International Journal of Technology Research and Management
ISSN (Online): 2348-9006
Vol 3 Issue 4 April 2016

Predicting Heart Disease using Regression based Artificial Neural Network

Dipanti Marothiya[1], Prof. Abhishek Raghuvanshi[2]
PG Scholar, Department of Information Technology, Mahakal Institute of Technology, Behind Air Strip, Dewas Road, Ujjain (M.P)-456001[1]
Associate Professor & Head, Department of Information Technology, Mahakal Institute of Technology, Behind Air Strip, Dewas Road, Ujjain (M.P)-456001[2]
Dipanti.m@gmail.com[1]

Abstract: *The data mining and their different applications are becomes more popular now in these days a number of large and small scale applications are developed with the help of data mining techniques i.e. predictors, regulators, weather forecasting systems and business intelligence. There are two kinds of model are available for namely supervised and unsupervised. The performance and accuracy of the supervised data mining techniques are higher as compared to unsupervised techniques therefore in sensitive applications the supervised techniques are used for prediction and classification. In this presented work the supervised learning based application is proposed and demonstrated. The proposed work is intended to demonstrate the data mining technique is disease prediction systems in medical domain. In order to perform this task the heart disease based data is selected for analysis and prediction. the proposed technique of heart disease prediction is a hybrid model of data mining which is combination of the two different predictive model namely regression analysis and the neural network. Here the regression analysis is used to identify the outliers of the data set and using the refined data set the neural network performs training. In addition of that for increasing the performance of the training of the system in terms of training time the validation set modeling is used. The implementation of the proposed working model is provided using the visual studio environment. After the implementation the performance of the system is estimated in terms of accuracy, error rate, memory consumption and time consumption. In addition of that a comparative study with the genetic neural network is also performed with the similar parameters. According to the experimental comparative study the proposed technique performs better as compared to similar data model.*

Keywords: *Heart disease prediction, Neural network, regression analysis, data mining, machine learning.*

1. INTRODUCTION

The data mining is a technique is analysing the data and extracting the essential patterns from the data. These patterns are used with the different applications for making decision making and prediction related task. The decision making and prediction is performed on the basis of the learning of algorithms. The data mining algorithms supports both kinds of learning supervised and unsupervised. In unsupervised learning only the data is used for performing the learning and in supervised technique the data and the class labels both are required to perform the accurate training. In supervised learning the accuracy is maintained by creating the feedbacks form the class labels and enhance the classification performance by reducing the error factors from the learning model.

The proposed work is intended to investigate these techniques in the application of the predictions. Therefore the heart disease prediction system is proposed to develop and implement. The proposed heart disease prediction system utilizes the aspects of the supervised learning for predicting the class labels of the input pattern of heart dataset samples. The proposed predictive data mining technique is being developed in the hybrid manner for predicting accurate class labels. Because the hybrid classifier includes the goodness of both kind of classifiers in the same place and improve the classification performance.

The data mining and their techniques offers to analyse the historical data patterns and learn using the learning over the previous patterns these techniques identify the upcoming patterns. That learning of the data mining techniques can be performed using the supervised techniques or in unsupervised approaches. The supervised learning techniques are known as the classification and prediction algorithms and the unsupervised algorithms are known as the clustering or categorization algorithms. The presented work is an effort for developing the supervised learning technique. The supervised learning techniques are developed in two major modules first the training of data model and then testing or classification of newly arrived patterns.

There are a number of applications exist which are developed in using the supervised learning approaches such as the stock market price forecasting, weather prediction systems and others. In this presented work the heart disease prediction system is introduced with the help of hybrid classification technique. The proposed technique involves the approaches of liner regression and the neural network for training and testing of the heart disease datasets. In this system the effort for reducing the neural network training time is also made. Using the validation set and separate training set the learning time is improved. Furthermore for justifying outcomes of the proposed hybrid neural network the comparative study among the previously available algorithm is also performed.

2. PROPOSED TECHNIQUE

To design a new system and performing the comparative study among both the models traditional genetic neural network [1] and the proposed regression based neural network the following simulation model is proposed as given in figure 1. According to the given diagram the input training set is produced in the initial steps.

