http://dx.doi.org/10.21172/ijiet.93.13

An Analysis Of Heart Disease Prediction Using Swarm Intelligence Algorithms

Lalitha Kumari Gaddala[1] & Dr. N. Naga Malleswara Rao[2]
[1]*Research Scholar, Acharya Nagarjuna University, Guntur, AP, India*
[2]*Professor, Dept. of IT, RVR & JC College of Engineering, Guntur, AP, India*

Abstract- Health diseases are increasing day by day due to life style and hereditary. In this aspect, heart disease is the most important cause of demise in the human kind over past few years. The objective of this paper is to predict the Heart Disease by applying Artificial Neural Network using swarm Intelligence algorithm. Swarm intelligence (SI) is relatively new interdisciplinary field of research. The Swarm-based algorithms have recently emerged as a family of nature-inspired, population-based algorithms that are capable of producing low cost, fast, and robust solutions to several complex problems. There are so many swarm intelligence algorithms for optimization like Group Search Optimization (GSO), Artificial Bee Colony (ABC), Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO) etc. This paper proposes Particle Swarm Optimization (PSO) is the most population Intelligence Algorithm and has good performance on optimization. This paper aims to predict the heart disease using Feed forward of Artificial Neural Network (ANN) to classifying patient as diseased and non diseased. We have evaluated our new classification approach via the well-known data sets .Finally, the performance was evaluated in terms of accuracy, sensitivity and specificity and also compare to other well-known data sets, it has been observed that these results are one of the best results compared with the results obtained from related previous studies.
Keywords: Swarm intelligence (SI), Group Search Optimization (GSO), Artificial Bee Colony (ABC), Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO), Artificial Neural Network (ANN), UCI-Machine Learning Repository.

I. INTRODUCTION

In the current competitive world, we require an efficient technique to summarize, analyze, present and maintain large datasets using data mining. This requires the knowledge of all data mining techniques in order to choose the best for desired datasets and these data mining techniques can answer the questions that traditionally were too time consuming to resolve. Research has shown that, data doubles every three years. While comparing the Data Mining with Artificial Neural Networks the performance is Highest, Computational Speed is fast and Complexity Level is high in Artificial Neutral Networks (ANNs) [1].

Artificial neural networks (ANNs) are networks of simple processing elements (called 'neurons') operating on their local data and communicating with other elements[2].ANN have been successfully solved many complex real world problem such as predicting future trends based on the huge historical data of an organization. ANN have been successfully implemented in all engineering fields such as biological modeling, decision and control, health and medicine, engineering and manufacturing, marketing, ocean exploration and so on[3].

The purpose of this paper is to present swarm intelligent technique-Particle swarm optimization (PSO) with feed forward neural network that can be used to provide solutions to the prediction of heart disease.

Bonabeau defined Swarm Intelligence as "*The emergent collective intelligence of groups of simple agents*" [4].A swarm is a large number of homogenous, simple agents interacting locally among themselves, and their environment, with no central control to allow a global interesting behavior to emerge. Swarm-based algorithms have recently emerged as a family of nature-inspired, population-based algorithms that are capable of producing low cost, fast, and robust solutions to several complex problems [5].

Swarm Intelligence (SI) can therefore be defined as a relatively new branch of Artificial Intelligence that is used to model the collective behavior of social swarms in nature, such as ant colonies, honey bees, and bird flocks. Although these agents (insects or swarm individuals) are relatively unsophisticated with limited capabilities on their own, they are interacting together with certain behavioral patterns to cooperatively achieve tasks necessary for their survival. The social interactions among swarm individuals can be either direct or indirect [6].

The typical swarm intelligence system has the following properties:

1) It is composed of many individuals.
(2)The individuals are relatively homogeneous.

(3)The interactions among the individuals are based on simple behavioral rules that exploit only local information that the individuals exchange directly or via the environment. **Examples** in natural systems of SI include ant colonies, bird flocking, animal herding, bacterial growth, and fish schooling [7].

