BioGecko

Vol 12 Issue 03 2023

ISSN NO: 2230-5807

Heart Disease Prediction Using Bioinspired Algorithms

Mrs. N. Sujata Kumari[1], K. Supraja[2], A. Mamatha Sri[3], N. Shivani[4]"

[1]Assistant Professor, Computer Science & Engineering, Sridevi Women's Engineering College, Hyderabad, India
[2]Computer Science & Engineering, Sridevi Women's Engineering College, B.Tech IV Year, Hyderabad, India
[3]Computer Science & Engineering, Sridevi Women's Engineering College, B.Tech IV Year, Hyderabad, India
[4]Computer Science & Engineering, Sridevi Women's Engineering College, B.Tech IV Year, Hyderabad, India

Abstract
The important organ in the human body is a human heart. The human body is used to do many important functions like body functioning, removal of waste product from our body. It there is any heart by pumping the blood throughout the body. It there is any heart disease or failure occurs it will be very risky for humans to live. Hence to avoid these risks, prediction of heart disease at beginning stage is necessary. To predict heart disease, many techniques are used but no accurate results are provided. Hence, to get accurate results in the prediction of heart disease, bio-inspired algorithms are used. There is different bio-inspired algorithm used to predict heart disease. Genetic algorithm BAT algorithm, BEE algorithm is used to get accurate results in the prediction of heart disease. Genetic algorithm gives correct results whereas BAT and BEE algorithm gives us accurate results.

Keywords: Heart disease prediction, Bio inspired algorithm, Dataset.

1. INTRODUCTION

The diagnosis of heart disease is a complex process, often requiring careful consideration of various factors. Such factors include smoking habits, family, history, cholesterol levels, body weight and blood pressure. To assess the risk of heart disease in a individual, these and other variables must be examined. Automated prediction systems are currently being developed to support cardiac diagnosis and precise treatments can be formulated with this information. The healthcare industry faces a major challenge today: providing the best quality services and accurate diagnoses. Heart disease continues to be aleading causeof the death round world, but they can be managed and controlled with proper detection in time. The key is to diagnose heart disease early. So that it can be appropriately. Human hearts have four chambers,connected to each other by valves that open and close to allow blood flow. When these valves don't open or close properly, it can lead to valvular diseases like regurgitation or stenosis. National estimates suggest that one in four women and one in five men will develop some form of heart disease over the course of their lives. To address these issues, machine learning-a subfield of data mining-provides an efficient way for the large-scale well-formatted data handling. ML techniques can be used from diagnosis to detection and even predicting different diseases in the medical field. Genetic Algorithm is a two-step process: first, generate class-based association rules from a training dataset and second, classify a test dataset using the predefined labels. Compared to traditional classification methods, this approach canbe more efficient and accurate. To tackle hard problems that cannot easily be solved by regular techniques, Genetic Algorithms (GAs) can often be used. These algorithms are especially beneficial for optimization tasks and when navigating complex search spaces. In India, Coronary Heart Disease (CHD) is epidemic and an alarming cause of disease, burden and death. Data from the global health registrar shows CHD as the major issue in many parts of India; for instance, one study on urban Chennai revealed 40% mortality due to CVD causes while another study on rural areas in A.P showed 30%. Given its prevalence, this paper proposes utilizing a

GA-based associative classification for the prediction of heart diseases.

2. LITERATURESURVEY

Jyothi et al, focuses on using different algorithms and combinations of attributes for the prediction of the heart attack using data mining. The classification ofdata mining, techniques like the Decision tree and Bayesian classification has similar accuracy[1]. [2] D. W. Bates et al, team works to develop the health care system which can be able to adopt electronic health records, which will increase the quality of clinical data. Different techniques are used for analyzing large quantities of data, and it ultimately leads to reduce the costs of health care, whereas in the same year Masethe et al, in this project the team aims implemented an algorithm which can be used in both natural andhybrid ways to enhance the prediction accuracy of heart disease.

[3] LokanathSarangi et al, to present a decision-making support system for heart disease diagnosis, the team proposed an approach based on MLP Back Propagation Neural Network. By applying two neural network algorithms on dataset where the results are shown based on the combination of MLPNN and RBFN architecture, whereas in the same year BoshraBahrami et al, team collect dataset of patients and then the data is pre-processed and applies four different classifications like J48 Decision Tree Classifier, KNN Classifier, SMO Classifier and Naive Bayes Classifier for the candidate's data set. After cross validation, optimum heart disease prediction is done.

