**Detection of Cardiovascular Disease using A.I & M.L - A Review**

Parthivi Thakore, Divyansh Johari

Student Researchers, Department of Advanced Computing, Poornima College of Engineering, Jaipur

**Abstract­**-In the 21st century according to stats, the risk of cardiovascular disease has become more common and the death rate caused by it is increasing way more. around 17.9 million lives are taken by CVDs each year. There are various reasons causing it but most importantly it is caused by not identifying it, hence it becomes a very important task for the human race to identify the CVDs and deal with the proper treatment so that the death dance caused by CVDs can be decreased and risk of it at an early age too. This work mainly aims to review and analyze various methods and approaches to detect the presence or absence of CVDs using AI and ML with accurate predictions. An artificial intelligence system for detecting heart disease from phonocardiogram (PCG) signals has been developed utilizing Artificial Neural Networks (ANN) algorithms and also by driving various A.I algorithm on the given electrocardiogram (ECG) data of the patients we can predict the absence or presence of CVDs.

**Keywords**— Cardiovascular disease, Artificial Intelligence, machine learning, classification, comparative analysis, PCG, ECG, Deep Learning.

**I. INTRODUCTION**-heart disease, another name for cardiovascular illness, is a serious global issue that has a big impact on people. According to a recent study, cardiac disorders were responsible for millions of deaths worldwide, or 31% of all fatalities. Medical research has shown that some risk factors increase a person's likelihood of developing heart disease (CVD). According to, some of these factors an unhealthy diet, nicotine use, depression, stress, excessive alcohol use, physical inactivity, inherited obesity, and age are the common causes of CVD. The World Health Organization has published several papers showing an increase in CVD-related deaths, which are primarily attributable to inadequate preventative actions despite rising risk factors.

The heart is one of many organs in the human body that provides blood supply through a function akin to a pump. A healthy heart is fundamental and necessary for human well-being. The leading cause of death in the modern period is cardiovascular disease (CVD), generally known as heart disease. The classification of associated disorders is a challenging undertaking that involves several biological markers and risk factors due to the highly complicated mechanism of the heart. Professionals in related fields employ cardiac physiological signals like the electrocardiogram(ECG) and phonocardiogram (PCG) to monitor or detect cardiovascular-related disorders. Now due to the heart’s complex structure and the death rate by the CVDs have grasped the attention of various researchers and scientists to find and perform various approaches, techniques, and methods that are required in order to detect the CVDs presence or absence of a patient. hence some scientists came up with the output of using AI and various Machine-learning techniques

There are a number of publications that proposes different techniques to apply to the patient’s data to find the collective research whether it is detecting Heart health by extracting its features through its sound that is by using a Phonocardiogram and classifying them using a Neural Network. the implementation of Deep-learning methods on Electrocardiogram data in order to predict CVD can also be applied.

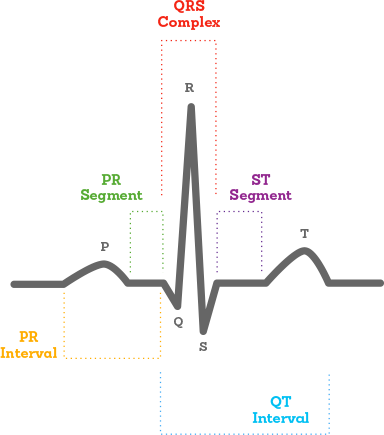
Mohamed and Raafat created a mathematical model in the late 1980s to use a limited set of parameters to characterize heart murmurs and sounds.

For the classification of cardiovascular disease, five alternative methods have been presented: support vector machine (SVM), K-nearest neighbor (K-NN), logistic regression (LR), decision tree (DT), and naive Bayes (NB). These methods were used to classify the various patient data in order to determine the presence of CVDs.

