1 Multilayer Perceptron Neural Networks

In order to learn more complex decision function the inputs are fed into a number of perceptrons nodes, each with its own set of weights and threshold [Bishop, 1997]. The outputs of these nodes are then input into another layer of nodes and so on, the

output of the final layer of nodes is the output of the network. Such a network is termed a multi-layer Perceptron (MLP) and the layers of nodes whose input and output are seen only by other nodes are termed hidden [Lippmann, 1987]. The connection weights are computed by means of a learning algorithm. There are different variants of back- propagation learning algorithms in the literature [Hagan, 1997].

2.2 Experimental Determination of Near Optimal MLP NN Classifier

Computer simulation is done using Neurosolution version 5.07 and MATLAB 7.3 [Neuro Dimension, 2007].The network is trained five times with different random initialization of connection weights so as to ensure true learning. Termination is when MSE is increased on CV set. It is also established from results that, the 90% training and 10% cross validation (normal tagging) data partition scheme provided best performances. 13-09-02 MLP NN configuration found outstanding. Performance found optimal for 09 neurons with regard to accuracy and MSE on train and CV dataset. It is clear that, transfer function of neurons in hidden layer as well as output layer should be hyperbolic-tangent (tanh).

Details about the various training algorithms and their parameters can be found in Figure 2. The choice of the optimal values was made as per the exhaustive experimentation for the training of the MLP NN for different values of these parameters. The MLP network should be trained using momentum algorithm for the best performance. Variable parameters of MLP NN are as shown in Table 1. Table 2 exhibits optimal parameter settings obtained for other neural networks. The designed classifier is evaluated on cross validation with regard to percent classification accuracy, sensitivity, specificity, area under the ROC curve, training time, and MSE.
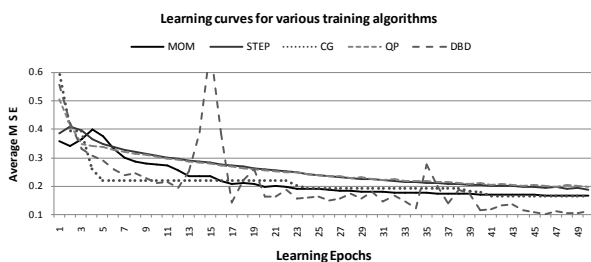
Figure 2 Comparison of different learning curves for the training of MLP NN

Table 1. Variable Parameters of MLP NN (13-09-02)

Parameter	Typical Range	Optimal values
Exemplars for training N	10% to 90%	90% (111)
Exemplars for cross validation	10% to 90%	10% (12)
Number of Epochs	1000 to 10000	1000
Number of hidden layers	1 to 3	1
Number of hidden neurons	2 to 100	09
Transfer function of neurons in hidden layer	Tanh, Sigmoid, Linear Tanh, Linear Sigmoid, Bias axon, Linear axon, Soft Max, Axon	Tanh
Transfer function of neurons in Output layer	Tanh, Sigmoid, Linear Tanh, Linear Sigmoid, Bias axon, Linear axon, Soft Max, Axon	Tanh
Supervised Learning Rule	Step, Momentum, Conjugate Gradient (CG), Quick Propagation (QP), Delta bar delta.	Momentum
Momentum Constant	0 to 1	0.7
Step Size at hidden and output layer (Learning Rate)	0 to 1	Hidden: 1.0 Output: 0.1
Training Time per Epoch per Exemplar		0.03603 msec
Number of free parameters, P (connection weights) N/P Ratio	I*H1 PE's + H1 PE's * Output PE's + (H1+ Output) PE's	P = 146 0.7603

Table 2 Variable Parameters of other NNs
Training exemplars N = 90%, Exemplars for cross validation = 10%, stopping criteria CV error increased

NN Model	Optimal values	connection weights (p)	N/p
SVM 13-111-02	Supervised learning epochs 1000, supervised learning rule momentum with momentum constant 0.7 and step size (learning rate) in hidden layer 0.93	2555	0.0434

