

Improved Bio-Inspired Algorithm Design for Prediction of Heart Diseases

Gumasa Vijay Kumar

Asst.Professor, Department of CSE, GCET, Hyderabad, vijay.gumasa@gmail.com

Thota Rakesh Kumar

*Asst.Professor, Department of CSE, GCET, Hyderabad,
thotarakeshkumar2012@gmail.com*

Pathlavath Rahul Das

Asst.Professor, Department of CSE, GCET, Hyderabad, rahuldaspathlavath@gmail.com

Abstract

Optimization methods are used to address dynamic, challenging, and robust problems to predict heart problems, the vast majority of machine learning algorithms are used. Machine learning for prediction frequently uses classification algorithms as one of its methodologies. Certain categorization techniques forecast accuracy within a reasonable range, while others might not. In this research, Bat and Genetic algorithms are utilized to streamline for various bioinspired algorithms that are used to forecast cardiac disease. Here, we use these bio-inspired algorithms to extract the important elements from the attribute for heart disease. The various classifiers are then updated with these retrieved features. In this study, we look at various bio-inspired algorithm strengthened using classifiers from Random Forest and SVM and compare the results using different metrics, precision, recall, accuracy, etc. Genetic Algorithm is best optimized among these Bio-Inspired Algorithms.

Keywords: *Heart condition forecast, genetic algorithm, bat colony optimization, and bio-inspired algorithms.*

I. Introduction

Most people throughout the world suffer from cardiac issues, which frequently result in death. Mortality statistics demonstrate that cardiac the greatest cause of death worldwide is sickness. Four out of every five fatalities are caused by according to cardiac attacks and stroke the WHO from cardiovascular diseases (CVD) [1]. Treatment for cardiac disorders is quite expensive in today's society. The heart is a vital organ that is crucial to our body's overall health. Blood is pumped all across the body by it. Heart is located at the middle partition of the chest. When cardiac disease manifests, It is

highly dangerous to human body and leads to death. So, Predicting the reasons that impact the heart is preferable. Obesity, stress, cholesterol, smoking, high blood pressure, diabetes, etc. all contribute to a number of cardiac issues. Heart disease's primary symptoms are irregular heartbeat, chest pain, limb numbness, anxiety, etc. Each of these symptoms varies from person to person. For improved outcomes, it is crucial to predict cardiac disorders using a smart strategy. In this study, we applied sophisticated optimization methods, such as heart disease prognosis algorithms based on biological inspiration. In order to address difficult problems, bioinspired algorithms are used. by

drawing inspiration from nature. These bio-inspired techniques are highly used in machine learning to solve new and robust problems. These techniques provide novel solutions for optimized and distributed control problems [2]. Providing the best answer grew more difficult due to the large amount of data. These days, bio-inspired algorithms are known to address challenging issues using creative methods. The goal of this study is to improve the dataset's properties for improved prediction. The attribute class labels are obtained through feature extraction utilising bio-inspired algorithms are evaluated for further processing. These class label values are produced by taking into account the remaining attribute values along with the retrieved features from the bio-inspired method. After that, the appropriate classifier is used on this dataset to estimate accuracy. The implementation of natural phenomena of the corresponding bio-inspired algorithm evaluates these class label values. The heart dataset was successfully applied to all four algorithms (Ant, Bat, Bee, and Genetic), and the results were improved.

II. RELATED WORK

The patient's data in healthcare is expanding daily with new medical data. Collecting relevant data from multiple patient data scenarios is crucial for disease prediction. This information is used by many methods and technologies to anticipate diseases. This paper's major goal is to examine patient data, extract key characteristics, and use bio-inspired algorithms to heart disease prediction. A strategy for predicting myocardial infarction using variations in ECG signals was reported by Padmavathi Kora [3]. The improved bat algorithm, a proposed approach, is used to predict myocardial infarction. But only 13 patient records were needed to build the better bat algorithm. The remaining 6 records are those of typical persons, leaving a total of 13, of which 7 are those of myocardial infarction

sufferers. Both the standard bat algorithm and the upgraded bat algorithm employ four techniques: Liebenberg-Marquardt neural network, Support Vector Machine, K-Nearest Neighbors, and Conjugate Gradient Neural Network. They came to the conclusion that the enhanced bat algorithm produces superior outcomes to the baseline bat algorithm.

By utilizing data mining techniques in conjunction with the Ant optimization algorithm, Animesh Dubey [4] was able to identify heart problems in their early stages (DMACO). They detected the risk level and took into account the ant's pheromone value. Every time the value of pheromones rises, the risk does as well. The detection rate of heart illness is then generated using The DMACO boosts pheromone intensity, according to research using approaches for data mining and the Ant algorithm. The publications already discussed make it is evident that all methods used to predict heart disease are hybrid classification techniques. However, the research in this study differs from the previously mentioned works in that two distinct optimization strategies were used in conjunction with two distinct classifiers. In this paper, we examine the effectiveness of integrating classifiers with optimization strategies.

