IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

CHRONIC HEART DISEASE PREDICTION USING ARTIFICIAL INTELLIGENCE

1st
Dr.S.VIJAYAKUMAR,M.Tech.,Ph.d.,
Electronics And Communication Engineering
Paavai Engineering College (Anna university Affiliated)
Namakkal, India

2nd KANNAN G
Electronics And Communication Engineering
Paavai Engineering College (Anna university Affiliated)
Namakkal, India

3rd NIRANJAN S
Electronics And Communication Engineering
Paavai Engineering College (Anna university Affiliated)
Namakkal, India

4th PRASANTHKUMAR P
Electronics And Communication Engineering
Paavai Engineering College (Anna university Affiliated)
Namakkal, India

***Abstract*-**The chronic heart disease finder (CHDF) is an AI-based system designed to assist healthcare professionals in the early detection and diagnosis of chronic heart diseases. By analyzing patient data such as medical history, symptoms, and diagnostic test results, the CHDF aims to accurately identify individuals at risk of or already affected by chronic heart conditions. The system leverages machine learning algorithms to process and interpret vast amounts of medical information, offering quick and reliable insights to facilitate timely intervention and personalized treatment plans. Ultimately, the CHDF strives to improve patient outcomes by enabling healthcare providers to identify, monitor, and manage chronic heart diseases more effectively through the power of artificial intelligence.

Keywords—Heart, AI, K Neighbors, Risk Factors, Algorithms

I. INTRODUCTION

Chronic heart disease, also known as coronary artery disease, is a pervasive and pressing health concern in real time. As a leading cause of morbidity and mortality worldwide, it poses a significant challenge to individuals, healthcare systems, and society as a whole. In real time, the prevalence of chronic heart disease continues to rise, driven by factors such as sedentary lifestyles, unhealthy dietary habits, and an aging population. The impact of chronic heart disease is felt daily, affecting the lives of countless individuals and placing a substantial burden on healthcare resources.

In real time, the complexities of chronic heart disease extend beyond the individual level to encompass broader public health, diagnostic, preventive, and management considerations From timely diagnosis and personalized treatment strategies to long-term disease management and resource allocation, chronic heart disease demands real-time attention and innovative solutions.

Real-time efforts to address chronic heart disease must therefore integrate cutting-edge medical advancements, technological innovations, and proactive public health initiatives to effectively combat this pervasive and evolving health challenge. By understanding the real-time dynamics of chronic heart disease and adapting interventions to the current landscape, we can work towards mitigating its impact and improving outcomes for individuals and communities.

II. LITERATURE REVIEW

With growing development in the field of medical science alongside machine learning various experiments and researches has been carried out in these recent years releasing the relevant significant papers.

[1] Purushottam, et, al proposed a paper “Efficient Heart Disease Prediction System “using hill climbing and decision tree algorithms. They used Cleveland dataset and preprocessing of data is performed before using classification algorithms. The Knowledge Extraction is done based on Evolutionary Learning (KEEL), an opensource data mining tool that fills the missing values in the data set. A decision tree follows top-down order. For each actual node selected by hill-climbing algorithm a node is selected by a test at each level. The parameters and their values used are confidence. Its minimum confidence value is 0.25. The accuracy of the system is about
86.7%.

[2] Santhana Krishnan. J, et, al proposed a paper "Prediction of Heart Disease Using Machine Learning Algorithms" using decision tree and Naive Bayes algorithm for prediction of heart disease. In decision tree algorithm the tree is built using certain conditions which gives True or False decisions. The algorithms like SVM, KNN are results based on vertical or horizontal split conditions depends on dependent variables.
But decision tree for a tree like structure having root node, leaves and branches base on the decision made in each of tree Decision tree also help in the understating the importance of the attributes in the dataset. They have also used Cleveland data set. Dataset splits in 70% training and 30% testing by using some methods. This algorithm gives 91% accuracy. The second algorithm is Naive Bayes, which is used for classification. It can handle complicated, nonlinear, dependent data so it is found suitable for heart disease dataset as this dataset is also complicated, dependent and nonlinear in nature. This algorithm gives an 87% accuracy.