**2. UNDERSTANDINGS & WORKINGS**

So, now to simplify the approach let us first understand more about ECGs and PCG and grasp the main understanding of the Conduction System of heart

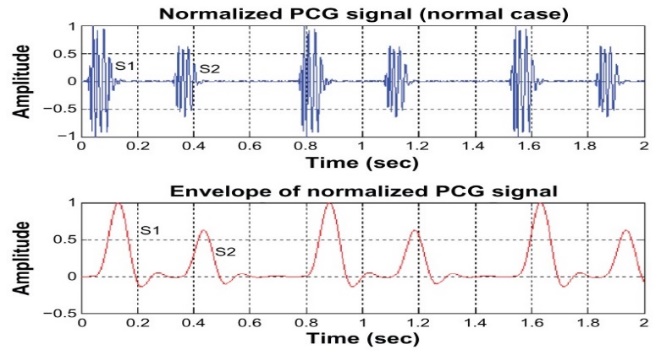
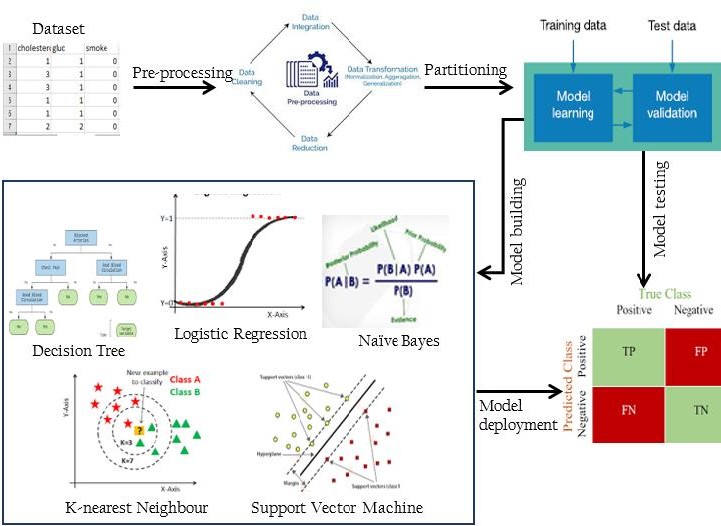
[I] Mainly ECG is a device that records the electrical signals produced by heart mainly by SA (Sino Atrial) Node {The Pacemaker} and AV(Atrio-ventricular) Node where SA node causes the upper heart chambers to contract while AV node to cause the Lower heart contract The ECG is made up of multiple electrodes that are wired to the instrument and to the patient's body. To detect impulses, each sensor detects a change in the electrical charge beneath the skin. An impulse travel quickly and is transmitted to the heart's surrounding cells. Figure 1 showed the electrocardiogram's typical waveform. So, by decoding this impulse we can find the abnormal behavior of the heart if it exists and then we can run our algorithm on it in order to predict the probability of CVDs risk.



According to the diagram you can easily spot three main areas of the ECG graph where the first one is knowns as the P wave depicting atrial depolarization i.e., contraction of atria, then comes the rotated V-shaped wave named QRS complex that depicts ventricular depolarization and atrial repolarization following it is T wave depicting ventricular repolarization i.e., the relaxation process of ventricles. Normally a healthy heart has about 60-100 BPM caused by the SA node and by 40-60 BPM AV node so we check the patient’s heart condition by interpreting this ECG graph by calculating the P waves, PR interval, QRS complex, hearts rhythm, heart rate (ECG should be 6-second strip) by applying 6-second method.



[II] Now for the same purpose PCG as referred to as phonocardiography where this graph gives us the sound wave of the cardiac cycle which is the contraction and relaxation of the heart that causes the murmurs sound which layman often refers to as the Lub-Dub sound of the heart.



These recorded sound tracks provide information regarding valve action and the efficiency of blood pumping under the body, assisting medical experts. Typically, the PCG approach divides the normal heartbeat into the S1 and S2 beats. We can determine the corresponding heart illnesses based on these recorded soundtracks. S1 sound depicts the closure of valves in the upper and lower chambers of the heart usually the frequency remains between 30-100 Hz range for a normal healthy heart while the S2 sound depicts the closure of the valves which has a frequency above 100Hz. So, this is the basic related field solution for finding the abnormalities in the heart in order to predict the probabilities of the risk of CVDs.

**3. Machine Learning Algorithms for the classification of cardiovascular diseases –**

The dataset for categorization is first collected as part of the approach for this study. One of the key jobs in data mining is classification, whose goal is to group documents into one or more classes or categories as a result, this work creates a useful technique for the dataset used to categorize cardiovascular disease. This is accomplished by assuming that there are only two classes: the positive class with uncertain outcomes and the negative class without unexpected findings.