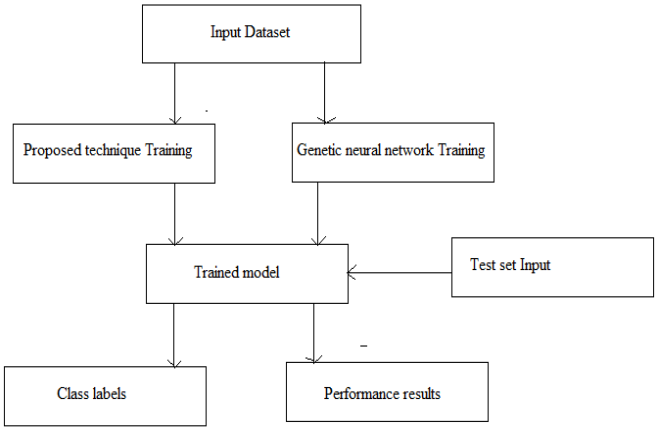


Figure 1: proposed simulation methodology

The input training set is accepted by the implemented supervised learning algorithms and the training of the algorithms are conducted. After training of the models the trained data model is prepared which can be used for identification of patterns. Therefore after training need to input a separate testing set, the test set is processed using the trained data model and their outcomes are recorded. There are two different outcomes are generated by the system first the test dataset's class labels and the classification performance according to the predictive outcomes of the implemented systems.

A. Proposed regression based neural network

In order to develop an improved classification technique there two key main aims of the classifier.

1. **Performance improvement of the neural network:** the performance improvement leads to improve the detection rate of the implemented classifier.

2. **Reducing the training time:** during the pattern learning a significant amount of time is consumed for training therefore that is required to regulate the training of the back propagation neural network.

In order to improve the performance of the traditional back propagation neural network the following improvements are made as demonstrated in figure 2.

Input training set: the supervised learning models needs to be train before making use for classification and prediction task. Therefore a training set of historical patterns are produced as input for performing training. The training set of data contains the different participating attributes and the predefined class labels for training. In the similar manner a training dataset is prepared as given in [1].

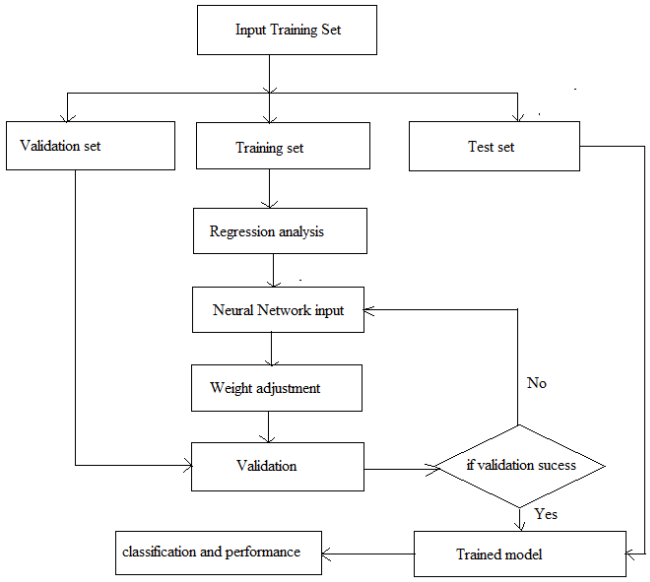


Figure 2: proposed classification technique

And using the defined data set the input data for the system is prepared. After input the training dataset is divided in three key parts as:

1. **Validation set:** the validation set is the 10% of the entire input training dataset. For preparing the validation set for the proposed working model the 10% of data is randomly selected. That dataset is used for validation of the learning model after 100 epoch cycles. If the model trained before the input epoch cycles then it reduces the training time of the neural network.

2. **Training set:** that is the 70% of the actual dataset input which is prepared by randomly selection of the entire input dataset. That set of data is actually used for training of the proposed neural network.

3. **Test set:** For the testing of the proposed working regression based classification technique the 30% of entire dataset is randomly selected for performing the testing of the developed data model. Based on the testing of the test set the performance of the trained classifier is measured.

Regression analysis: Regression analysis is a statistical method of finding relationship between the data attributes. In linear analysis the data is modelled using linear approximation function. The key advantage of working with the regression analysis is that, it allows additional inputs and outputs relevant to statistical analysis. The outcomes of the linear regression is a least-squares estimator, lower confidence bounds, upper confidence bounds, residuals, matrix of intervals, and the statistics (i.e. the R2 statistic, the F statistic and p value, and error variance).

Basically in simple linear regression analysis, initially dataset is a set of n points, in this context an independent variable x_i, and two parameters, β_0 and β_1 are used for class predations thus the relationship is developed linearly. And these n points are going to fit in a straight line therefore the:

Straight line can be defined using $y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1, \dots, n$……. (1)

Adding a term in x_i^2 to the preceding regression gives:

Parabola: $y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \varepsilon_i \quad i = 1, \dots, n$ ..... (2)

This is still linear regression; although the expression on the right hand side is quadratic in the independent variablex_i, it is linear in the parameters$\beta_0, \beta_1 \; and \; \beta_2$. In both cases, ε_i is an error term and the subscript i indexes a particular observation. Given a random sample from the population, we estimate the population parameters and obtain the sample linear regression model:

$y'_i = \beta_0 + \beta_1 x_i$..... (3)

The residual, $e_i = y_i - y'_i$ is the difference between the value of the dependent variable predicted by the modely'_i, and the true value of the dependent variable y_i. One method of estimation is ordinary least squares. This method obtains parameter estimates that minimize the sum of squared residuals.

