Improving Heart Disease Data Classification Using Group Search Optimization

Babu M.[1*], Ramaraj N.[2] and Rajagopalan S.P.[1]
1. GKM College of Engineering and Technology, Chennai, Tamil Nadu, INDIA
2. Vignan's University, Guntur, Andhra Pradesh, INDIA
*babumunirathinam67@gmail.com

Abstract

Cardiovascular disease is one of the major causes of death across the globe and the detection of the disease at an early stage is of utmost importance. Of late, due to the development of technology, early and accurate detection of the disease has become possible and this in turn can help in reducing the risk of the disease and in reducing the number of patients affected. At the same time, diagnosis of the disease is the most critical as well as a complicated problem which has to be accomplished in an accurate and efficient manner. Heart disease data can be analyzed using a neural network approach. The primary goal of data mining is to find the relation between data as well as predicting the result. The prime data-mining technique to classify a provided set of input data is classification. Several practical issues which we face in our day-to-day life in various multifaceted disciplines from business to medicine can be tackled by classification approach. Classification process can be made more efficient by adopting parallel approach in the training phase. Optimization technique can be used for better classification and to improve accuracy. To optimize the weight of the Artificial Neural Network (ANN) structure, Group Search Optimizer (GSO) technique is used.

Keywords: Heart Disease Prediction, data mining techniques, Artificial Neural Network (ANN) structure Group Search Optimizer (GSO) technique.

Introduction

Blood is circulated throughout the body through a muscular organ, which is the heart. Heart disease has become one of the most important causes of death and the death rate has increased in a very short span of time. Diseases associated with the heart are typically called as Cardio Vascular Diseases (CVD). This kind of disease affects population from developing countries rather than developed countries.

Safety and accuracy in diagnosis is the prime factor in healthcare practice as improper diagnosis can at times prove fatal. So, the detection of heart disease is a multifaceted issue that should be free from false assumptions and unpredictable side effects. Heart disease as mentioned earlier is not a single disease but many diverse conditions affecting the circulatory system

An important role is played by data mining in disease prediction. In the medical world, data mining is utilized widely in the prediction and diagnosing of diseases like heart illness, lung cancer, breast cancer etc. The main problem is extraction of information from the patient regarding their symptoms, which in turn will help in the correct diagnosis of the disease condition. Of late, many investigative tests are available which provide the exact information needed in the diagnosis and treatment of the disease. Large scale data which are available by the investigation can be classified accordingly. Data classification is the method of splitting dataset into two or more different classes. The classification is done based on the properties of dataset like number of attributes, instances, values and dataset.

Artificial neural network (ANN) structure resembles that of the brain. A centralized control is lacking in the neural network since all the interconnected processes are modified accordingly with the flow of information. Neural network can solve multifaceted problems and this remains the chief advantage of ANN. Back Propagation (BP) protocol is utilized to solve problems and is designed in a manner so as to decrease error between actual and desired output and to adjust the weight of the ANN and to reduce the bias. Unlike the conventional methods, ANN is capable of handling indistinct functional relations during the learning stage itself.

Particle Swarm Optimization (PSO) algorithm is usually utilized for solving global optimization problems and this is appropriate for handling nonlinear, nonconvex design spaces with dis-continuities as well as is robust with fast convergence characteristics. The suggested method employs PSO trained ANN (NN-PSO) which is able to tackle the problem of prediction of heart diseases.

This technique merges the global scheme of enhanced opposition-based PSO with local searching capacity of conventional back-propagation protocol (BPA) with momentum term. The opposition-based as well as arbitrary perturbation techniques are 2 diversification elements of the protocol. Time variant social as well as cognitive elements enhance the capacity of the protocol to search. Constriction factor is one more variable which ensures convergence. The common problem with large datasets is over-fitting, that is acquiring more than the essential specifications of during training.

Related works

The most effective technique in the diagnosis of disease is based on artificial intelligence. Rao et al[3] used a hybridized protocol that is a combination of GSO as well as ABC, for enhancing the training process of the NN for diagnosing heart diseases efficiently. Initially, a population was obtained for training the neural network. Hybrid algorithm operation was used to identify the appropriate member who fits the criteria to train this network. Fitness of all members are found and the members are categorized for performing the hybrid operations. After performing all the procedures on the members, a new set is chosen and the procedure is repeated until a stable one for producer operations is got. The weight value of the producer is chosen for training the NN so as to detect heart diseases.