II. BACKGROUND AND RELATED WORK

A. Background –

Doctors will diagnose Heart Failure based on Medical and Family Histories, a Physical Exam and Test Results of a Patient. Almost all the doctors are predicting heart disease by learning and experience. The diagnosis of disease is a difficult and tedious task in medical field. Diagnosis of Heart disease from various factors or symptoms is a multi-layered issue which may lead to false presumptions and unpredictable effects. Only human intelligence alone is not enough for proper diagnosis. A number of difficulties will arrive during diagnosis, such as

(a) less accurate results
(b) less experience
(c) time dependent performance
(d) Knowledge up gradation is difficult [8].

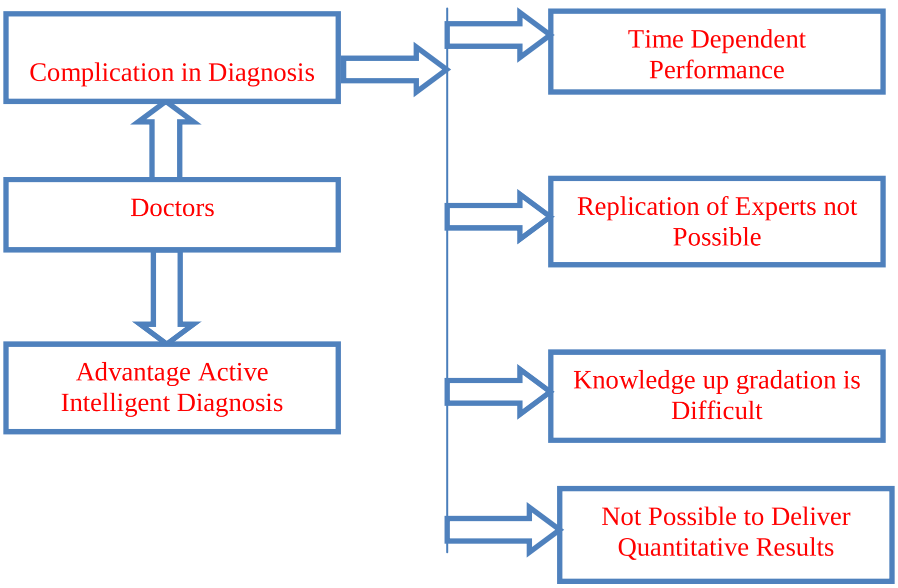


Fig-1: Block Diagram of Complexity in Diagnosis with Doctor

B. Related Work –

Neural Networks have been developed rapidly since around 1985 and are now used widely. The feed forward neural network was the first and simplest type of artificial neural network devised. These are now widely used in classification problems [9].Consider a Feed Forward neural Network with single Hidden Layer denoted by N-h-N, Where N is the Number of units in the Input and Output Layers, and h is the number of units in the Hidden Layer. The Input Layer Units are Fully Connected to the hidden Layer units which are in turn fully connected to the output units [10].

