Implementation of Some Bio-Inspired Algorithms in Prediction of Heart Diseases: A Review

Raghunath Satpathy[1]
[1]*Department of Biotechnology, MITS Engineering College, Rayagada, Odisha, India*

***Abstract*—** Heart constitutes as one of the foremost important organ of the human body. The natural mechanism of operation of a heart is very complex and failure by any means (disease) is risky to human lives. Various technologies have been found to be effective in the heart diagnosis systems, however, to avoid any kind of medical error and unwanted results, computer-based diagnostic systems are preferred. Recent biomedical research literature emphasizes about is much interest from the scientific researchers in implementing the human intelligence, in health care industry. Various data mining techniques have been used to make clinical decision support systems, to get accurate results on the basis of information collected by researches from the study. To facilitate this, several computational algorithms can be implemented for the effective prediction of heart disease. This review highlights about different bio-inspired algorithms and their implementation specifically for the heart disease prediction.

***Keywords*—** Heart disease prediction; bio inspired algorithms, genetic algorithms, ant colony optimization, particle swarm optimization, artificially bee colony optimization

I. INTRODUCTION

Worldwide basis humans are affected by many types of life threatening diseases, among of this, the heart disease has received more attention. Heart disease basically causes the injury of the heart and the blood vessels. Therefore, the heart syndrome is a most important reason for mortality and death for people in most of the countries all over the world. According to one survey in 2008 approximately 17.3 million people died from heart diseases (7.3 million deaths were due to coronary heart disease and 6.2 million were due to stroke), that corresponds to the 30% of all global deaths had occurred. Recently, many types of research in the medical industry have been able to identify risk factors of heart diseases (Table I), however, more contribution is necessary to use this knowledge to reduce the risk of deaths. The significant mortality rate caused by the heart disease throughout the globe need for the development of new heart disease prediction methods [1-2-3-4]. These systems allow patients to calculate the heart disease risks. There are many factors of heart disease that affecting the structure or function of the heart. Some of the common symptoms of this disease are feeling discomfort in the chest area (chest pain), shortening of breath, excess sweating, dizziness, etc., however, the above symptoms differ from person to person [5]. So, this might be challenging for the doctors to predict heart disease accurately. Therefore, it is important to utilize automated innovations in heart disease prediction in order to help specialists to determine quicker to have higher precisely. The medical diagnosis helps to identify the symptoms and causes of this disease. The diagnostic procedures produce information about the different variations of the disease. Researchers have developed many techniques such as data mining, soft computing and optimization techniques for diagnosing heart disease [6-7-8].

TABLE 1 RISK FACTORS ASSOCIATED WITH THE HEART DISEASE

Cause	Explanation
Smoking	The chemicals present in the tobacco smoke harm blood cells also damage the function of heart and the structure and function of blood vessels ultimately leads to heart attack
Obesity	Extra weight increases risk for heart function by damaging in the main pumping chamber (left ventricle), which prevent it from filling sufficiently between beats.
High Blood Pressure	High blood pressure or hypertension is a widely increase the risk of the walls of our blood vessels walls becoming over stretched and injured
High blood cholesterol	Presence of the high amount of cholesterol (a fat-like substance) in the blood, decreases the walls of arteries, thereby creating it narrowed and blood flow to the heart is slowed down or blocked
Poor Diet	Choosing the right foods can make a difference in living a long and healthy life or helps to face with the diseases like a heart attack or heart disease
Physical Inactivity	Lack of exercise is a risk factor for developing coronary artery disease and also increases the risk for high blood pressure causes heart diseases

II. ROLE OF ALGORITHMS AND FEATURE SELECTION METHODS IN HEART DISEASE PREDICTION

Proper healthcare management system of a country having a strong impact on the economy as well as daily life of the people. Therefore, the data mining through various bio-inspired algorithms has been developed to explore hidden reasons from a large volume of clinical disease data sets for making the scientific prediction (Fig 1). One of the challenging procedure is to analyze the disease data sets because these are large in quantities to make scientific predictions.So prediction system by implementation of algorithms will be helpful for the early diagnosis of several diseases by selecting the correct parameters [9-10-11]. Therefore, the data mining through various bio-inspired algorithms have been derived by different diagnostic procedures can be massive and high dimensional in nature. Feature selection is the method of selecting a feasible subset of features from the original set of candidate features. So proper feature selection method is applied to datasets to identify the significant features and discard irrelevant or redundant features and to reduce the time and the resource

usage. [12-13-14]. The features must be selected in such a way that, it should enhance the accuracy of the prediction of the disease (Table II). Prediction and classification methods from the disease datasets have been well studied in the recently published literature, but the accuracy level of acceptance models are still to be improved.