Metkari & Pradhan[4] proposed a new approach using GA as well as ANN. To improve the accuracy of independent classifiers, discretization method is used. Genetic algorithm is used because it gives effective classification of heart disease datasets by performing global search in complex, large and multi-modal landscapes and thus provides optimum solution. The goal of the paper is to increase the accuracy in diagnosis of heart disease using the proposed approach as the approach strengthened the classifier in providing improved accuracy and efficiency in data mining.

Due to extensive research, a large volume of medical data is available of late which acts as the source forthe prediction of useful and hidden facts in almost all medical problems. This helps the medical practitioners make timely diagnosis, thus preventing irreparable side effects. ANN has been very helpful in providing best accuracy on medical data. Durairaj & Revathi [5] predicted the existence of heart disease using Back Propagation Multilayer Perception of ANN.

He[6] proposed a new ANN training protocol based on an improved GSO. He proposed to substitute the Gaussian random walk GSO with Levy Flight (LGSO) that proved to enhance the efficiency and accuracy of research sources in unpredictable environments. The first step in this process is for evaluating the enhanced GSO with LGSO on a set of five optimization benchmarks. The LGSO protocol should be applied to the variables of three-layer feed-forward artificial neural network, which includes connection weight as well as biases. Cleveland heart disease classification problem as well as the sunspot number forecasting problem are utilized for assessing the performance of LGSO-trained ANN (LGSOANN). LGSOANN shows as having better convergence as well as generalization performance in both the problems, when compared to the other machine learning methods suggested in literature.

Another method was presented by Yaghini et al[7] which emphasized on the ability of meta-heuristics and greedy gradient based protocols for obtaining a hybridized enhanced opposition based as well as a BP protocol with momentum term. Opposition based learning as well as arbitrary perturbation helps population diversification in iterations. Usage of time varying parameters and constriction factor improves the search capacity of generic PSOA and ensures particle convergence. Over-fitting could be prevented by a novel cross validation technique. Efficacy of the suggested technique was contrasted with other ANN training protocols on other benchmarks.

Methodology

In the proposed method, Cleveland database which was obtained from internet is used for comparison. Cleveland database classifies the samples as normal and abnormal in terms of heart disease, whereas this work uses a different classification to optimize ANN and accuracy, as normal person, 1st, 2nd stroke, as well as end of life. The ANN comprises three layers which are the input, hidden, as well as output layers. In the preparation of methodology, 80% of the dataset is employed for training function and the remaining 20% is utilized for the purpose of authentication of the proposed model. Each interconnected layer is assigned with random weights. Here, the optimization algorithm is employed to obtain the class of heart disease. In this optimization process GSO with ANN performs better than genetic algorithm with ANN.

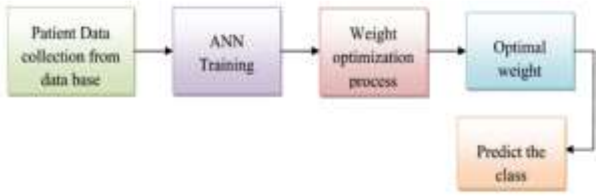

Patient Data collection from data base
ANN Training
Weight optimization process
Optimal weight
Predict the class

Figure 1: Block diagram for overall process

Cleveland database

Database Description: The database was got from Cleveland dataset which is a public data set available on the Internet. Cleveland data set is concerned with the classification of individuals into normal or abnormal with regard to heart diseases. While the databases have 76 raw attributes, only 14 of them are actually used. The output field is a 2 bit value that represents 4 distinct classes as class 0- normal person, 1-1st stroke, 2- 2nd stroke as well as 3- end of life.

Data Representations: Quantity of instances: 414.
Quantity of features: 14 as well as a class feature.

Class: Class 0: Normal Person, Class 1: 1st stroke, Class 2: 2nd stroke, Class 3: end of life

Table 1
Feature Information and class

Feature	Name	Feature	Name
A1	Age	A8	thalach
A2	Sex	A9	exang
A3	Cp	A10	oldpeak
A4	trestbps	A11	slope
A5	chol	A12	ca
A6	Fbs	A13	thal
A7	restecg	A14	num

Dataset for classification process: As per statistics from the WHO, heart diseases are the most common cause of death worldwide. A significant component for classifying heart diseases through usage of ANN is the selection of data. Data is got from 4 distinct data sets of UCI, Centre for Machine Learning and Intelligent Systems.