Learning curves for various training algorithms

— MOM — STEP ········ CG ----- QP — — DBD

Jordan 13-16-02	Topology one, context unit time 0.7, transfer function at context unit Integrator axon, Supervised learning epochs 1000, hidden neurons 16with Linear tanh transfer function at hidden and Soft max at output, supervised learning rule momentum with momentum constant 0.7 and step size (learning rate) in hidden layer 0.1, output layer 0.1	258	0.4302
PCA 13-11-02	Principal components 04, learning rule Sangers full, Supervised learning epochs 1000, Unsupervised learning epochs 100, learning rate starts at 0.01 and decay to 0.001, hidden neurons 11 with Tanh transfer function at hidden and Axon at output, supervised learning rule momentum with momentum constant 0.7 and step size (learning rate) in hidden layer 1.0, output layer 0.1	178	0.623
Modular 13-08-02	Architecture one, Supervised learning epochs 1000, hidden neurons 08 with Axon transfer function at hidden and Tanh at output, supervised learning rule momentum with momentum constant 0.7 and step size (learning rate) in hidden layer 0.01, output layer 0.1	130	0.853
GFF 13-07-02	Supervised learning epochs 1000, hidden neurons 07 with Tanh transfer function at hidden and output, supervised learning rule momentum with momentum constant 0.7 and step size (learning rate) in hidden layer 0.1, output layer 0.1	114	0.973
SOFM 13-10-02	Supervised learning epochs 1000, Unsupervised learning epochs 100, learning rate starts at 0.01 and decay to 0.001, hidden neurons 10 with Linear Tanh transfer function at hidden and Tanh output, supervised learning rule momentum with momentum constant 0.7 and step size (learning rate) in hidden layer 1.0, output layer 0.1, Row 05, column 05, start radius 1, final radius zero, neighborhood shape Squared Kohonen Full	162	0.685
RBF 13-25-02	Gaussian cluster centers 25, competitive learning metric Euclidean, competitive unsupervised learning rule Conscience full, Supervised learning epochs 1000, Tanh transfer function at hidden and output, supervised learning rule momentum with momentum constant 0.7 and step size (learning rate) in output layer 1.0	402	0.276

2.3 Dimensionality Reduction using Principal Component Analysis

Reduction in dimensionality of input space and hence the network can be achieved by Principal Component Analysis (PCA). PCA is performed using XLSTAT2008. Experimentation is done using Pearson (n), Pearson (n-1), Covariance (n-1), Covariance (n), Spearman, Kendall and Polychonic types, out of which Pearson (n) rule is found best. Eigenvalues and variation of average classification accuracy on number of principal components as inputs reflects the quality of the projection from 13 to 10 dimensions.

Table 3 displays various performance measures of MLP NN on different datasets with respect to normal and diseased heart instances. MSE is included in following tables only as a matter of record, since small MSE does not necessarily imply good generalization with unseen data. From performance comparison Table 4 of proposed MLP NN technique with other NN models, it implies that the MLP NN as a classifier for this work possesses more learning ability than the other NN's.

Table 3 Performance Measures of MLP NN Classifiers

Data sets	% Classification Accuracy		MSE	ROC Analysis Area under ROC and its convex hull	%Sensitivity ± SD	%Specificity ± SD
	Average ± SD	Overall ± SD				
13:09:02MLP 90% train data	98.95 ± 1.48	98.02 ± 1.17	0.0026 0.0106	0.94428, 0.9942	97.90 ±1.24	100
10% CV data Three Fold CV	98 ± 2.83	96.67 ± 4.56	0.0141 0.053	0.89818, 0.9312	96 ± 5.48	100
	95.42 ± 4.5	86 ± 3.5	0.081 0.100	0.81671, 0.9711	90.07 ± 2.3	89.00± 3.5
10:05:02MLP 90% train data	94	95		0.9024, 0.9366	93	95
10% CV	91.6	91.45		0.88	90±1.0	92.9±5

data Three Fold CV	±4.5	±5.5		44, 0.9123	5	
	90	85		0.8326, 0.8625	74	81

2.4 Selection of Error criterion

Normally Euclidian or L_2 norm is used. However when the problem incorporates very high degree of nonlinearity different error norms could be examined for their suitability in computation of error between output of NN model and the desired output. To select the correct error criterion various norms has been tested for MLP NN and results depicts that L_2 norms provided highest classification accuracy on test, CV and train data as well as minimum MSE.

2.5 Multifold Cross Validation (Leave-N-Out)

Proposed MLP NN is trained using leave –n-out cross validation technique so as to ensure that its performance does not depend on specific data partitioning scheme. In this cross validation rows are shifted by a factor n which depends on data partitioning percentage of train and cross validation. Number of runs change as per shift, (for Shifts, 12, 20, 25, 30, 37; runs are 10, 7, 5, 4, and 3 respectively). Classification accuracies of 85 to 92 % and area under curve approaching unity on many validation sets for MLP NN are observed.

