Implementing Bio Inspired Algorithms for Prediction of Heart Disease

Dileep P 1, Revathy P 2

1 Associate Professor, Department of Computer Science and Engineering
2 Assistant Professor, Department of Computer Science and Engineering

1 Malla Reddy College of Engineering and Technology, Kompally, Hyderabad, India.
2 Narsimha Reddy Engineering College, Kompally, Hyderabad, India.

Abstract
Caused by both environmental and genetic factors, the prevalence of health problems is steadily rising. In this regard, heart disease has emerged as the leading killer of humans in recent decades. In this research, we employ an Artificial Neural Network trained with a swarm Intelligence approach to the problem of cardiac disease prediction. Swarm intelligence (SI) is a cutting-edge academic subfield that spans traditionally separate disciplines. Swarm-based algorithms are a new class of population-based algorithms that take their cues from nature and can efficiently solve several difficult problems with minimal resources. Group Search Optimization (GSO), Artificial Bee Colony (ABC), Ant Colony Optimization (ACO), and Particle Swarm Optimization (PSO), to name a few, are only a few of the swarm intelligence methods for optimization. This study argues that Particle Swarm Optimization (PSO) is the best performing and most popular Intelligence Algorithm for optimization problems. The purpose of this research is to use the Feed forward method of Artificial Neural Networks (ANNs) to distinguish between patients with and without cardiac disease. We have tested our novel categorization method on the industry-standard datasets.

I. INTRODUCTION
Artificial neutral networks (ANNs) outperform Data Mining on every measure, including performance, computational speed, and complexity level [1]. Networks of basic processing components (called "neurons") that use local data and communicate with one another make up artificial neural networks (ANNs) [2]. Many difficult problems in the real world have been handled with the help of ANN, including the prediction of future trends using a company’s massive data archive. Successful applications of ANN have been made in every branch of engineering, including biology, medicine, health care, manufacturing, marketing, oceanography, and decision science, to name a few[3]. In order to aid in the prediction of cardiovascular illness, this research introduces a swarm intelligence approach called Particle swarm optimization (PSO) combined with a feed forward neural network.

As Bonabeau put it, "The emergent collective intelligence of groups of simple agents" [4] is what we mean when we talk about "Swarm Intelligence." A swarm consists of a huge number of similar, basic agents that interact locally with one another and their surroundings, with no overarching leadership or guidance allowing for intriguing global behavior to arise.
In recent years, a new class of population-based algorithms called swarm-based algorithms has arisen [5] that can efficiently solve a wide variety of complicated problems with minimal resources and time investment. Therefore, Swarm Intelligence (SI) may be characterized as a relatively new branch of AI that is used to represent the collective behavior of social swarms found in nature. Examples of such swarms include ant colonies, honey bee hives, and bird flocks. Individually, these agents (insects or swarm members) are not very clever or capable, but they interact in predictable ways to accomplish tasks that are essential to their survival. Individuals in a swarm may have direct or indirect social interactions with one another [6].

The following is the outline for this paper. Work in this area is discussed in Section 2. In Section 4, we explore the theoretical foundations of our methodology and the theory-driven model we developed to identify dissatisfaction. Section 5 provides a conclusion and summary of this study.

**RELATED WORKS**

In response to pupil displeasure, a typical swarm intelligence system will generate and show individualized encouragement messages. Mindspark [15] is a commercially available, massive-scale ITS for teaching mathematics, and we have implemented and assessed our strategy for dealing with student dissatisfaction. With the theory-driven approach [13] included into Mindspark, we can now monitor student displeasure with the ITS in real time. Based on the identified causes of dissatisfaction and data from Mindspark log files, we created inspirational messages to address students' feelings of helplessness. We developed and presented the inspirational messages using attribution theory [16], complimenting students' efforts [17], and demonstrating empathy [4]. Also, we programmed an algorithm to trigger the display of the inspirational words whenever it was determined that the learner was experiencing frustration. After putting our method into practice, we counted how many times each session ended in frustration. Displaying the inspirational quotes significantly decreased the number of reported moments of irritation (p < 0.05).