Classification Algorithm

Artificial Neural Networks (ANN): ANN is typically utilized as a tool for the resolution of several decision modeling problems. ANN is non-parametric and does not make guesses regarding the sharing of data and hence, is excellent at allowing the data to speak for itself. This makes it a perfect option for modeling multi-faceted medical issues when huge datasets of important clinical information are in hand. There are three input layers in ANN, they are: input, intermediate (known as the hidden layer) and output layers. Many hidden layers may be present between the other two layers.

- **Input –** Behavior of input units denote raw data which is provided to the network.
- **Hidden –** Behavior of hidden units are decided by activity of input units as well as the weights on links between input as well as hidden units.
- **Output –** Behavior of output units relies on that of hidden units as well as weights between hidden as well as output units.

Hidden layers accept data from the input layer. Input values are modified through some weight values and the novel value is then sent to output layer which will again be altered by a certain weight from the link between hidden as well as output layer. Output layer processes the information obtained from the hidden layer and generates an output. Outputs are then processed by activation functions [8]. ANNs are flexible as well as adaptive in nature, learning as well as adjusting with various internal or external stimuli. ANNs are utilized in sequence as well as patterns recognition systems, data processing and even in modeling. Figure 2 shows the structure of ANN.

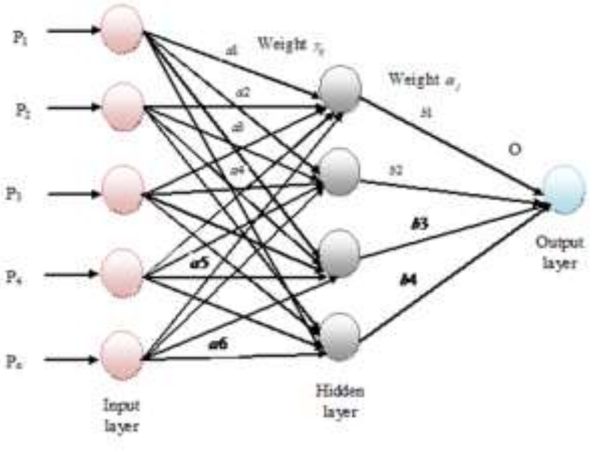


Figure 2: ANN structure

In general, the ANN comprises of three layers which are input, hidden as well as the output layers. All the layers comprise a certain quantity of neurons. Here, every neuron in the input is linked with the hidden layer neurons while those are in turn linked with the output layer with an arbitrary weight. Random weights are assigned to each interconnected layer. In the current work, PSOA is included for supporting ANN training. ANN classifier trained with PSOA is utilized for the prediction of heart diseases.

Naïve Bayes (NB): NB Classifier [9] method has its basis in the Bayesian theorem and is especially appropriate when dimensionality of inputs is great. In spite of its simplicity, it outperforms other protocols. Bayesian classification denotes a supervised learning technique and a statistical technique for classification. It is assumed that there is an underlying probabilistic model that permits the capturing of uncertainty regarding the model in a principled fashion through determination of probabilities of the outputs. It is capable of solving diagnostic as well as predictive issues.

Given training data X, posterior probability of H, P(H|X), that is based on Bayes' principle

$$P(H|X)=P(X|H)P(H)/P(X) \quad (1)$$

Proocol: The NB protocol is based on Bayes' principle as mentioned by (1). The steps in the protocol are given below:
1. All data samples are denoted by n dimensional features vector, X = (x1, x2….. xn), indicating n measures on the sample from n features, which are A1, A2, …, An.
2. Let there be m classes, C1, C2……Cm. For an unknown instance, X, the classifier makes the prediction that X is a part of a class having greatest posterior probability, if and only if:

$P(C_i / X) > P(C_j / X)$ for all 1<=j<=m $and\ j != i$

This maximizes $P(C_i / X)$. The class C_i for which $P(C_i / X)$ is made maximum is known as maximum posteriori hypothesis. By Bayes theorem,
Because P(X) is same for every class, only $P(X / C_i)\mathrm{P}(C_i)$ is to be maximized. If the class prior probabilities are unknown, then it is presumed that the classes have equal probability, i.e., P(C1) = P(C2) = …..= P(Cm), and hence $P(X / C_i)$ will be made maximum. Else $P(X / C_i)\mathrm{P}(C_i)$ will be made maximum.
It is to be noted that the class prior probabilities can be predicted by $\mathrm{P}(C_i)$ = si/s , wherein Si refers to the number of training instances of class C_i while s refers to the total quantity of training instances on X, i.e., the naive probability designates an unknown sample X to class C_i .

