Enhancing Cardiovascular Disease Prediction Accurancy Through Optimized Machine Learning Technique

[1]Rajan Raman, [2]Dr. Dinesh Kumar Sahu

[1]M. Tech. Scholar, [2]Professor
Department of Computer Science and Engineering
Sarvepalli Radhakrishnan University, Bhopal

Abstract

Health is not everything, however, everything else is nothing without health. Then and now, people are trying out ways which can increase the longevity of life. Still, technology is far away from achieving this goal of reducing the mortality rate. However, for a start, the beginning steps have been done. Advancing technology and its influence on peoples‟ life is already leading to healthy lifestyles. Healthy living habits, routines, proactive health monitoring, and early detection of diseases lead to increased expectancy of life. Today, the world is adopting Internet of Things in its daily uses. There are various wearable technological devices that have been developed to monitor/measure different health attributes. The criticality of health diseases are more crucial and can even lead to vulnerable consequences if it is not detected at an earlier stage. The techniques such as electronic health records, body area networks are emerged to continuously monitor and diagnose patient‟s health conditions through the projection of medical sensors and wearable devices across human bodies. Since the data generated from the body area networks are continuous and tremendous in volume, the machine learning techniques are used for efficient health data classification processes. However, health data classification is the most challenging process as it needs to be executed accurately with an earlier prediction of heart diseases.

Keywords- *Particle swarm optimization, Neural Network, Classification, Heart Disease*

I. INTRODUCTION

Healthcare is the perpetuation of one‟s state of good health through timely prevention, diagnosis, and treatment of body parts‟ impairments or diseases. According to an old saying, ”Health is Wealth”, a healthy body is the most precious thing a person could possess. However, in today‟s busy world, people are neglecting the fact that health is one of the most important aspects of life, that without health nothing is useful. So, there is a requirement of something that consequently keeps a check on one‟s health status [1]. In this era, nothing could better serve for this purpose than the technology itself. Today, Internet of Things (IoT) has become an essential contributor to the healthcare sector with its salient features such as connectivity, sensing, expressing, safety and intelligence [2]. IoT reflects the connectivity of physical devices on the internet, which is allowed to communicate with each other for sharing information [3].

Besides interacting with each other, they can be monitored and controlled remotely. Data is collected from various sensors like fetal monitors, electrocardiograms (ECG), blood glucose sensors, hence keeping the health records. Thus they prove to help solve the significant problems in the healthcare domain such as unavailability of healthcare facilities and improper treatments that are provided in rural areas [4]. It should also be beneficial in treating the illness of the elderly and the patients suffering from chronic diseases [5]. Therefore, these IoT devices collect information for the real-time processing of a person‟s health. Using this data, it is suggested if there is a need for the person to avail any health-related facilities or bring a change in his/her lifestyle, by using the Data Mining and Machine Learning techniques. Data mining is a field in which we apply different preprocessing techniques like Statistical Transformation [6], Normalization [7], Standardization [8] and some of Machine Learning operations from the available dataset to get some useful patterns. These patterns are then used in extracting knowledge from the available datasets. Hence, this process is also called Knowledge Discovery in Databases (KDD). The primary purpose of data mining application in healthcare systems is to develop an automated tool for identifying and disseminating relevant healthcare information [9]. Machine Learning, on the other hand, is a field of Computer Science in which the human train machines with a given training dataset and teach it how to perform on a given dataset according to the requirements and then provide results on the input test dataset. It has numerous applications in healthcare, as it can be helpful in disease prediction at an early stage so that proper precautions could be taken [10]. It gives reduced healthcare costs, so it is affordable for a wider demographic [11].

