



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 12, Issue 3, March 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Disease Prediction using Machine Learning Algorithm for Smart Healthcare System

Govind Singh¹, Prof. Suresh S. Gawande²

M. Tech. Scholar, Department of Electronics and Communication, Bhabha Engineering Research Institute, Bhopal, India¹

Guide, Department of Electronics and Communication, Bhabha Engineering Research Institute, Bhopal, India²

ABSTRACT: Heart disease is a major cause of death worldwide. About 31% of global deaths occur due to this disease. According to the WHO report, approximately 17.9 million people worldwide die each year due to this disease. Early detection of heart disease can reduce the chance of the disease progressing to a more severe stage by providing appropriate treatment. Coronary angiography is the most precise method of detecting the disease; nevertheless, it is a painful and expensive procedure, and patients do not prefer it in the early stages. As a result, a non-invasive and reliable method that can be employed as a screening tool for disease diagnosis is required. Machine learning has influenced every element of human life in the modern era, and it has a huge impact in the health care sector also. Machine learning (ML) based decision support systems can be used to diagnose heart disease by evaluating the clinical parameters of a patient. Various ML techniques have been investigated in this paper, and a system has been designed for the diagnosis of heart disease, with an emphasis on coronary artery disease. Relevant features were selected using the Extra Tree feature selection method and classification was done by XGBoost (extreme gradient boosting) classifiers.

KEYWORDS: Diabetes Mellitus, ML, DT, GB, NN, Accuracy

I. INTRODUCTION

As of now, Heart Disease (HD) is considered as a significant justification for the expanded death rate. In light of the review announced by World Heart Federation Report, it is expressed that 33% of the demise rate can be diminished on account of before distinguishing proof of Heart Disease. Essential side effects of Heart Disease are chest torment, breathing issues, neck torment, jaw, throat, upper stomach, or back. Likewise, barely any restricting variables that help with limiting the dangers of HD are controlled Blood Pressure (BP), low cholesterol, abstaining from smoking, and routine activities. In the greater part of the cases, the HD couldn't be distinguished until a cardiovascular failure, or stroke happens. Thus, it is expected to notice the cardiovascular boundaries and examine with specialists [1, 2]. The innovative improvements in information and registering have empowered the clinical space to accumulate and save persistent clinical information, which supports vital clinical choices [3].

The information which is put away may be analyzed to make fundamental clinical choices that could incorporate conclusion, line of medicines, expectation, and picture examination. The information accessible in the medical care framework are rich. DM strategies go about as a significant job in settling profoundly nonlinear expectation and grouping as well as mind boggling issues over the new many years. In this manner, it is likely to fabricate a model that could anticipate the nonattendance or presence of HD relying upon various side effects of heart-related highlights. It is an imperative need of any undertaking of infection expectation to unequivocally portion the undesirable and solid patient [4]. Else, a sound patient could under unnecessary treatment with an aftereffect of misclassification. It is profoundly vital for foresee any event of sickness precisely.

HD or cardiovascular disease is a chronic illness that threatens all over the world. Disease like a scourge-bubonic plague, yellow fever, and smallpox are not as critical as cardiovascular disease; it strikes down an imaginary fraction of the population without any symptoms but causes prolonged disability. Many kinds of research analyze the reason for cause of HD, diagnosing the disease, treatment for ailments, and preventing them happens rapidly. An explosive extension for understanding the structure and activities of cardiovascular system as well as the capability of evaluating

the parameters in alive human being, sometimes in terms of models which needs penetration of the skin with improved accuracy, by non-invasive models [5].

Heart is one of the significant organs present in the human body. It is a muscle that is made with 4 chambers divided by valves and two divisions. Every half consists of single chamber termed as chamber and one called ventricle. The atria (plural term of chamber) assembles blood, and ventricles agreement to move blood from heart. The right 50% of the heart siphons oxygen, i.e., blood, which has least measure of oxygen to lungs where platelets acquire most extreme oxygen. Then, new oxygenated blood moves from lungs to the left chamber and left ventricle [6]. Left ventricle siphons the new oxygenated blood to substitute organs and different tissues of the body. The oxygenated blood promotes more energy and health for human body [7]. HD is common for a wide range of diseases, anomalies, and limitations which influence the heart as well as the blood vessels. It is prone to males and females in US and crosses millions in Americans where they have myocardial infarctions [8].