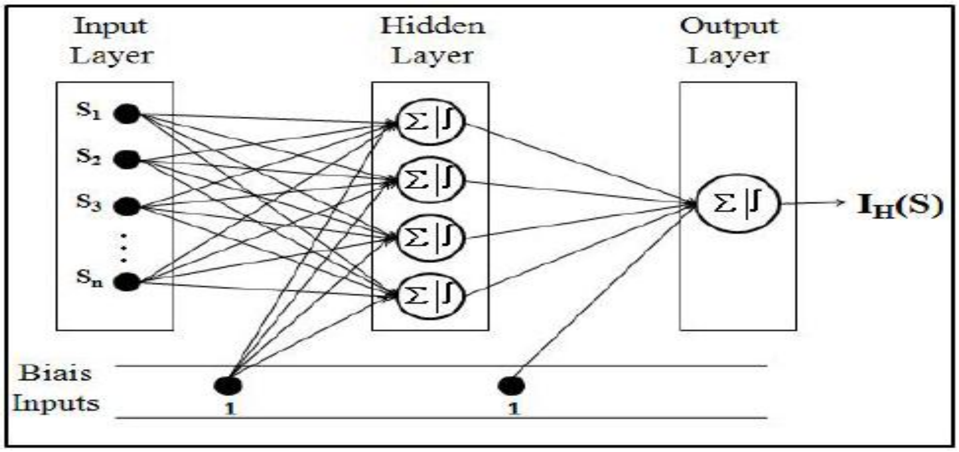


Fig-2: The Feed Forward Neural Network Architecture

C. The Choice of PSO Algorithm-

Particles swarm optimization (PSO) algorithm, originally developed by Kennedy and Eberhart[11][12], is a method for optimization on metaphor of social behavior of flocks of birds and/or schools of fish. Each particle is interconnected in particles swarm optimization. Particles can imitate the nature, to achieve performance improvement. Similar to genetic algorithms (GAs), the PSO is also an optimizer based on population. PSO can solve many nonlinear function minimization, no-differentiable, discontinuous and multi model problems, game theory so it has been used to solve science and engineering problems [13].In PSO, a solution is represented as a particle, and the population of solutions is called a swarm of particles. PSO algorithms are easy to implement and achieve global optimal solutions with high probability.

The basic PSO algorithm is shown in given Figure:

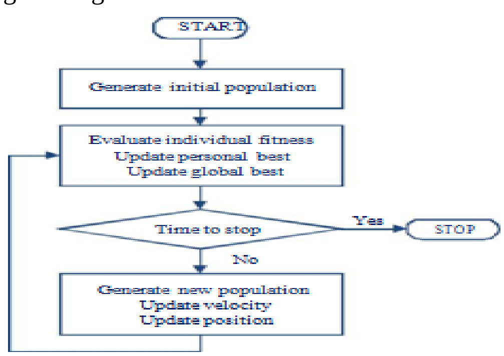


Figure : Flowchart for particle swarm optimization algorithm.

In PSO the population of candidate solutions is represented by a swarm of particles. Each particle is a point in the n–dimensional search space. The i^{th} particle in the swarm is represented by its current position pi, and its current velocity vi (both xiand vi are n–dimensional vectors). PSO tries to find the optimal solution to the problem by moving the particles and evaluating the fitness of the new position. A particle's position is updated by the following equation [14].

$$v_i(t+1) = v_i(t) + (c_1 \times rand() \times (p_i^{best} - p_i(t))) + (c_2 \times rand() \times (p_{gbest} - p_i(t)))$$

A particle's position is updated using:

$$p_i(t+1) = p_i(t) + v_i(t)$$

Where,

$\mathbf{v_i}$ **(t+1)**, is the new velocity for the i^{th} particle.

c1 and c2 are the Weighting Coefficients for the Personal best and Global best positions respectively.

$\mathbf{p_i(t)}$ is the i^{th} particle position at time t

$\mathbf{p_i^{best}}$is the i^{th} particle's best known position

http://dx.doi.org/10.21172/ijiet.93.13

$\mathbf{p_g^{best}}$is the best position known to the swarm
The **rand ()** Function generate a uniformly random variables € [0, 1] variants on this updateequation [15].

Elements used in PSO are:
Before working with the PSO, we have to know about the elements used in the PSO. First of all, we shall overview the brief concepts of the PSO elements.

- **Particle**---We can define the particle as Pi for real numbers.
- **Fitness Function**---Fitness Function is the function used to find the optimal solution. Usually it is an objective function.
- **Local Best**---It is the best position of the particle among its all positions visited so for.
- **Global Best**---The position where the best fitness is achieved among all the particles visited so for.
- **Velocity Update**---Velocity is a vector to determine the speed and direction of the particle. Velocity is updated by the equation (I).
- **Position Update**---All the particles try to move toward the best position for optimal fitness. Each particle in PSO updates their positions to find the global optima. Position is updated by equation (2)

Each particle is initialized with a random position and velocity. Each particle is then evaluated for fitness value. Each time a fitness value is calculated, it is compared against the previous best fitness value of the particle and the previous best fitness value of the whole swarm, and the personal best and global best positions are updated where appropriate. If a stopping criterion is not met, the velocity and position are updated to create a new swarm. The personal best and global best positions, as well as the old velocity, are used in the velocity update [16].
To improve the performance of PSO, researches modified the PSO in different ways.