[3] Sonam Nikhar et al proposed paper " Prediction of Heart Disease Using Machine Learning Algorithms" their research gives point to point explanation of Naïve Bayes and decision tree classifier that are used especially in the prediction of Heart Disease.IoT-based solar grass cutting robots are a promising solution for lawn maintenance, offering a range of benefits, including reduced carbon emissions, lower maintenance costs. IoT-based solar grass cutting robot is a sustainable and efficient solution for lawn maintenance, and its potential for further development makes it an exciting prospect for the future of gardening and landscaping. The project aims to use renewable energy sources like solar electricity and to operate a cutter equipped with various accessories and to cut and gather lawn grass. The DC motor, powered by a battery, where the charge is kept via a solar panel, has a spiral shaped grass cutting blade. The solar panel is attached to the structure and charges if the system does not run, transferring charges to the battery through the circuit. During the day and night, the solar grass cutter uses both. Some analysis has been led to think about the execution of prescient data mining strategy on the same dataset, and the result decided that Decision Tree has highest accuracy than Bayesian classifier.

[4] Aditi Gavhane et al proposed a paper "Prediction of Heart Disease Using Machine Learning", in which training and testing of dataset is performed by using neural network algorithm multi-layer perceptron. In this algorithm there will be one input layer and one output layer and one or more layers are hidden layers between these two input and output layers. Through hidden layers each input node is connected to output layer. This connection is assigned with some random weights. The other input is called bias which is assigned with weight based on requirement the connection between the nodes can be feedforwarded or feedback.

[5] Avinash Golande et al, proposed "Heart Disease Prediction Using Effective Machine Learning Techniques" in which few data mining techniques are used that support the doctors to differentiate the heart disease. Usually utilized methodologies are k-nearest neighbour, Decision tree and Naïve Bayes. Other unique characterization-based strategies utilized are packing calculation, Part thickness, consecutive negligible streamlining and neural systems, straight Kernel self-arranging guide and SVM (Bolster Vector Machine).

[6] Lakshmana Rao et al, proposed "Machine Learning Techniques for Heart Disease Prediction" in which the contributing elements for heart disease are more. So, it is difficult to distinguish heart disease. To find the seriousness of the heart disease among people different neural systems and data mining techniques are used.

[7] Abhay Kishore et alproposed "Heart Attack Prediction Using Deep Learning" in which heart attack prediction system by using Deep learning techniques and to predict
the probable aspects of heart related infections of the patient Recurrent Neura System are used. This model uses deep learning and data mining to give the best precise model and least blunders. This paper acts as strong reference model for another type of heart attack prediction models

[8] Senthil Kumar Mohan et al, proposed "Effective Heart Disease Prediction Using
Hybrid Machine Learning Techniques" in which their main objective is to improve exactness in cardiovascular problems. The algorithms used are KNN, LR, SVM, NN
to produce an improved exhibition level with a precision level of 88.7% through the prediction model for heart disease with hybrid random forest with linear model (HRFLM).

[9] Anjan N. Repaka et al, proposed a model stated the performance of prediction for two classification models, which is analyzed and compared to previous work. The experimental results show that accuracy is improved in finding the percentage of risk prediction of our proposed method in comparison with other models.

[10] Aakash Chauhan et al, proposed "Heart Disease Prediction using Evolutionary Rule Learning". Data is directly retrieved from electronic records that reduce the manual tasks. The amount of services were decreased and shown major number of rules helps within the best prediction of heart disease. Frequent pattern growth association mining is performed on patient's dataset to generate strong association.

II.PROBLEM STATEMENT

The major challenge in heart disease is its detection. There are instruments available which can predict heart disease but either it are expensive or are not efficient to calculate chance of heart disease in human. Early detection of cardiac diseases can decrease the mortality rate and overall complications. However, it is not possible to monitor patients every day in all cases accurately and consultation of a patient for 24 hours by a doctor is not available since it requires more sapience, time and expertise. Since we have a good amount of data in today's world, we can use various machine learning algorithms to analyze the data for hidden patterns. The hidden patterns can be used for health diagnosis in medicinal data.

III .PROPOSED METODOLOGY

The working of the system starts with the collection of data and selecting the important attributes. Then the required data is preprocessed into the required format. The data is then divided into two parts training and testing data. The algorithms are applied and the model is trained using the training data. The accuracy of the system is obtained by testing the system using the testing data. This system is implemented using the following modules.
1.) Data Collection
2.) Data Pre-Processing
3.) Feature Selection and Engineering

4.) Model Selection
5.) Training the Model
6.) Model Evaluation
7.) Interpretability and Explainability
8.) Integration with Clinical Workflow
9.) Validation and Clinical Trials
10.) Continuous Improvement

IV. EXISTING SYSTEM

Heart disease is even being highlighted as a silent killer which leads to the death of a person without obvious symptoms. The nature of the disease is the cause of growing anxiety about the disease & its consequences. Hence continued efforts are being done to predict the possibility of this deadly disease in prior. So that various tools & techniques are regularly being experimented with to suit the present-day health needs. Machine Learning techniques can be a boon in this regard. Even though heart disease can occur in different forms, there is a common set of core risk factors that influence whether someone will ultimately be at risk for heart disease or not. By collecting the data from various sources, classifying them under suitable headings & finally analyzing to extract the desired data we can conclude. This technique can be very well adapted to the do the prediction of heart disease. As the well-known quote says "Prevention is better than cure", early prediction & its control can be helpful to prevent & decrease