II. SYMPTOMS OF HEART DISEASE

As dyspnea is often caused by strenuous exertion in healthy, well-conditioned subjects and by gradual exertion in those people are normal but unaccustomed to exercise, it should be considered as abnormal if it happens at remaining stage of physical activity and unexpectedly cause this symptom. It is associated with any HD. It occurs only at rest and not while exertion is almost invariable. A functional origin for dyspnea is while heightened awareness of breathing is accomplished by extended stabbing pain cardiac apex region or prolonged for more than 2 hours, dull chest pain, and associated with complexity of obtaining sufficient air to the lungs [9].

2.1 Chest Pain or Discomfort

Though chest pain is one of the cardinal malfunctioning of HD, it is not easy to analyze the origination is from the heart or variety of non-cardiac reasons. Some of the signs denote chest pain. People assume that chest pain is heart attack. If alternate constraints are the cause of chest pain, then cardiac disease is most common and dangerous, whereas the sign of heart pain could be treated. It is applied for defining the pain, pressure, choking, numbness, and alternate discomfort in the chest, neck, or upper abdomen which is related with pain in jaw, head, arms and so on. It could retain for seconds to days or even a week, which is to be predicted. The definition of chest pain very vague, and several medicinal conditions arise from HD that produce symptoms [10].

It have described the syncope as “a loss of consciousness, leads to reduced perfusion of the brain.” Since (2017) defined as “Syncope is a immediate and temporary fainting.” Then it is a general sign that many people experience it once in a lifetime, and it is not a critical disease. But, sometimes syncope represents a dangerous, life-threatening condition; if syncope happens, it should be analyzed and treated. The cause of syncope is divided into levels: neurologic, metabolic, vasomotor, and cardiac. Cardiac syncope is a disease that leads to unpredicted death. Heart palpitations It is a common symptom for HD and defined as “an unpleasant awareness of the rapid pumping of the heart.” It might be caused by various disorders which change the cardiac rhythm. Many people complain about palpitations, which is defined as "skipping of heartbeat" of rapid as well as uneven heartbeats. Several people with palpitations are constrained with cardiac arrhythmia which is assumed to be abnormal heart rhythms. There are different types of arrhythmias that lead in palpitations; however random causes [11].

These symptoms are common for all diseases. Fatigue could be assumed as an inability to function normally. Somnolence signifies sleep or worst sleep. A patient many falls to sudden asleep in daytime. Indicates HD, a non-specific indicator denotes the malfunctioning of other organs of body. Likewise, light headedness and giddiness, fatigue and lethargy require a common treatment to begin pinning down a specific cause. Somnolence occurs as the nocturnal sleep anomalies like, restless, leg syndrome, or insomnia [12].

III. MACHINE LEARNING

In today's era, all smart applications are using the features of machine learning for better performance. ML aids in the discovery of hidden information in raw data that normal programs are unable to provide. ML has become an integrated component of all research domains in recent years. Machine learning provides the ability for machines to learn automatically. It is used to give the machine the same decision-making ability as humans [13]. Machine performance increases with experience, and the machine gains experience from data that acts as a mechanism for imparting training to the machine.

3.1 Supervised Learning

In this type of learning, desired outputs are given for inputs in the training data. The machine learns by analyzing the training data and finding the relationships between inputs and outputs. Supervised learning can solve two types of

problems: classification and regression. In classification, data is classified into two or more predetermined classes. In regression, the value of some variables has to be predicted [14].

3.2 Unsupervised Learning

In unsupervised learning, unlabeled data is used to train the machine, and the patterns in the data have to be recognized by the system. The system processes the data to obtain a summary of the data. Clustering falls under the category of unsupervised learning. Clustering means that objects are combined into different groups. Similar objects are combined into a group, which is known as a cluster [15].

3.3 Reinforcement Learning

In reinforcement, learning focus is on performing the best actions to achieve a goal. An agent is responsible for performing actions. The agent acts by learning the environment in which the actions are to be performed. To learn the environment, the agent maintains an internal state. The agent learns the environment by interacting with the environment. The agent learns about its environment through the use of a reward function. Depending upon the actions performed by the agent, it gets a positive reward or negative reward. The agent tries to increase the positive rewards and decrease the negative rewards. In reinforcement learning, there is no need for human experts with knowledge of the application domain. A self-driving car is an application in which reinforcement learning is used [8, 9].

