Decision Support System for Cardiovascular Heart Disease Diagnosis using Improved Multilayer Perceptron

Sunila
CSE Dept GJUS&T,
Hisar, Haryana

Prabhat Panday
Add. Director of Higher Education, Rewa
M.P

Nirmal Godara
GCW Hisar
Haryana

ABSTRACT
Medical science industry has huge amount of data, but most of this data is not mined to find out hidden information in data. Diagnosing of heart disease is one of important issue to develop medical decision support system which will help the physicians to take effective decision.This paper discusses standardMultilayer perceptron algorithm and analyzes its behavior.This paper proposes an Improved Multilayer perceptron algorithm which divides datasets into multiple subsets Then MLP algorithm was called individually for each subset and results obtained from different subsets are combined using voted combiner with majority probability rule. FinallyPerformance of these techniques is measured through sensitivity, specificity, accuracy and ROC.Improved MLP approach significantly outperforms MLP approach in overall execution time.

Experimental Result shows that Improved MLP algorithm gives better results than MLP algorithm.In our study 10-fold cross validation method is used to measure the unbiased estimate of the model. Cleveland,Hungarian and Switzerland datasets are used for empirical comparisons.

Keywords
heart disease, artificial neural network,
Multilayer perceptron, supervised learning,

1. INTRODUCTION
Decision Support System (DSS) are a specific class of computerized information system that assists totake decisions. A properly designed DSS is an interactive software-based system intended to help decision makers compile useful information from raw data, documents, personal knowledge to identify and solve problems and make decisions.

A Decision Support System can be defined as an interactive data processing and display system which is used to assist in the concurrent decision-making process, and which also conforms to the following characteristics:
It is sufficiently user-friendly to be used by decision-makers in person.
It displays its information in a format and terminology which is familiar to its users.
It is selective in its provision of information and avoids exposing its users to information overload.

Decision Support Systems evolved early in the era of distributed computing. According to Keen and Seott Morton (1978), the concept of decision support has evolved from two main areas of research: the theoretical studies of organizational decision making in the late 1950s and early 1960s, and the technical work on interactive computer systems, mainly carried out in the 1960s. IN the middle and late 1980s, executive information systems (EIS), group decision support systems (GDSS), and organizational decision support systems (ODSS) evolved from the single user and model-oriented DSS. Beginning in about 1990s, data warehousing and on-line analytical processing (OLAP) began broadening the realm of DSS. As the millennium approached, new web-based analytical applications were introduced. DSS has endless possibilities that they can be used anywhere and anytime, for its decision making needs.

Detecting a disease from several factors or symptoms is a many-layered problem that also may lead to false assumptions with often unpredictable effects. Therefore, the attempt of using the knowledge and experience of many specialists collected in databases to support the diagnosis process seems reasonable [1]. Diagnostic decision support is still very much an art for physicians in their practices today due to lack of quantitative tools. A medical diagnostic DSS is a computer program that contains all relevant medical domain knowledge about a certain medical domain and generates a differential diagnosis on the basis of individual patient findings. A medical diagnostic DSS may be extremely useful because it is able to improve the accessibility of expert knowledge and patient information, resulting in quality improvement of the diagnostic process, increase of efficiency and reduction of costs [1].

Cardiovascular disease (CVD) refers to any condition that affects the heart. Many CVD patients have symptoms such as chest pain (angina) and fatigue, which occur when the heart isn't receiving adequate oxygen. As per a survey nearly 50 percent of patients, however, have no symptoms until a heart attack occurs. A number of factors have been shown to increase the risk of developing CVD. Some of these are [2]:Low level of HDL(good) cholesterol.

- Family history of cardiovascular disease
- High levels of LDL (bad) cholesterol
- Hypertension
- High fat diet
- Lack of regular exercise
- Obesity

With so many factors to analyze for a diagnosis of heart disease, physicians generally make a diagnosis by evaluating a patient's current test results. Previous diagnoses made on patients with the same results are also examined by physicians. These complex procedures are not easy. Therefore, a physician must be experienced and highly skilled to diagnose heart disease in a patient.

Data mining has been heavily used in the medical field, to include patient diagnosis records to help identify best practices. The difficulties posed by prediction problems have resulted in a variety of problem-solving techniques. For example, data mining methods comprise artificial neural networks and decision trees, and statistical techniques include linear regression and stepwise polynomial regression [2].

