INTERNATIONAL JOURNAL OF PROGRESSIVE RESEARCH IN ENGINEERING MANAGEMENT AND SCIENCE (IJPREMS)

www.ijprems.com
editor@ijprems.com

Vol. 03, Issue 04, April 2023, pp : 572-577

e-ISSN : 2583-1062

Impact Factor : 5.725

HEART DISEASE PREDICTION

K. Karthik[1], U. Chandra Sekhar[2], L. S. N. V. Lakshmi Kanth[3], U. V. Ramesh[4],

[1,2,3]CSE, B.Tech, Aditya Engineering College, Surampalem,

[4]Guide, M. Tech, Aditya Engineering College, Surampalem, A.P, India

ABSTRACT

Health disorders are on the rise due to inherited and lifestyle factors. In this regard, during the past few years, heart disease has been the main factor in the extinction of the human race. This project's goal is to use artificial neural networks and the swarm intelligence algorithm to forecast heart disease. An interdisciplinary study area that is still developing is swarm intelligence (SI). Swarm-based algorithms are a new class of population-based, nature-inspired algorithms that can provide quick, reliable, and affordable solutions to a variety of challenging issues. The three most popular intelligence algorithms are proposed in this project, which also performs well in terms of optimisation. Through the use of genetic, BEE, and BAT algorithms, this initiative seeks to forecast the development of heart disease by dividing patients into unhealthy and healthy groups. We used the well-known data sets to assess our novel classification methodology.

Keywords: BAT Algorithm, BEE algorithm, Genetic Algorithm.

1. INTRODUCTION

Most of the time, a complicated mix of clinical and pathological data is required for the diagnosis of cardiac disease; this complexity results in high medical expenses, which have an adverse effect on the standard of care. According to WHO statistics, heart disease was the top cause of death in developing nations by 2010, killing one-third of the world's population. According to a research by the American Heart Association, one-third of adult Americans have one or more types of heart disease. In order to translate biological information into clinical practise and to explain biological phenomena from clinical data, computational biology is frequently used. One of the major achievements of computational biology is the identification of biomarkers for heart disease. For diagnostic purposes, this procedure entails the creation of a predictive model and the integration of various data and knowledge types. Additionally, this procedure calls for the design and fusion of several statistical analysis and data mining approaches. Around the world, heart illnesses have become one of the leading causes of death. It may contribute to major cardiovascular events including heart attack and stroke. The body will shut down and a person will pass away very quickly if the heart ever stops working and stops pumping blood. Since heart disease affects a person in a way that makes it difficult for patients to get better as quickly as possible, risk assessment is essential to identifying possibilities for prevention. Due to the wide range of odd symptoms and indicators, it is a very heterogeneous and complex condition that is challenging to discover and diagnose at the appropriate time. Numerous lives are lost as a result of hospital misdiagnosis and misinterpretation. Unfortunately, a variety of circumstances might affect and make it more difficult to detect heart irregularities, which can lead to a delayed or incorrect diagnosis. Even for specialists, who commonly consider accurate techniques to locate all the risk variables and deliver a clear conclusion in a certain time period, diagnosing heart disease can occasionally be challenging due to many unknown risk factors. People's lives are lost due to doctors' misdiagnosis and lack of understanding..

2. LITERATURE REVIEW

[1] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). Effective heart disease prediction using hybrid machine learning techniques. IEEE access, 7, 81542-81554.

The numerous research methodologies taken into account in this work for the categorization and prediction of heart disease using ML and DL are extremely accurate in determining the effectiveness of these methods.According to the findings of this article, ANN has been found to provide the highest accuracy when compared to earlier approaches. Heart illness is predicted using the back propagation multilayer perceptron (MLP) of ANN. NN, DT, SVM, and Naive Bayes are utilised to find patterns in the data of heart disease patients collected from the UCI laboratory. The employment of this strategy was ineffective. With ease and effectiveness, the innovative methods provided here lower costs and enhance heart disease prediction. The employment of this strategy was ineffective. The fresh ideas given here In the training phase, this approach takes into account the heart cycles with varying start locations from the electrocardiogram (ECG)signals. In the patient's testing phase, CNN can provide features with a range of positions. Performance and accuracy are compared between the outcomes using these algorithms. 86.8% accuracy for F-measure is achieved by the suggested hybrid technique, competing with the other existing methods.

[2] Guo, C., Zhang, J., Liu, Y., Xie, Y., Han, Z., & Yu, J. (2020). Recursion enhanced random forest with an improved linear model (RERF-ILM) for heart disease detection on the internet of medical things platform. IEEE Access, 8, 59247-59256.