3. STATISTICAL CLASSIFIER BASED DSS

Software used for implementing this model is XLSTAT 2008. It is based on branched tree with various rules and goals. The classification tree has been created by applying different methods, measures and depth of the tree. Depth of tree is varied from 5 to 10 in step of one. Table 5 displays performances on conventional statistical approaches. From the results it is clear that DA statistical classifier is able to diagnose heart diseases with classification accuracy only up to 70.91 %.

Table 4 Comparative results of all NNs on CV data

NN Model	Performances on test on CV data							
	% Classification Accuracy		% Sensitivity ± SD	% Specificity ± SD	Area under ROC curve	Train time/ epoch/ exemplar msec	N/P ratio	MSE
	Average ± SD	Overall ± SD						
MLP	98 ±2.83	96.67± 4.56	96±5.48	100	0.89818	0.03603	0.7603	0.0026
MLP DR	93.6± 4.5	91.45± 5.5	90±1.05	92.9 ± 5	0.8844	0.015	1.6567	0.0084
SVM	97±4.47	95±7.45	94±8.94	100	0.9014	0.1891	0.0434	359
Jordan	89.8± 0.27	80±4.56	79.60	100	0.8845	0.09	0.430	0.0208
PCA	88±2.74	86.67± 4.57	86±5.48	100	0.8490	0.0491	0.623	0.0054
Modular	83±2.74	78.33± 4.56	76±5.48	100	0.8490	0.0063	0.853	0.131
GFF	81±2.74	80.67± 4.57	81 ±5.48	100	0.8467	0.04504	0.973	0.103
SOFM	75	91.67	100	50	0.8085	0.0065	0.685	0.0431
RBF	72±2.74	86.67± 4.57	94±5.48	50	0.8490	0.0409	0.276	0.106

To what extent the MLP NN classifier is able to correctly classify the exemplars is the most important criterion for its proper evaluation. In order to confirm whether the proposed model is really consistently capable of near optimum classification, different 54 data partitions sets (forward, reverse tag, differential learning, different split ratios etc) are used to train the classifier. When estimated MLP NN is evaluated on testing instances, average classification accuracy is seen to varying between 82 to 98 %, highest than SVM and Jordan NN on different split ratios.

As per the confusion matrices it was found that the MLP neural classifier has the advantage of reducing misclassifications among the neighborhood classes compared to other NN classifiers and provided consistent classification accuracy over 10 runs for both, normal and diseased instances.

Table 5 Performance Measures from Classification and Regression Tree

Significance level 5 %, split threshold 5%, maximum tree depth 08, rule based model

Performance Measure	CART-Method and Measure				Discriminant Analysis
	CHAID Likelihood	EX CHAID Likelihood	CART Ginni	QUEST	
Accuracy (%)	66.66	66.66	65.12	58.33	**70.91**
Specificity (%)	50	50	50	50	75
Sensitivity (%)	70	70	60	60	81.82
AUC(ROC)	0.674				0.774

4. RESULTS AND DISCUSSION

Table 6
Performance Comparison of Proposed Technique with Others on Same Dataset

Previous Technique	Performances % Accuracy, train time, error rate, sensitivity, specificity	References
Neuro Fuzzy RBF NN	Angina 79 %	Sandhu S. (1989)
Proposed Techniques **MLP NN (13:09:02)** For class Normal and Diseased	Average Accuracy 98±2.83% , Overall 96.67±4.56 % on test, 98.95±1.48 % on train, sensitivity 96±5.48%, specificity 100 %, 0.03603 msec, AUC (ROC) 0.89818, MSE 0.0026	
MLP DR (10:05: 02) For class Normal and Diseased	Average Accuracy 91.6±4.5 % Overall 91.45 ±5.5 % on test, Average 94%, Overall 95% on train, sensitivity 90±1.05%, specificity 92.9±5%, 0.01517 msec, AUC (ROC) 0.8844, MSE 0.084	
Discriminant Analysis Classification Tree	Overall 70.91 % on test, 80.41 % on train sensitivity 75%, specificity 81.82 % AUC (ROC) 0.774 Overall 66.66 % on test, sensitivity 70 %, specificity 50 % %, AUC (ROC) 0.674	

From performance comparison of proposed technique with others on same dataset as shown in Table 6 it is proved that proposed MLP NN Classifier clearly outperforms earlier researchers' techniques as well as statistical methods. Published studies, on Switzerland heart disease database, report only 79 % classification accuracy from previous related research. With chosen optimal parameters of MLP NN, when it is trained using 5 runs and tested over cross validation (unseen data sets) five times, the average (and best respectively) classification of 98±2.83 % (and 100%), 96.67±4.56% overall accuracy, sensitivity 96±5.48, specificity 100% are achieved which shows consistent performance than other NN models. The system runs in a 0.03603 millisecond in the environment of Intel Pentium 4 PC with 2.4 GHz CPU and 1 GB DDRAM. Thus proposed MLP NN Classifier surpasses the existing methods.