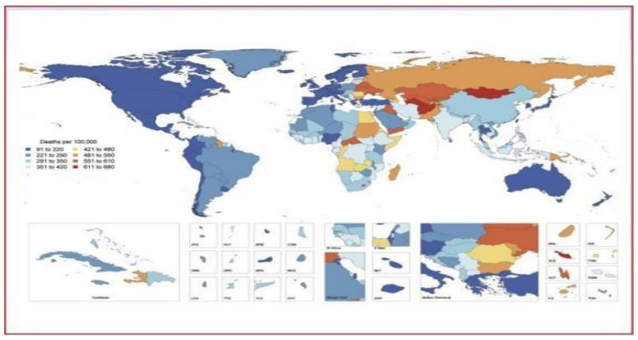

Fig. 1: Global Map Standardized Death Rate of CVD in 2017

The variety of modern imaging, clinical methodology for diagnosis of heart disease cost is too high. Primary symptoms associated with the cardiac disease include chest discomfort, dyspnoea, fatigue, edema, palpations, and syncope, cough, haemoptysis, and cyanosis are additional examples. Heart disease death rate structure of the year 2017 is showed on Fig 1.

II. LITERATURE REVIEW

Niloy Biswas et al. [1], nearly 17.9 million people die from cardiovascular disease, accounting for 32% of deaths worldwide. It is a worldwide concern today. However, the good news is that early treatment can reduce the mortality rate from heart disease. Early detection is important. This study aims to develop a potential machine learning model for cardiac prediction. Early-stage disease using various feature selection techniques to identify important features. There were three different approaches Chi-square, ANOVA, mutual information, etc. were applied for feature selection and the selected feature subsets were labeled. They are SF1, SF2, and SF3 respectively. Next, six different machine learning models, such as B. Logistic regression (C1) and support vectors To find this out, machines (C2), K-Nearest Neighbors (C3), Random Forests (C4), Naive Bayes (C5), and Decision Trees (C6) were applied. Most optimistic model and best-fitting feature subset. In the end, Random Forest turned out to be the most optimistic. Performance of SF3 feature subset (accuracy 94.51%, sensitivity 94.87%, specificity 94.23%, area under the ROC curve 94.95) (AURC) and 0.31 log loss. The performance of the applied model and the selected features indicate that the proposed model is this. It has great potential for clinical use to predict early cardiac disease at low cost and in a short time.

Karna et al. [2], Cardiovascular diseases (CVDs) are killing about 17.9 million people every year. Early prediction can help people to change their lifestyles and to endure proper medical treatment if necessary. The data available in the healthcare sector is very useful to predict whether a patient will have a disease or not in the future. In this research, several machine learning algorithms such as Decision Tree (DT), Discriminant Analysis (DA), Logistic Regression (LR), Naïve Bayes (NB), Support Vector Machines (SVM), K-Nearest Neighbors (KNN), and Ensemble were trained on Cleveland heart disease dataset. The performance of the algorithms was evaluated using 10-fold cross-validation without and with Principal Component Analysis (PCA). LR provided the highest accuracy of 85.8% with PCA by keeping 9 components and Ensemble classifiers and attained an accuracy of 83.8% using a Bagged tree with PCA by keeping 10 components.

M. Ganesan et al. [3], latest advancements in field of IoT and sensing technologies can be employed for online healthcare services. The gigantic quantity of information is being formed through the IoT devices in the medical field and cloud computing techniques have been used to manage the massive amount of data. To avail good service to the user using the online healthcare services, a fresh Cloud as well as IoT based Healthcare application to monitor in addition to diagnose serious diseases is developed. In this study, an efficient framework is utilized for heart disease is created utilizing the UCI Repository dataset as well as the healthcare sensors to predict the public who suffer from heart disease. Moreover, classification algorithms are used to classify the patient data for the identification of heart disease. In the training phase, the classifier will be trained using the data from benchmark dataset. During the testing phase, the actual patient data to identify disease is used to identify the presence of disease. For experimentation, a benchmark dataset is tested using a set of classifiers namely J48, logistic regression (LR), multilayer perception (MLP) and support vector machine (SVM). The simulation results ensured that the J48 classifiers shows superior performance in terms of different measures such as accuracy, precision, recall, F-score and kappa value.