V. PROPOSED METHODOLOGY

The data for 500 people was collected from surveys done by the American Heart Association. Most of the heart disease patients had many similarities in the risk factors. The TABLE I below shows the identified important risk factors and the corresponding values and their encoded values in brackets, which were used as input to the system. Developing a chronic heart disease finder using AI involves several key steps and methodologies. Here's a proposed methodology for building such a system:

1. Data Collection:
Gather a large and diverse dataset of medical records, including patient history, diagnostic tests, imaging results, and other relevant medical information related to chronic heart disease. Ensure the data is representative of different demographics, regions, and types of heart disease to make the AI model more robust and effective.

2. Data Preprocessing:
Clean the data to remove any inconsistencies, missing values, or errors that may affect the performance of the AI model. Normalize and standardize the data to ensure that it is in a consistent format and scales across different features.

3. Feature Selection and Engineering:
Identify and select the most relevant features from the dataset that are indicative of chronic heart disease. This may involve consulting with medical experts to ensure that important indicators are not overlooked. Create new features or transform existing ones to better represent the underlying patterns in the data that could help in detecting chronic heart disease.

4. Model Selection:
- Evaluate various AI techniques such as machine learning algorithms (e.g., logistic regression, random forests, support vector machines) and deep learning models (e.g., convolutional neural networks, recurrent neural networks) to determine the most suitable approach for detecting chronic heart disease. Consider the trade-offs between model complexity, interpretability, and performance to select the most appropriate model for the task.

5. Training the Model:
Split the dataset into training, validation, and test sets to train and evaluate the performance of the AI model. Use the training data to teach the model to recognize patterns associated with chronic heart disease and adjust the model parameters to minimize prediction errors.

6. Model Evaluation:
Assess the performance of the AI model using metrics such as accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve (AUC-ROC). the model on an independent test set to ensure that it generalizes well to new, unseen data.

7. Interpretability and Explainability:
Implement methods for interpreting and explaining the AI model's predictions, particularly in the context of medical decision-making. This is crucial for gaining the trust of healthcare professionals and patients.

8. Integration with Clinical Workflow:
Integrate the AI model with electronic health record systems or other clinical platforms to streamline the process of chronic heart disease detection within the healthcare setting. Ensure that the AI system complies with relevant healthcare regulations and standards for patient data privacy and security.

9. Validation and Clinical Trials:
Conduct validation studies and clinical trials to assess the real-world performance and impact of the AI-based chronic heart disease finder in clinical practice. Collaborate with healthcare providers to validate the system's effectiveness and potential benefits for patients and clinicians.

10. Continuous Improvement:
Continue to monitor and improve the AI model's performance over time by incorporating new data, feedback from healthcare professionals, and advancements in AI research and technology.

TABLE I
RISK FACTORS VALUES AND THEIR ENCODINGS

	Risk Factors	Values
1	Sex	Male (1), Female (0)
2	Age (years)	20-34 (-2), 35-50 (-1), 51-60 (0), 61-79 (1) , >79 (2)
3	Blood Cholesterol	Below 200 mg/dL - Low (-1) 200-239 mg/dL - Normal (0) 240 mg/dL and above - High (1)
4	Blood Pressure	Below 120 mm Hg- Low (-1) 120 to 139 mm Hg- Normal (0) Above 139 mm Hg- High (-1)
5	Hereditary	Family Member diagnosed with HD -Yes (1) Otherwise –No (0)
6	Smoking	Yes (1) or No (0)
7	Alcohol Intake	Yes (1) or No (0)
8	Physical Activity	Low (-1) , Normal (0) or High (-1)
9	Diabetes	Yes (1) or No (0)
10	Diet	Poor (-1), Normal (0) or Good (1)
11	Obesity	Yes (1) or No (0)
12	Stress	Yes (1) or No (0)
Output	Heart Disease	Yes (1) or No (0)

sketchbook: a standard place to store your programs (or sketches). The sketches in your sketchbook can be opened from the File > Sketchbook menu or from the Open button on the toolbar. The first time you run the Arduino software, it will automatically create a directory for your sketchbook.

VII. RESULT

The result of the chronic heart disease finder using AI can have a significant impact on healthcare by improving the early detection, intervention and patient outcomes. Here is the summary of the potential results and benefits that can be achieved early detection, improved accuracy, personalized risk assessment, enhanced treatment planning, reduction in healthcare cost, streamlined workflow, decision support for healthcare professionals, patient empowerment and adaptation.