Anaconda Jupyter Notebook, a Python 3 application, is used to implement the algorithms. The datasets are split into training and test sets after pre-processing. Most researchers choose a 70:30 ratio (70% for training and 30% for testing) since the system produces more optimum and accurate results the more data allotted to training. As a consequence, the 70:30 partitioning ratio is employed.

**3.1.Dataset-** The Kaggle online repository was utilized to obtain the dataset used to assess and compare the algorithms employed in this study. The dataset, which comprises of 77,000 patient clinical trial records collected by hospitals for cardiovascular disorders, has three input components: Objective (realistic information), Examination (outcomes of medical investigation), and Subjective (data obtained from a patient). Eleven attributes total, including one target variable with the label "(Absence or Presence) for diagnosis," four objective features, four examination features, three subjective features, and four other attributes, make up the dataset. Table 1 provides a brief summary of the cardiovascular disease dataset collected for this study's analysis.

Table 1 Dataset Description

|  |  |  |  |
| --- | --- | --- | --- |
|  | **Attributes** | **Input Features** | **Data Type/ Description** |
| 1 | Age | Objective features | Int / days |
| 2 | Height | Int / centimeters |
| 3 | Weight | Float/ kilograms |
| 4 | Gender | Categorical code 1:  male, 2: female |
| 5 | Systolic blood pressure | Examination features | Int/ |
| 6 | Diastolic blood pressure | Int/ |
| 7 | Cholesterol | 1: normal, 2: above normal, 3: well above normal |
| 8 | Glucose | 1: normal, 2: above normal, 3: well above normal |
| 9 | Smoking | Subjective features | Binary |
| 10 | Alcohol | Binary |
| 11 | Physical activity | Binary |
| 12 | Cardiovascular | Target | Presence or absence of CVD / target variable. |

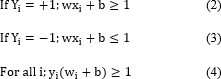
**3.2.Classification Methods-**

A description of the machine learning methods used in this research follows Five well-known classification models—Decision Tree, K-Nearest Neighbor, Logistic Regression, Naive Bayes, and Support Vector Machine—have been developed, and their prediction accuracy has been compared.. Numerous research contrasted data mining techniques using various parameter settings. The majority of these earlier research concluded that these techniques outperformed their statistical equivalents because they were less restricted by presumptions and produced better categorization outcomes. A few of these techniques are covered in brief.

**3.2.1.Linear Regression -** One of the most popular machine learning techniques for assessing multivariate regression problems in the medical field is logistic regression (LR). Using a continuous independent variable that aids in both the diagnosis and prediction of illnesses, LR is used to anticipate the outcome of a dependent variable. It is a method for discriminating between categories that use the input vector to extract important statistical data points from the model or forecast data trends. In the LR, the dependent variable is a binary variable that only accepts data that is coded as 0 (yes, success, etc.) or 1. (no, failure, etc.). Calculating the log chances of an event is the basic goal of an LR analysis. As shown mathematically, LR calculates multiple linear regression functions as follows:



**3.2.2.Support Vector Machine**: SVM was first introduced in 1992 by Boser, Guyon, and Vapnik as a binary classification method. Since then, a number of classification and regression problems have been added to its list of uses. Without adding prior knowledge, its generalization performance makes it a good classifier even with a lot of data. Data with separable classes may be classified by determining the optimum hyperplane that maximizes the margin between classified classes. SVM is modeled as a set of finite-dimensional vector spaces, where each dimension represents a distinct characteristic of an item. The technique for dealing with high-dimensional space issues has proven to be effective. Due to its processing efficiency on huge datasets, SVM has demonstrated great performance for illness prediction in the medical field in recent years. The main goal is to reduce generalization mistakes and develop it as a supervised learning system for regression and classification applications.

The SVM is mathematically expressed as:

The equation shows "w" as a weight and a vector, and "X" as a vector point. Therefore, the data in (3) and (4) must be constantly larger than zero and below zero, respectively, in order to distinguish the data in (2). SVM selects the hyperplane with the greatest distance between it and all other potential hyperplanes.