Priyan Malarvizhi Kumar et al. [4], among the applications enabled by the Internet of Things (IoT), continuous health monitoring systems are particularly important. Wearable sensor devices used in IoT health monitoring systems continuously generate large amounts of data. The data generation rate of IoT sensor devices is very fast. Therefore, the amount of data in IoT-based health monitoring systems is also very large. To solve this problem, this paper proposes a scalable his three-layer architecture for storing and processing large amounts of data from wearable sensors. Tier-1 focuses

on data collection from wearable IoT sensor devices. Tier-2 uses Apache HBase to store large amounts of data from wearable IoT sensors in cloud computing. Additionally, Tier-3 uses Apache Mahout to develop a logistic regression-based heart disease prediction model. Finally, ROC analysis is performed to identify the most important clinical parameters for the development of heart disease.

Prabal Verma et al. [5], in recent years, m-healthcare applications that utilize the Internet of Things (IoT) have become established. Multidimensional capabilities and real-time services. These applications provide a platform to millions of people People regularly receive health updates for a healthier lifestyle. Introduction of IoT devices in healthcare. The environment activated several features of these applications. Big data generated by IoT Healthcare devices are analyzed in the cloud instead of relying solely on limited storage. Handheld device computing resources. In this context, cloud-centric and IoT-based Proposes a disease diagnosis framework for m- health monitoring that predicts potential diseases Along with its severity. Key terms are defined to generate user-centered health measurements By exploring computer science concepts. Architectural prototypes for smart students Healthcare is designed around use cases. Results are calculated after health processing Measurements in specific contexts. Our case studies include systematic health data from the student perspective Generated using UCI dataset and medical sensors to predict students with different disease severities. Diagnostic schemes are applied using various state-of-the-art classification algorithms and the results are presented. Calculated based on accuracy, sensitivity, specificity, and F value. The experimental results are The proposed methodology outperforms baseline disease prediction methods.

M.Ganesan et al. [6], internet of Things (IoT) patterns are best practices for solving common problems in the IoT paradigm. Due to its scope and scale, IoT consists of many patterns that IoT practitioners use to design and build ubiquitous smart objects for various IoT use cases. Currently, finding appropriate classification schemes for these patterns remains a hurdle for many IoT architects and practitioners, and existing classification schemes are either arbitrary, incomplete, or contain overlapping classification domains. I'm using. In this article, we conduct a survey of some well-known IoT patterns and their classification schemes and systematically review the literature to identify commonalities and areas of emphasis in the IoT pattern literature. We recommend shifting your focus to using collaborative IoT pattern languages and ontologies to organize your IoT patterns. Structural health monitoring (SHM) is a technology that ensures the safety and reliability of structures through continuous real-time monitoring. IoT-based sensors have become a popular solution for implementing SHM systems, and research in this area is important to improve the accuracy and reliability of his SHM systems.

R.Rajaduari et al. [7], to identify the challenges and opportunities for further development of SHM systems based on IoT sensors, it is necessary to consider the current state-of-the-art. This study presents a study to comprehensively review SHM with a focus on IoT sensors. Second, we classify current structural monitoring methods and discuss the advantages and disadvantages of the methods. Third, an analysis is performed and the results are compared with methods of civil structural monitoring. Finally, we discuss and summarize the main features of this method and finally suggest some directions for future research.

Amin Khatami et al. [8], recognizes a classifier model of sending highlight incorporation joined with back-disposal of different data sets specifically Arrhythmia, and coronary illness datasets and ECG datasets. Experiential outcomes focus on that the element choices improved order procedures precisely and downsized the amount of admissions. In this paper the given arrhythmia dataset gives the upgraded presentation of 78 rate with condensed conspicuously decreased from highlights of nineteen. The following data set gives the upgraded exhibition of 85 rate and the amount of highlights cut back to four. The above Investigation exhibits that reduct highlights upgrade the classifier execution.