Therefore the regression analysis is made using the function:

$[\beta_0, \beta_1, R, R_i, stat] = regress(D)$..... (4)

Where β_0 is used for least-squares estimator, confidence intervals are defined by β_1. R is stands for residuals and matrix of intervals is given by R_i. $stat$ is used for statistics.

Neural network input: the back propagation neural network is works in three major steps

1. **Initialization:** in this phase the neural network is initialized therefore the neural network need to define the number of neurons in the input layer, number of hidden layers and the output layer's neurons. In addition of that the network needs to have some additional parameters on which the training of the network is depends such as the number of epoch cycles and learning rates. Using the dimensions of the neural network the network is initialized with the random weights and in further using the adjustments of weights the network is trained.

2. **Weight estimation:** in this phase using the input values the weights of the network are computed. The computed weights are forwarded in next neuron layers and combined weights are generated at the end of output layer. The generated weights of the output layer are treated with the activation function and the outcome of the network is estimated.

3. **Error computation and weight correction:** in this phase the output of neural network is compared with the predefined class labels or the actual outcomes of the network if the difference among the compute output and actual outcomes are higher enough then the error correction using the weight estimation and update is performed.

A traditional neural network model is defined as follows:

Here are some situations where a BP NN might be a good idea [22]:

- A large amount of input/output data is available, but you're not sure how to relate it to the output.
- The problem appears to have overwhelming complexity, but there is clearly a solution.
- It is easy to create a number of examples of the correct behavior.
- The solution to the problem may change over time, within the bounds of the given input and output parameters (i.e., today 2+2=4, but in the future we may find that 2+2=3.8).
- Outputs can be "fuzzy", or non-numeric.

The implementation of neural network is defined in two phases' first training and second prediction: training method utilizes data and designs the data model. By this data model next phase prediction of values is performed [23].

Training:

1. Prepare two arrays, one is input and hidden unit and the second is output unit.

2. Here first is a two dimensional array W_{ij} is used and output is a one dimensional array Y_i.

3. Original weights are random values put inside the arrays after that the output is given as**.**

$$x_j = \sum_{i=0} y_i W_{ij}$$

Where, y_i is the activity level of the j^{th} unit in the previous layer and W_{ij} is the weightof the connection between the i^{th} and the j^{th} unit.

4. Next, action level of y_i is estimated by sigmoidal function of the total weighted input.

$$y_i = \left[\frac{e^x - e^{-x}}{e^x + e^{-x}}\right]$$

When event of the all output units have been determined, the network calculates the error (E) given in equation**.**

$$E = \frac{1}{2}\sum_i (y_i - d_i)^2$$

Where, y_i is the event level of the j^{th} unit in the top layer and d_i is the preferred output of the j_i unit.

Calculation of error for the back propagation algorithm is as follows:

- Error Derivative (EA_j) is the modification among the real and desired target:

$$EA_j = \frac{\partial E}{\partial y_j} = y_j - d_j$$

- Error Variations is total input received by an output changed

$$EI_j = \frac{\partial E}{\partial X_j} = \frac{\partial E}{\partial y_j} X \frac{dy_j}{dx_j} = EA_j y_j (1 - y_i)$$

- In Error Fluctuations calculation connection into output unit is required:

$$EW_{ij} = \frac{\partial E}{\partial W_{ij}} = \frac{\partial E}{\partial X_j} = \frac{\partial X_j}{\partial W_{ij}} = EI_j y_i$$

- Overall Influence of the error:

$$EA_i = \frac{\partial E}{\partial y_i} = \sum_j \frac{\partial E}{\partial x_j} X \frac{\partial x_j}{\partial y_i} = \sum_j EI_j W_{ij}$$

Weigh adjustment: in the proposed model a small change after this process is made, therefore after weights adjustment the model is cross validated using the validation set and the approximation is made is the model trained properly or not. The validation process of the proposed model is given in next section.

Validation: the data which separated at the input step as the validation set is used in this phase. Therefore after the 10% of actual training cycles the validation is performed with the help of the validation dataset. If during the validation of the data model the 90% of data is correctly identified then the system stops the training and the next pattern is used further training of the system. To understand the validation process the for example the total epoch cycles are 100 then after the 10 cycles the validation set on the trained model is applied if 90% of the validation data is identified correctly then the model stop the training and the next pattern from the training set is used for further training.

Trained model: after successfully validation of the proposed data model. That is expected the data model is trained successfully and now it can be used for classification and prediction. Therefore in this phase the test set is applied over the trained model and their performance and class labels are predicted.