It is difficult, however, to compare the accuracy of the techniques and determine the best one because their performance is data-dependent. A few studies have compared data mining and statistical approaches to solve prediction problems. The comparison studies have mainly considered a specific data set or the distribution of the dependent variable.

This paper is organized as follows: A brief overview of related work will be given in section 2. MLP and Improved MLP prediction models are presented in section 3.section 4 describes about data source. Section 5 and section 6 contains results and conclusion of our study.

2. BACKGROUND

Up to now, several studies have been reported that have focused on cardio vascular disease diagnosis. These studies have applied different approaches to the given problem and achieved high classification accuracies, of 77% or higher, using the datasets taken from the UCI machine learning repository. Here are some examples:

- Robert Detrano's experimental results showed correct classification accuracy of approximately 77% with logistic-regression derived discriminate function [3].
- Zheng Yao applied a new model called R-C4.5 which is based on C4.5 and improved the efficiency of attribution selection and partitioning models. An experiment showed that the rules created by R-C4.5s can give health care experts clear and useful explanations [4].
- Resul Das introduced a methodology that uses SAS base software 9.13 for diagnosing heart disease. A neural networks ensemble method is at the center of this system [5].
- Colombetet al. evaluated implementation and performance of CART and artificial neural networks comparatively with a LR model, in order to predict the risk of cardiovascular disease in a real database [6].
- EnginAvci and Ibrahim Turkoglu study an intelligent diagnosis system based on principle component analysis and ANFIS for the heart valve diseases [7].
- Imran Kurt , MevlutTure , A. TurhanKurum compare performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease [8].
- The John Gennari's CLASSIT conceptual clustering system achieved a 78.9% accuracy on the Cleveland database [9].
- A medical diagnosis decision system (MDDSS) based on SVM has been established to intellectively diagnose 4 types of acid-base disturbance by Lei Guoet al. [14].
- A rule-based decision support system was presented by Tsipouras M.G. et al. [15] for the diagnosis of coronary artery disease.Ten fold cross validation was employed and the average sensitivity and specificity obtained was 80% and 65% respectively.
- A multi-layer perceptron based decision support system is developed by Hongmei Yana et al. [16] to support the diagnosis of heart diseases.
- A decision support system that classifies the Doppler signals of the heart valve to two classes (normal and abnormal) was presented by EmreComaket al. [17] to support the cardiologist. They aimed to develop their previous work by using least-squares support vector machine (LS-SVM) classifier instead of ANN.
- A computational model based on a multilayer perceptron (MLP) neural network withthree layers was employed by Hongmei Yan et al. [31] to develop a decision support system for the diagnosis of five major heart diseases.. The experimental results have shown that the adopted MLP-based decision model can achieve high accuracy level (63.6-82.9%) on the classification of heart diseases, qualifying it as a good decision support system deployable in clinics.

3. CVD PREDICTION MODELS

3.1MLP Neural Network

Artificial neural networks (ANNs) are commonly known as biologically inspired, highly sophisticated analytical techniques, capable of modeling extremely complex non-linear functions. ANNs are analytic techniques modeled after the processes of learning in the cognitive system and the neurological functions of the brain and capable of predicting new observations from other observations (on the same or other variables) after executing a process of so-called learning from existing data. One of popular ANN architecture is called multi-layer perceptron (MLP) with back-propagation (a supervised learning algorithm). The MLP is known to be a powerful function approximator for prediction and classification problems. It is arguably the most commonly used and well-studied ANN architecture. Given the right size and the structure, MLP is capable of learning arbitrarily complex nonlinear functions to arbitrary accuracy levels. The MLP is essentially the collection of nonlinear neurons (perceptrons) organized and connected to each other in a feedforward multi-layer structure. Fig 2 shows MLP feed forward Neural Network.

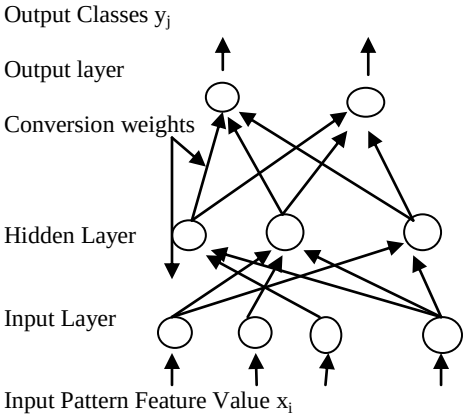


Fig 2: MLP

There is one neuron in the input layer for each predictor variable. In the case of categorical variables, *N*-1 neurons are used to represent the *N* categories of the variable.

