National Conference on Engineering Innovations and Solutions (NCEIS – 2018)
International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 4 | Issue 6 | ISSN : 2456-3307

Decision Making for Heart Disease Detection Using Hybrid Neural Network-Particle Swarm Optimization Algorithm

Ranjitha P*1, Vanishri Arun2
*1Software Engineering, J S S Science and Technology University
2Assistant Professor, J S S Science and Technology University

ABSTRACT

Cardiovascular disease is one of the most rampant causes of death around the world and was deemed as a major illness in Middle and Old ages. Coronary artery disease, in particular, is a widespread cardiovascular malady entailing high mortality rates. Angiography is, more often than not, regarded as the best method for the diagnosis of coronary artery disease; on the other hand, it is associated with high costs and major side effects. Much research has, therefore, been conducted using machine learning (genetic algorithm) and data mining to seek alternative modalities. Accordingly, we herein propose a highly accurate hybrid method for the diagnosis of coronary artery disease. As a matter of fact, the proposed method is able to increase the performance of neural network through enhancing its initial weights using Particle Swarm optimization Algorithm (PSO) that suggests better weights for neural network. Making use of such methodology, we can improve accuracy, sensitivity and specificity rates on Z-Alizadeh Sani dataset.

Keywords: Artificial Neural Network, Particle swarm optimizer, Coronary Artery disease, Evolutionary algorithm.

I. INTRODUCTION

Cardiovascular disease (CVD) is a class of diseases that involve the heart or blood vessels. Cardiovascular disease includes coronary artery diseases (CAD) such as angina and myocardial infarction (commonly known as a heart attack). Other CVDs include stroke, heart failure, hypertensive heart disease, rheumatic heart disease, cardiomyopathy, heart arrhythmia, congenital heart disease, valvular heart disease, carditis, aortic aneurysms, peripheral artery disease, thromboembolic disease, and venous thrombosis.

The heart is a muscle that acts like a pump to move blood throughout the body. To function properly, the heart must receive oxygen. Oxygen is supplied to the heart by the coronary (heart) arteries that wrap around the surface of the heart. When coronary artery disease (CAD) is present, blood flow through the arteries can be reduced. When this happens, the heart muscle may not receive enough oxygen, and chest pain (called angina) may be felt. CAD is caused by the build-up of fatty substances, such as cholesterol, that collect along the lining of the coronary arteries, in a process known as atherosclerosis. This referred to as a “plaque”, “lesion”, “blockage” or “stenosis”. This means that there is a narrowing in the artery caused by a build-up of substances which may eventually block the flow of blood. Because the coronary arteries supply oxygen rich blood to the heart, untreated blockages can be very serious and can lead to a heart attack (myocardial infarction) or even death. Over the course of a person's lifetime many influences can cause one or more of your coronary arteries to become narrowed or blocked.

The present study, given the risks of invasive diagnostic procedure such as angiography and auspicious experiences in the field of data mining, attempts were made to propose a model for identifying coronary arteries disease. The suggested detection model, based on artificial neural networks and Particle swarm optimizer algorithms, can detect coronary artery disease based on clinical data without the need for invasive diagnostic methods. Although ANNs are powerful, designing a proper network can be a very tough task and the more complex a desired dynamics are, the more difficult the design of the network becomes. Many researchers have sought to automate ANN design process by using computer programs. They have used algorithms that explore various combinations of network parameters (size, topology, connection weights, etc.) and the most suitable networks.

II. RELATED WORK

Evolutionary algorithms (EAs) are optimization and search methods, based on Darwinian evolution. EAs are especially useful for Finding global optima of functions which have many locally optimal solutions, because in comparison with traditional gradient-based search methods, EAs have more chances to escape from local optima. EAs are independent of gradient signals and are thus suitable for handling problems where such information is not available. EAs have been frequently used to carry out various tasks regarding 1[1].

Distribution of age and sex was different in the two groups. Subjects with disease had a mean age of 53.3 years and 84% were males, whereas those without disease had a mean age of 49.1 years and 51% were males. Because of these differences and the subsequent determination that the risk associated with cigarette smoking and serum cholesterol level differed with both age and sex, discriminant analyses were performed in nine age-sex subgroups[2].

Heart disease is the leading cause of death in the world over the past 10 years. Researches have been using several data mining techniques in the diagnosis of heart disease. Most of these systems have successfully employed Machine learning methods such as Naïve Bayes and Support Vector Machines for the classification purpose[3].