Classification and performance: according to the test set input the model is evaluated for their performance. Thus the amount of accurately classified patterns, the misclassified patterns and the consumed resources in terms of memory consumption and the training time is computed and reported.

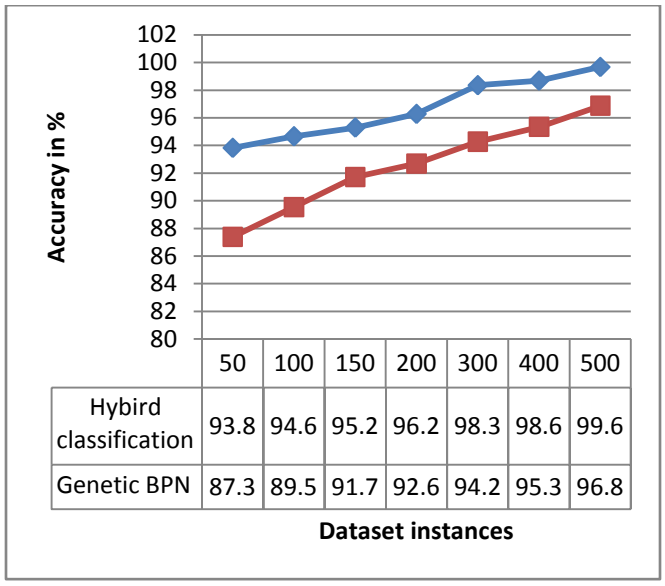
3. RESULTS ANALYSIS

The given section provides the study about the proposed classification algorithm and the comparative performance study among the implemented classifiers in different performance factors. The performance outcomes and the estimated analysis are provided in this chapter.

A. Accuracy

The accuracy is a measurement of the data model for finding the amount of correctly classified data using the input samples. The performance of the algorithm in terms of accuracy can be evaluated using the following formula.

$$accuracy\ \% = \frac{total\ correctly\ classified\ data}{total\ input\ datasets} X100$$



	50	100	150	200	300	400	500
Hybird classification	93.8	94.6	95.2	96.2	98.3	98.6	99.6
Genetic BPN	87.3	89.5	91.7	92.6	94.2	95.3	96.8

Figure 3: accuracy

The performance of the proposed hybrid classifier and the traditional genetic back propagation neural network is compared using the figure 3. In this diagram the X axis shows the training samples in the dataset and the Y axis shows the obtained accuracy in terms of percentage. The results of both the classifiers are demonstrating the different behaviour of classification aspects, in the traditionally implemented classifier the performance of the classification is reduces as the amount of training instances are increases. On the other hand the performance of the proposed classification technique is increases as the amount of training samples are increase. Thus the proposed classifier performs more effectively as compared to traditional manner of classification. For analysing the results in the statistical manner the mean accuracy of both the classifiers are computed and their difference in performance is reported using the figure 4. In this diagram the mean performance of both the method in terms of accuracy is demonstrated using the Y axis and the X axis contains the implemented methods for making comparative performance study. According to the obtained performance the proposed classifier is producing approximately 96% of accurate results and the traditional classifier produces the 92.53 % of accurate results. Thus the proposed classification technique is much efficient and accurate as compared to the traditional technique of data classification.