Heart disease is a group of illnesses that affect people's hearts and veins. Depending on the precise form of cardiac disease, there are various symptoms. Numerous things affect this, including ageing, diabetes, smoking, being overweight, eating junk food, and so forth. Numerous factors that either cause or aggravate heart disease have been found. In order to forecast heart illnesses, the Enhance Random Forest with an improved linear model (RERF-ILM) approach is suggested in this research. In the field of medicine, ANN has been used to make predictions with the highest accuracy. Furthermore, the three crucial IoMT components of accuracy, dependability, and repeatability must always take precedence. In the field of medicine, ANN has been used to make predictions with the highest accuracy. Additionally, accuracy, dependability, and repeatability are three crucial factors that must always take precedence. The author of this study has considered five different approaches to analysing performances and accuracy.

3. PROPOSED SYSTEM

In this study, we use Bio Inspired 4 features optimising algorithms, such as Genetic Algorithm, Bat, Bee, and ACO, to identify heart disease from the dataset. Since I can't use the heart disease dataset to create the ACO method, which is designed in Python to discover the shortest path, I'm using the Genetic, Bat, and Bee algorithms instead.

Hardware Requirements:

- System: Intel i5 Processor.
- Hard Disk: 500 GB.
- Ram: 8 GB

Software Requirements:

- Operating System: Windows 10
- Languages: Python.
- Libraries: Sklearn, Pandas, NumPy
- Software: Jupyter Notebook
- Dataset

4. ALGORITHMS USED

- **Bat Algorithm**:

A BAT algorithm (BA) is a heuristic algorithm that performs global optimisation by mimicking the echolocation behaviour of bats. Due to the BA's outstanding performance, many optimisation issues are solved with this method. BA employs a frequency-tuning technique to broaden the population's diversity of solutions. An initial population of bats with each having a position and a velocity in the search space makes up the basis of the Bat Algorithm. By adjusting their position and velocity in accordance with their own fitness values and the population's best fitness value, bats look for the global optimum.The Bat Algorithm can be used for various optimization problems, such as function optimization, feature selection, clustering, and classification. In machine learning, the Bat Algorithm can be used for model selection and hyperparameter optimization. BAT algorithm takes the dataset as a input and predict the disease from the dataset and accuracy will displayed on the screen.

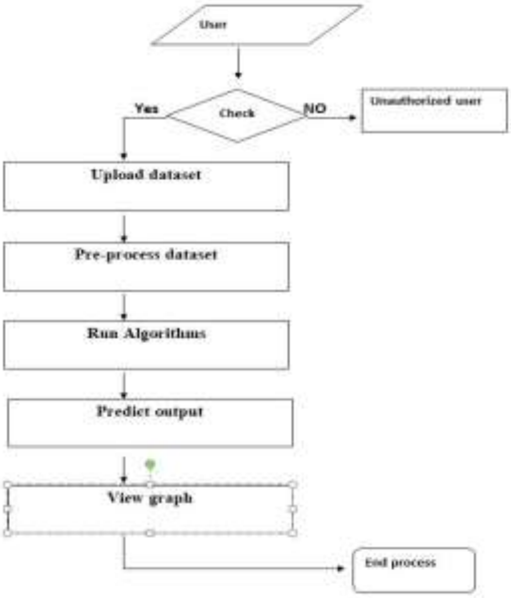
- **Bee Algorithm:**

Swarm intelligence programme called the Bee programme was motivated by honeybees' foraging habits. A variety of optimisation issues can be resolved with this metaheuristic technique. The Bee Algorithm starts with an initial population of bees, each with a position and a fitness value in the search space. The bees search for the global optimum by exploiting the local search space and exploring the global search space. The BEE algorithm is that some measure of distance between the solutions is defined. Due to its effectiveness, the BEE is frequently utilised in a variety of optimisation situations.BEE algorithm takes the dataset as a input and predict the disease from the dataset and accuracy will displayed on the screen.

- **Genetic Algorithm**:

The concepts of natural selection and genetics served as inspiration for the metaheuristic algorithm known as the genetic algorithm (GA). To identify the best answer to a problem, a population-based algorithm models the course of evolution.. Starting with a population of potential solutions, the algorithm begins. Each potential solution is represented by a genotype or chromosome, which is a collection of binary or real-valued values. The quality of a solution is measured by a fitness function, which evaluates how well the solution solves the problem.The GA proceeds through a series of iterations, or generations. The algorithm creates a new population of potential solutions in each generation by using selection, crossover, and mutation operators on the population. The selection operator favours solutions with higher fitness values, while the crossover and mutation operators produce new solutions by combining and modifying existing solutions. The GA has been used in a variety of optimization problems, including function optimization, feature selection, and parameter tuning. It has been shown to be an effective algorithm for finding high-quality solutions to difficult optimization problems, especially when the search space is large and complex.