**II. PROPOSED SYSTEM ARCHITECTURE**

**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 so far.

**Global Best**—The position where the best fitness is achieved among all the particles visited so far.

**Velocity Update**—Velocity is a vector to determine the speed and direction of the particle. Velocity is updated by the equation (1).

**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).

The goal of this research is to use Bio Inspired 4 features optimization algorithms such the Genetic Algorithm, Bat Algorithm, Bee Algorithm, and ACO to identify heart disease in a dataset. Since the Python-based ACO approach I'm using here to solve the Traveling Salesman Problem by finding the shortest route cannot be applied to the heart disease dataset, I'm instead using the Genetic, Bat, and Bee algorithms.

In order to improve the prediction accuracy of classification algorithms, bio-inspired algorithms are designed to optimize the features used in datasets for training these algorithms. Since some datasets may contain irrelevant values within the dataset, optimizing algorithms can remove these features (attribute values) from the dataset. If an attribute is judged to be irrelevant to the dataset after being run through these optimized techniques, it will be eliminated.

To implement this algorithms I am using Heart disease dataset which contains 14 attributes.
and 4 class labels where 0 refers to No heart Disease and 1 refers to stage 1 disease and 2 and 3 refers stage 3 and 4 disease.

Below are some values from dataset to train algorithms

age,sex,cp,trestbps,chol,fbs,restecg,thalach,exang,oldpeak,slope,ca,thal,class

63.0,1.0,1.0,145.0,233.0,1.0,2.0,150.0,0.0,2.3,3.0,0.0,6.0,0
67.0,1.0,4.0,160.0,286.0,0.0,2.0,108.0,1.0,1.5,2.0,3.0,3.0,2
67.0,1.0,4.0,120.0,229.0,0.0,2.0,129.0,1.0,2.6,2.0,2.0,7.0,1
37.0,1.0,3.0,130.0,250.0,0.0,0.0,187.0,0.0,3.5,3.0,0.0,3.0,0

First records contains dataset column names and remaining records are the values of dataset. In last column we have class values as 0, 2, 1 and 3 as disease stage.

Test dataset also contains record values but it will not have class labels and application will apply that test values on train dataset to predict it class labels. Some values from test dataset:

age,sex,cp,trestbps,chol,fbs,restecg,thalach,exang,oldpeak,slope,ca,thal

63.0,1.0,1.0,145.0,233.0,1.0,2.0,150.0,0.0,2.3,3.0,0.0,6.0
67.0,1.0,4.0,160.0,286.0,0.0,2.0,108.0,1.0,1.5,2.0,3.0,3.0
67.0,1.0,4.0,120.0,229.0,0.0,2.0,129.0,1.0,2.6,2.0,2.0,7.0

In above test dataset we can see there is no class name and application will predict it. All this files are available inside ‘heart_dataset’ folder.
In that ‘heart_dataset’ folder I kept dataset URL and information of dataset for references.

III. RESULTS AND DISCUSSION
The results obtained after executing the implementation code is shown from Fig.1 to Fig.9.
To run this project double click on ‘run.bat’ file to get below screen
In above screen click on 'Upload Heart Disease' button and upload heart disease dataset. See below screen.

In above screen uploading dataset file, after uploading will get below screen.
Now click on ‘Run Genetic Algorithm’ button to run genetic algorithm on dataset and to get its accuracy details. While running this algorithm you can see black console to see feature selection process, while running it will open empty windows, you just close all those empty windows except current window.

In above screen for GA accuracy, precision and recall we got 100% result. Now click on ‘Run Bat’ algorithm button to get its accuracy.
In above screen for BAT we got 45% accuracy, now click on ‘Run BEE Algorithm’ button to get BEE accuracy.

In above screen for BEE we got 38% accuracy, now click on ‘Upload & Predict Test Data’ button to upload test data and to predict it class.
In above screen I am uploading test file which contains test data without class label, after uploading test data will get below screen.

In above screen application has predicted disease stages. Now click on ‘Accuracy Graph’ button to view accuracy of all algorithms in graph format.
IV. FUTURE SCOPE AND 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.

REFERENCES


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


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