IV. PROPOSED METHODOLOGY

Highlight choice is a course of making a subset of elements intended to further develop order results and execution. Include choice is utilized for an assortment of reasons, like information decrease, diminishing time intricacy, and further developing execution and precision. Include determination frequently alludes to choosing the best and ideal elements connected with the objective classes. The principal objective of this review is to find the best important elements for a given dataset while superfluous highlights are disposed of to further develop the arrangement precision. The general plan of the proposed model is displayed in Fig. 1. Unimportant elements frequently increment the mistake rates and, to diminish the characterization blunder rate, includes that are exceptionally useful are chosen for arrangement.

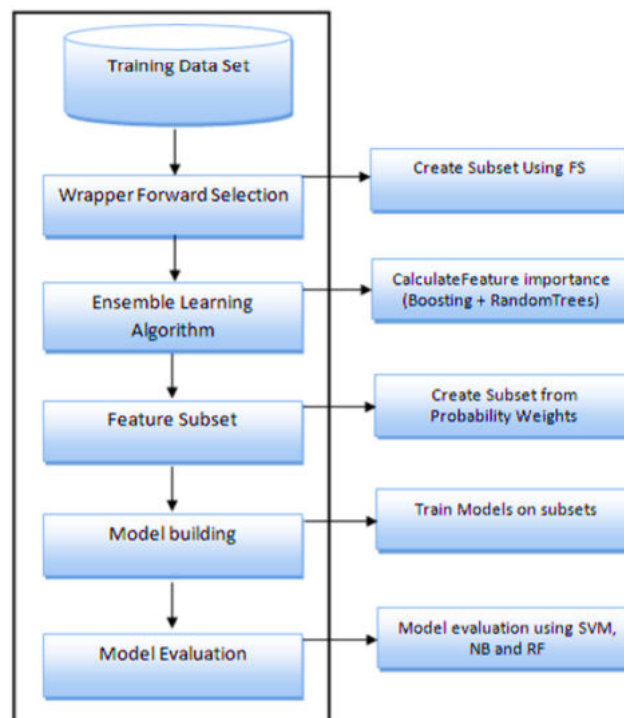


Fig. 1: Schematic Representation of proposed model

For X dataset with Y highlights, FS makes a Y-layered space of SY in which a subspace I from SY_i includes the most significant elements that make sense of the objective class. The normal element choice incorporates the channel,



installed, and covering strategies. Covering strategies advance the advancing iteratively through an avaricious pursuit and use calculations for rating models. The rating models guarantee that the determinations are in line with the calculations' exhibitions. The hunt system frames a stage of various highlights Y in the element space SY and halting standard C is utilized to quit delivering subsets. The closest neighbor is utilized as the base classifier for choosing the best component subset in view of the better presentation. The subsequent subset is utilized to construct irregular trees to additionally stay away from high relationship between's the factors and accomplish consensus.

The most remarkable learning troupe tree model depends on grouping trees that form trees through 2 irregular advances. The initial step takes bootstrap tests and develops trees on each single element. The leaf hub comprises of highlights and its relating individuals in the middle of the road hubs. The subsequent advance adds another tree to the lower part of the current trees. Each tree is a sort of class inclination to the objective class. Trees that have the most class inclinations are grouped into the objective class, separately. The troupe classifier is constructed utilizing arbitrary characterization trees, through which individual trees are developed for every variable by joining different bootstrapped tests B_t from the preparation set. Let $X = (x_1, x_2, \dots, x_d)$ be a vector of highlights. Each component x_1 is an irregular variable in the information vector X . Allow y to be the order of sickness ($y = 0$ shows nonattendance and $y = 1$ demonstrates presence). Presently includes from X foresee y through outfits of the classifiers E with the end goal that $E = (E_1(X), \dots, E_k(X))$. Each $E_l(X)$ is a choice tree and we mean the hyper boundaries of a choice tree for the classifier $E_l(X)$ as $P_k = (P_{k1}, \dots, P_{kn})$. The choice tree classifier is given by $E_k(X) = E(X|P_k)$. Each tree in view of hyper boundaries P_k projects votes to classy in the information vector X , the class with most elevated vote gains.