Zhang, Shuai et al. [9], shows a helpful neural organization troupe (CNNEs) to prepare the individual neural organizations in a gathering model. This model utilizes a valuable way to deal with improve the exactness measures to decide the concealed hubs in an individual neural organization. Each neural organization in a multi- layered model is prepared independently in a steady way utilizing the negative connection ideas. The utilization of hurtful relationship learning strategies keeps up the property of variety over various layers of the neural organizations. This technique is widely tried across different AI issues, for example, malignancy identification, diabetes, coronary illness forecast, and letter acknowledgment datasets. The outcomes show it furnishes a helpful neural organization model with a further extent of speculation capacity.

Zhiyong Wang et al. [10], this framework is uncommonly intended to manage UCI apparatus heart data sets. This framework utilizes the choice tree based calculation to recognize the fundamental ascribes that help viable determination and meds. The yield information determined in fluffy guideline base. Fluffy estimate is utilized for yield deduction. Molecule swarm improvement calculation enhances the created fluffy master framework to yield a higher precision of 93.27%. The noteworthy favorable position of the framework is that the yield models delivered by the fluffy master frameworks are effectively deciphered in contrast with the other order draws near.

III. TYPES OF CARDIAC DISEASE

There are several categories of heart diseases. Figure 2 shows the various types of heart disease based on clinical conditions. These categories are broadly classified as myocardial infarction, heart failure, heart arrhythmia, angina pectoris, cardiomyopathy, atrial fibrillation based on their clinical evidence. Heart disease has many features, which affect the function or structure of the heart [8].

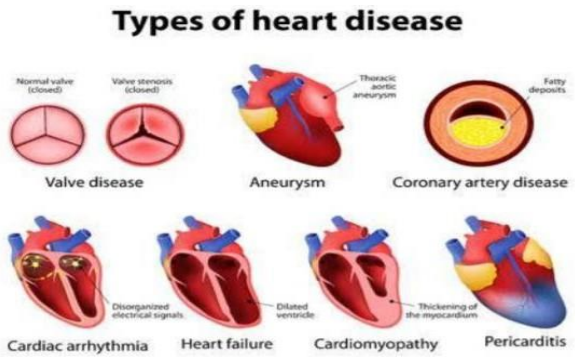

Fig. 2: Types of Cardiac Disease

Coronary Artery Disease

The coronary artery disease is discomfort induce by depleted circulation of blood .The depletion supply in arteries will damage the vein and produce the discomfort to the regular systolic and diastolic function of the heart [9].

Acute myocardial infarction

Clinical name for a cardiac arrest is acute myocardial infarction. A cardiac arrest is a condition that fatty substances present in the blood value affect the rate of flow which results tissue damage on arteries. The blockage arteries may not be able to supply the oxygenated blood supply to the body which will result in the dysfunction to other organs. Figure 3 explains a type of heart arrest caused by intense pressure [10].

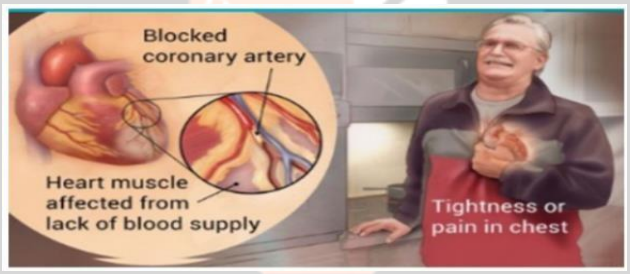

Fig. 3: Acute Myocardial Infarction

Chest Pain (Angina)

Clinical name of chest pressure is Angina. It is predominant medical attention need emergency treatment for the patients. Patients has to treated with ventilators immediately if we experience this type of discomfort. Due to the poor supply of blood flow will cause the pressure on the blood walls and affect the blood vessels. Which will creates pressure on the blood vessals results chest pain. Figure 4 shows typical angina caused in the coronary vessel. Stable angina is the condition causes in peritoriam. Irregular blood flow between the peritoris walls. The main reasons of unstable angina are lifestyle modification, behavioural habits. Figure 5 shows typical unstableangina caused in the coronary vessel [11].