M. Chen et al, team worked on the Emotion Communication System by HMI System using many methods of emotion recognition. Heart attacks are the leading cause of mortality worldwide, including in South Africa, and their possible prevention depends on early identification. Biological prediction is the method of making a choice based on knowledge from the past. It is employed to study historical events and make early illness predictions. Medical professionals provide data that contain a wealth of hidden information, but it is not correctly or effectively used to make predictions. The analysis converts the wasted data into a set of values for modelling using different data mining techniques as a result. Data mining is the process of extracting relevant information from a large number of databases.Data mining techniques may be used to predict many illnesses in the medical profession.[4].

[5] J. Wang et al, the team designedmanaging of data framework for the health system using BSN, Cloud managing system and remote hospital end. The result shows the most appropriate bandwidth estimating algorithm for underling the health application.

Bo Jin et al, Data Mining has numerous applications in a number of industries such as business, health care, scientific and engineering. In this paper specific emphasis is provided by using data mining, along with Genetic Algorithm, to accurately diagnose heart disease based on a subset of attributes. To achieve this objective 14 features are narrowed down to 6 using GA, and then3 classification techniques, namely Navie bays, decision tree and clustering were used for predicting the diagnosis. The results demonstrated that the decision tree method was the most efficient technique when incorporating features election and had high model construction time. Naviebayes out performed overall consistently after the attribute's reduction with similar model construction time before and after. Additionally, classification by clustering performed poorly compared to other methods. Further research can be done applying fuzzy learning models to evaluate the intensity of cardio vascular diseases. We begin with a brief history of big data as well as a look at some of the technologies that surround it, including cloud storage, the IOT, and Hadoop. We then zero in on the four stages of the big data value chain: creation, the collection, storage, and analysis. We provide an overview of the context, outline the technological difficulties, and summarize the most recent developments at each stage. Enterprise management, the Internet of the whole world, online social activities, biological applications, collective knowledge, and the smart grid are just few of the examples of real-world big data use cases that we explore in the end. The goal of these analyses is to provide readers a bird's-eye

perspective of this fascinating field. Open issues and potential next steps are discussed to round out this overview [6].

[7] Senthil Kumar Mohan et al,psychologicaldata is monitored in real time by pervasive sensors to provide remote telemedicine in the health system that has both medical and nonmedical applications, including store-and-forward data services for later offline study by appropriate professionals. Unfortunately, the speed and accuracy of the large system are not taken into account by the present the health systems, whereas with the team [8] Real-time data from emergencies should be maintained centrally and sent to far-flung medical facilities. In addition, the dispersed datacenter cannot provide a highly efficient distant real-time service since the patient data is split across several servers. In order to assist a cloud, broker allocates compute nodes and networks efficiently, we present a possibility bandwidth paradigm in a healthcare cloud system. This brokering approach takes into account the PHR cloud location protocol, and plans the real-time signaling with minimal data exchange between hosts. The broker employs different bandwidth evaluation strategies to foretell the foreseeable demand for broadband in the health setting. The results of the simulations, together with the verified effectiveness of the inserted service, demonstrate the efficacy of our concept.

[9] JaishriWankhede et al, A significant difficulty in the study of clinical data analysis is the estimation of cardiovascular disease (CVD). In this research, an effective feature-based prediction of heart disease is created. Data cleansing, the transformation of data, missing value attribution, and data normalization are the initial steps in the pre-processing of the data. Then, in the feature identification phase, the decision-making function-based stochastic salp swarm (DFCSS) method is employed to choose the best features. The upgraded Elman artificial neural network (IENN) is then given the selected attributes to classify the input. Here, the most suitable weight value of the IENN is determined using the sailfish optimizing (SFO) algorithm. Heart disease is a major problem in the world. To get rid of this disease the proposed system used here is the genetic algorithm. Before many algorithms and techniques are used but the stages of heart disease is not predicted correctly and the accurate results are not displayed. Hence bio inspired algorithms are very useful in predicting the stages of heart disease of the person. The common risk found in people affected by heart disease participants compared to people who don’t have heart disease expresses the intensity of the relationship with the disease, and the progressive association (concurrent with the risk factor) reflects the relevance of the risk factor.