Predicting heart disease use inputs from complex tests conducted in labs. Developing a system which will predict heart based on the risk factors[4]. Artificial Bee Colony (ABC) Optimization for the Prediction of Coronary Heart Disease. This probabilistic algorithm has performed better to produce the best accuracy rate. The intelligent behaviour algorithm can be recommended for solving highly complicated optimization problems[5].

III. METHODOLOGY

A. **Dataset**

The present research used Z- Alizadeh Sani dataset, containing information on 303 patients, 216 of whom suffered from CAD. Fifty-four features were collected for each patient.

Table 1 Dataset

Feature type	Feature name
Demographic	Age Sex BMI DM(Diabetics Mellitus) HTN(Hypertension) Current Smoker EX-smoker FH(Family History) Obesity CRF CVA Airway diseases

	Thyroid Diseases CHF DLP(Dyslipidemia)
Symptoms and examination	BP PR(Pulse rate) Edema Weak peripheral pulse Lung rales Systolic murmur Diastolic murmur Typical Chest pain Dyspnea Function class Atypical Nonanginal chest pain Exertional chest pain Low th Ang
ECG	Rhythm Q wave St elevation St depression T inversion LVH Poor R-Wave Progression
Laboratory and echo	Fbs, Cr, TG, LDL, HDL, BUN, ESR, HB, K, NA, WBC, Lymph, Neut, PLT, EF, Region with RWMA,VHD.

A. Architecture

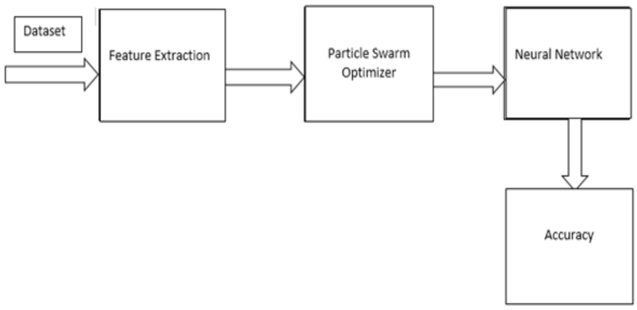

Figure 1. Block Diagram

Z-Alizadeh Sani dataset is considered for the work. Where the dataset contains information on 303 patients, 216 of whom suffered from CAD. Fifty-four features were collected for each patient. Firstly dataset is send for pre-processing where noise and missing data or record are removed. For feature selection, weight by SVM and principal component analysis (PCA) were used. In weight by SVM, attribute weights are the coefficients of the normal vector of a linear SVM .PCA converts a set of correlated variables into a smaller number of uncorrelated variables employing an orthogonal transformation. Later the selected feature are passed for the particle swarm optimization algorithm for the generation of global best weight. Where global weight are given as input for neural network for training and testing.

B. Workflow

The evolutionary algorithm approach is not the only approach for stochastic optimization of ANNs. In 1995, Kennedy and Eberhart introduced particle swarm optimization (PSO), which is a stochastic population-based search method inspired by social behaviour of animals such as birds. It is known that PSO is also successful for optimization of ANNs and in this case it produces better results than GAs. PSO excels in global search and compare to back propagation (BP) algorithm, which is a very common gradient based method for training the connection weights of ANNs, PSO shows faster convergence. Compared to evolutionary algorithms, PSO is faster in approaching the optima. PSO consists of a swarm of particles where each particle has its own position and velocity. Each particle is initialized randomly at the beginning of the algorithm and the heuristics update the position and velocity in the latter stages of algorithm.

The pseudo code of PSO for global minimization problems is as follows:

Let P = fp1; p2; p3; : : :png be set of particles where each pi is of d dimensions;