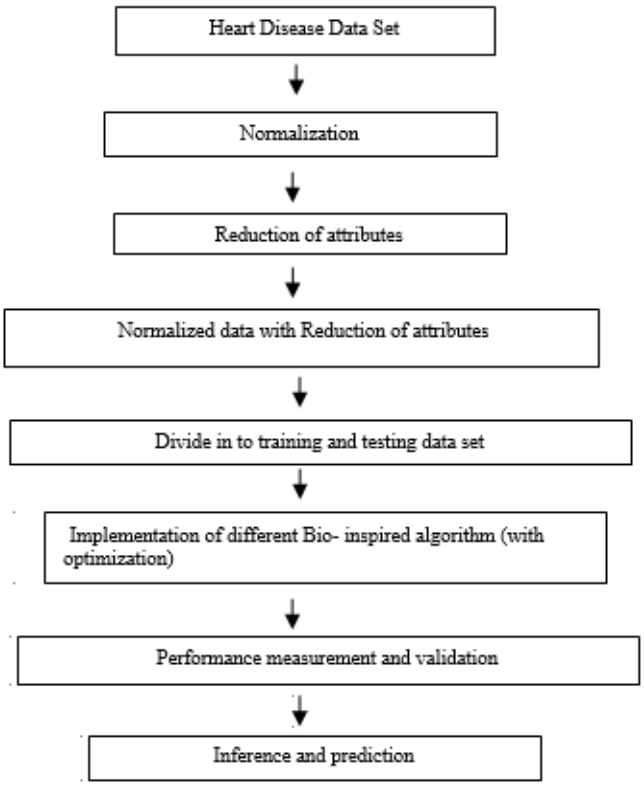

Fig.1 General scheme of implementation of nature inspired algorithms in heart disease dataset

TABLE II CLINICAL FEATURES USED IN THE CLASSIFICATION OF HEART DISEASE PREDICTION SYSTEMS FROM THE UCI DATA SET (OBTAINED FROM https://archive.ics.uci.edu/ml/datasets/heart+disease)

Feature	Description
Age	Age in years
Sex	Male, female
Cp	Chest pain type Value 1: typical angina Value; 2: atypical angina Value 3:non–anginal pain; 4: asymptomatic
Trestbps	Patient's resting blood pressure in mm Hg at the time of admission to the hospital 5
Chol	Serum cholesterol in mg/dl
Fbs	Boolean measure indicating whether fasting blood sugar is greater than 120 mg/dl (1 = True; 0 = false)
Restecg	Electrocardiographic results during rest
Thalach	Maximum heart rate attained
Exang	Boolean measure indicating whether exercise induced angina has occurred
Oldpeak	ST depression brought about by exercise relative to rest
Slope	The slope of the ST segment for peak exercise
Ca	Number of major vessels (0–3) coloured by fluoroscopy
Thal	The heart status (normal, fixed defect, reversible defect
Predicted attribute	Either 0 or 1

III. ROLE OF ALGORITHMS AND FEATURE SELECTION METHODS IN HEART DISEASE PREDICTION METHODS

Many models have been used to detection of the heart disease along with the most significant factor that contributes for the cause of the disease [15-16-17]. Several popular algorithms and their implementation process have been discussed in the below section.

A. Genetic Algorithms

Genetic algorithms (GA) concept was first introduced by John Holland in 1975 and widely used to solve basically the searching and the optimization problems. GA uses the fundamentals of genetics principles for problem-solving. As an emerging field of artificial intelligence, a genetic algorithm (GA) being heuristic nature, widely used for disease prediction problems [18-19-20-21-22]. The technique generates optimized solutions using techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. Each solution generated in the Genetic algorithms is called a chromosome. Each chromosome is made up of genes, which are the individual elements (alleles) that represent the problem. The collection of chromosomes is called a population.

Basically, there are three genetic operators are used for generating new strings. The functions of genetic operators are as follows:

- Selection: The process of selection basically deals with the probability of survival of the fittest i, e more fit chromosomes are chosen to survive.
- Crossover: This operation is performed by selecting a random gene along the length of the chromosomes and swapping all the genes after that point. The crossover operators may be of single point, two point or multipoint types.
- Mutation: Mutation alters the new solutions and searches for the better solution.