5. FLOWCHART



6. RESULTS AND SCREENSHOTS

Fig 6.1 Output screen

In above screen click on ‘Upload Heart Disease button and upload heart disease dataset. See below screen

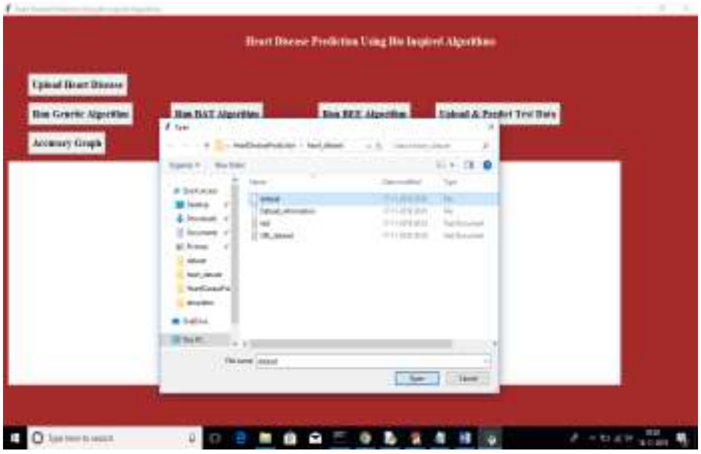


Fig 6.2 Uploading dataset

To execute the genetic algorithm on the dataset and obtain details about its accuracy, click the "Run Genetic Algorithm" button now. While running this algorithm u can see black console to see feature selection process, while running it will open empty windows, u just close all those empty windows except current window

Fig 6.3 genetic algorithm

In above screen for GA accuracy, precision and recall we got 95% result. To determine its accuracy, press the 'Run Bat' button now.

Fig 6.4 BAT algorithm

In above screen for BAT we got 62% accuracy, now, select the 'Run BEE Algorithm' button to get BEE accuracy.

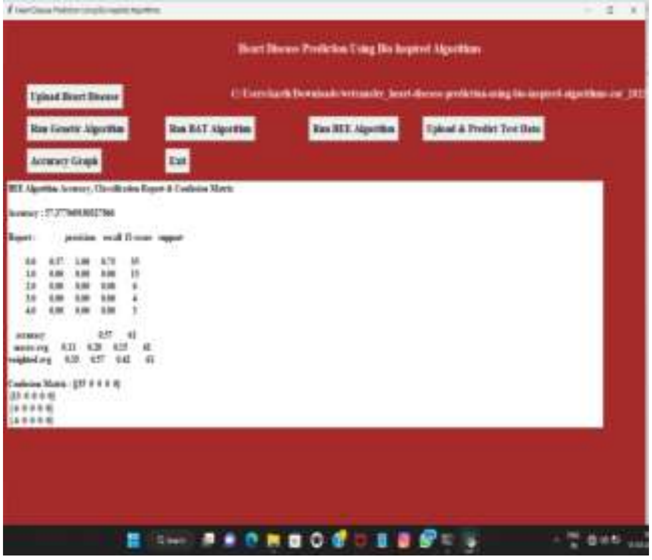


Fig 6.5 BEE algorithm

In above screen for BEE we got 57% accuracy, now click on ‘Upload & Predict Test Data’ button to upload test data and to predict it class

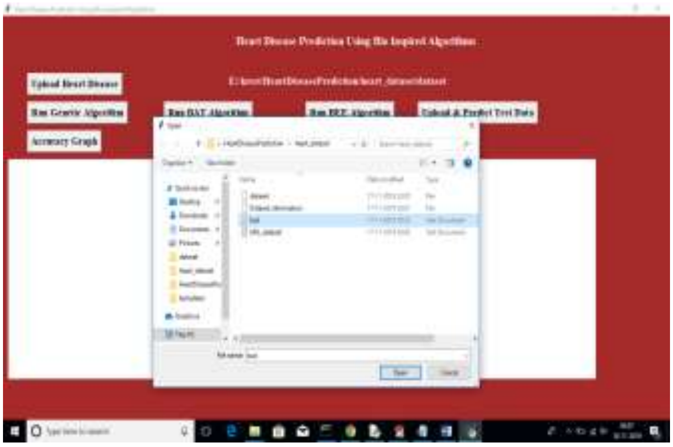


Fig 6.6 uploading test and predicting dataset

In above screen I am uploading test file which contains test data without class label, after uploading test data will get below screen