Back Propagation (BP): BP is a popular technique to train ANNs and is utilized along with an optimization method like gradient descent. The method calculates gradient of loss functions with respect to every weight in the network.

Provided the quantity of nodes in the input, hidden as well as output layers n, k, m correspondingly, the overall quantity of input instances is x_{pi} that implies the P instance's ith input value, v_{ki} denotes the ith node of input to hidden layer of the kth node weight, ω_{jk} implies the node weight from hidden of the k to the output layer of the j. To make it convenient, threshold is connection weights, as well as the outputs of hidden layer node k is:

$$z_{pk} = f(net_{pk}) = f\left(\sum_{i=0}^{n} v_{ki}x_{pi}\right) \tag{2}$$

Output layer nodes for the node j:

$$y_{pj} = f(net_{pj}) = f\left(\sum_{i=0}^{n} w_{jk}z_{pk}\right) \tag{3}$$

wherein, standard sigmoid function is chosen as incentive function:

$$f(x) = \frac{1}{1+e^{-x}} \tag{4}$$

The global error function may be given as in equation (5):

$$E = \sum_{p=1}^{P} E_P = \frac{1}{2}\sum_{p=1}^{p}\sum_{j=1}^{m}(t_{pj} - y_{pj})^2 \tag{5}$$

where E_P, is error of the sample p, t_{pj} is the ideal result. The adjustment formulae of weights are thus.

The weight adjustment equation of output layer neurons:

$$\Delta\omega_{jk} = \eta\sum_{p=1}^{p}(t_{pj} - y_{pj}).y_{pj}(1-y_{pj}).\mathrm{z}_{pk} \tag{6}$$

wherein, η represents the learning rate, in which general range is 0.1 - 0.3.

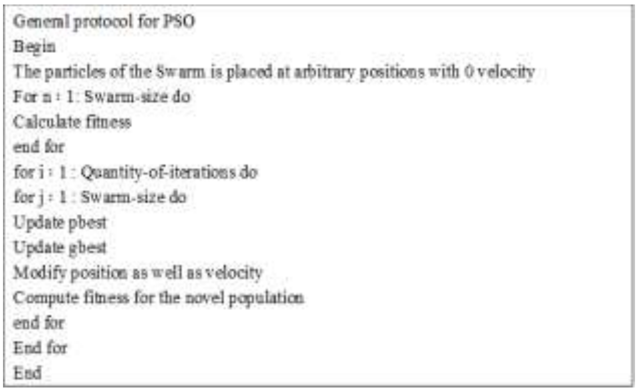
The weight adjustment equation of hidden layer neurons:

$$\Delta v_{ki} = \eta\sum_{p=1}^{p}\left(\sum_{j=1}^{m}\delta_{pi}\omega_{jk}\right)z_{pk}(1-z_{pk})x_{pi} \tag{7}$$

The notion of BP is that the learning procedure may be split into two phases: the first one is the forward propagation procedure and input data is given via a layer by layer processing every hidden layer as well as actual output value of every unit of y_{pj} is computed. The 2nd phase is the reverse procedure, wherein if the output layer does not obtain the anticipated output value, then layer by layer recursive calculation of difference of error between actual as well as anticipated output is done. Gradient descent technique alters weights of Δv_{ki}, $\Delta\omega_{jk}$, ensuring the overall error function is minimal [10].

C4.5: A decision tree is a powerful tool utilized for classifications as well as predictions[11]. Decision tree produces rules that may be inferred by humans and utilized in knowledge systems like databases. C4.5 refers to a protocol for constructing decision trees. It is an expansion of ID3 protocol and is designed by Quinlan. It modifies trained trees into sets of if-then rules. It deals with both discrete as well as continuous features. C4.5 is a popular protocol and it constructs decision trees from a set of training data utilizing the idea of information entropy. It is additionally called a statistical classifier.