Latha Parthiban and R. Subramanian in 2007, [23] presented an Intelligent Heart Disease Prediction System using Genetic Algorithm. In this method, the fuzzy inputs adapted with a modular neural network to evaluate a rapid and accurate complex functions. The model further in combination with the neural network and genetic algorithm holds good for the diagnosis of the presence of the disease. Similarly, a framework for decision support system is developed for the analysis of medical data [24]. A genetic-based neural network approach is used to predict the severity of the cardiovascular disease by [25]. In this work, the weights for the neural network are determined using the genetic algorithm and the pre-processed data was classified into five classes based on the severity of the disease using back propagation algorithm. A novel algorithm that combines KNN with genetic algorithm was used for effective prediction and classification systems and diagnosis of heart disease [26-27]. Another proposed heart disease prediction system is developed using the neural network and genetic algorithm to train the network with uses the global optimization techniques. For this prediction, system

12 parameters such as sex, age, blood cholesterol were used to predict the heart disease up to 98% accuracy [28].

B. Ant Colony optimization (ACO) algorithm

Ant colony optimization (ACO) was first proposed by Dorigo et al. [29]. The principles behind the algorithm are about the activity of the real ants that are capable to find the direct path from a source of food and to their nest. The ant colony optimization algorithm (ACO) method uses a probabilistic technique for solving computational problems. This method uses the random movement of ants upon finding food return to their colony while laying down pheromone trails. However, if other ants find such a path, then they stop the random traveling, but follow the trail to find the food. As time goes on, the pheromone starts to evaporate, hence reducing its attractive power. The time period that is taken by an ant to travel along the path and back is directly proportional to the reduction in pheromone concentration. So a short path is frequently adopted by the ants that provide the higher pheromone density rather than the longer ones. A condition whenever there is no evaporation of pheromones then, the paths chosen by the first ants can be considered as attractive to the previous one that follows [30].

An efficient data mining algorithm has been implemented by Dubey et al. 2014 [31] with the aim to prevent the risk of heart disease this in the earlier stages. This method generated support is used as a weight of the symptom which will be the initial pheromone value of the ant. The risk level is identified by finding the max pheromone value, the increasing detection rate of the heart disease was observed. A solution based on ant colony optimization and tailored for the case of Bayesian classifiers, that ultimately enhances the predictability of heart disease [32]. Another ACO based techniques implemented by ACO will be used for pattern classification in which classification is based on fuzzy rules in which the fuzzy rules are used to control the influence of pheromone values in ACO. For update the pheromone rule that improves the quality of each rule [33].

C. Particle Swarm optimization Algorithm

Particle swarm optimization (PSO) has been used to solve numerous optimization problems developed by Kennedy and Eberhart in 1995 [34]. This is a population-based optimization method stimulated by social behavior of bird flocking. PSO consists of a swarm of particles and each particle exists in at a position in the search space. The fitness of each particle represents the quality of its position. The particles fly through the search space with a certain velocity and it is based upon its own best position. Currently, by using this principle, several different heart disease prediction systems have been developed [35-36]. A multi-layer feed forward neural network (MLFFNN) optimized with particle swarm optimization is adopted for the heart disease prediction at the early stage by using the clinical record [37]. In another work, the PSO optimized neural network (PSONN) were implemented for optimization are the number of hidden neurons, momentum factor, and learning rate to enhance the efficiency [38].

D. Artificial Bee Colony Optimization (ABC) Algorithm

Artificial Bee Colony (ABC) algorithms are basically developed by considering the intelligent behavior of real honey bee colonies. In a typical bee colony, usually, three categories of bees are found such as employed bees, onlooker bees, and scout bee. The work of the employed bees are they try to improve source for their food which is analogous to the feasible solution for the optimization problem corresponds to the initialization of the algorithm. These employed bees are usually involved in the foraging of food sources and move position information about them to the hive. Similarly, the onlooker bees search around the solutions of employed bees by considering information shared by employed bees. The function of the scout bee is, if an employed bee could not improve self-solution in a certain this employed bee becomes a scout bee. After a new solution is produced for this scout bee, the scout bee becomes employed bee. The main advantage in ABC allows the results to converge to the optimal solution quickly and also it is simple and easy to implement This algorithm is iterative in nature and solves the problems in four levels, named as initialization level, employed bee level, onlooker bee level, and scout bee level [39-40].

