AN EFFECTIVE ARCHETYPE DESIGN OF HEART DISEASE ANTICIPATION USING OPTIMIZATION TECHNIQUES

P.Deepika 1

1 Research Scholar & Assistant Professor, PG & Research Department of Computer Science, Hindusthan College of Arts and Science, Coimbatore.

S.Saranya 2

2 Research Scholar & Assistant Professor, Department of Computer Applications, Hindusthan College of Arts and Science, Coimbatore.

Dr.S.Sasikala 3

3 Associate Professor, Department of Computer Applications, Hindusthan College of Arts and Science, Coimbatore.

DOI: https://doi.org/10.36713/epra3747

ABSTRACT

Heart Disease is one of the most prevalent diseases that can lead to reduce the life span of human beings now a days. The initial finding of abnormal heart conditions is important to detect heart problems and avoid sudden cardiac death. Heart Disease are also known as Cardiovascular Disease which happen due to unhealthy lifestyle, smoking, alcohol and high intake of fats which may cause hypertension, high blood pressure, diabetes. They are caused by disorders of heart and blood vessels, and includes coronary heart disease (heart attacks). Data mining plays a major role in the construction of an intellectual prediction model for healthcare system to detect Heart Disease (HD) using patient data sets, which support doctors in diminishing mortality rate due to heart attack. In this review, we focus the novel and unique aspects of cardiovascular disease health and the methodologies used to predict the CVD.

KEYWORDS: Data mining, Classification, Particle Swarm Optimization, Cardiovascular Disease (CVD).
1. INTRODUCTION
Cardiovascular Disease (CVD) is increasing daily in this modern world. According to World Health Organization, an estimated 17 million people die each year from cardiovascular disease, particularly heart attacks and strokes[1]. For example, in a country such as the United States Cardiovascular Diseases are considered the main cause of death. As reported by America Heart Association, about 525,000 people suffer an early heart attack and about 210,000 people suffer repeatable ones, each year[2]. According to the reports of Center for Disease Control (CDC), almost more than half of deaths in 1999, occurred before any emergency services and hospital treatments to be done on the patient[3]. It is, therefore, necessary to record the most important symptoms and health habits that contribute to CVD. Various tests are performed prior to diagnosis of CVD, including auscultation, ECG, blood pressure, cholesterol and blood sugar etc. These tests are often long and when a patient’s condition may be critical and he or she must start taking medication immediately. So it is important to prioritize the test[4].

Data mining produces different techniques to discover the hidden patterns from huge amount of data. The heart disease diagnosis approaches in data mining are used to detect the heart disease from the patient’s medical data. The clinical data contains the information produced by the patient that may contain incomplete and interrelated data[5]. Classification is an effective data mining technique with wide ranging applications to classify the various kinds of data used in nearly every field in our life[6].

2. OPTIMIZATION ALGORITHMS
An optimization algorithm is a technique which is executed iteratively by comparing several solutions till an optimum or satisfactory solution is found. With the introduction of computers, optimization has become a part of computer aided design activities. The optimization algorithms are broadly divided into three categories[17]:
1. Biology based algorithms
2. Physical based algorithms
3. Geography based algorithms

1. Biology Based Algorithms: it is an important part of computational sciences, which are derived from or based on analogy to natural evolution and biological activities[18]. The Biology based algorithms are broadly classified into two types:

   i) Evolution based optimization algorithms
   ii) Swarm Based optimization algorithms

i) Evolution Based Algorithms (EAs):
Evolution Based algorithms are stochastic search methods that mimic the process of biological evolution and / or the social behavior of species. The species surpasses others by learning, adaptation, and evolution. There are number of evolution based algorithms, such as Genetic algorithm (GA), Evolutionary Programming (EP), Evolutionary Strategy (ES), Differential Evolution Algorithm (DE), Harmony Search Algorithm (HS).

ii) Swarm Based Algorithms:
A swarm is a great 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 are recently emerged as a family of nature-inspired, population based algorithms. They are capable of producing low cost, fast, and robust solutions to serve complex problems. Swarm is used to model the collective behavior of social nature such as ant colonies, honey bees, and bird flocks etc. The Swarm based algorithms are Artificial Immune System (AIS), Particle Swarm Optimization (PSO), Bacteria Foraging Optimization Algorithms (BFO), Cuckoo Search Algorithm (CS), Artificial Bee Colony Algorithm (ABC), Ant Colony Optimization (ACO), Coral Reef Optimization Algorithm (CRO), Teaching-Learning Based Optimization Algorithm (TLBO), Firefly Algorithm (FA), Shuffles Frog Leaping Algorithm (SFLA), Pigeon Inspired Optimization (PIO).