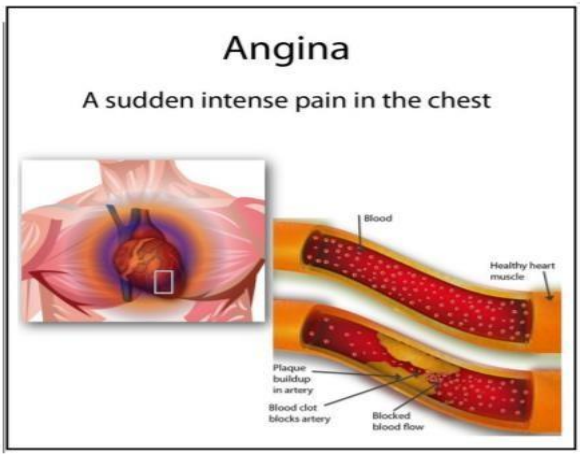

Fig. 4: Angina

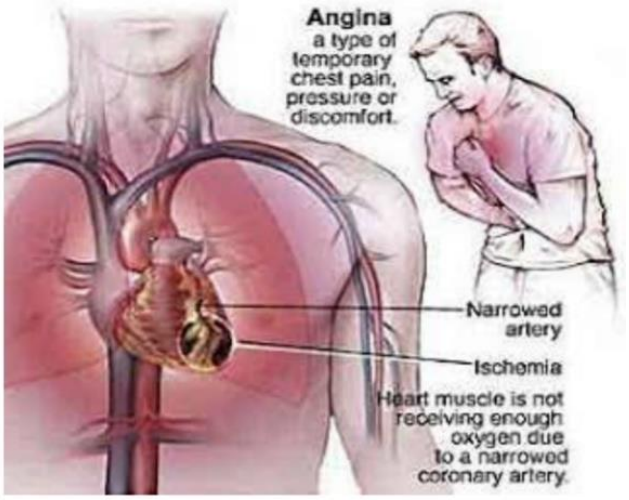

Fig. 5: Unstable Angina

IV. METHODS

1. **Particle Component Analysis**:-The notation is used in Particle swarm optimization are as follows:
Xid :Component in dimension d of the ith particle of swarm
Vid :The particle velocity of particle I in dimension d PBi : the best position achieved so far by particle i GB : The best global best position
C1,C2 : Constant weight factors W: The inertia weight
r1,r2: Random factors in [0,1] interval
Vmin : The minimum velocity value of particle Vmax: The maximum velocity value of particle yi : The fitness value of particle i
PSO is a population base stochastic optimization technique inspired by the social behavior of swarm, such as bird flocking or fish schooling, to obtain a promising position to achieve certain objectives. The PSO algorithm works by having a population (called a swarm) of candidate solution (called particles).Each particle in a population has a fitness value computed from a fitness function and each particle has a position, and move based on an updated velocity according to few simple formula. The movements of the particles are guided by their own best known position in the search space as well as the entire swarm‟s best known position. The particle movements are directed by the position vector and velocity vector of each particle.

2. Meachine Learing
The main property of an ML is its capability to learn. Learning or preparing is a procedure by methods for which a neural system adjusts to a boost by making legitimate parameter modifications, bringing about the generation of wanted reaction. Learning in an ML is chiefly ordered into two classes as [9].

- Supervised learning
- Unsupervised learning

Supervised Learning
Regulated learning is two stage forms, in the initial step: a model is fabricated depicting a foreordained arrangement of information classes or ideas. The model developed by investigating database tuples portrayed by traits. Each tuple is expected to have a place with a predefined class, as dictated by one of the qualities, called to have a place with a reclassified class, as controlled by one of the traits called the class name characteristic. The information tuple are dissected to fabricate the model all things considered from the preparation dataset.

Unsupervised learning
It is the kind of learning in which the class mark of each preparation test isn't knows, and the number or set of classes to be scholarly may not be known ahead of time. The prerequisite for having a named reaction variable in preparing information from the administered learning system may not be fulfilled in a few circumstances.
Data mining field is a highly efficient techniques like association rule learning. Data mining performs the interesting machine-learning algorithms like inductive-rule learning with the construction of decision trees to

development of large databases process. Data mining techniques are employed in large interesting organizations and data investigations. Many data mining approaches use classification related methods for identification of useful information from continuous data streams.