The pseudo code of original PSO is [17]:

```
Initialize the population randomly
While (Population Size)
{
         Loop
        Calculate fitness
        If fitness value is better from the best fitness value (pbest) in history then
        Update p_best with the new pbest
        End loop
        Select the particle with the best fitness value from all particles as gbest
         While maximum iterations or minimum error criteria is not attained
        {
                For each particle
                 Calculate particle velocity by equation (1)
                Update particle position according to equation (2)
                 Next
         }
}
```

An Analysis on the Advantages of PSO:
1. PSO has no overlapping and mutation calculation. The search can be carried out by the speed of the particle. During the development of several generations, only the most optimist particle can transmit information onto the other particles, and the speed of the researching is very fast.
2. The calculation in PSO is very simple. Compared with the other developing calculations, it occupies the bigger optimization ability and it can be completed easily
3. PSO adopts the real number code, and it is decided directly by the solution. The number of the dimension is equal to the constant of the solution [18]

III. EXPERIMENTS

Metrics: Evaluated Performance of PSO Algorithm utilizes following Metrics: complexity, accuracy, sensitivity, specificity given in *Karalolis et al.(2010)* these metrics are common in medical applications are discussed below[19] :

True Positive (TP): It denotes the number of Heart Disease Patients classified correctly by the Hybrid Neural Network

True Negative (TN): It denotes the number of patients not having heart disease correctly classified by the system.

False Positive (FP): It denotes the number of healthy patients wrongly classified as a heart disease patient by the system

False Negative (FN): It denotes the number of healthy patients classified as a heart disease patient by the system

Confusion Matrix

Result of the diagnostic test		Physician diagnosis	
		Positive	Negative
Classifier Result	Positive	TP	FP
	Negative	FN	TN

In this work, the data-set is adopted from UCI Machine Learning Repository, Cleveland Database, Hungary Database [20] which is a medical data-set related to heart disease. Cleveland database used to classify person into four classes (Normal(0), First Stroke (1), Second Stroke (2) and End of life (3)) regarding to the heart disease. Cleveland, Hungary contains 76 attributes. Among all the 76 attributes, 14 attributes are taken for experimentation. The performance was evaluated in terms of accuracy, sensitivity, specificity of PSO using different data bases.
The 14 attributes are given below:

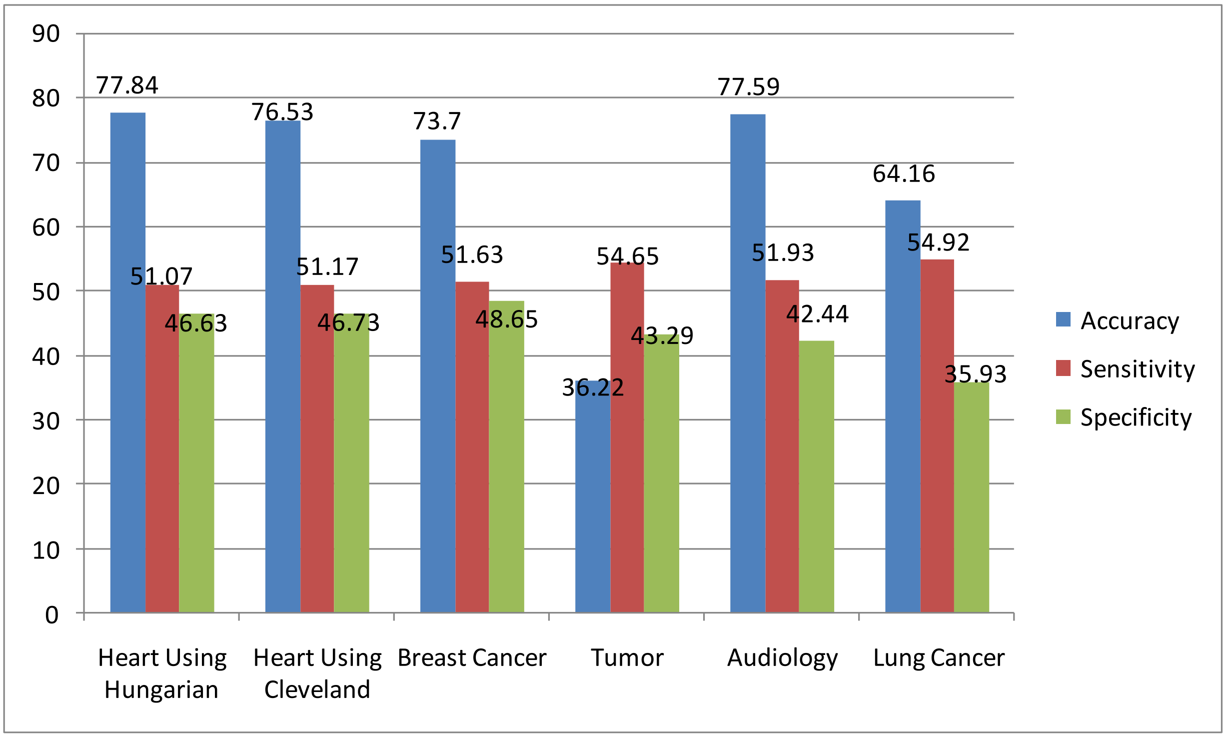
Table-1: Description of Data Attributes

S.NO	NAME	ATTRIBUTE DISCRIPTION
1	Age	Age in Years
2	Sex	sex (1 = male; 0 = female)
3	Cp	chest pain type -- Value 1: typical angina -- Value 2: atypical angina -- Value 3: non-anginal pain -- Value 4: asymptomatic
4	trestbps	resting blood pressure (in mm Hg on admission to the 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 -- Value 0: normal -- Value 1: having ST-T wave abnormality (T wave inversions and/or ST elevation or depression of > 0.05 mV) -- Value 2: showing probable or definite left ventricular hypertrophy by Estes' criteria
8	thalach	maximum heart rate achieved
9	exang	exercise induced angina (1 = yes; 0 = no)
10	Oldpeak	ST depression induced by exercise relative to rest
11	Slope	the slope of the peak exercise ST segment -- Value 1: upsloping -- Value 2: flat -- Value 3: downsloping
12	Ca	number of major vessels (0-3) colored by flourosopy
13	Thal	3 = normal; 6 = fixed defect; 7 = reversable defect
14	num (the predicted attribute)	diagnosis of heart disease (angiographic disease status) -- Value 0: < 50% diameter narrowing -- Value 1: > 50% diameter narrowing (in any major vessel: attributes 59 through 68 are vessels)

Table-2: Proposed algorithm performance results with different data Sets

Classification of Data Sets	Accuracy	Sensitivity	Specificity
Heart Using Hungarian	77.84	51.07	46.63
Heart Using Cleveland	76.53	51.17	46.73
Breast Cancer	73.7	51.63	48.65
Tumor	36.22	54.65	43.29
Audiology	77.59	51.93	42.44
Lung Cancer	64.16	54.92	35.93

Table-2 describes the performance analysis of proposed PSO-ANN using different data sets.



Resulted Graph describes the performance analysis of proposed PSO-ANN using different data sets.

IV. CONCLUSION

In this work we use the PSO technique as a training algorithm for ANN to predict the heart diseases. After applying the PSO, We found that compare to different diseases was able to improve the accuracy, sensitivity and specificity. Based on these results it can be shows that the proposed system is able to good performance in the category of optimization.