2) Physical Based Algorithm:
Physical Based optimization algorithms are the heuristic algorithms which mimic the physical properties or physical behavior of the matter are in the same philosophy as the laws of physics. These algorithms are classified as Simulated Annealing (SA), Gravitational Search Algorithm (GSA), Chaotic Optimization Algorithm (COA), Intelligent Water Drop algorithms (IWD), Magnetic Optimization Algorithm (MOP).

3) Geography Based Algorithms:
Geography Based Algorithms are the meta-heuristic algorithm which generate random solution in the geographical search space. This algorithm is classified as Tabu Search Algorithm (TS) and Imperialistic Competition Algorithm (ICA).
3. LITERATURE REVIEW

Youness Khourdifi et al.[7] proposed the Fast Correlated-Based Feature Selection (FCBF) method to filter redundant features in order to improve the quality of heart disease classification. They performed a classification based on various classification algorithms such as Naïve Bayes, K-Nearest Neighbor (K-NN), Support Vector Machine (SVM), Random Forest and a Multilayer Perceptron. Artificial Neural network was optimized by Particle Swarm Optimization (PSO) combined with Ant Colony Optimization (ACO) approaches. Their proposed mixed approach is applied to heart disease dataset. The proposed optimized model by FCBF, PSO and ACO achieved an accuracy score of 99.65% with K-NN and 99.6% with Random Forest.

Lalitha Kumari Gaddala et al.[8] used various Swarm Optimization techniques such as Swarm Intelligence (SI), Group Search Optimization (GSO), Artificial Bee Colony Optimization (ABC), Ant Colony Optimization (ACO) and Particle Swarm Optimization (PSO). They proposed that the Particle Swarm Optimization was the most population intelligent algorithm and had a good performance of optimization. They predict heart disease using Feed Forward of Artificial Neural Network (ANN) to classify patient as diseased and non-diseased. Finally the performance was evaluated in terms of Accuracy, sensitivity and specificity and compared to well-known other data sets.

A.Sheik Abdullah[9] proposed a model using Particle Swarm Optimization with J48 algorithm. The proposed model give reduced set of features using the feature selection algorithm along with improved prediction accuracy. The developed model had the functionalities such as evaluating the risk factors related to coronary Heart Disease. The experimental results showed that the prediction accuracy using PSO-J48 provided an improved results and reduced number of features than that of the other optimized algorithms.

Azhar Hussian Alkeshuosh et al[10] used Particle Swarm Optimization based algorithm to predict Heart disease. First the Random Rules are encoded and then they were optimized based on their accuracy using PSO Algorithm. The algorithm is compared with the Decision Tree based on C4.5 algorithm in UCI Repository of Machine Learning Databases. The experimental results showed the PSO achieved higher predictive accuracy and much smaller rule list than C4.5. Based on the average accuracy, the accuracy of PSO method is 87% and the accuracy of C4.5 is 63%. By using PSO, one can extract effective classification rules with acceptable accuracy. They concluded that PSO algorithm in rule production had good performance for rule discovery on continues data.

Majid Ghonji Feshki et al[11] used feature ranking on effective factors of disease related to Cleveland clinic database and by using Particle Swarm Optimization (PSO) as well as Neural Network Feed Forward Back Propagation, 13 effective factors reduced to 8 optimized features in terms of cost and accuracy. The valuation of selected features of classified methods also showed that PSO method with Neural Network of Feed Forward-Back Propagation had the best accurate criteria of the rate of 91.94% on these features.