III. PROPOSED SYSTEM

To Crack all the above discussed problems, we must associate both the ordered and unstructured collection of data. It gives the best results to measure the disease prediction. To complete the missing data, we use latent factor method in this research. Then we predict the key diseases in a specific region to form the structured data. From the structured data, by using the respective bio algorithm we extract the useful information. After extracting the features, we update the dataset and then apply the classifier on the updated dataset. In this research we are using bat and genetic algorithms as the feature extracting methods

and SVM and Random Forest are the classifiers. Finally, we get a conclusion that which extracted method is the best and which classifier is the best. All these inferences are drawn once we implement these algorithms on the data set only.

Bat Algorithm

It is a meta heuristic algorithm and highly used for solving optimization problems. It was created based on how echolocation is used by micro bats to communicate [16]. The majority of the time, bats use echo-based location to find food and move about. Bats having exceptionally modern feeling of hearing, they emanate sound that skip off the item in their way, sending reverberation back to the bats. This feature is called echolocation or Bio sonar. By using these echo variations, bats can find out the size of the object and how far it is, how fast they are travelling and even their texture. Change their velocity, frequency, and sound in accordance when the bats are looking for food. This dynamic behavior of bats is controlled by using frequency-tuning technique and parameters of bat algorithm are tuned to regulator the balance between exploration and exploitation. In order to find prey, bats fly at random with a velocity of v_i and a position of x_i , employing different wavelengths, different loudness levels and frequency ranges of $f \in [f_{min}, f_{max}]$. The wavelength is automatically adjusted for target proximity, as well as a range of $[0, 1]$ for the pulse emission rate. Loudness value is in the range of $[A_0, A_{min}]$. The frequency-tuning algorithm in this case assigns the bat a frequency at random between $[f_{min}, f_{max}]$.

Genetic Algorithm

John Holland presented Genetic Algorithm in 1975 to solve optimization problems. It is a Metaheuristic algorithm and comes under the class of evolutionary algorithms. It is a nature selection inspired algorithm and highly used to solve search and optimization complications.

Genetic algorithm provides accurate solutions by using three different operators namely crossover, selection and mutation, these three operators are biologically inspired. A population of solutions is created via a genetic algorithm [18]. A chromosome is a solution in that population. Group of genes formed as a chromosome and each gene in a chromosome is represented as 0's and 1's of a string. Like in natural genetics these solutions are taken again and again to produce new children and these are assigned to fitness value [19]. In this it generates multiple generations until the stop criterion is met.

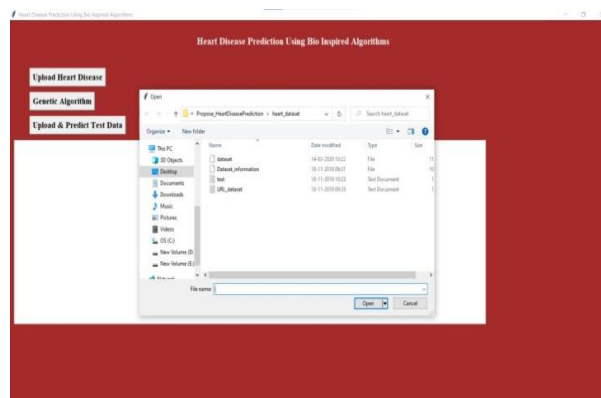
IV. Screenshots

In this project, we want to detect heart disease from dataset using Bio Inspired 4 features optimizing Algorithms such as Genetic Algorithm, Ant Algorithm, Bee Algorithm and Bat Algorithm.

In above test dataset we can see there is no class name and application will predict it. All these files are available inside 'heart dataset' folder.

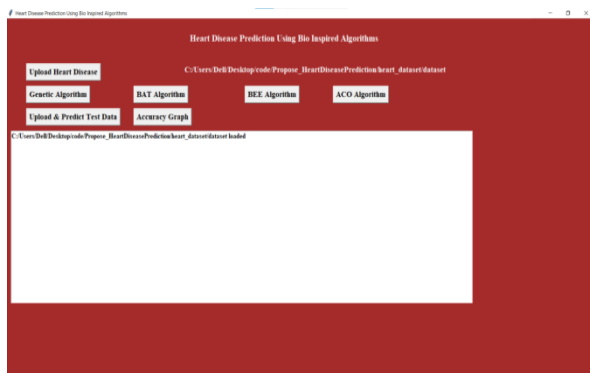
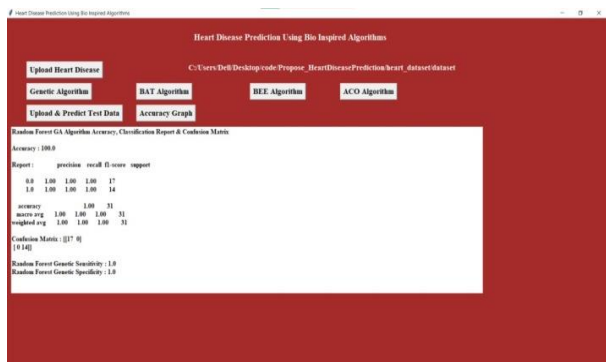
In that 'heart dataset' folder I kept dataset URL and information of dataset for references.