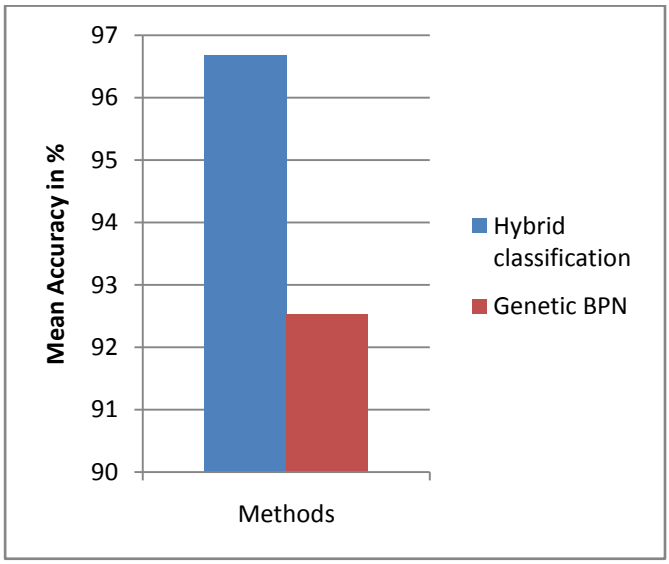


Figure 4: mean accuracy

B. Error rate

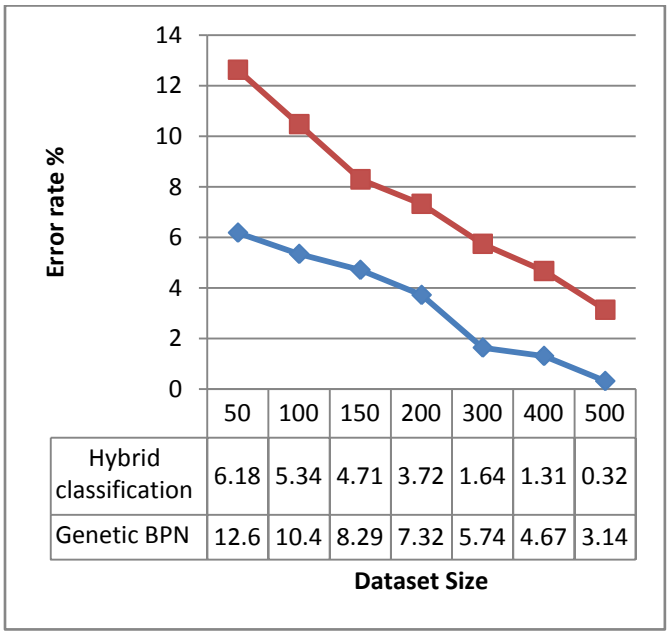
The error rate of the classifier provides the estimation about the misclassified samples during the testing of the trained classifier. The evaluation of error rate can be performed using the following formula.

$$error\ rate\ \% = \frac{total\ misclassified\ samples}{total\ input\ samples} X100$$

Or

$$error\ rate\ \% = 100 - accuracy\ \%$$

The comparative error rate of the proposed and traditional classification technique is provided using the figure 5. The given figure includes the X axis to show the size of training samples and the Y axis shows the amount of misclassified patterns in terms of percentage. According to the demonstrated results the error rate of the proposed classifier is reduces as the amount of training instances are increases in the database. On the other hand the error rate of the traditional scheme is increases as the amount of data for learning is increases. Thus the proposed classifier is improving the outcomes of the classification with increasing the learning patterns. In order to understand the performance of the classification more clearly the mean error rate percentage is evaluated and reported using the figure 6. In this figure the amount of error rate produced by the algorithms are demonstrated using Y axis and X axis shows the methods implemented with the system.



	50	100	150	200	300	400	500
Hybrid classification	6.18	5.34	4.71	3.72	1.64	1.31	0.32
Genetic BPN	12.6	10.4	8.29	7.32	5.74	4.67	3.14

Figure 5: error rate

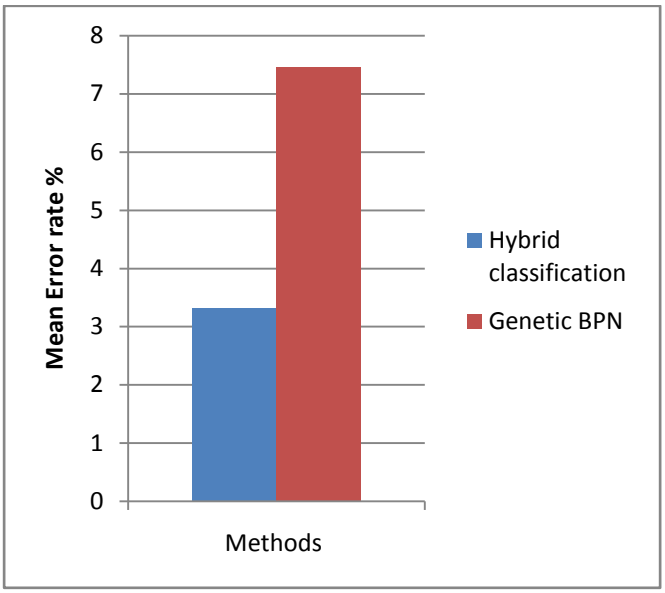
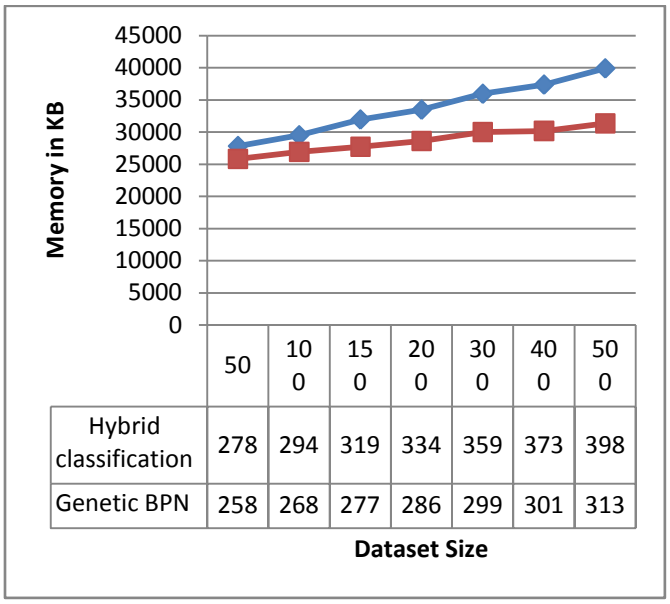


Figure 6: mean error rate

According to the obtained results the from the mean error rate percentage the proposed hybrid classifier produces more effective and improving performance as compared to the traditional classification technique.