Input Layer —A vector of predictor variable values ($x_1...x_p$) is presented to the input layer. The input layer standardizesthese values so that the range of each variable is -1 to 1. The input layer distributes the values to each of the neurons in the hidden layer. In addition to the predictor variables, there is a constant input of 1.0, called the *bias* that is fed to each of the hidden layers; the bias is multiplied by a weight and added to the sum going into the neuron.

Hidden Layer —Arriving at a neuron in the hidden layer, the value from each input neuron is multiplied by a weight (w_{ji}), and the resulting weighted values are added together producing a combined value u_j. The weighted sum (u_j) is fed into a transfer function, σ, which outputs a value h_j. The outputs from the hidden layer are distributed to the output layer.

Output Layer — Arriving at a neuron in the output layer, the value from each hidden layer neuron is multiplied by a weight (w_{kj}), and the resulting weighted values are added together producing a combined value v_j. The weighted sum (v_j) is fed into a transfer function, σ, which outputs a value y_k. The *y* values are the outputs of the network[12].

$Y_j = f(\Sigma W_{ji} X_i)$

The back-propagation algorithm can be employed effectively to train neural networks; it is widely recognized for applications to layered feed-forward networks, or multi-layer perceptrons. The back-propagation algorithm is capable of adjusting the network weights and biasing values to reduce the square sum of the difference between the given output (X) and an output values computed by the net (X ') with the aid of gradient decent method as follows:

The back-propagation algorithm consists of four steps:

1. Compute how fast the error changes as the activity of an output unit is changed. This error derivative (EA) is the difference between the actual and the desired activity.

$$EA_j = \frac{\partial E}{\partial yj} = y_j - dj$$

2. Compute how fast the error changes as the total input received by an output unit is changed. This quantity (EI) is the answer from step 1 multiplied by the rate at which the output of a unit changes as its total input is changed.

3. Compute how fast the error changes as a weight on the connection into an output unit is changed. This quantity (EW) is the answer from step 2 multiplied by the activity level of the unit from which the connection emanates.

$$EI_j = \frac{\partial E}{\partial xJ} = \frac{\partial E}{\partial yj} \times \frac{\partial yj}{\partial xj} = EA_j y_j(1-y_j)$$

$$EW_{ij} = \frac{\partial E}{\partial Wij} = \frac{\partial E}{\partial xJ} \times \frac{\partial xj}{\partial Wij} = EI_j y_i$$

4. Compute how fast the error changes as the activity of a unit in the previous layer is changed. This crucial step allows back propagation to be applied to multilayer networks. When the activity of a unit in the previous layer changes, it affects the activities of all the output units to which it is connected. So to compute the overall effect on the error, we add together all these separate effects on output units. But each effect is simple to calculate. It is the answer in step 2 multiplied by the weight on the connection to that output unit.

$$EA_i = \frac{\partial E}{\partial yi} = \sum_j \frac{\partial E}{\partial xj} \times \frac{\partial xj}{\partial yi} = \sum_j EIjWij$$

By using steps 2 and 4, we can convert the EAs of one layer of units into EAs for the previous layer. This procedure can be repeated to get the EAs for as many previous layers as desired. Once we know the EA of a unit, we can use steps 2 and 3 to compute the EWs on its incoming connections

3.2 Improved MLP Neural Network

MLP algorithm was improved by dividing its training dataset on multiple subsets. Then MLP algorithm was called individually for each subset and results obtained were combined using voted combiner with majority probability rule. Finally results were validated on testing dataset. Steps of Improved MLP in WEKA are as:

(1) On the WEKA screen for Improved MLP algorithm adda new field to capture number of training subset in program. Get separate training and testing data subset using 10 fold cross validation technique.

(2) In program standard classes related to core capabilities and instance are added.

(3) Declare MLPMod class which inherits properties of standard class ‘Classifier’ declare various fields to be used on screen of MLP program.

(4) Ensure that there is some value of number of training on screen subsets. Get and set values of fields entered

(5) In build classifier, delete records from training data which has missing class value. Create an object using voted with majority probability rule to combine result of MLP classifier applied on training subset of data.Divide training data into parts on basis of number of subset entered on input screen. Make different subsets of dataset on basis of stratification. This will ensure that all subsets are of similar kind. Create a loop which will be run as many times as number of data subset. On each loop pass, call MLP by passing values on screen.

(6) Check if desired EA has to be achieved or not. If yes than go to step 7 else go to step 5.

(7) End Training subset loop of step 6.