Najmeh Ghadiri Hedeshi et al.[12] proposed a new ensemble PSO-based approach to extract a set of rules for diagnosis of coronary artery disease. The new presented boosting mechanism considers the cooperation between generated fuzzy if-then rules using PSO meta-Heuristic. They evaluated their new classification approach using well known Cleveland data set. The experimental results gained 92.59% classification accuracy, 94.37% specificity and 90.51% sensitivity values for diagnosis of CAD which are acceptable results.

Jairam P.Kelwade et al.[13] proposed a hybridized Particle Swarm Optimization and Radial Basis Function Neural Network (RBFN) model to improve cardiac arrhythmia prediction accuracy. The performance of RBFN is more sensitive to parameter such as spread. In the training process RBFN, the spread was optimized using PSO. In the proposed method a set of linear and non-linear features are extracted from the RR interval time series datasets, which are derived from the MIT-BIH arrhythmia database. The proposed hybrid approach could be effective for improving the performance evaluation of RBFN classifier. It is also observed that the proposed approach is more accurate in predicting more number of datasets of normal and abnormal heart patients.
Animesh Kumar Paul et al.[14] proposed DMS-PSO system to select the critical attributes and assist the diagnosis of heart disease. They showed that the combination of the Fuzzy logic and DMS-PSO can offer more effective systems of medical diagnosis with improved system accuracy. From the experimental analysis, the FDSS exhibits a relatively high performance when compared with existing system.

Chaitanya Suvarna et al.[15] proposed a prediction algorithm with the help of Data Mining and Optimization Techniques. In their research they used Constricted Particle Swarm Optimization model to produce the better accuracy. They proved that CPSSO performs better than some of the techniques used for prediction. From the experimental results, it can be concluded that the best approach to use with particle swarm optimization is to utilize the construction factor approach.

Indu Yekkala et al.[16] analyzed various ensemble methods (Bagged tree, Random Forest, and AdaBoost ) along with Feature Subset Selection method- Particle Swarm Optimization(PSO), to accurately predict the occurrence of Heart Disease for a particular patient. Their proposed approach used PSO as a feature selection method to reduce the least ranked feature. Then applied ensemble methods as a classifier to reduce misclassification rate and to improve the classification performance. From the experimental results, it had been proved that the learning accuracy can be significantly improved by using Bagged Tree Ensemble Classifier on PSO. This model will help medical practitioners to accurately predict and early diagnosis of heart disease using a subset of features.

B.Subanya et al.[19] used a meta-heuristic algorithm to determine the optimal feature subset with improved classification accuracy in cardiovascular disease diagnosis. Swarm Intelligence(SI) based Artificial Bee Colony(ABC) Algorithm is ultimately used to find the best features in the disease prediction. To evaluate the fitness of ABC, Support Vector Machine Classification is used. The performance of the proposed algorithm was validated against Cleveland Heart Disease dataset from UCI machine learning repository. The experimental results showed that, ABC-SVM performed better than Feature Selection with reverse ranking.

Yuanning Liu [20] proposed a design of modified Multi-Swarm PSO (MSPSO) to solve discrete problems, which consists of a number of sub-swarms and a multi-swarm scheduler that can check and control each sub-swarm using the rules. For Feature Selection problems, they proposed an Improved Feature Selection(IFS) method by integrating MSPSO, Support Vector Machines (SVM) with F-score method. The performance of the proposed method was compared with that of the standard PSO based, Genetic Algorithm based and the Grid search based methods on 10 benchmark datasets taken from UCI machine learning and StatLog dataset. Their results showed that IFS methods performed better than the other methods.

Sana Bharti et al.[21] proposed Particle Swarm Optimization for Feature Reduction Artificial Neural Network was used for classifying the patient as diseased and non-diseased. The network was trained under two conditions. One is before applying PSO and the other is after applying PSO on training data. The performance of the network is evaluated by using the techniques of Regression, Performance plot, Confusion Matrix and ROC value. It was evident that the performance of the whole network was increased after the application of PSO for Feature Reduction. From the experiment, it was observed that PSO algorithm was very effective in pre-processing medical data. The Pre-processed data using PSO algorithm increases overall network performances.