V. SIMULATION RESULTS

Results as follow

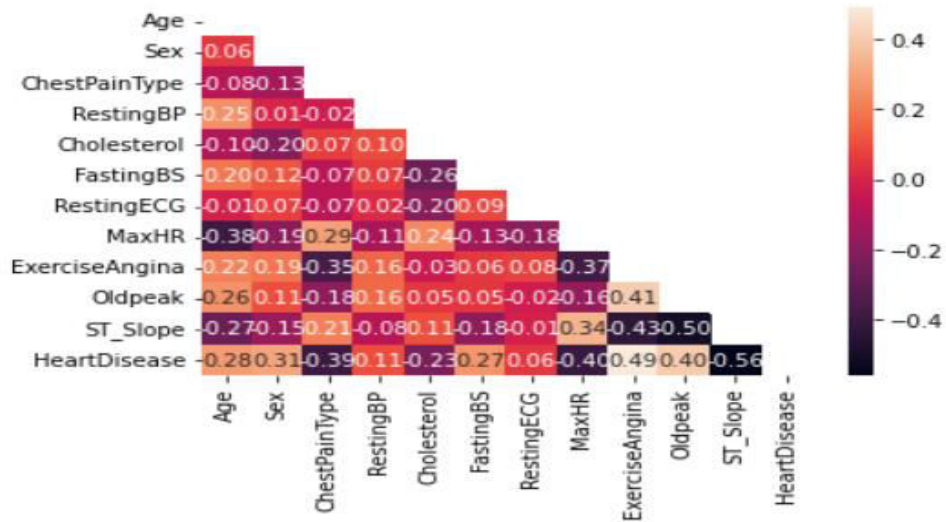
1. Load data and convert it into a pandas data frame

| | Age | Sex | ChestPainType | RestingBP | Cholesterol | FastingBS | RestingECG | MaxHR | ExerciseAngina | Oldpeak | ST_Slope | HeartDisease |
|---|-----|-----|---------------|-----------|-------------|-----------|------------|-------|----------------|---------|----------|--------------|
| 0 | 40 | M | ATA | 140 | 289 | 0 | Normal | 172 | N | 0.0 | Up | 0 |
| 1 | 49 | F | NAP | 160 | 180 | 0 | Normal | 156 | N | 1.0 | Flat | 1 |
| 2 | 37 | M | ATA | 130 | 283 | 0 | ST | 98 | N | 0.0 | Up | 0 |
| 3 | 48 | F | ASY | 138 | 214 | 0 | Normal | 108 | Y | 1.5 | Flat | 1 |
| 4 | 54 | M | NAP | 150 | 195 | 0 | Normal | 122 | N | 0.0 | Up | 0 |

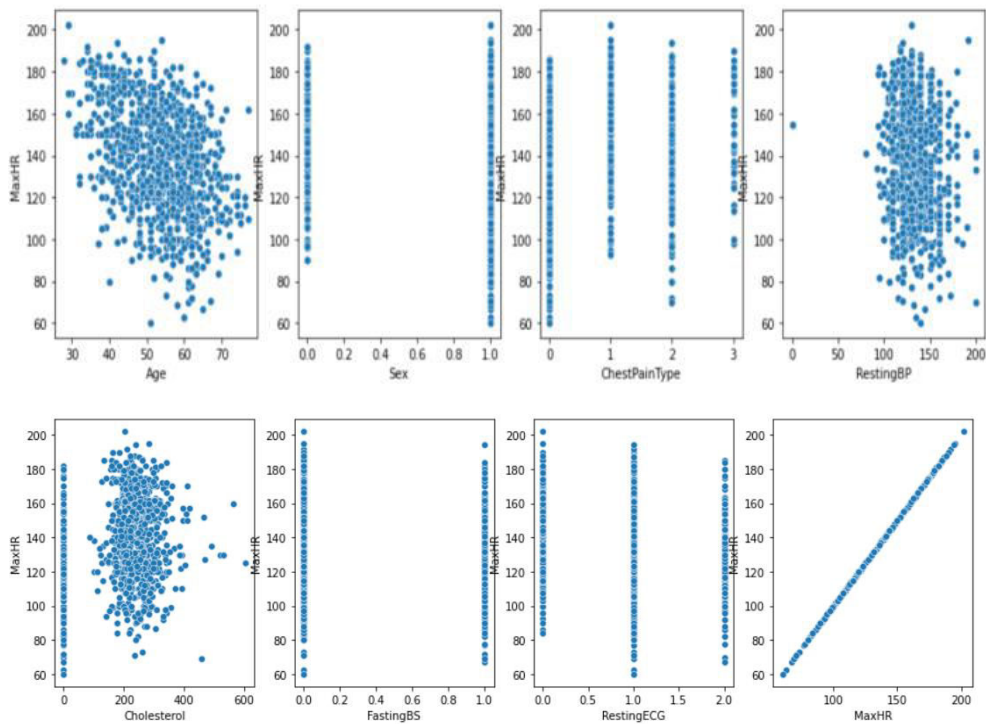
2. Label encoding for changing string value to Numeric