VIII. LIMITATIONS

It's important to note that the success of the chronic heart disease finder using AI will depend on the thorough validation of the model in real-world clinical settings, its integration into healthcare workflows, and the collaboration and trust of healthcare professionals. Additionally, ongoing monitoring and evaluation will be crucial to ensure that the AI system continues to deliver accurate and beneficial results while adhering to ethical and regulatory standards for patient safety and privacy.

IX .CONCLUSION

Heart diseases are a major killer in India and throughout the world, application of promising technology like machine learning to the initial prediction of heart diseases will have a profound impact on society. The early prognosis of heart disease can aid in making decisions on lifestyle changes in high-risk patients and in turn reduce the complications, which can be a great milestone in the field of medicine. The number of people facing heart diseases is on a raise each year. This prompts for its early diagnosis and treatment. The utilization of suitable technology support in this regard can prove to be highly beneficial to the medical fraternity and patients. In this paper, the seven different machine learning algorithms used to measure the performance are SVM, Decision Tree, Random Forest, Naïve Bayes, Logistic Regression, Adaptive Boosting, and Extreme Gradient Boosting applied on the dataset. The expected attributes leading to heart disease in patients are available in the dataset which contains 76 features and 14 important features that are useful to evaluate the system are selected among them. If all the features taken into the consideration, then the efficiency of the system the author gets is less. To increase efficiency, attribute selection is done. In this n features have to be selected for evaluating the model which gives more accuracy. The correlation of some features in the dataset is almost equal and so they are removed. If all the attributes present in the dataset are taken into account, then the efficiency decreases considerably. All the seven machine learning methods accuracies are compared based on which one prediction model is generated. Hence, the aim is to use various evaluation metrics like confusion matrix, accuracy, precision, recall, and f1-score which predicts the disease efficiently. Comparing all seven the extreme gradient boosting classifier gives the highest accuracy of 81%.

ACKNOWLEDEMENT

We would like to show our gratitude to the Paavai Engineering College and thank teaching and non-teaching staff of Department of ECE. Also, thanks to our parents and friends who all are directly or indirectly supported for this research.

REFFERENCES

[1] Soni J, Ansari U, Sharma D & Soni S (2011). Predictive data mining for medical diagnosis: an overview of heart disease prediction. International Journal of Computer Applications, 17(8), 43-8

[2] Dangare C S & Apte S S (2012). Improved study of heart disease prediction system using data mining classification techniques. International Journal of Computer Applications, 47(10), 44-8.

[3] Ordonez C (2006). Association rule discovery with the train and test approach for heart disease prediction. IEEE Transactions on Information Technology in Biomedicine, 10(2), 334-43.

[4] Shinde R, Arjun S, Patil P & Waghmare J (2015). An intelligent heart disease prediction system using k-means clustering and Naïve Bayes algorithm. International Journal of Computer Science and Information Technologies, 6(1), 637-9.

[5] Bashir S, Qamar U & Javed M Y (2014, November). An ensemble-based decision support framework for intelligent heart disease diagnosis. In International Conference on Information Society (i-Society 2014) (pp. 259-64). IEEE. ICCRDA 2020 IOP Conf. Series: Materials Science and Engineering 1022 (2021) 012072 IOP Publishing doi:10.1088/1757-899X/1022/1/012072 9

[6] Jee S H, Jang Y, Oh D J, Oh B H, Lee S H, Park S W & Yun Y D (2014). A coronary heart disease prediction model: the Korean Heart Study. BMJ open, 4(5), e005025.

[7] Ganna A, Magnusson P K, Pedersen N L, de Faire U, Reilly M, Ärnlöv J & Ingelsson E (2013). Multilocus genetic risk scores for coronary heart disease prediction. Arteriosclerosis, thrombosis, and vascular biology, 33(9), 2267-72.

[8] Jabbar M A, Deekshatulu B L & Chandra P (2013, March). Heart disease prediction using lazy associative classification. In 2013 International MutliConference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s) (pp. 40- 6). IEEE.

[9] Brown N, Young T, Gray D, Skene A M & Hampton J R (1997). Inpatient deaths from acute myocardial infarction, 1982-92: analysis of data in the Nottingham heart attack register. BMJ, 315(7101), 159-64.

[10] Folsom A R, Prineas R J, Kaye S A & Soler J T (1989). Body fat distribution and self-reported prevalence of hypertension, heart attack, and other heart disease in older women. International journal of epidemiologyy, 18(2), 361-7.