Durairaj. M et al.[23] proposed a technique of preprocessing the dataset and using Particle Swarm Optimization (PSO) algorithm for Feature Reduction. After applying PSO, the accuracy for prediction is tested. It was observed from the predictions, a potential result of 83% accuracy in the prediction. The performance of the PSO algorithm is then compared with Ant Colony Optimization (ACO) algorithm. The experimental results showed that the PSO algorithm was better than ACO.
### 4. REVIEW ANALYSIS

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Purpose</th>
<th>Techniques used</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Youness Khourdifi et al[7]</td>
<td>2019</td>
<td>Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization</td>
<td>Artificial Neural Network optimized by Particle Swarm Optimization combined with Ant colony optimization</td>
<td>99.65% with K-NN and 99.6% with Random Forest</td>
</tr>
<tr>
<td>Lalitha Kumari Gaddala et al[8]</td>
<td>2018</td>
<td>An Analysis of Heart Disease Prediction Using Swarm Intelligence Algorithms</td>
<td>Artificial Neural Network(ANN) with Swarm Intelligence, Ant Colony Optimization, Artificial Bee Colony Optimization and Group Search Optimization and Particle Swarm Optimization</td>
<td>Evaluated results with Accuracy, Sensitivity and Specificity compared with many data sets.</td>
</tr>
<tr>
<td>A.Sheik Abdullah[9]</td>
<td>2012</td>
<td>A Data mining Model to Predict and Analyze the Events Related to Coronary Heart Disease using Decision Tree with Particle Swarm Optimization for Feature Selection</td>
<td>Particle Swarm Optimization with J48 Algorithm with Feature Selection Algorithm</td>
<td>Improved Accuracy compared by without Feature Selection Results.</td>
</tr>
<tr>
<td>Azhar Hussain Alkeshuosh et al[10]</td>
<td>2017</td>
<td>Using PSO Algorithm for Producing Best Rules in Diagnosis of Heart Disease</td>
<td>Particle Swarm Optimization(PSO)</td>
<td>PSO provided more accuracy than C4.5 algorithm. PSO produced 87% accuracy. C4.5 produced 63%.</td>
</tr>
<tr>
<td>Majid Ghonji Feshki et al[11]</td>
<td>2016</td>
<td>Improving the Heart Disease Diagnosis by Evolutionary Algorithm of PSO and Feed Forward Neural Network</td>
<td>Particle Swarm Optimization along with Neural Network Feed Forward Back Propagation</td>
<td>Produced accuracy of 91.94%.</td>
</tr>
<tr>
<td>Najmeh Ghadiri Hedeshi et al[12]</td>
<td>2011</td>
<td>An Expert System Working upon an Ensemble PSO-Based Approach for Diagnosis of Coronary Artery Disease</td>
<td>New Ensemble Particle Swarm Optimization(PSO)</td>
<td>92.59 % Accuracy, 94.37% Specificity, 90.51 % Sensitivity produced.</td>
</tr>
<tr>
<td>Jairam P.Kelwade et al[13]</td>
<td>2016</td>
<td>Prediction of heart abnormalities using Particle Swarm Optimization in Radial Basis Function Neural Network</td>
<td>Hybridized Particle Swarm Optimization and Radial Basis Function Neural Network(RBFN)</td>
<td>Produced more accuracy in predicting more number of datasets of normal and abnormal heart patients.</td>
</tr>
<tr>
<td>Animesh Kumar Paul et al[14]</td>
<td>2015</td>
<td>Fuzzy Membership Function Generation using DMS-PSO for the Diagnosis of Heart</td>
<td>DMS based Particle Swarm Optimization combined with Fuzzy</td>
<td>FDSS exhibits a relatively high performance when</td>
</tr>
<tr>
<td>Authors</td>
<td>Year</td>
<td>Title</td>
<td>Description</td>
<td>Comparison</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------</td>
<td>------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------------</td>
<td>------------</td>
</tr>
<tr>
<td>Chaitanya Suvarna et al[15]</td>
<td>2017</td>
<td>Efficient Heart Disease Prediction System using Optimization Techniques</td>
<td>Constricted Particle Swarm Optimization (CPSO)</td>
<td>CPSO performed better for prediction than the existing methods.</td>
</tr>
<tr>
<td>Indu Yekkala et al[16]</td>
<td>2017</td>
<td>Prediction of Heart Disease using Ensemble Learning and Particle Swarm Optimization</td>
<td>Various Ensembles (Bagged tree, Random Forest, and AdaBoost) along with Feature Subset Selection method- Particle Swarm Optimization(PSO)</td>
<td>Proved that the learning accuracy can be significantly improved by using Bagged Tree Ensemble Classifier on PSO.</td>
</tr>
<tr>
<td>B.Subanya et al[19]</td>
<td>2014</td>
<td>Feature Selection using Artificial Bee Colony for Cardiovascular Disease Classification</td>
<td>Artificial Bee Colony (ABC) with Support Vector Machine(SVM)</td>
<td>Provided better accuracy than Feature Selection with reverse ranking in Classification.</td>
</tr>
<tr>
<td>Yuanning Liu [20]</td>
<td>2011</td>
<td>An Improved Particle Swarm Optimization for Feature Selection</td>
<td>Modified Multi- Swarm PSO(MSPSO) with Improved Feature Selection (IFS)</td>
<td>MSPSO with IFS method performed better than the other existing methods.</td>
</tr>
<tr>
<td>Sana Bharti et al[21]</td>
<td>2015</td>
<td>Analytical Study of Heart Disease Prediction Comparing With Different Algorithms</td>
<td>Particle Swarm Optimization(PSO), Genetic Algorithm, Artificial Neural Network</td>
<td>PSO provides better results than the other methods.</td>
</tr>
<tr>
<td>Sivagowry.S et al[22]</td>
<td>2014</td>
<td>PSO - An Intellectual Technique for Feature Reduction on Heart Malady Anticipation Data</td>
<td>Particle Swarm Optimization for Feature Reduction</td>
<td>The Performance of the whole network was increased after the application of PSO for Feature Reduction.</td>
</tr>
<tr>
<td>Durairaj. M et al[23]</td>
<td>2015</td>
<td>Feature Diminution by Using Particle Swarm Optimization for Envisaging the Heart Syndrome</td>
<td>Particle Swarm Optimization for Feature Reduction</td>
<td>PSO provided 83% of accuracy. It was higher than ACO algorithm.</td>
</tr>
</tbody>
</table>