**3.2.3.K-Nearest Neighbor** - The K-Nearest Neighbor (K-NN) technique classifies cases based on how closely they resemble one another. When a case is new at a certain location, its distance from each model case—which is calculated as the nearest neighbor and is the most comparable to the approach that suggests the case—indicates the case. In this manner, the case is added to the output of neighbors that are closest to it. A new input class label is predicted by the K-NN algorithm, which bases its prediction on how similar the new input is to samples of its input from the training set. The K-NN classification output is poor if the new input is identical to the training set's samples.

The following mathematical formula calculates the Euclidean distance between two points, where p:x\*x=R is a function that yields the distance between the two locations x(xx, x′).



**3.2.4.Decision Tree** : A supervised learning strategy called a DT is applied mostly to classification-related problems. DT classifies the data using decision rules derived from the training data by assessing the entropy and gain of information. a classification tree structure where each node represents a property The root node will be the primary node, followed by the child nodes. The leaf nodes then represent the decision's result. It works well with both categorical and continuous characteristics. With DT, the population is split into two or more groups based on important predictors. The entropy for each feature is computed as the first step in DT. The dataset was then divided depending on the variables and predictors, either with high data gain or low entropy. The two phases are followed by the remaining qualities, as mentioned.



While "l" referred to a response variable module count, "qk" is the ratio of the count of the kth class procedures to the overall count of models.

**3.3.Evaluation Method -** To determine the effectiveness of the algorithm, specific metrics such as the f1 score, precision, recall, and accuracy are utilized as the basis for the assessment criteria. We also keep track of how long each algorithm takes to train. The classification algorithm's output is displayed in the confusion matrix, and this information forms the basis for further parameter calculations. As a consequence, by comparing predicted values with actual values, the confusion metrics may analyze a model's correctness and establish whether a classification algorithm commonly assigns items the erroneous labels. The parameters below give a quick description of the values for the confusion matrix and its representation as seen in Table 2.

True positive (TP) circumstances occur when the datapoint's actual class and projected class are both 1.

* False positive (FP) scenarios occur when a data point's expected class is 0 but its actual class is 0.
* False negative (FN) scenarios occur when a data point's real class is 0 and its anticipated class is 1.
* True negative (TN) circumstances occur when a data point's real class is 1 and its anticipated class is 0, respectively.

Table II- Confusion Matrix

|  |  |  |  |
| --- | --- | --- | --- |
|  | | **Actual value** | |
| ***Classified as absence*** | ***Classified as presence*** |
| **Predicted value** | *Absence* | TP | FN |
| *Presence* | FP | TN |

**3.3.1.F1 score**

When an f1-score achieves its highest value at 1 and its poorest score at 0, it is a function that is understood as a weight of recall average and precision. The f1-score formula is as follows:



**3.3.2.Recall**

Recall measures how much pertinent data is recovered from any machine learning system. The capacity to locate all connected events in the data is the main focus. The recall is represented by the equation below:



**3.3.3.Precision**

Being precise means being exact and correct. Precision conveys the sense of incidents that were accurately expected. It measures the proportion of genuine positives among all positives and quantifies forecasts that fall into the positive category as follows:



**3.3.4.Accuracy**

An important metric for describing an algorithm's performance is accuracy. It establishes the threshold at which an algorithm can accurately forecast both positive and negative instances and is quantified by the following formula:

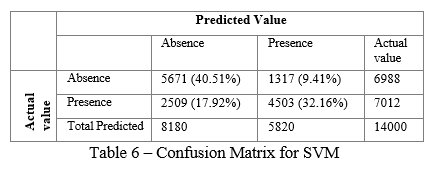
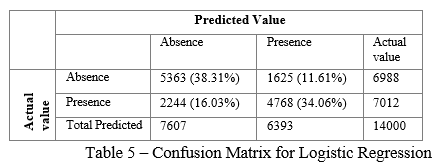
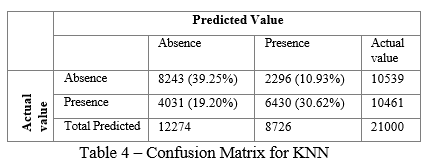
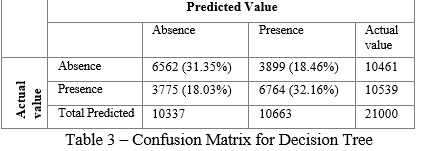


**3.4.Result of the algorithm –**

Five artificial intelligence techniques were used in this work to classify the cardiovascular disease dataset: decision trees (DT), logistic regressions (LR), K-nearest neighbors (KNN), and support vector machines (SVM). The dataset consists of 70,000 samples and 11 attributes. The classes can be examined both with and without the disease.