(8) Finally combine result of all classifiers and get final result.

Results applied on heart disease datasets showed that performance of Improved MLP algorithm was better than MLP algorithm.

4. DATA SOURCE

To compare these data mining classification techniques Cleveland, Hungarian and Switzerland cardiovascular disease datasets from UCI repository are used. The Cleveland dataset has 14 attributes and 303 instances. Hungarian dataset has 14 attributes and 294 instances. The Switzerland dataset has 14 attributes and 123 instances. Table 1 below lists attributes of these datasets:

Table 1: Attributes of Cardiovascular datasets

No.	Name	Description
1	Age	Age in years
2	Sex	1 = male, 0 = female
3	Cp	Chest pain type (1 = typical angina, 2 = atypical angina, 3 = non-anginal pain, 4 = asymptomatic)
4	Trestbps	Resting blood sugar (in mm Hg on admission to hospital)
5	Chol	Serum cholesterol in mg/dl
6	Fbs	Fasting blood sugar > 120 mg/dl (1 = true, 0 = false)
7	Restecg	Resting electrocardiographic results (0 = normal, 1 = having ST-T wave abnormality, 2 = left ventricular hypertrophy)
8	Thalach	Maximum heart rate
9	Exang	Exercise induced angina
10	Oldpeak	ST depression induced by exercise relative to rest
11	Slope	Slope of the peak exercise ST segment (1 = upsloping, 2 = flat, 3 = downsloping)
12	Ca	Number of major vessels colored by fluoroscopy
13	Thal	3 = normal, 6 = fixed defect, 7 = reversible defect
14	Num	Class (0 = healthy, 1 = have heart disease)

5. RESULTS

These data mining classification model were developed using data mining classification tool Weka version 3.6. MLP and Improved MLP were applied on datasets.

A distinguished confusion matrix was obtained to calculate sensitivity, specificity and accuracy. Confusion matrix is a matrix representation of the classification results. Table 2 shows confusion matrix.

Table 2: Confusion Matrix

	Classified as Healthy	Classified as not healthy
Actual Healthy	TP	FN
Actual not healthy	FP	TN

The upper left cell denotes the number of samples classifies as true while they were true (i.e., TP), and the lower right cell denotes the number of samples classified as false while they were actually false (i.e., TN). The other two cells (lower left cell and upper right cell) denote the number of samples misclassified. Specifically, the upper right cell denoting the number of samples classified as false while they actually were true (i.e., FN), and the lower left cell denoting the number of samples classified as true while they actually were false (i.e., FP).

Below formulae were used to calculate sensitivity, specificity and accuracy:

Sensitivity = TP / (TP + FN)

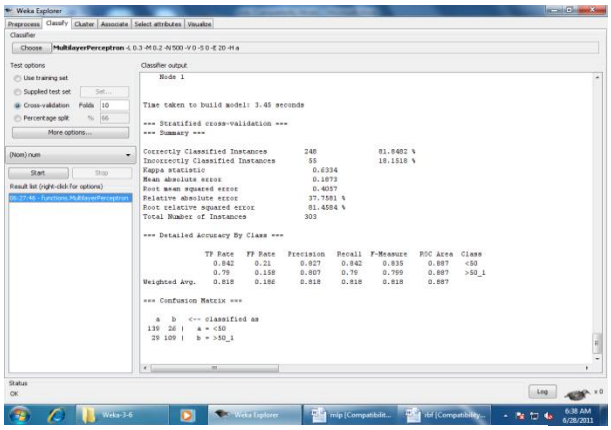
Specificity = TN / (TN + FP)

Accuracy = (TP + TN) / (TP + FP + TN + FN)

We have divided the dataset into different number of subsets. When we took number of subset equal to 0 than Improved MLP works as MLP. As we increases number of subsets than there is gradually increase in performance of Improved MLP. We have achieved maximum performance at number of subsets equal to 15. After that there is gradually decrease in performance which showed that no of samples are not stratified at higher value of number of subsets as we had 303 records in our dataset.

Table 3 and 4 below shows confusion matrix for MLP and Improved MLP Techniques for Cleveland dataset.Table 5 and 6 below show confusion matrix of MLP and Improved MLP Techniques for Hungarian dataset.Table 7 and 8 below show confusion matrix of MLP and Improved MLP Techniques for Switzerland dataset.