V. CONCLUSION

This paper provides the need for machine learning techniques in health data mining and management processes. In particular, the dissertation provides special attention to heart disease prediction and diagnosis measures. This due to the reason that heart diseases are one of the major concern across several countries and can even lead to death if not treated at an earlier stage. The major thesis objectives are achieved successfully and specifically we have accomplished three tasks:

- First, we formulate the problems and difficulties of machine learning classification models across healthcare applications.
- We solve the issues across healthcare systems, we formulated algorithms that deal with those issues in an effective manner.
- We evaluated the proposed approach with real-time datas.

REFERENCES

[1] Niloy Biswas, Md Mamun Ali, Md Abdur Rahaman, Minhajul Islam, Md. Rajib Mia, Sami Azam, Kawsar Ahmed, Francis M. Bui, Fahad Ahmed Al-Zahrani and Mohammad Ali Moni, "Machine Learning-Based Model to Predict Heart Disease in Early Stage Employing Different Feature Selection Techniques", Hindawi BioMed Research International Volume 2023, Hindawi 2023.

[2] Karna Vishnu Vardhana Reddy, Irraivan Elamvazuthi, Azrina Abd Aziz, Sivajothi Paramasivam and Hui Na Chua, "Heart Disease Risk Prediction using Machine Learning with Principal Component Analysis", International Conference on Intelligent and Advanced Systems (ICIAS), IEEE 2021.

[3] M.Ganesan and Dr. N. Sivakumar, "IoT based heart disease prediction and diagnosis model for healthcare using machine learning models", International Conference on System, Computation, Automation and Networking (ICSCAN), IEEE 2019.

[4] Priyan Malarvizhi Kumar, Usha Devi Gandhi, "A novel threetier Internet of Thingsnarchitecture with machine learning algorithm for early detection of heart diseases", Computers and Electrical Engineering, Vol.65, pp. 222–235, 2018.

[5] Prabal Verma, Sandeep K. Sood, "Cloud-centric IoT based disease diagnosis healthcare framework", J, Parrallel Distrib. Comput., 2018.

[6] M.Ganesan, Dr.N.Sivakumar, "A Survey on IoTrelated Patterns",International Journal of Pure and Applied Mathematics,Volume 117 No. 19, 365-369, 2019.

[7] R.Rajaduari, M.Ganesan,Ms.Nithya "A Survey on Structural Health Monitoring based on Internet of Things"International Journal of Pure and Applied Mathematics,Volume 117 No. 18, 389-393, 2020.

[8] Amin Khatami, AbbasKhosravi, C. L. (2021), „Medical image analysis using wavelet transform and deep belief networks", Journal of Expert Systems With Applications 3(4), 190–198. Zhang, Shuai, Y.-L. S. A. (2017), „Deep learning based recommender system: a survey and new perspectives", Journal of ACM Computing Surveys 1(1), 1–35.

[9] Zhiyong Wang, Xinfeng Liu, J. G. (2016), „Identification of metabolic biomarkers in patients with type-2 diabetic coronary heart diseases based on metabolomic approach", 6(30), 435–439.

[10] Ashwini Shetty, Naik, C. (2022), „Different data mining approaches for predicting heart disease", International journal of innovative research in science, engineering and technology 3(2), 277–281.

[11] Berikol, B. and Yildiz (2023), „Diagnosis of acute coronary syndrome with a support vector machine", Journal of Medical System 40(4), 11–18.

[12] Chebbi, A. (2023), „Heart disease prediction using data mining techniques", International journal of research in advent technology 25(3), 781–794.

[13] Cheng-Hsiung Wenga, Tony Cheng-Kui Huang, R.-P.H. (2024), „Disease prediction with different types of neural network classifiers", Journal of Telematics and Informatics (4), 277–292.