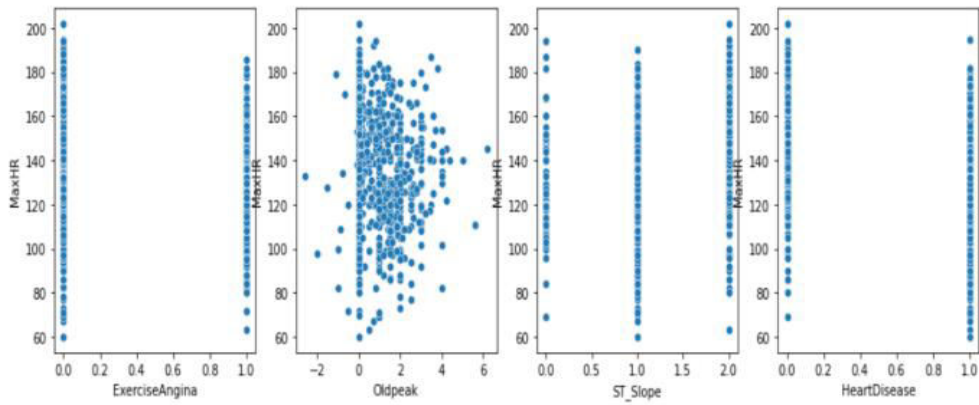
| | Age | Sex | ChestPainType | RestingBP | Cholesterol | FastingBS | RestingECG | MaxHR | ExerciseAngina | Oldpeak | ST_Slope | HeartDisease |
|---|-----|-----|---------------|-----------|-------------|-----------|------------|-------|----------------|---------|----------|--------------|
| 0 | 40 | M | ATA | 140 | 289 | 0 | Normal | 172 | N | 0.0 | Up | 0 |
| 1 | 49 | F | NAP | 160 | 180 | 0 | Normal | 156 | N | 1.0 | Flat | 1 |