An ABC based algorithm was developed by İsmail Babaoğlu et al. [41] that utilizes the clinical information of coronary artery disease of patient and utilizing as the training followed by classification by using k-nearest neighbor algorithm. This resulted as an alternative classifier for diagnosis of coronary artery disease. Similarly, a binary artificial bee colony (BABC) algorithm is used to find the best features in the disease identification. The fitness of the BABC system is evaluated using Naive Bayesian method [42]. Subanya, & Rajalaxmi also evaluated the fitness of BABC by using K–Nearest Neighbor (KNN) method [43]. The machine-learning algorithms have been implemented in Bee Colony Optimization (BCO) are used to predict and diagnoses the heart diseases by entering the symptoms by the user [44]. In another work the fuzzy systems are used to predict the heart disease prediction accuracy is computed with Artificial Bee Colony Optimization with improved manner [45].

IV. CONCLUSION

A foremost major challenge facing healthcare organizations such as hospitals, medical centers is to provide disease diagnosis facility services at affordable costs. In the recent developing countries including India have the rate of cardiovascular disease is more and also it is predicted that by 2020 coronary heart disease will be leading cause of death in adult Indians. Therefore, a decision support system is required to predict the heart disease. Several data mining, machine learning, artificial intelligence, pattern recognition, soft computing and optimization techniques have been proposed for heart disease diagnosis. The major aim of implementing these algorithms is to determine the attribute which contributes towards the diagnosis of disease to facilitate the risk management. The current chapter discusses specifically the implementation of four different bio-inspired algorithms

used for heart disease diagnosis purpose. Many of such algorithm is to be implemented both in natural and in their hybridized manner to enhance the prediction accuracy of the heart disease.

V. REFERENCES

[1] Kim, Jaekwon, Jongsik Lee, and Youngho Lee. "Data-mining-based coronary heart disease risk prediction model using fuzzy logic and decision tree." Healthcare informatics research, vol.21, no. 3, pp. 167-174, 2015.

[2] Mythili, T., Dev Mukherji, Nikita Padalia, and Abhiram Naidu. "A heart disease prediction model using SVM-Decision Trees-Logistic Regression (SDL)." International Journal of Computer Applications, vol. 68, no. 16, pp. 11-15, 2013.

[3] Palaniappan, Sellappan, and Rafiah Awang. "Intelligent heart disease prediction system using data mining techniques." In Computer Systems and Applications, 2008. AICCSA 2008. IEEE/ACS International Conference on, pp. 108-115. IEEE, 2008.

[4] Lloyd-Jones, Donald, Robert J. Adams, Todd M. Brown, Mercedes Carnethon, Shifan Dai, Giovanni De Simone, T. Bruce Ferguson et al. "Heart disease and stroke statistics—2010 update." Circulation 121, no. 7 (2010): e46-e215.

[5] Whooley, Mary A., Peter de Jonge, Eric Vittinghoff, Christian Otte, Rudolf Moos, Robert M. Carney, Sadia Ali et al. "Depressive symptoms, health behaviors, and risk of cardiovascular events in patients with coronary heart disease." Vol. 300, no. 20, pp.2379-2388, 2008.

[6] Rajkumar, Asha, and G. Sophia Reena. "Diagnosis of heart disease using datamining algorithm." Global journal of computer science and technology, vol.10, no. 10, pp.38-43, 2010.

[7] Soni, Jyoti, Ujma Ansari, Dipesh Sharma, and Sunita Soni. "Predictive data mining for medical diagnosis: An overview of heart disease prediction." International Journal of Computer Applications, vol.17, no. 8, pp.43-48, 2011.

[8] Nawi, Nazri Mohd, Rozaida Ghazali, and Mohd Najib Mohd Salleh. "The development of improved back-propagation neural networks algorithm for predicting patients with heart disease." In International Conference on Information Computing and Applications, pp. 317-324. Springer, Berlin, Heidelberg, 2010.

[9] Masethe, Hlaudi Daniel, and Mosima Anna Masethe. "Prediction of heart disease using classification algorithms." In Proceedings of the world Congress on Engineering and computer Science, vol. 2, p. 2224. 2014.

[10] Srinivas, K., B. Kavihta Rani, and A. Govrdhan. "Applications of data mining techniques in healthcare and prediction of heart attacks." International Journal on Computer Science and Engineering (IJCSE), vol. 2, no. 02, pp.250-255, 2010.

[11] Xing, Yanwei, Jie Wang, and Zhihong Zhao. "Combination data mining methods with new medical data to predicting outcome of coronary heart disease." In Convergence Information Technology, 2007. International Conference on, pp. 868-872. IEEE, 2007.