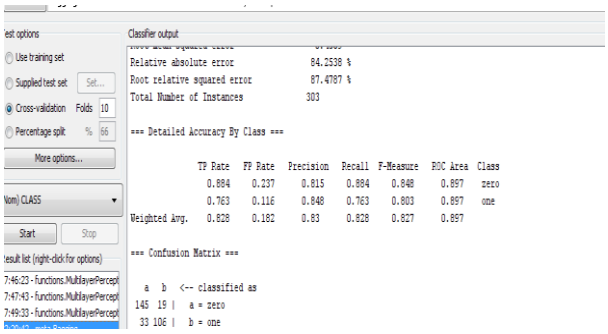
Table 3: MLP Confusion Matrix(Cleveland dataset)



```
=== Confusion Matrix ===

   a   b   <-- classified as
 138  26 |   a = zero
  38 101 |   b = one
```

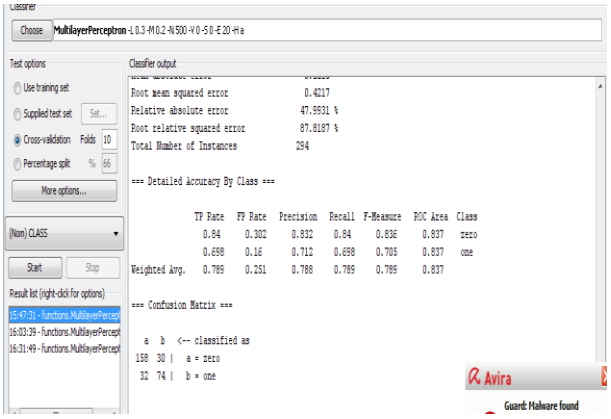
Table 4: Improved MLP Confusion Matrix(Cleveland dataset)



```
=== Confusion Matrix ===

   a   b   <-- classified as
 145  19 |   a = zero
  33 106 |   b = one
```

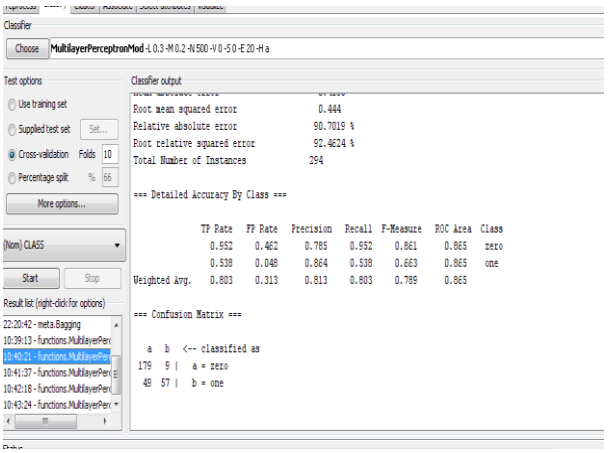
Table 5: MLP Confusion Matrix(Hungarian dataset)



```
=== Confusion Matrix ===

   a   b   <-- classified as
 158  30 |   a = zero
  32  74 |   b = one
```

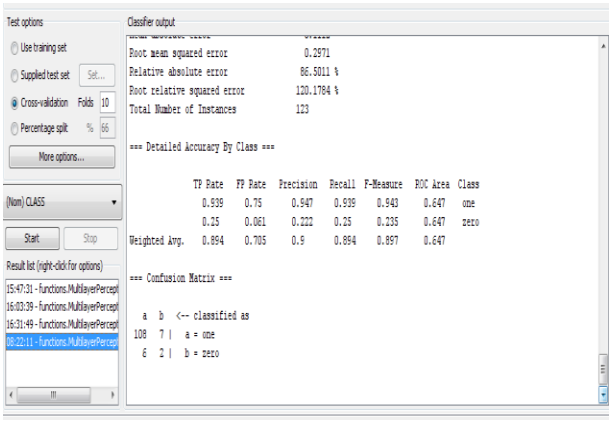
Table 6: Improved MLP Confusion Matrix(Hungarian dataset)



```
=== Confusion Matrix ===

   a   b   <-- classified as
 179   9 |   a = zero
  49  57 |   b = one
```