pid = fpi1; pi2; : : : ; pidg. Each particle has its own

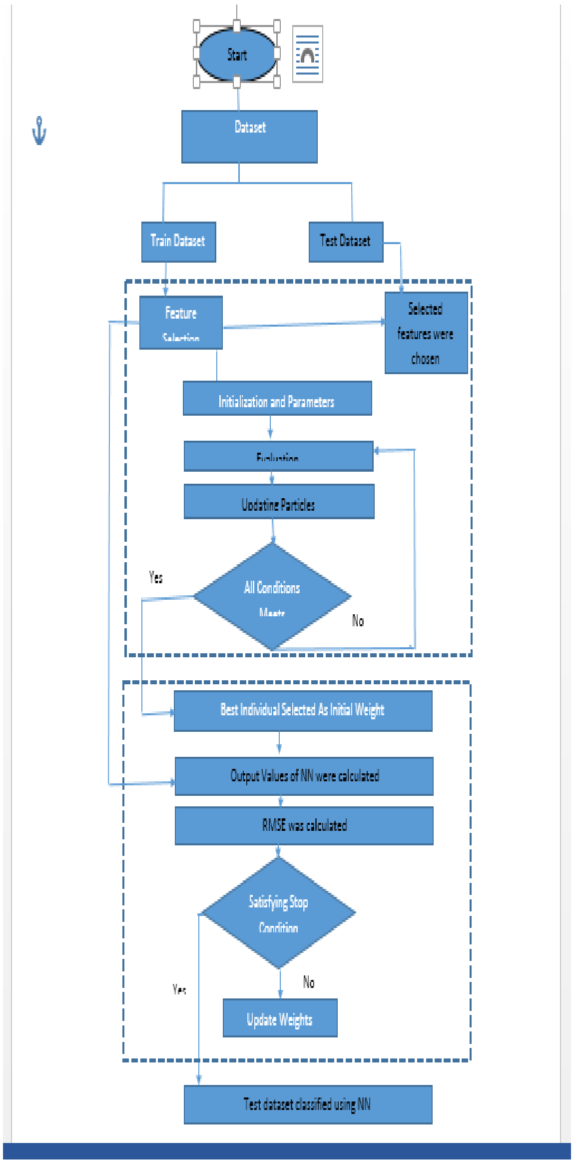

Figure 2. Workflow

individual velocities vi i.e.,

V = fv1;v2;v3; : : : vng.

Start

Initialize positions and velocities randomly of permissible range for each particle.

While convergence criteria is not met

DO

For each particle

Calculate the fitness value

If the fitness value of the particle is less than pibest (old) then

Update pibest to present value

End if

End

Update gbest with the particle that provides best fitness value of all the particles in

the swarm

For each particle

For each dimension

vid = w _ vid(old) +c1*rand*(pibestd _pid)+c2*rand*(gibestd _pid) // update

particles velocity

pid = pid (old)+vid // updating particles position

end for

end for

END

IV. RESULT DISCUSSION

A. Preprocessing and Feature extraction

The dataset contains 54 features for each record. Firstly all the missing data are removed. The values of all 54 features are converted into numeric. Where 'Yes' is considered as 1 ,'No' is considered as 0.For the feature extraction Principal component analysis and weighted svm[5] are mainly used. The Fig 4 shows the extracted features for training purpose.

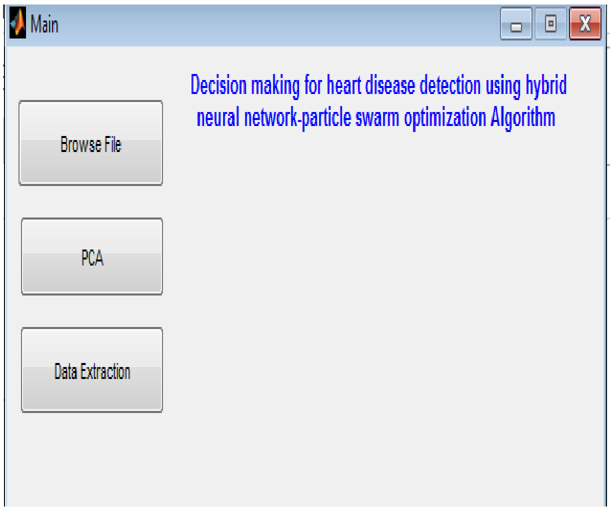

Figure 3. PCA Feature Extraction

Feature
Typical chest pain
Atypical
Age
Nonanginal
DM
Tinversion
FH
Region RWMA
HTN
TG
PR
Diastolic murmur
Current smoker
Dyspnea
ESR
BP
Function class
Sex
FBS
St depression
St elevation
Q wave

Figure 4. Selected features

B. Traning and testing phase

Artificial neural network (ANN) serves the objective providing a model which has the ability to relate very complex input and output datasets. This ANN model works extremely well for very complex data sets which are normally very difficult to predict using mathematical modelling (equations). The ANN is a network of neuron connected among themselves through weights and biase. Once the structure of the ANN is formed then the next task is to train the network.

Training of the networks means finding the optimum values of various weights and biases of the network. Normally, various types of techniques are used to find the suitable values of weights and biases of the ANN. I this work, optimum training of the network have been obtained through particle swarm optimization (PSO).