3. Find Correlation among data

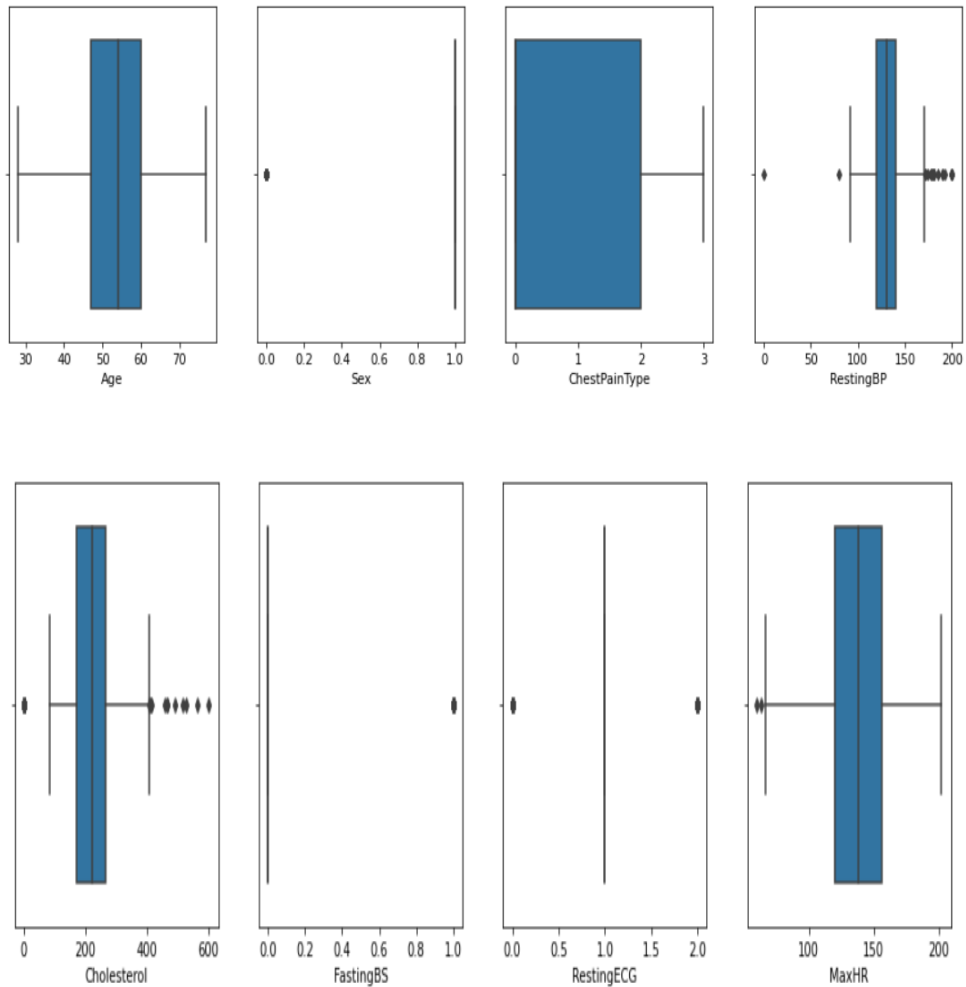


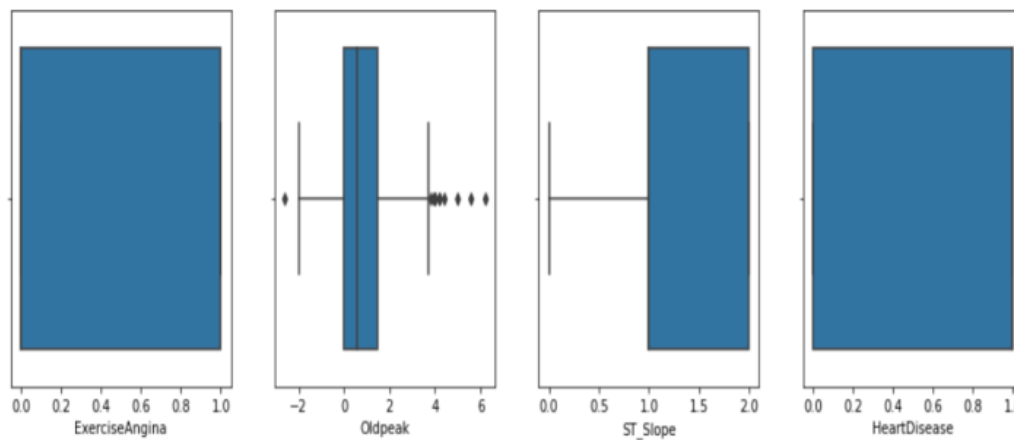
4. Visualization





5. Outlier detection in data





VI. CONCLUSION

Medical diagnosis in clinical practice involves careful physical and history examinations. Various methods of DM were aimed at Heart Disease and have been used to distinguish the Heart Disease patients to identify various Heart Disease sub-kinds and to compute Heart Disease severity. At a later stage, if the heart failure is detected, the methods of DM can be useful wherever the therapeutic advantages for aid and survival prospect are constrained as it enables for timely prediction of morbidity, the risk of readmission and mortality. Recently, ML algorithms have been employed to classify the data in the medical field. The data complexity and quantity needs to be examined and managed to transform the efficient and accurate Heart Disease diagnosis. ML techniques act as a major role in resolving highly nonlinear prediction and classification as well as complex problems over the recent decades. Therefore, it is probable to build a model that might predict the absence or presence of Heart Disease depending on different symptoms of heart-related features.