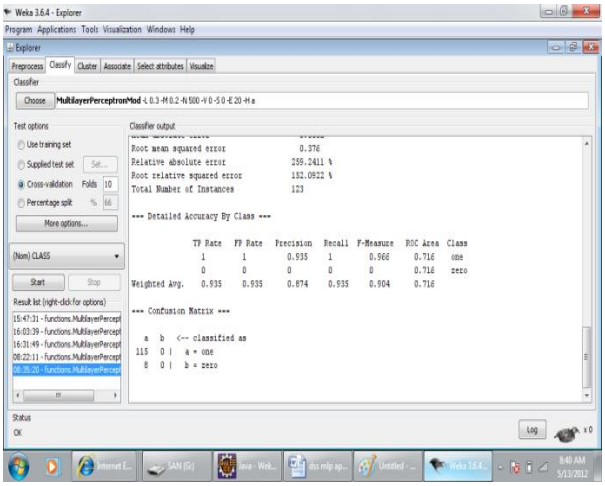
Table 7: MLP Confusion Matrix(Switzerland dataset)




```
=== Confusion Matrix ===

   a   b   <-- classified as
 108   7 |   a = one
   6   2 |   b = zero
```

Table 8: Improved MLP Confusion Matrix(Switzerland dataset)




```
=== Confusion Matrix ===

   a   b   <-- classified as
 115   0 |   a = one
   8   0 |   b = zero
```

Table 9 shows sensitivity, specificity and accuracy of MLP and Improved MLP for Cleveland dataset. Table 10 showssensitivity, specificity and accuracy of MLP and Improved MLP techniques for Hungarian dataset.Table 11shows sensitivity, specificity and accuracy ofMLP and Improved MLP techniques for Switzerland dataset.

Table 9: MLP and Improved MLP Results on Cleveland Dataset.

	Sensitivity	Specificiy	Accuracy
MLP	84.14%	72.66%	78.87%
Improved MLP	88.41%	72.66%	82.8%

Table 10: MLP and Improved MLP Results on Hungarian Dataset

	Sensitivity	Specificiy	Accuracy
MLP	84.04%	69.81%	78.91%
Improved MLP	95.21%	53.77%	80.73%

Table 11: MLP and Improved MLP Results on Switzerland Dataset

	Sensitivity	Specificiy	Accuracy
MLP	93.91%	25.0%	89.43%
Improved MLP	100.0%	0.0%	93.49%

A Receiver Operating Characteristic (ROC) space is defined by False Positive Rate and True Positive Rate which shows relative trade-off between true positive and false positive.

True Positive Rate = TP / (TP + FN)

False Positive Rate = FP / (FP + TN)

ROC value 1.0 represents 100% True Positive Rate and no False Positive Rate which will be ideal case.

Table 12 shows TPR,FPR,ROC and F Measure ofMLP and Improved MLP techniques for Cleveland dataset. Table 13 shows TPR,FPR,ROC and F Measure ofMLP and Improved MLP techniques for Hungarian dataset.Table 14 shows TPR,FPR,ROC and F Measure ofMLP and Improved MLP techniques for Switzerland dataset.

Table 15,Table 16 and Table 17 show graphical representation of TPR,FPR ROC and F Measure for Cleveland, Hungarian and Switzerland datasets respectively.

Table 12: TPR, FPR, ROC and F measure (Cleveland dataset)

	True Positive Rate	False Positive Rate	ROC	F Measure
MLP	0.781	0..221	0.859	0.788
Improved MLP	0.828	0.182	0.897	0.827

Table13: TPR, FPR, ROC and F measure(Hungarian dataset)

	True Positive Rate	False Positive Rate	ROC	F Measure
MLP	0.789	0.201	0.837	0.781
Improved MLP	0.803	0.251	0.865	0.795

Table 14: TPR, FPR, ROC and F measure(Switzerland dataset)

	True Positive Rate	False Positive Rate	ROC	F Measure
MLP	0.894	0.705	0.647	0.897
Improved MLP	0.935	0.663	0.716	0.904

Table 15: Graphical representation of TPR, FPR, ROC and F Measure(Cleveland dataset)

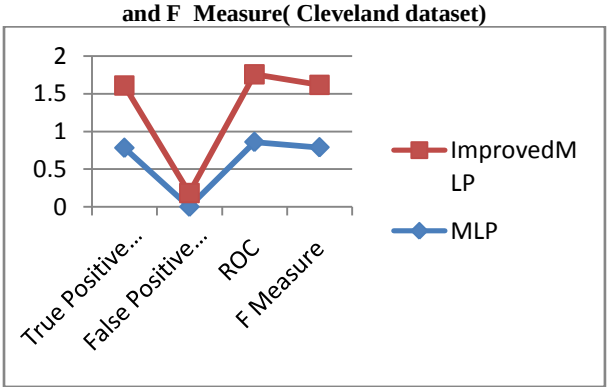


Table 16: Graphical representation of TPR, FPR, ROC and F Measure(Hungarian dataset)