Particle Swarm Optimization (PSO): PSO is a famous population-based optimization method. It takes into consideration a set of potential solutions (called as particles) in D-dimensional search space [12]. All particles are related to a fitness value which predicts the particle's capacity to attain the goal. First, the particles are situated arbitrarily in search space while swarm flies within the search space for reaching optimum fitness value. Particles have a corresponding 'pbest' value that accounts for the best solution reached so far by the particle. The best solution reached by a particle in the swarm is called 'gbest' (global best). All the positions as well as velocities of the particles are set arbitrarily. After each iteration, fitness of particles are computed and required alterations are done to the positions as well as velocities of all particles for moving them to optimum fitness.



```
General protocol for PSO
Begin
The particles of the Swarm is placed at arbitrary positions with 0 velocity
For n : 1: Swarm-size do
Calculate fitness
end for
for i : 1 : Quantity-of-iterations do
for j : 1 : Swarm-size do
Update pbest
Update gbest
Modify position as well as velocity
Compute fitness for the novel population
end for
End for
End
```

Group Search Optimizer (GSO): GSO is a population-based optimization protocol as well as utilizes producer-scrounger model as well as animal scanning method. Producer-scrounger as a design of optimal search scheme owes its inspiration to animal searching behaviour as well as group living theories. Two foraging methods which are producing (Searching of food) as well as scrounging (combining resources discovered by others) are adopted by the protocol. So as to not be forced into local minimum, GSO utilizes ranger foraging method. The population of the GSO protocol is referred to as a group and all the individuals are known as members. Three types of members are present in the group: producer, which searches for food, scrounger which joins resources found by others, and ranger, which utilizes random walk strategy for arbitrarily spread resources. At every iteration, member that found the best resource remains as producer, few except the producer are scroungers while the rest are rangers [13.]

The diagrammatic representation of the producer scanning field is shown in the figure 4 below.

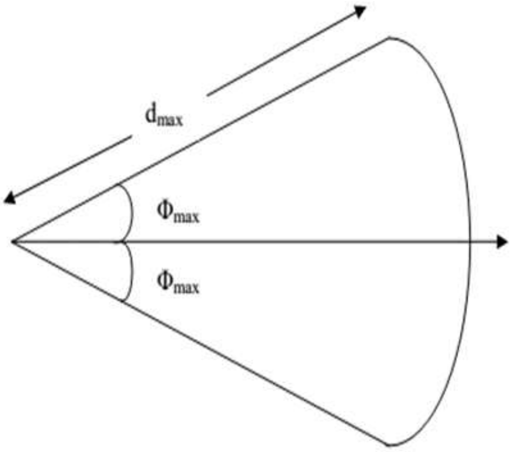


Figure 4: Producer scanning field

When the searching procedure in GSO is executed, the scrounger or the ranger will also have the chances to discover a better location than the present producer or other members during the time the producer fails to discover the better location. A scrounger or a ranger that has found the better location will be the producer in the next searching session and the producer as well as the other members in the prior searching session will now carry out the scrounging activity.

In n-dimension search space, the ith member at kth searching bout (iteration), has a current position $X_i^k \in \mathbb{R}^n$ a head angle $\varphi_i^k = (\varphi_{i1}^k, ..., \varphi_{(n-1)}^k) \in \mathbb{R}^{n-1}$ as well as a head direction $D_i^k(\varphi_i^k) = (\mathrm{d}_{i1}^k, ..., \mathrm{d}_{in}^k) \in \mathbb{R}^n$ that may be computed from φ_i^k through a Polar to Cartesian coordinate alteration:

$$\mathrm{d}_{i1}^k = \prod_{p=1}^{n-1} \cos(\varphi_{ip}^k)$$

$$\mathrm{d}_{ij}^k = \sin(\varphi_{i(j-1)}^k).\prod_{p=1}^{n-1} \cos(\varphi_{ip}^k)$$

$$\mathrm{d}_{in}^k = \sin(\varphi_{i(\mathrm{n}-1)}^k) \tag{8}$$

For accuracy as well as convenience of computations, in GSO, it is presumed that there is solely 1producer at every iteration. It has its basis in studies that proposed that bigger the group, smaller the proportion of informed individuals required for guiding the group with improved precision. Most basic joining policy, that presumes that every scrounger joins the resource discovered by the producer, is utilized.