C. Memory usages

The amount of main memory required to successfully execute the algorithms is known as the memory consumption of the algorithms. The given figure 7 shows the comparative performance of both the implemented classifiers. In the given diagram the X axis shows the number of training input samples produced for the training to the data models and the Y axis shows the amount of main memory consumed by the implemented algorithms. According to the obtained results the amount of memory consumption in the proposed data modeling is higher as compared to traditional technique because the proposed classifier needs to process the data using both the classifiers.



	50	100	150	200	300	400	500
Hybrid classification	278	294	319	334	359	373	398
Genetic BPN	258	268	277	286	299	301	313

Figure 7: memory usage

In order to understand the memory usage difference among both the classification technique the mean memory consumption is demonstrated using the figure 8 in this diagram the X axis shows the amount of instances of the data used for training and the Y axis shows the amount of main memory consumed during evaluation of data.

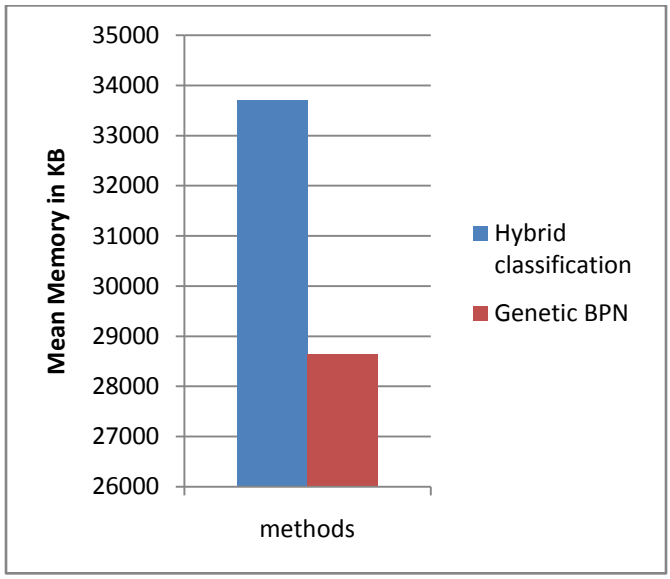
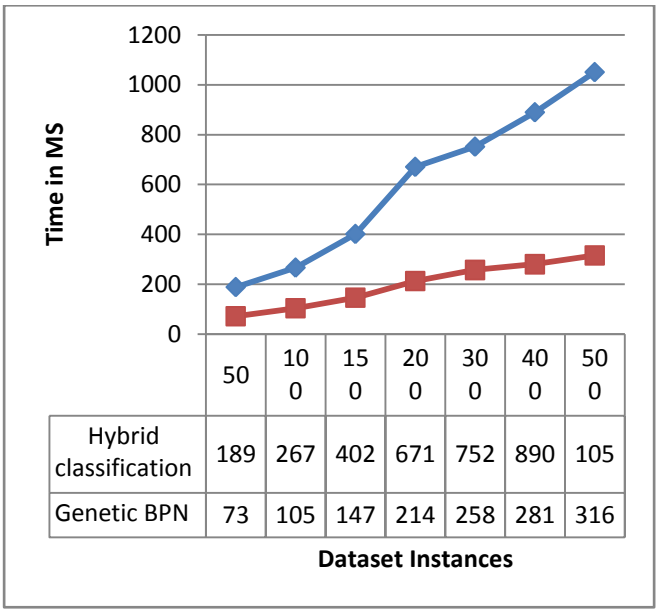


Figure 8: mean memory consumption

D. Time consumption

The amount of time required to process the data using the proposed algorithm is termed here as the time consumption of the system. The comparative time consumption of both the data models during the training is demonstrated using figure 9. In this diagram the X axis contains the amount of data used for training and the Y axis shows the amount of time required to process the data samples. According to the obtained results the proposed technique consumes higher time as compared to the traditional classifier. The proposed scheme utilizes the back propagation neural network and learning of this algorithm is an iterative process thus the amount of time is higher as compared to the genetic neural network.



	50	100	150	200	300	400	500
Hybrid classification	189	267	402	671	752	890	105
Genetic BPN	73	105	147	214	258	281	316

Figure 9: time consumption

In order to understand the difference among both the technique's performance the mean time consumption of both the algorithms are computed. According to the obtained performance the proposed technique consumes more time as compared to the traditional technique. Thus the proposed model is a time consuming model for training time.