5. CONCLUSION

The dimensionally reduced MLP neural network method has also proved to be reliable for implementing quantitative prognosis of angina in patients with heart failure. Additional studies with larger numbers of patients are required to better assess the usefulness of artificial neural networks. It is observed that 13-09-02 MLP NN is fastest network, simple in design and synthesis, lowest average MSE, highest accuracy and ROC analysis is perfect approaching unity. Significant (16.67 %) reduction in connection weights and (58.36 %) reduction in training time are achieved with PCA dimension reduction. These issues have not been addressed in previous other researchers' studies [Akhbardeh, 2005].

From the design of neural networks in this work, it is evident that MLP NNs required a compact architecture as compared to other NNs, in terms of number of hidden nodes required for the near optimal classifiers. Thus, the number of free parameters (weights and biases) required for the designed MLP NN is sufficiently lower than other. This simplicity and compactness in the structure indicates the feasibility of MLP NN for the online implementation, and the hardware implementation [Reyneri, 2003].

Whenever new research findings are listed in journals, this DSS can be retrained to accommodate new knowledge. This binary heart disease classifier can be used to assist the physicians to detect angina for preliminary diagnosis, while examining the patients. Thus they can attempt perfection in the diagnosis of heart diseases. Optimal MLP NN model developed can be implemented in FPGA. Appropriate interfaces could be developed for the interaction between patients, computer and FPGA NN based DSS.

6. REFERENCES

[1] Bonow, Libby, Mann, Zipes, 2006. Heart Disease: a textbook of cardiovascular medicine, eighth edition, Saunders, Elsevier.

[2] Mathers C. D., Lopez A., Stein D. 2004. Deaths and disease burden by cause: Global burden of disease estimates by World Bank Country Groups.

[3] Principe J., Euliano, N., Lefebvre, C. 1999. "Neural and adaptive systems: fundamentals through simulations", John Wiley and Sons.

[4] Murphy, P. M. and Aha, D. W. 2004. UCI Machine Learning Databases Repository Irvine C. A. : University of California, Department of Information

and Computer science, ftp://ftp.ics.uci.edu/pub/machine-learning-databases/heart/

[5] Haykin, S. 2007. "Neural network: A Comprehensive foundation", Pearson Prentice Hall, New Delhi.

[6] Bose, N.K. and Liang, P. 2001. "Neural network fundamentals with graphs, algorithms, and applications", Tata McGraw- Hill publishing company Ltd. New Delhi.

[7] Tokan Fikret , Türker Nurhan and Yıldırım Tülay, 2006. "ROC Analysis as a Useful Tool for Performance Evaluation of Artificial Neural Networks" Lecture Notes in Computer Science, Publisher Springer Berlin / Heidelberg, ISSN 0302-9743 (Print) 1611-3349 (Online), Volume 4132/2006, Book Artificial Neural Networks – ICANN.

[8] Bishop C. 1997. "Neural Networks for pattern Recognition", Oxford University Press, New York.

[9] Lippmann, R.P. 1987. "An introduction to computing with neural nets", IEEE ASSP Magazine, pp.4-22.

[10] Hagan, M.T., Demuth H.B., Beale M.H. 1997. "Neural Network Design, PWS publishing", Boston, MA.

[11] Neurosolution version 5.07, 2007. Neuro Dimension, Inc. Gainesville, Florida, (USA),

[12] MATLAB: Neural Networks Toolbox users guide, The Math Works, Inc., Natick, MA, 2008.

[13] Akhbardeh A., Junnila, S., Koivuluoma, M., Koivistoinen T., Varri A., The heart disease diagnosing system based on force sensitive chair's measurement, biorthogonal wavelets and neural networks, *Proceedings of IEEE International Conference on Advanced Intelligent Mechatronics,* pp. 676 – 681, Digital Object Identifier 10.1109/AIM.2005.1511060.

[14] Reyneri, L.M. 2003. "Implementation Issues of Neuro- fuzzy hardware: going towards HW/SW co design", IEEE Trans. on Neural Networks, vol. 14, no.1, pp.176-194.