Group Search Optimizer Artificial Neural Network (GSOANN): In GSO-based training protocol (GSOANN) all members of the population are vectors having connection weights as well as bias values. With no loss of generality, W1 is denoted as the connection weight matrix between input as well as hidden layers, Θ_1 is the bias term to the hidden layer, W2 is between hidden as well as output layer, while Θ_2 is the bias term to the output layer. The ith member in the population may be denoted thus: $X_i = [\mathrm{W}_1^i\ \Theta_1^i\ \mathrm{W}_2^i\ \Theta_2^i]$. The fitness function designated to the ith individual is the least-squared error function given in equation (9):

$$F_i = \frac{1}{2}\sum_{p=1}^{P}\sum_{k=1}^{K}(d_{kp} - y_{kp}^i)^2 \tag{9}$$

where y_{kp}^i denotes the kth calculated output in (13) of ANN for the pth sample vector of the ith member; P represents the total quantity of sample vectors; while d_{kp} refers to anticipated output in the kth output node. It is observedthat minimization of error function is not the same as maximization of generalization [36]. The error on training set may be driven to a very small value by minimizing the error function, however, as a side effect, sometimes the over-fitting problems will occur, which may result in a large generalization error. Hence, for improving ANN performance, earlier stopping strategy is suggested. Error rate of validation set is monitored at the time of training. If validation error rises for a certain set of iterations, training stops.

Results and Discussion

This section details the analysis on the Cleveland dataset with regard to sensitivity, specificity, accuracy as well as time taken for execution with techniques such as NB, ANN, BPP, ANN-PSO structure optimized and GA-GSO. Table 2 shows the summary of results. Figure 5 to 8 shows the results of classification accuracy, specificity, sensitivity and F measure respectively. Figure 9 shows the best fitness of PSOANN and GSOANN.

From table 2 and figure 5 it is seen that the classification accuracy of GSO-ANN-BP performs better by 8.88% than NB, by 17.6% than C4.5, by 11.11% than ANN-BP and by 4.13% than PSO-ANN-BP.

From table 2 and figure 6 it is observed that the specificity of GSO-ANN-BP performs better by 8.86% than NB, by 17.91% than C4.5, by 11.33% than ANN-BP and by 4.04% than PSO-ANN-BP.

Table 2
Summary of Results

	Naïve Bayes	C4.5	ANN-BP	PSOANN-BP	GSO-ANN-BP
Classification Accuracy	0.837	0.7667	0.8185	0.8778	0.9148
Specificity	0.8365	0.7638	0.81605	0.87785	0.91405
Sensitivity	0.8325	0.7633	0.8175	0.8742	0.91335
F measure	0.8341	0.76355	0.81665	0.8757	0.9137