Gowri Shankar Nayak et al, The use of analysis of data in medicine is increasing as it helps to better understand diagnoses, improve research techniques, and arrange the necessary equipment supply based on the significance of emerging illnesses. In order to analyse the current data and make the best predictions for the future, artificial intelligence provides software solutions. For the categorization of cardiac disease, a computational model is susceptible of using many data processing techniques. A particular type of data is a specific focus of this study. From training and test data, the categorization enables us to create a prediction model. In order to create a better model with detailed data that contains the identical data, these data is filtered by an algorithm for classification which uses a mix of mathematical methods and computer procedures. We must employ the optimization approach to analyze the available data and forecast the best outcomes. When a dataset contains a wide variety of characteristics, few of these are not useful and leads to less accurate results. As a result, the primary goal of this investigation is to get together all the methods to enhance feature selection and classification in order to get better diagnose heart disease. In this paper, the selecting behaviourof critical characteristics in heart disease is optimised using a multinational aggressive method with a great algorithm approach. With regard to feature selection for genetic and other optimisation techniques, this approach can offer a more ideal solution. The dataset was split into two sets, a training set and a testing set, after data pre-processing. On the contrary hand, this danger is reduced when this element is removed or improved. The finding shows that the uniquealgorithm that has been

suggested can deliver a reliable health tracker for the early identification of heart disease [10].

3. METHODOLOGY

We have utilized python operations to do heart disease classification using data obtained from the UCI repository after reviewing the outcomes from the current approaches. It gives you a graphical interface for exploring data, creating models, and doing predictive analytics. The first step in using a genetic algorithm is cleaning the raw data andfollowed by feature selection on assessing the efficacy of data cleaning, categorization, and modeling. The results precision is boosted further by using the BAT and BEE algorithms. To reduce the number of dimensions in a set of data, one chooses a mathematical representation that allows for the correlation of most, but not all, of the variation included in the data. Data for a given job or issue may have several properties or dimensions, not all of which will have the same bearing on the final result. Over fitting, which produces subpar outcomes, is areal possibility when working with a high number of qualities or features. Thus, dimensionality reduction should be given serious thought before proceeding with the construction of any model. Extraction of features and feature selection are the two most common approaches to dimensionality reduction. Natural selection, the mechanism that drives biological evolution, is the inspiration for the genetic algorithm, a technique for tackling limited and unconstrained optimization problems. The genetic algorithm iteratively improves upon a pool of candidate answers. The emetic algorithm performs each step in a choice parents from the existing population and puts them to work reproducing the future generation. Throughout time, thepopulace "evolves" to a better and better state of affairs. The genetic algorithm may be used to address optimization issues that are not am enable to more traditional optimization techniques, such as those were the functional form is discontinuous, non-differentiable, stochastic, or extremely nonlinear. Theevolutionary algorithm may be used to solve mixed integer programming problems, in which some variables must take on integer values. Formalized Crossing Operators for Bio-Inspired Genetic Algorithms in Robotics. When it comes to optimizing robotic applications, genetic algorithms are commonly used. For such life-or-death systems, explicitly proving the validity of genetic algorithms is crucial.

System Architecture

It covers four System Components:

- Dataset and Data Analysis
- Model Selection and Implementation,
- Data Comparison, and
- Data Visualization.

The Dataset and Data Analysis module involves obtaining the dataset and then performing analysis to prepare the data formodel selection and implementation. The Model Selection and Implementation module selects the most appropriate model for the dataset, implements it, and evaluates the performance. The Data Comparison module enables comparison of different models to select the most suitable one.