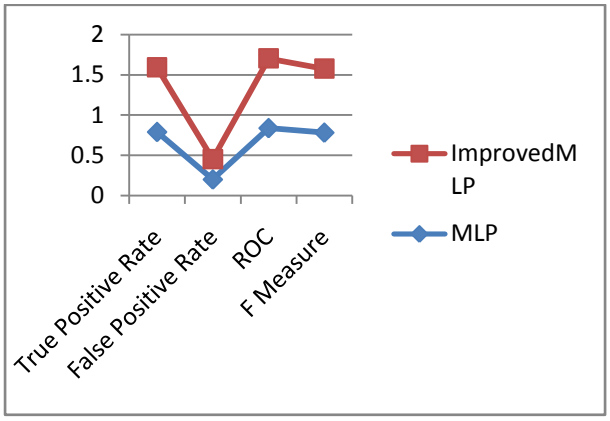
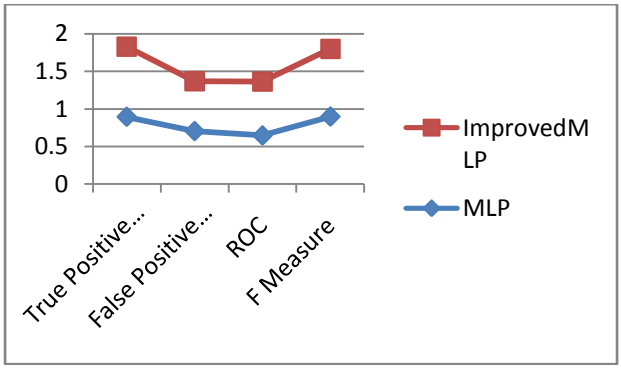


Table 17: Graphical representation of TPR, FPR, ROC and F Measure(Switzerland dataset)



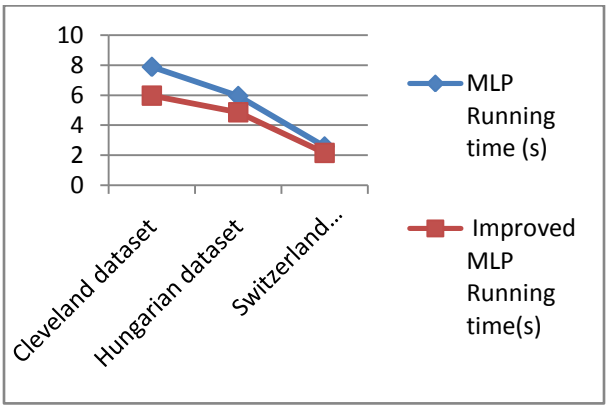
Results of above three simulated experiments show that in all parameters such as TPR,FPR,ROC,F Measure and Accuracy Improved MLP approach out performs MLP. ROC value is also closer to1.0.

Table 18 shows Running Time of MLP and Improved MLP on (Cleveland, Hungarian and Switzerland datasets).Table 19 shows graphical representation of Running Time of MLP and Improved MLP on Cleveland, Hungarian and Switzerland datasets.

Table 18: Running Time of MLP and Improved MLP on (Cleveland, Hungarian and Switzerland datasets)

	Cleveland dataset	Hungarian dataset	Switzerland dataset
MLP Running time (s)	7.89	5.95	2.59
Improved MLP Running time(s)	5.97	4.86	2.16

Table 19:Graphical Representation ofRunning Time for MLP and Improved MLP on (Cleveland, Hungarian and Switzerland datasets)



Results of above simulatedexperiments showthat ImprovedMLP approach significantlyoutperforms MLP approach in overall execution time. As size of dataset increases this effect becomes more significant.So the Improved algorithm proposed in this paper is feasible.

6. CONCLUSION

In this paper MLP and Improved MLP data mining classification techniques are used to predict cardiovascular disease in patients. Here an independent evaluation is done using 10V fold cross validation .Improved MLP model have more prediction power than MLP. Our results showed that Improved MLP algorithm gives better performance than MLP on parameters TPR,FPR,F Measure, ROC, sensitivity and Accuracy. Efficiency of MLP is not too much high. Results of above three simulated experiments show that in all parameters such as TPR,FPR,ROC,F Measure and Accuracy Improved MLP approach out performs MLP. ROC value is also closer to1.0.