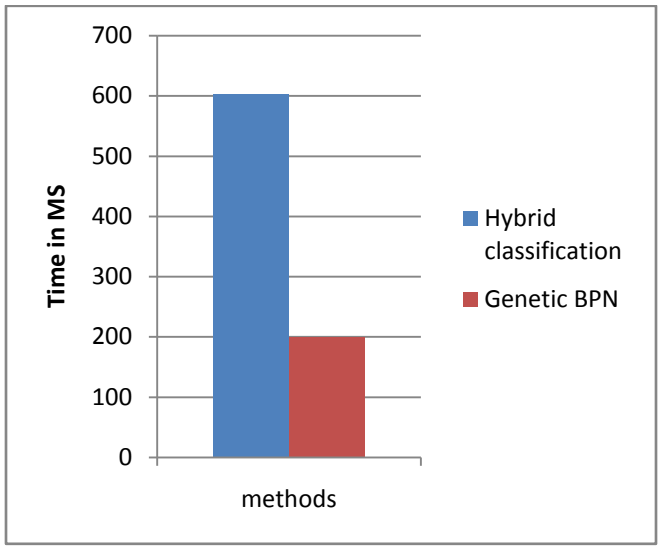


Figure 10: mean time consumption

4. CONCLUSION

The main aim of the performed study is to design and develop an efficient and accurate prediction and classification model for heart disease prediction. Therefore the given chapter provides the summary of the entire work performed and the future work on the basis of the extension possibilities are presented in this chapter.

A. Conclusion

The data mining is helpful for analysing the data, when the manually analysis of the data is not feasible then the data mining techniques are applied for analysis. The data mining techniques are the computer based algorithms which identify the relationship among the data and extraction of the similar pattern data on which they are trained. Basically the data mining techniques are analyse data in two different manner in first the training of the algorithm is not required which is called as the unsupervised learning techniques and the techniques in which the training before use of the algorithm is necessary is known as the supervised learning technique.

In this proposed work the supervised learning technique is demonstrated and implemented. The proposed learning technique is usages the historical heart disease data as training input and after the training from the previous samples the testing of the model is performed for similar pattern detection. The proposed supervised learning model combines the goodness of two different supervised learning approaches i.e. regression analysis and the neural network. Therefore the proposed technique is a supervised hybrid classification technique. In this model the input data is divided in three parts in first the training set, validation set and third the testing set. The training set is processed using the regression analysis technique to find the outliers exist on the training patterns. After removal of outliers the refined data set is used with the neural network for training. During the training of neural network the validation set is used to ensure the learning of the classifier this helps to improve the classification accuracy and the time consumption in neural network learning. Finally the test set is applied for discovering the performance of classification.

The implementation of the proposed technique is performed with the help of visual studio technology. Additionally the performance estimation of the proposed data model is performed with the help of accuracy, error rate, time consumption and the memory consumption. In addition of that for comparing the outcomes of the proposed technique the genetic neural network is also implemented and compared with the proposed model. The summary of the implemented data model's performance is given using table 6.1.

Table 6.1 performance summary

S. No.	Parameters	Proposed	Genetic neural network
1	Accuracy	High	Low
2	Error rate	Low	High
3	Memory consumption	Low	High
4	Time consumption	Low	High

While According to the obtained performance the proposed technique provides efficient and accurate results as compared to the similar classification techniques i.e. genetic neural network. Therefore the proposed predictive algorithm is adoptable for heart disease prediction and similar applications.

B. Future work

The proposed work for modeling and improvement in classification and prediction techniques is prepared successfully. This technique is helps to reduce the training time of the neural network and improve the learning by pre-analysis of the data. Therefore the proposed technique can be used when the accuracy and computational performance is required. Therefore in near future the proposed model is enhanced with the help of more optimization algorithm and the initialization of neural network by which the cost of weight adjustment is also reduced significantly.

REFERENCES

[1] Syed Umar Amin, Kavita Agarwal, Dr. Rizwan Beg, "Genetic Neural Network Based Data Mining in Prediction of Heart Disease Using Risk Factors", Proceedings of 2013 IEEE Conference on Information and Communication Technologies (ICT 2013), 978-1-4673-5758-6/13/$31.00 © 2013 IEEE.

[2] Data Mining: What is Data Mining?, http://www.anderson.ucla.edu/faculty/jason.frand/teacher/technologies/palace/datamining.htm.

[3] Data Mining - Applications & Trends, http://www.tutorialspoint.com/data_mining/dm_applications_trends.htm.

[4] Mahak Chowdhary, Shrutika Suri and Mansi Bhutani, "Comparative Study of Intrusion Detection System", 2014, IJCSE All Rights Reserved, Volume-2, Issue-4.