[12] Amin, Syed Umar, Kavita Agarwal, and Rizwan Beg. "Genetic neural network based data mining in prediction of heart disease using risk factors." In Information & Communication Technologies (ICT), 2013 IEEE Conference on, pp. 1227-1231. IEEE, 2013.

[13] Ordonez, Carlos, Edward Omiecinski, Levien De Braal, Cesar A. Santana, Norberto Ezquerra, Jose A. Taboada, David Cooke, Elizabeth Krawczynska, and Ernest V. Garcia. "Mining constrained association rules to predict heart disease." In Data Mining, 2001. ICDM 2001, Proceedings IEEE International Conference on, pp. 433-440. IEEE, 2001.

[14] Devi, S. Kiruthika, S. Krishnapriya, and Dristipona Kalita. "Prediction of Heart Disease using Data Mining Techniques." Indian Journal of Science and Technology, vol. 9, no. 39, 2016. DOI: 10.17485/ijst/2016/v9i39/102078.

[15] Pattekari, S. A., & Parveen, A. (2012). Prediction system for heart disease using Naïve Bayes. International Journal of Advanced Computer and Mathematical Sciences, 3(3), 290-294.

[16] Dangare, Chaitrali S., and Sulabha S. Apte. "Improved study of heart disease prediction system using data mining classification techniques." International Journal of Computer Applications, vol.47, no. 10, pp. 44-48, 2012.

[17] Anbarasi, M., E. Anupriya, and N. C. S. N. Iyengar. "Enhanced prediction of heart disease with feature subset selection using genetic algorithm." International Journal of Engineering Science and Technology, vol. 2, no. 10, pp.5370-5376, 2010.

[18] Tang, Zhonghua, and Qin Liao. "A new class based associative classification algorithm." Imecs, vol. 2007, pp. 685-689, 2007.

[19] Goldberg, David E., and John H. Holland. "Genetic algorithms and machine learning." Machine learning,vol. 3, no. 2,pp.95-99, 1988.

[20] Ibrahim, SP Syed, K. Chandran, and J. Christopher. "An evolutionary approach for rule set selection in a class based associative classifier." Eur. J. Sci. Res, vol.50, no. 3, pp. 417-425, 2011.

[21] Picek, Stjepan, and Marin Golub. "On the efficiency of crossover operators in genetic algorithms with binary representation." In Proceedings of the 11th WSEAS International Conference on Neural Networks. 2010.

[22] Jabbar, M. Akhil, Bulusu Lakshmana Deekshatulu, and Priti Chandra. "Heart disease prediction system using associative classification and genetic algorithm." arXiv preprint arXiv: 1303.5919 (2013).

[23] Parthiban, Latha, and R. Subramanian. "Intelligent heart disease prediction system using CANFIS and genetic algorithm." International Journal of Biological, Biomedical and Medical Sciences, vol.3, no. 3,pp. 157-160, 2008.

[24] Amma, NG Bhuvaneswari. "Cardiovascular disease prediction system using genetic algorithm and neural network." In Computing, Communication and Applications (ICCCA), 2012 International Conference on, pp. 1-5. IEEE, 2012.

[25] Shanthi, D., G. Sahoo, and N. Saravanan. "Evolving connection weights of artificial neural networks using genetic algorithm with application to the prediction of stroke disease." International Journal of Soft Computing, vol. 4, no. 2, pp. 95-102, 2009.

[26] Deekshatulu, B. L., and Priti Chandra. "Classification of heart disease using k-nearest neighbor and genetic algorithm." Procedia Technology, vol.10, pp.85-94, 2013.

[27] Jabbar, M. A., B. L. Deekshatulu, and Priti Chandra. "An evolutionary algorithm for heart disease prediction." Communications in Computer and Information Science. Springer Verlag, vol. 292, pp.378-389, 2012.

[28] Waghulde, Nilakshi P., and Nilima P. Patil. "Genetic neural approach for heart disease prediction." International Journal of Advanced Computer Research, vol. 4, no. 3, pp.778, 2014.

[29] Dorigo, Marco, Vittorio Maniezzo, and Alberto Colorni. "The ant system: An autocatalytic optimizing process." (1991).

[30] Goss, Simon, Serge Aron, Jean-Louis Deneubourg, and Jacques Marie Pasteels. "Self-organized shortcuts in the Argentine ant." Naturwissenschaften, vol.76, no. 12, pp.579-581, 1989.