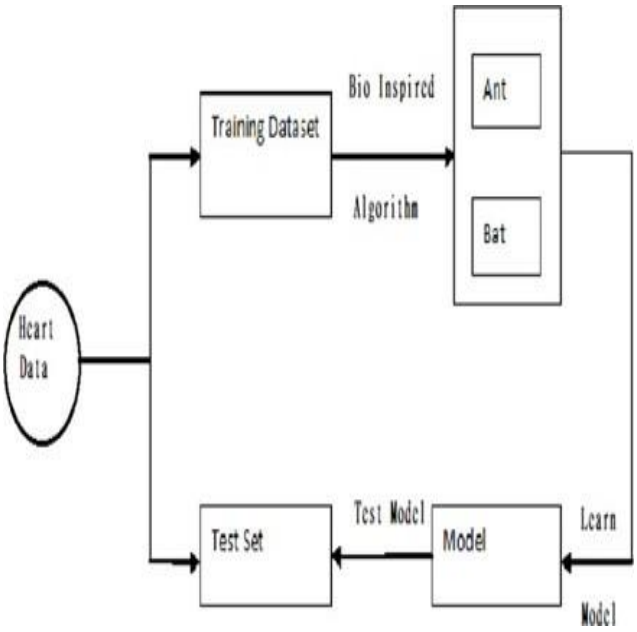

Fig 1: System Architecture

Finally, the Data Visualization module helps visualize the data for better insights. Dataanalysisisanimportantpart of extractingmeaningfulinput from data,and it is used to help makebetter business decisions. In order to do this, firstly admin has to login and upload the dataset for training for a machine learning approach. From there, model selection is performed and a final machine learning a model is selected and implemented on the dataset. The data isthen tested with various parameters andcompared with the user's test data, beforefinally being visualized in graphical form to help understand patterns.

Algorithm

John Holland initially proposed the idea of genetic algorithms in 1975, and they are now commonly utilized toresolveessentiallysearchingand optimization issues. The basic concepts of genetics are applied using genetic algorithms to solve problems. A genetic algorithm is a heuristic character makes it popular for solving challenges involving illness prediction inthe rapidly developing field of artificialintelligence.Theprocessusesmethodsincluding passing on, selection, mutation,anda cross over that are inspired by evolutionary processes to provideoptimized solutions. A chromosome is thename given to each outcome produced bygeneticalgorithms. A population is the name given to the set of chromosomes.

The following are the duties of geneticoperators:

- Selection: More fit chromosomal are chosen to survive, therefore selection essentially deals with the likelihood that the fittest will survive.
- Crossover: For this procedure, a randomgene is chosen along the chromosome'slength, and all the genes that follow areswitched.Onepoint,twopoints,ormultiple point crossover operators are allpossible.
- Mutation: Mutation modifies novel solutions and seeks out improved ones.

An application of the genetic algorithm:

Step1: Representing the solution variables is the first step.
Step 2: Probabilistic Initialization of the Solution Variable Set of potential values.
Step 3: Determine the goal function for each value of the solution variables.
Step 4: Show the best objective function value.
Step5: Using four primary processes, generate a fresh set of potential values forthe solutionvariables.
Step6:Repeatsteps3,4,and5 repeatedly until thereis no more room for enhancement of the best measurement ofthe objectivefunction.

Algorithm:
Create randomly Population P of Popsize individuals
While(stop_condition)FALSE
P*=empty
While (Size of(P*) is not equal to Popsize)
Select p1 and p2 from P
Crossover p1, p2 to obtain h1 and h2 Mutate h1 and h2 to obtain c1 and c2
If[Dis(p1,c1)+Dis(p2,c2)]<[Dis(p1,c2)+Dis (p2,c1)]
Ifc1<p1then P*= P*union{c1}elseP*
=P*union{p1}
Ifc2<p2 then P*=P*union{c2}elseP*
=P*union{p2}
Else
Ifc1<p2 then P*= P*union{c1}elseP*
=P*union{p2}
Ifc2<p1 then P*= P*union{c2}elseP*
=P*union{p1}
EndWhile

4. RESULT

The goal of this paper is to use Bio inspired3 featured optimization algorithms such as genetic algorithm, bat algorithm and bee algorithm to identify the stage of heart disease in a dataset. In this case, we are implementing the Genetic, Bat, and Bee algorithms since the ACO method, which was designed in Python to solve the Traveling Salesman Problem and discover the shortest route, is incompatible with the heart disease dataset. The graph shows the comparison of 3 algorithms used in this such as genetic algorithm, bat algorithm and bee algorithm. The genetic algorithm has more accuracy compared to bat and bee algorithms. Here bat and bee algorithms are used to prove that genetic algorithm gives more accurate results than bat and bee algorithms. The bar graph is represented using confusion matrix values. These values are predicted using recall and precision values.
Graph shows how well they perform by comparing the accuracy of each algorithm in the form of the bar graph. The graph's x-axis shows the name of the algorithms, while the y-axis shows accuracy levels.