**Table 1: Review Report of Optimization Algorithms in Heart Disease Identification**

**5. FINDINGS FROM THE SURVEY**

Prediction of Heart Disease by using Machine Learning Techniques are the Challenging task in Health care. CVD continues to be the leading cause of death for women in all over the world. So our future work based on the Cardiovascular Disease for women those who are working in various sectors. Particle Swarm Optimization is the powerful method for predicting heart disease compared with other methods. In future, we focus the Optimization Techniques for producing the better results of Predicting women who is affecting Heart Disease.
6. ARCHETYPE DESIGN FOR THE PROPOSED WORK

7. CONCLUSION AND FUTURE WORK

The main objective of this work is to study various optimization algorithms that can be used to predict the Cardiovascular Disease. Optimization Algorithms are more effective in Heart Disease Classification. Particle Swarm Optimization is one the powerful algorithm which plays an important role in Heart Disease Prediction. In this study, we conclude that using Particle Swarm Optimization for Heart Disease Identification will produce better results than the other existing methods. In our future work, we will use Particle Swarm Optimization with the combination of some other classification algorithms to predict the Heart Disease of women.

REFERENCES

11. Majid Ghonji Feshki and Omid Sojoodi Shijani,” Improving the Heart Disease Diagnosis by Evolutionary Algorithm of PSO and Feed Forward Neural Network” Artificial Intelligence and Robotics (IRANOPEN), 2016, 2016, pp. 48-53.


16. Indu Tekala, Sunanda Dixit and M.A.Jabbar, “Prediction of Heart Disease using Ensemble Learning and Particle Swarm Optimization” International Conference on Smart Technologies for Smart Nation 2017.