This paper presented a simple and efficient way to improve execution time of MLP by increasing accuracy .So the proposed Improved MLP method is feasible. In future we will try to analyze Improved MLP on large datasets and enhance the performance of MLP. Limitation of our study: Drawing all conclusions from data that are in part collected for decision making should be done with caution. Also newly identified cases still need to be validated to confirm their positivity.

7. REFERENCES

[1] Frank Lemke, Johann-Adolf Müller, "Medical Data Analysis Using Self-Organizing Data Mining Technologies", Systems Analysis Modelling Simulation, Vol 43, No: 10, pp: 1399-1408, 2003.

[2] Khemphila, A; Boonjing, V. "Comparingperformance of logistic regression, decision trees and neural networks for classifying heart disease patients".Proceedings of International Conference on Computer Information System and Industrial Management Applications 2010, p 193 – 198.

[3] Detrano, R.; Steinbrunn, W.; Pfisterer, M. (1987). "International application of a new probability algorithm or the diagnosis of coronary artery disease". American Journal of Cardiology, Vol. 64, No. 3, 1987, p 304-310.

[4] Yao, Z.; Lei, L.; Yin, J. (2005). "R-C4.5 Decision tree model and its applications to health care dataset". Proceedings of International Conference on Services Systems and Services Management 2005, p 1099-1103.

[5] Das, R.; Abdulkadir, S. (2008). "Effective diagnosis of heart disease through neural networks ensembles". Elsevier, 2008.

[6] Colombet, I.; Ruelland, A.; Chatellier, G.; Gueyffier, F. (2000). "Models to predict cardiovascular risk: comparisoof CART, multilayer perceptron and logistic regression". Proceedings of AMIA Symp 2000, p 156-160.

[7] Avci, E.; Turkoglu, I. (2009). "An intelligent diagnosis system based on principle component analysis and ANFIS for the heart valve diseases". Journal of Expert Systems with Application, Vol. 2, No. 1, 2009, p 2873-2878.

[8] Kurt, I.; Ture, M.; Turhan, A. (2008). "Comparing performances of logistic regression, classification and regression tree, and neural networks for predicting coronary artery disease". Journal of Expert Systems with Application, Vol. 3, 2008, p 366-374.

[9] Gennari, J. (1989). "Models of incremental concept formation". Journal of Artificial Intelligence, Vol. 1,1989, p 11-61.

[10] Cohen, W. (1995). "Fast effective rule induction". Proceedings of International Conference on machine Learning 1995, p 1-10.

[11] Chau, M.; Shin,D. (2009). "AComparative Study of Medical Data Classification Methods Based on Decision Tree and Bagging Algorithms". Proceedings of IEEE International Conference on Dependable, Autonomic and Secure Computing 2009, p 183-187.

[12] Patil, S.; Kumaraswamy, Y. (2009). "Intelligent and effective Heart Attack prediction system using data mining and artificial neural networks". European Journal of Scientific Research, Vol. 31, 2009, p 642-656.

[13] Han, J.; Kamber, M. (2006). "Data Mining Concepts and Techniques". 2nd Edition, Morgan Kaufmann,SanFrancisco.

[14] Lei Guo, Youxi Wu, Weili Yan, XueqinShen, Ying Li, "Research on Medical Diagnosis Decision Support System for Acid-base Disturbance Based on Support Vector Machine", proc. of the IEEE 27th Annual International Conference of the Engineering in Medicine and Biology Society, pp: 2413-2416, 2006.

[15] Tsipouras M.G., Exarchos, T.P., Fotiadis D.I., Kotsia A., Naka A. and Michalis L.K.,"A Decision Support System for the Diagnosis of Coronary Artery Disease", 19th

IEEE International Symposium on Computer-Based Medical Systems, pp: 279-284, 2006.

[16] Hongmei Yana, YingtaoJiangb, Jun Zhenge, ChenglinPengc, and Qinghui Lid, "A multilayer perceptron-based medical decision support system for heart disease diagnosis", Expert Systems with Applications, Vol 30, No: 2, pp: 272-281, 2006.

[17] EmreComak, AhmetArslan, _brahimTürkoglu, "A decision support system based on support vector machines for diagnosis of the heart valve diseases", Computers in Biology and Medicine, Vol 37, Issue 1, pp: 21-27, January, 2007.

[18] Hongmei Yan, Jun Zheng, Yingtao Jiang, ChenglinPeng, Qinghui Li, "Development of a decision support system for heart disease diagnosis using multilayer perceptron", Proceedings of the 2003 International Symposium on Circuits and Systems, Vol 5, pp: 709-712, 2003.