[31] Dubey, Animesh, Rajendra Patel, and Khyati Choure. "An efficient data mining and ant colony optimization technique

(DMACO) for heart disease prediction." International Journal of Advanced Technology and Engineering Exploration (IJATEE), vol.1, no. 1, pp.1-6, 2014.
[32] Bouktif, Salah, Eileen Marie Hanna, Nazar Zaki, and Eman Abu Khousa. "Ant Colony Optimization Algorithm for Interpretable Bayesian Classifiers Combination: Application to Medical Predictions." PloS one,vol. 9, no. 2, (2014): e86456.Doi: https://doi.org/10.1371/journal.pone.0086456.
[33] Kavitha S, and Nithya M "An Efficient Cardiovascular Disease prediction through Pheromone Based ACO with Hybrid Fuzzy logic." International Journal of Scientific Research, vol.5, no.8, doi: 10.15373/22778160.
[34] Kennedy, James. "Particle swarm optimization." In Encyclopedia of machine learning, pp. 760-766. Springer US, 2011.
[35] Swati Sharma and Dr. Sukhvir Singh " Heart Disease Diagnosis using Genetic and Particle Swarm Optimization", International Journal of Engineering Research & Technology, Vol. 3, no. 8, pp. 1499-1503, 2014.
[36] Chitra, R., and V. Seenivasagam. "Risk Prediction of Heart Disease Based on Swarm Optimized Neural Network." In Proceedings of International Conference on Computer Science and Information Technology, pp. 707-714. Springer, New Delhi, 2014.
[37] Durairaj, M and Sivagowry, SA "Survey on Particle Swarm Optimization and Rough Set Theory in Feature Selection for Heart Disease Prediction", vol.4, no.3, pp. 87-92, 2015.
[38] Deivanai, M, "Link Prediction for Heart Diseases using Heuristic Approach", Doi: 10.6084/ijact.v0i0.442, 2012.
[39] Karaboga, Dervis. "Artificial bee colony algorithm." scholarpedia, vol.5, no. 3, 6915, 2010.
[40] Karaboga, Dervis, and Bahriye Basturk. "On the performance of artificial bee colony (ABC) algorithm." Applied soft computing, vol. 8, no. 1, pp. 687-697, 2008.
[41] Babaoglu, Ismail, Mustafa Servet Kiran, Erkan Ülker, and Mesut Gündüz. "Diagnosis of coronary artery disease using artificial bee colony and k-nearest neighbor algorithms." International Journal of Computer and Communication Engineering, vol. 2, no. 1 pp.56, 2013.
[42] Subanya, B., and Rajalaxmi, R. R. "Artificial bee colony based feature selection for effective cardiovascular disease diagnosis." International Journal of Scientific & Engineering Research, vol.5, no. 5, pp. 606-612, 2014.
[43] Subanya, B., and Rajalaxmi, R. R. "A Novel Feature Selection Algorithm for Heart Disease Classification" International Journal of Computational Intelligence and Informatics, Vol. 4, no. 2, pp. 117-124, 2014.
[44] Goyal, M., Kaur, M., and Kaur, K." Heart Expert System using Bee Colony Optimization (BCO) Algorithm and Biogeography Based Optimization (BBO) Algorithm", vol.4, no. (Spl.) 1, pp. 81-83, 2013.
[45] Sowmya, N., and Vijayabhanu, R., "Artificial Bee Colony (ABC) Optimization for the Prediction of Coronary Heart Disease". International Journal of Trend in Research and Development, vol. 2, no. 5, pp. 427-432, 2015.

Raghunath Satpathy received his M.Sc in Botany (with specialization in Biotechnology) from Berhampur University science, India; (Post M.Sc) Advanced P.G Diploma in Bioinformatics from the University of Hyderabad and M. Tech. degree in Biotechnology from VIT University, Vellore India. He has submitted the thesis for the degree of the doctor of philosophy (Ph.D.) in Biotechnology of Sambalpur University, Odisha India. Currently, he is working as an Assistant professor in M.I.T.S Engineering College Odisha, India. He is having more than 9 years of experiences in both teaching and research. His current research interests include computational analysis of protein structure and environmental applications, data mining, biological database and analytical tool development. He is the recipient of prestigious Adarsh Vidya Saraswati Rashtriya Puraskar (A national award of excellence from Global Management council, Ahmedabad, India) and Senior Educator and Researcher award (A national award from National foundation for entrepreneurship development (NFED), Coimbatore, India).