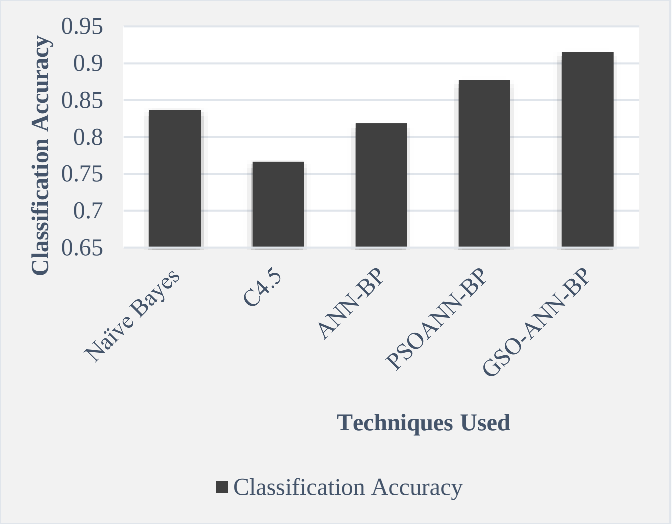


Figure 5: Classification Accuracy

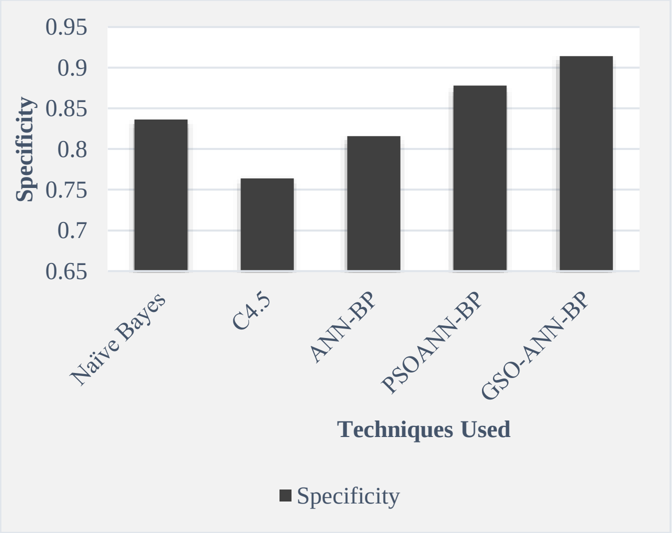


Figure 6: Specificity

From table 2 and figure 7 it is observed that the sensitivity of GSO-ANN-BP performs better by 9.3% than NB, by 17.89% than C4.5, by 11.08% than ANN-BP and by 4.38% than PSO-ANN-BP.

From table 2 and figure 8 it is observed that the F Measure of GSO-ANN-BP performs better by 9.11% than NB, by 17.9% than C4.5, by 11.22% than ANN-BP and by 4.25% than PSO-ANN-BP.

From figure 9 it is observed that the best fitness of PSOANN method convergence at iteration number 390, where GSOANN method convergence at iteration number 250.

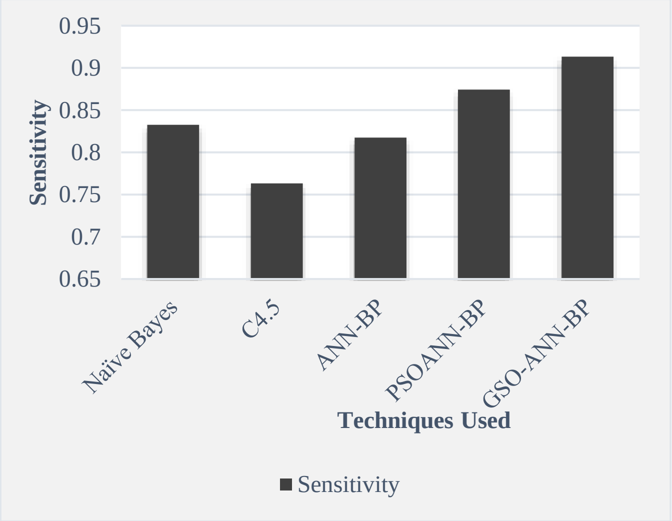


Figure 7: Sensitivity

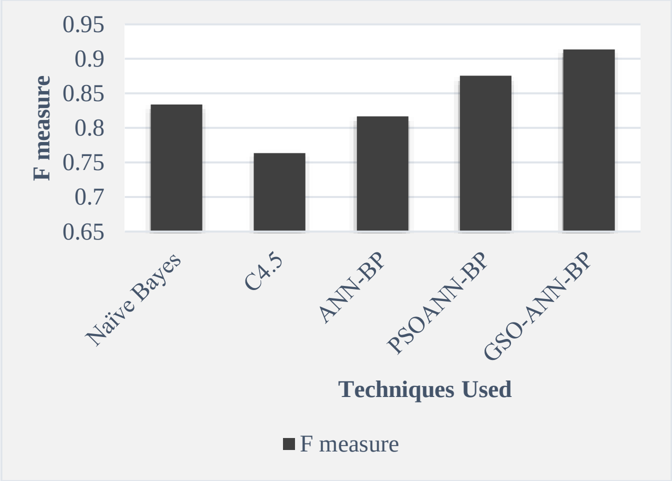


Figure 8: F Measure

Conclusion

Heart diseases are a severe threat and typically occur when arteries that provide oxygen and blood to the heart are entirely blocked or made narrow. There is a vast amount of data produced in medical organizations however it is not appropriately utilized. Classification problem of designating various observations into various disjointed groups has an important role to play in making business decisions among others. ANN is the mathematical simulation of biological neurons responsible for human brain functioning. ANN model is structured with inter-connected computational neurons utilized for executing mathematical mapping at the time of learning procedure. Outcomes prove that the classification accuracy of GSO-ANN-BP outperforms NB by 8.88%, C4.5 by 17.6%, ANNBP by 11.11% and PSOANNBP 4.13%. The best fitness of PSOANN technique converges at 390thiteration while GSOANN technique converges at 250th iteration.