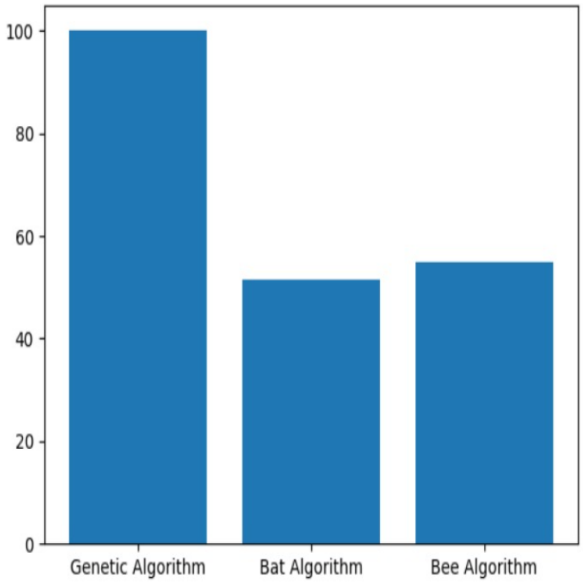

Fig 2: Accuracy graph

In the above figure, the accuracy graph is displayed. The accuracy results of genetic algorithm is high when compared to bat algorithm and bee algorithm. Hence the genetic algorithm is the best algorithm to use than bat and bee algorithm to predict the stage of heart disease. The graph is displayed using confusion matrix, recall and precision values. Hence the genetic algorithm gives accurate results compared to bat and bee algorithm.

5. CONCLUSION

Three bio-inspired algorithm modeling methodologies were used to create a method for predicting heart disease. This work extracts the individual's medical information that leads to deadly cardiovascular disease from the dataset that comprises patient medical history, including chest x-rays, hospitalizations, and medications,symptoms, blood sugar, blood pressure, cholesterol levels, maximal heart rate and so forth. If a patient has already been diagnosed with heart disease, this heart disease detection method may help. In this construction of the specified model, the Genetic, BAT, and BEE algorithms are used. Our model has perfect precision. More training data increases the likelihood that the model will correctly determine whether or not a specific individual has heart disease. In addition, we find that the BAT algorithm achieves 51% accuracy, whereas the BEE algorithm achieves just 53%. The ultimate goal of the genetic algorithm is perfect accuracy. This leads us to believe that the Genetic Algorithm is superior to the BAT and BEE methods in terms of accuracy.

6. REFERENCES

1. Soni, Jyoti, Ujma Ansari, Dipesh Sharma, and SunitaSoni. "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.
2. D. W. Bates, S. Saria, L. Ohno-Machado, A. Shah, and G. Escobar, ‘‘Big data in health care: Using analytics to identify and manage high-risk and high-cost patients,’’ Health Affairs, vol. 33, no. 7, pp. 1123–1131, 2014.
3. LokanathSarangi, Mihir Narayan Mohanty, SrikantaPattnaik, “An Intelligent Decision Support System for Cardiac Disease Detection”, IJCTA, International Press 2015.
4. M. Chen, P. Zhou, and G. Fortino, ‘‘Emotion communication system,’’ IEEE Access, vol. 5, pp. 326337,2017,doi:10.1109/ACCESS.2016.2641480
5. J. Wang, M. Qiu, and B. Guo, ‘‘Enabling real-time information service on telehealthsystem

over cloud-basedbigdata platform,''J.Syst.Archit.vol.72,pp.69–79, Jan.2017.

6. Bo Jin ,Chao Che, Zhen Liu, Shulong Zhang, Xiaomeng Yin, And Xiaopeng Wei, "Predicting the Risk of Heart Failure With EHR Sequential Data Modeling" ,IEEE Access 2018.
7. Senthil Kumar Mohan, ChandrasegarThirumalai, and Gautam Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", IEEE Access 2019.
8. ApurbRajdhan, Milan Sai, Avi Agarwal, Dr. PoonamGhuli, "Heart Disease Prediction using Machine Learning". Vol. 9 Issue 04, April-2020.
9. JaishriWankhede, "Efficient heart disease prediction-based on optimal feature selection using DFCSS and classification by improved Elman-SFO", 2020.
10. Gowri Shankar Nayak, "Heart disease prediction by using novel optimization algorithm: A Supervised learning perspective", 2021.