Fig 6.7 testing and prediction output

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

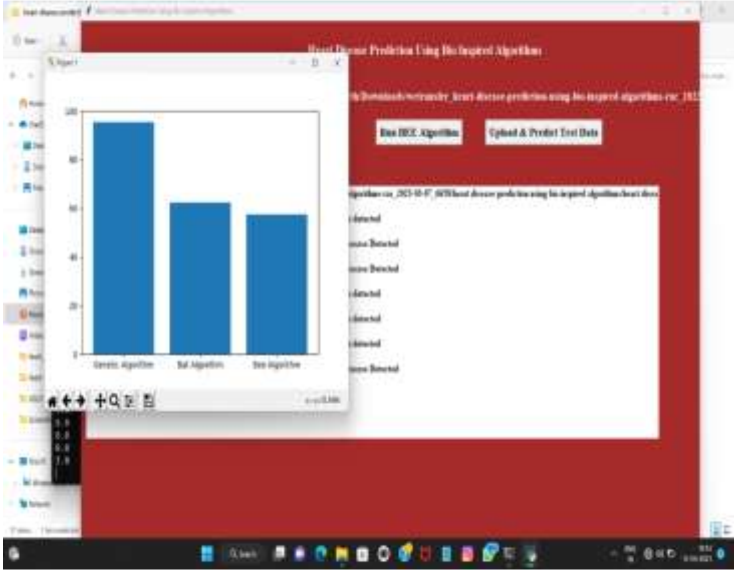


Fig 6.8 Accuracy Graph

7. ADVANTAGES

- High Accuracy.
- High performance
- High feature compatibility

8. CONCLUSION AND FUTURE SCOPE

• Offering illness diagnosis facility services at reasonable prices is a top problem for healthcare organisations like hospitals and medical centres. Heart disease is more prevalent in recent developing nations like India, and it is expected that by 2020, coronary heart disease would be the main cause of death for Indian adults. So, to forecast cardiac disease, a decision support system is needed. For the diagnosis of cardiac disease, a number of data mining, machine learning, artificial intelligence, pattern recognition, soft computing, and optimisation techniques have been proposed.

• The main goal of implementing these algorithms is to identify the attribute that helps in disease diagnosis to make risk management easier. The implementation of four different bio-inspired algorithms used to diagnose cardiac disease is the focus of the current chapter. To improve the heart disease prediction accuracy, many of these algorithms must be used, both naturally and by hybridization..

9. REFERENCES

[1] Wu R, Peters W, Morgan MW. The next generation clinical decision support: linking evidence to best practice. J Healthc Inf Manag, 2002;16:50-5.

[2] Thuraisingham BM. A Primer for Understanding and applying data mining. IT Professional 2000;1:28-31.

[3] Rajkumar A, Reena GS. Diagnosis of heart disease using datamining algorithm. Global Journal of Computer Science and Technology 2010;10:38-43.

[4] Anbarasi M, Anupriya E, Iyengar NCHSN. Enhanced prediction of heart Disease with feature subset selection using genetic algorithm. International Journal of Engineering Science and Technology 2010;2:5370-76.

[5] Palaniappan S, Awang R. Intelligent heart disease prediction system using data mining techniques. International Journal of Computer Science and Network Security 2008;8:343-50.

[6] Baxt WG. Application of artificial neural networks to clinical medicine. Lancet 1995;346:1135-38.

[7] UCI Machine Learning Repository [homepage on the Internet]. Arlington: The Association; 2006 [updated 1996 Dec 3; cited 2011 Feb 2]. Available from: http://archive.ics.uci.edu/ml/datasets/Heart+Disease

[8] Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y. R., & Suraj, R. S. (2021, January). Heart disease prediction using hybrid machine learning model. In 2021 6th international conference on inventive computation technologies (ICICT) (pp. 1329-1333). IEEE.

[9] Lakshmanarao, A., Swathi, Y., & Sundareswar, P. S. S. (2019). Machine learning techniques for heart disease prediction. Forest, 95(99), 97.

[10] Kumar, K. K., Chaduvula, K., & Markapudi, B. R. (2020). A Detailed Survey on feature extraction techniques in image processing for medical image analysis. European Journal of Molecular & Clinical Medicine, 7(10), 2020.

[11] Kumar, K. K. (2021). An efficient image classification of malaria parasite using convolutional neural network and ADAM optimizer. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 12(2), 3376-3384.

[12] Indira, D. N. V. S. L. S., Babu, C. S., Kumar, K. K., & Rao, C. V. (2021). A comprehensive review on sentiment analysis techniques and machine learning libraries in image processing. Annals of the Romanian Society for Cell Biology, 4260-4267.