RF Classifier Screens



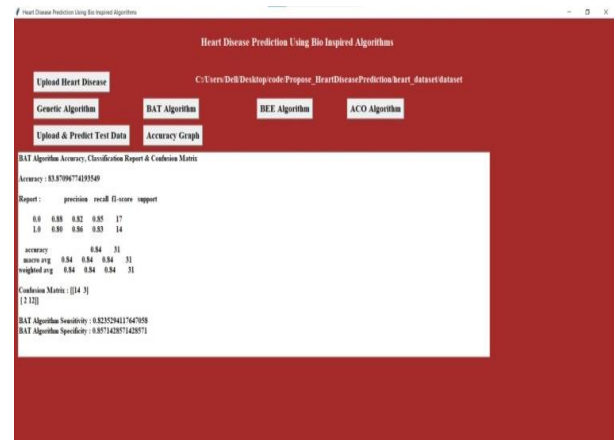
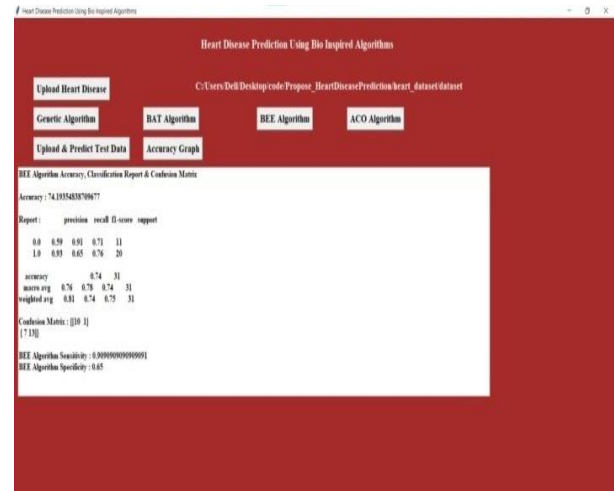


In above upload the heart illness dataset by selecting the "Upload Heart Disease" button on the screen. Then we will get the below screen.



Click the "Genetic Algorithm" button to run the genetic algorithm on the dataset and check its accuracy parameters. You can see a dark console while this algorithm is running to see the feature selection process. Empty windows will open as it runs; you should close all but the current one. From the screen for Random Forest Genetic Algorithm accuracy, precision and recall we got 100% result. To get

the accuracy of bat algorithm just click on it for accuracy.



From the screen for Random Forest Bat Algorithm accuracy, we got 83.87% result. To get the accuracy of bee algorithm just click on it for accuracy.

References

- [1]Youness Khourdifi, Mohamed Bahaj.: Heart Disease Prediction and Classification Using Machine Learning Algorithms Optimized by Particle Swarm Optimization and Ant Colony Optimization, International Journal of Intelligent Engineering and Systems (2019).
- [2]Darwish A.: Bio-inspired computing: Algorithms review, deep analysis, and the

scope of applications. Future Computing and Informatics Journal (2018).

- [3]Padmavati Kora, Kalva Sri Ramakrishna: Improved Bat algorithm for the detection of myocardial infarction in Springer Plus 4(1):666(2015).
- [4]Animesh Dubey, 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).
- [5] Dataset, UCI Machinelearning Repository[online];<https://archive.ics.uci.edu/ml/datasets/Heart+Disease>
- [6] Prakash Rao et al.: An Efficient Approach for Detection of Heart Attack Using Noble Ant Colony Optimization Concept of Data, IJESRT(2018)
- [7] Marco Dorigo.: Ant colony optimization.Scholarpedia, 2(3):1461(2007).
- [8] Ibraheem Nasiruddin, Abdul Quaiyum Ansari, Sapna Katiyar.: Ant Colony Optimization: A Tutorial Review, Conference Paper (2015).
- [9] Hinduja. R, Mettildha Mary. I, Ilakkiya .M, Kavya. S.: CAD Diagnosis Using PSO, BAT, MLR and SVM. International Journal of Advance Research, Ideas and Innovations in Technology (2017).
- [10] sotiris kotsiantis, P.E. Pintelas, I.D. Zaharakis.: Machine learning: A review of classification and combining techniques, Artificial Intelligence Review 26(3):159-190(2006).