[5] Mrs. Pradnya Muley, Dr. Anniruddha Joshi, "Application of Data Mining Techniques for Customer Segmentation in Real Time Business Intelligence", International Journal of Innovative Research in Advanced Engineering (IJIRAE) ISSN: 2349-2163, Issue 4, Volume 2 (April 2015).

[6] Ghazaleh Khodabandelou, Charlotte Hug, Rebecca Deneckere, Camille Salinesi, "Supervised vs. Unsupervised Learning for Intentional Process Model Discovery", Business Process Modeling, Development, and Support (BPMDS), Jun 2014, Thessalonique, Greece. pp.1-15, 2014.

[7] Importance of Predictive Analytics in Business, http://www.orchestrate.com/blog/importance-of-predictive-analytics-in-business.

[8] David A. Dickey, N. Carolina State U., Raleigh, NC, "Introduction to Predictive Modeling with Examples", Statistics and Data Analysis, SAS Global Forum 2012.

[9] Hand, Manilla, & Smyth, "Descriptive Modeling", http://www.stat.columbia.edu/~madigan/DM08/descriptive.ppt.pdf.

[10] K.Jayavani, "STATISTICAL CLASSIFICATION IN MACHINE INTELLEGENT", ISRJournals and Publications, Volume: 1 Issue: 1 18-Jul-2014.

[11] Chaitrali S. Dangare, Sulabha S. Apte, "Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", International Journal of Computer Applications (0975 – 888) Volume 47– No.10, June 2012.

[12] Jyoti Soni, Ujma Ansari, Dipesh Sharma, Sunita Soni, "Predictive Data Mining for Medical Diagnosis: An Overview of Heart Disease Prediction", International Journal of Computer Applications (0975 – 8887) Volume 17– No.8, March 2011.

[13] Shadab Adam Pattekari and Asma Parveen, "Prediction System for Heart Disease Using Naive Bayes", International Journal of Advanced Computer and Mathematical Sciences ISSN 2230-9624. Vol 3, Issue 3, 2012, pp 290-294.

[14] N. Aditya Sundar, P. Pushpa Latha, M. Rama Chandra, "Performance Analysis of Classification Data Mining Techniques Over Heart Disease Data Base", International Journal of Engineering Science & Advanced Technology, Volume-2, Issue-3, 470 – 478.

[15] R. Thanigaivel, Dr. K. Ramesh Kumar, "Review on Heart Disease Prediction System using Data Mining Techniques", Asian Journal of Computer Science and Technology (AJCST) Vol.3.No.1 2015 pp 68-74.

[16] M.I. López, J.M Luna, C. Romero, S. Ventura, "Classification via clustering for predicting final marks based on student participation in forums", Proceedings of the 5th International Conference on Educational Data Mining.

[17] Neeraj Shah, Valay Parikh, Nileshkumar Patel, Nilay Patel, Apurva Badheka, Abhishek Deshmukh, Ankit Rathod, James Lafferty, "Neutrophil lymphocyte ratio significantly improves the Framingham risk score in prediction of coronary heart disease mortality: Insights from the National Health and Nutrition Examination Survey-III", International Journal of Cardiology, 2013 Elsevier Ireland Ltd. All rights reserved.

[18] P.K. Anooj, "Clinical decision support system: Risk level prediction of heart disease using weighted fuzzy rules", Journal of King Saud University – Computer and Information Sciences (2012) 24, 27–40.

[19] Nicholas P. Tatonetti, Patrick P. Ye, Roxana Daneshjou, and Russ B. Altman, "Data-Driven Prediction of Drug Effects and Interactions", Published in final edited form as: Sci Transl Med. 2012 March 14; 4(125): 125ra31. doi:10.1126/scitranslmed.3003377.

[20] Peter C. Austin, Jack V. Tu, Jennifer E. Ho, Daniel Levy, and Douglas S. Lee, "Using methods from the data mining and machine learning literature for disease classification and prediction: A case study examining classification of heart

failure sub-types", Published in final edited form as: J Clin Epidemiol. 2013 April; 66(4): 398–407. doi:10.1016/j.jclinepi.2012.11.008.

[21] Suman Bala, Krishan Kumar, "A Literature Review on Kidney Disease Prediction using Data Mining Classification Technique", IJCSMC, Vol. 3, Issue. 7, July 2014, pg.960 – 967.

[22] Shweta Pandey, Prof. Megha Mishra, "Cryptanalysis of Feistel cipher using Back propagation Neural Network", International Journal of Emerging Technology and Advanced Engineering Website: www.ijetae.com (ISSN 2250-2459, Volume 2, Issue 3, March 2012).

[23] Pratik Gite, Sanjay Thakur, "An Effective Intrusion Detection System for Routing Attacks in MANET using Machine Learning Technique", International Journal of Computer Applications (0975 – 8887) Volume 113 – No. 9, March 2015.