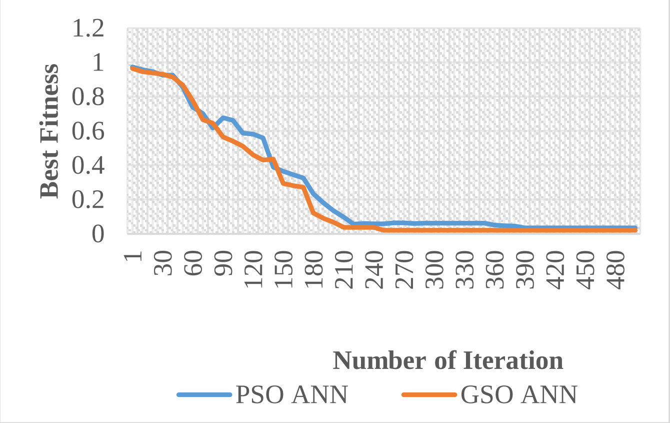


Figure 9: Best Fitness

References

1. Deepthi S. and Ravikumar A., Computation Methods for the Diagnosis and Prognosis of Heart Disease, *International Journal of Computer Applications*, **95(19)**, 5-9 **(2014)**

2. Kavitha K.S., Ramakrishnan K.V. and Singh Manoj Kumar, Modeling and design of evolutionary neural network for heart disease detection, *IJCSI International Journal of Computer Science Issues*, **7(5)**, 272-283 **(2010)**

3. Rao B.S., Rao K.N. and Setty S.P., An approach for heart disease detection by enhancing training phase of neural network using hybrid algorithm, In Advance Computing Conference (IACC), 2014 IEEE International, IEEE, 1211-1220 **(2014)**

4. Metkari M. and Pradhan M., Improve the Classification Accuracy of the Heart Disease Data Using Discretization, *International Journal of Innovative Research in Advanced Engineering*, **10(2)**, 1-5 **(2015)**

5. Durairaj M. and Revathi V., Prediction of Heart Disease Using Back Propagation MLP Algorithm, *International Journal of Scientific & Technology Research*, **4(08)**, 235-239 **(2015)**

6. He S., Training Artificial Neural Networks Using Lévy Group Search Optimizer, *Multiple-Valued Logic and Soft Computing*, **16(6)**, 527-545 **(2010)**

7. Yaghini M., Khoshraftar M.M. and Fallahi M., A hybrid algorithm for artificial neural network training, *Engineering Applications of Artificial Intelligence*, **26(1)**, 293-301 **(2013)**

8. Mokashi A.R., Tambe M.N. and Walke P.T., Heart Disease Prediction Using ANN and Improved K-Means, *Heart Disease*, **4(4) (2016)**

9. Medhekar D.S., Bote M.P. and Deshmukh S.D., Heart disease prediction system using naive Bayes, *Int. J, Enhanced Res. Sci. Technol & Eng*, **2(3)**, 1-5 **(2013)**

10. Duan Z., Wang Y. and Xing Y., Sound Quality Prediction of Vehicle Interior Noise under Multiple Working Conditions Using Back-Propagation Neural Network Model, *Journal of Transportation Technologies*, **5(02)**, 134 **(2015)**

11. Banu M.N. and Gomathy B., Disease Predicting System Using Data Mining Techniques, *International Journal of Technical Research and Applications*, **1(5)**, 41-45 **(2013)**

12. Yang X.S., A new metaheuristic bat-inspired algorithm. In Nature inspired cooperative strategies for optimization (NICSO 2010), Springer Berlin Heidelberg, 65-74 **(2010)**

13. Shen H., Zhu Y., Niu B. and Wu Q.H., An improved group search optimizer for mechanical design optimization problems, *Progress in Natural Science*, **19(1)**, 91-97 **(2009)**.