V. REFERENCES

[1] Aartikaushal, manshishukla, Comparative Analysis to Highlight Pros and Cons of Data Mining Techniques-Clustering, Neural Network and Decision Tree, International Journal of Computer Science and Information Technologies, Vol. 5 (1) , 2014, 651-656 , ISSN NO :0975-9646

[2] W.S. mcculloch, W. Pitts, A logical calculus of ideas immanent in nervous activity, Bull. Math.Biophys. 5 (1943) 115-133.

[3] Kuldipvora,shrutiyagnik, A Survey on Backpropagation Algorithms for Feed forward Neural Networks, ISSN: 2321-9939.

[4] Bonabeau E, Dorigo M, Theraulaz G. Swarm Intelligence: From Natural to Artificial Systems. Journal of Artificial Societies and Social Simulation.1999; 4: 320.

[5] Hazem Ahmed, Janice Glasgow ,Swarm Intelligence: Concepts, Models and Applications Technical Report 2012-585, Queen University, School of Computing Queen's University Kingston, Ontario, Canada K7L3N6, February 2012

[6] M. Belal, J. Gaber, H. El-Sayed, and A. Almojel, Swarm Intelligence, In Handbook of Bioinspired Algorithms and Applications. Series: CRC Computer & Information Science. Vol. 7.Chapman & Hall Eds, 2006.ISBN 1-58488-477-5.

[7] Lalitha Kumari Gaddala ,A Vanamala Kumar,ysurekha,Clinical Support System to predict Heart Diseases using ABC Algorithm of Artificial Neural Networks, International Journal of Innovations in Engineering and Technology (IJIET), Volume 7 Issue -2 ,August 2016 , ISSN: 2319 – 1058.

[8] A. T. Sayad, P. Halkarnikar, Diagnosis of Heart Disease Using Neural Network Approach, International Journal of Advances in Science Engineering and Technology, ISSN: 2321-9009 Volume- 2, Issue-3, July-2014

[9] B.D.RIPLEY, A Text Book of Neural Networks and Related methods for Classification, University of Oxford, UK

[10] P. Srikala, Shaik Umar Faruq, Neural Network based Image Compression with Lifting Scheme and RLC , International Journal of Reseach In Engineering and Technology, ISSN : 2319-1163

[11] Russell Eberhart, James Kennedy, " A New Optimizer using Particle Swarm Theory, IEEE ,1995

[12] Chunkai Zhang, Huihe Shao, " Particle Swarm Optimization in Feed Forword Neural Network in Artificial Neural Networks in Medical Biology, Springer-2000

[13] Haufen Yang, You Yang, Dejian Kong, Dechun Dong, Zuyuan Yang, Lihui Zhang,An Improved particle swarm optimization algorithm. Http://ieeexplore.ieee.org/abstract/document/6818010/authors?Ctx=authors

[14] Muhammad Rashid , Abdul raufbaig, Improved Opposition-Based PSO for Feed Forword Neural Network Training, http://ieeexplore.ieee.org/abstract/document/5480380/authors

[15] Jason Brownlee, Text Book - Clever Algorithms: Nature-Inspired Programming Recipes

[16] Vorataskachitvichyanukul , Comparison of Three Evolutionary Algorithms: GA, PSO, and DE , Industrial Engineering & Management Systems, Vol 11, No 3, September 2012, pp.215-223. Http://dx.doi.org/10.7232/iems.2012.11.3.215

[17] Muhammad Imran , rathiahhashima ,Noor elaizaabdkhalidb , An Overview of Particle Swarm Optimization Variants, Procedia Engineering ,Volume 53, 2013, Pages 491-496,ELSEVIER, Http://www.sciencedirect.com/science/article/pii/S1877705813001823?Via%3Dihub

[18] Qinghai Bai, Analysis of Particle Swarm Optimization Algorithm, Computer and Information Science, Vol 3 , No.1, February 2010

[19] A Text Book Of Fuzzy Systems:Concepts, Methodologies, Tools and Applications, Information Resource Management, USA

[20] UCI, Machine Learning Repository, Cleveland Database. Https://archive.ics.uci.edu/ml/datasets/Heart+Disease