The following seven steps have been used to train ANN using PSO.

Step 1) Collect dataset.

Step 2) Create input data and output data from the dataset.

Step 3) PSO algorithm parameters gets initialized.

Step 4) Weights gets generated.

Step 5) Training using PSO and neural network .

Step 6) Check the Network

Step 7) Testing

```
****************************************
 RUN fval ObFuVa
     1 0.0964526   0.0965
Elapsed time is 163.723091 seconds.
Final nn model is net_f

err =

    0.0965

Trained ANN net_f is ready for the use
```

Figure 5. Result of traning phase

The trained ANN is **'net_f'**. Once the training is completed the regression plot will be displayed. Figure 6 shows the regression plot of the trained ANN (net_f). From this figure, it is observed that regression coefficient R is 0.77978.

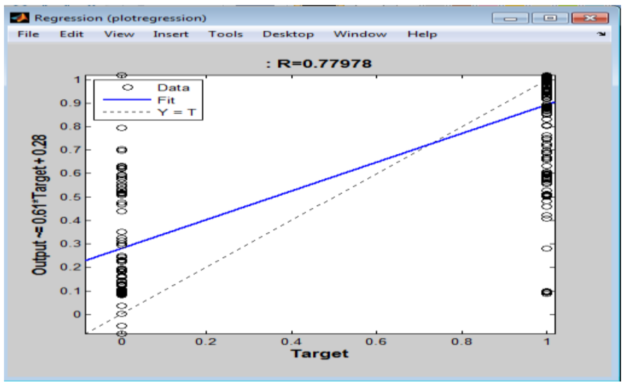

Figure 6. Regression Plot

IV. CONCLUSION

We proposed a new hybrid method to augment the perfor- mance of neural network. The method put forth can ameliorate the performance of neural network as concerns CAD detection. Specifically, using this method, CAD can be detected without angiography which can help eliminate high costs and major side effects. Meanwhile, other versions of neural networks can be tested and compared. Parameters like learning rate, and momentum factor can also be optimized for this work. Finally, new data with some other features must be checked by this algorithm.

V. REFERENCES

1. G. Parthiban, S.K.Srivatsa," Applying Machine Learning Methods in Diagnosing Heart Disease for Diabetic Patients", International Journal of Applied Information Systems (IJAIS) Volume 3– No.7, August 2012 – www.ijais.org.
2. V.Arun ,"Disease classification and prediction using principal component Analysis and Ensemble classification framework", International journal of control Theory and Application, Vol 10 ,2017,ISSN:0974-5572.
3. Sowmya N. and Dr. R .Vijayabhanu ," Artificial Bee Colony (ABC) Optimization for the Prediction of Coronary Heart Disease", International Journal of Trend in Research and Development, Volume 2(5), ISSN 2394-9333 www.ijtrd.com
4. Jamal Salahaldeen Majeed Alneamy and Rahma Abdulwahid Hameed Alnaish " Heart Disease Diagnosis Utilizing Hybrid Fuzzy Wavelet Neural Network and Teaching Learning Based Optimization Algorithm", Hindawi Publishing Corporation Advances in Artificial Neural Systems Volume 2014, Article ID 796323
5. Sowmya N. and Dr. R .Vijayabhanu." Intelligent Heart Diseases Prediction System Using A New Hybrid Metaheuristic Algorithm", International Journal of Engineering Research & Technology (IJERT) Vol. 1 Issue 8, October - 2012 ISSN: 2278-018
6. Murat Koklu and Yavuz Unal, "Analysis of a Population of Diabetic Patients Databases with Classifiers" , International Science Index, Biomedical and Biological Engineering Vol:7, No:8, 2013 waset.org/Publication/16219
7. Prof(Dr.) Nirali Nanavati ,Aashna Agarwal," Association rule mining using hybrid GA-PSO for multi-objective optimization", 978-1-5090-0612-0/16/ ©2016 IEEE .
8. Majid Ghonji Feshki, Omid Sojoodi Shijani," Improving the Heart Disease Diagnosis by Evolutionary Algorithm of PSO and Feed Forward Neural Network", 978-1-5090-2169-7 ©2016 IEEE.
9. R Geetha Ramani, Lakshmi Balasubramanian,, Shomona Gracia Jacob," Automatic Prediction of Diabetic Retinopathy and Glaucoma through Retinal Image Analysis and Data Mining Techniques", 978-1-4673-2322-2/12 ©2012 IEEE.