REFERENCES

- [1] Romany Fouad Mansour, Adnen El Amraoui, Issam Nouaouri, Vicente García Díaz , Deepak Gupta, and Sachin Kumar, "Artificial Intelligence and Internet of Things Enabled Disease Diagnosis Model for Smart Healthcare Systems", *IEEE Access* 2021.
- [2] G. Muhammad, M. S. Hossain, and N. Kumar, "EEG-based pathology detection for home health monitoring," *IEEE J. Sel. Areas Commun.*, vol. 39, no. 2, pp. 603610, Feb. 2021,
- [3] A. A. Mutlag, M. K. A. Ghani, M. A. Mohammed, M. S. Maashi, O. Mohd, S. A. Mostafa, K. H. Abdulkareem, G. Marques, and I. de la Torre Díez, "MAFC: Multi-agent fog computing model for healthcare critical tasks management," *Sensors*, vol. 20, no. 7, p. 1853, Mar. 2020.
- [4] M. S. Hossain and G. Muhammad, "Deep learning based pathology detection for smart connected healthcare," *IEEE Netw.*, vol. 34, no. 6, pp. 120125, Nov. 2020.
- [5] M. S. Hossain G. Muhammad and A. Alamri "Smart Healthcare Monitoring: A Voice Pathology Detection Paradigm for Smart Cities" *Multimedia Systems* vol. 25 no. 5 pp. 565-75 Oct. 2019
- [6] V Krishnapraseeda, M S Geetha Devasena, V Venkatesh and A Kousalya, "Predictive Analytics on Diabetes Data using Machine Learning Techniques", 7th International Conference on Advanced Computing and Communication Systems (ICACCS), pp. 458-463, IEEE 2021
- [7] Valasapalli Mounika, Devi Sree Neeli, Gorla Suma Sree, Parimi Mourya and Modala Aravind Babu, "Prediction of Type-2 Diabetes using Machine Learning Algorithms", International Conference on Artificial Intelligence and Smart Systems (ICAIS), pp. 167-173, IEEE 2021
- [8] I.K. Mujawar, B.T. Jadhav, V.B. Waghmare and R.Y. Patil, "Development of Diabetes Diagnosis System with Artificial Neural Network and Open Source Environment", International Conference on Emerging Smart Computing and Informatics (ESCI), pp. 778-784, IEEE 2021
- [9] Cecilia Saint-Pierre; Florencia Prieto; Valeria Herskovic; Marcos Sepúlveda, "Team Collaboration Networks and Multidisciplinarity in Diabetes Care: Implications for Patient Outcomes", *IEEE Journal of Biomedical and Health Informatics*, Vol. 14(1), pp. 319-329, 2020.

- [10] M. S. Hossain and G. Muhammad, "Emotion-aware connected healthcare big data towards 5G," *IEEE Internet Things J.*, vol. 5, no. 4, pp. 23992406, Aug. 2018.
- [11] M. Pham, Y. Mengistu, H. Do, and W. Sheng, "Delivering home healthcare through a cloud-based smart home environment (CoSHE)," *Future Gener. Comput. Syst.*, vol. 81, pp. 129140, Apr. 2018.
- [12] A. Kaur and A. Jasuja, "Health monitoring based on IoT using raspberry PI," in *Proc. Int. Conf. Comput., Commun. Autom. (ICCCA)*, Greater Noida, India, May 2017, pp. 13351340.
- [13] U. Satija, B. Ramkumar, and M. Sabarimalai Manikandan, "Realtime signal quality-aware ECG telemetry system for IoT-based health care monitoring," *IEEE Internet Things J.*, vol. 4, no. 3, pp. 815823, Jun. 2017.
- [14] O. S. Alwan and K. Prahald Rao, "Dedicated real-time monitoring system for health care using ZigBee," *Healthcare Technol. Lett.*, vol. 4, no. 4, pp. 142144, Aug. 2017.
- [15] P. Kakria N.K. Tripathi and P. Kitipawang "A Real-Time Health Monitoring System for Remote Cardiac Patients Using Smartphone and Wearable Sensors" *t'l. J. Telemedicine and Applications* vol. 2015 2015.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor
Impact Factor: 8.379



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



www.ijircce.com

Scan to save the contact details