35021 of the data samples are marked as missing, whereas 334979 are marked as suffering from cardiovascular illness. Training and testing sets are split up into the data in the ratio of 70:30.

The confusion matrix of the DT, LR, KNN, and SVM prediction results is shown in Tables 1 through 5.



Hence the best-performing algorithm according to the Confusion Matrix, F1 scores and precision is examined and selected for identifying the CVD in an individual.

**4.Deep Learning Methods applied to ECG**

One of the key objectives for the use of AI to ECG interpretation is the creation of thorough, human-like interpretation skills. Since the development of the digital ECG more than 60 years ago, efforts have been made to give a computer-generated interpretation of the ECG that is rapid, accurate, and comprehensive. The problem seems to be solved since ECG interpretation is a fairly constrained application of pattern recognition to a small dataset. Early digital ECG interpretation software was capable of quickly identifying fiducial sites, taking precise measurements, and defining common measurable anomalies. Modern technology has advanced past these rule-based methods to find patterns in vast amounts of labeled ECG data.

Several organizations have created AI-driven algorithms, and some of these algorithms are already being deployed in a few confined clinical contexts. Many single-lead ECG datasets have been used in research to create CNNs, which have subsequently been applied to 12-lead ECGs. For instance, one team used a CNN to sort 2 million tagged single-lead ECG recordings from the Clinical Outcomes in Digital Electrocardiology study into six distinct types of abnormalities on the 12-lead ECG. This study demonstrated the viability of the approach, but its widespread use or external validation in other 12-lead ECG datasets is still to come. In related research using single-lead ECGs, another team found that CNN’s could make some diagnoses more accurately than practicing cardiologists. It remains to be seen, nevertheless, if this method will result in 12-lead ECG interpretation software that is therapeutically relevant. A CNN was created for the multilabel identification of 21 different heart rhythms based on the 12-lead ECG in an assessment that was published in 2020 utilizing a training and validation dataset of >80,000 ECGs from >70,000 patients. The cardiologists' committee's consensus labels served as the reference standard. The best network performed much better than a single cardiologist interpretation in a test dataset of 828 ECGs, matching the gold standard labels in 80% of the ECGs. The model's sensitivity, specificity, and mean area under the curve (AUC) receiver operating characteristic scores were 98%, 87%, and 99%, respectively.

The team of researchers developed a thorough ECG-interpretation infrastructure using their own dataset of >8 million ECGs done for clinical purposes (all of which have been labeled by experienced ECG readers and are connected to the relevant electronic health record). They showed that a CNN has good diagnostic performance and can recognize 66 distinct codes or diagnosis labels. Recently, Researchers created a unique technique that translates ECG characteristics into ECG codes and text strings using a transformer network and a CNN to extract ECG features. By providing information in a same manner and using comparable language, this approach produces a model output that is more like that of a human ECG reader. It also makes sense of related codes, , preventing the display of opposing or mutually conflicting interpretations that a human reader would not present. This approach will be especially important as our reliance on ECG data collected by cutting-edge, consumer-facing apps that are greatly scalable increases. For instance, AI-ECG algorithms have been applied to single-lead ECG traces obtained from mobile, smartwatch-enabled recordings for the detection of AF. The democratisation of ECG technology will cause the volume of signals that need to be interpreted to increase quickly, maybe faster than the rate at which human ECG readers can handle them. These autonomous, consumer- or patient-facing models are projected to be essential for telehealth technology. They could also make it possible to build essential lab spaces with the capacity to store and process massive quantities of data.

The signal quality produced with these devices can vary, as seen in the wristwatch research cited above, and AI-ECG may be less able than human expert over-readers to classify the heart rhythm utilising inferior tracings. Similar to this, another study discovered that a deep neural network built using ECG recordings from smartwatches performed well for passive detection of atrial fibrillation (AF) in comparison to AF diagnosed from 12-lead ECGs, but that performance was noticeably less reliable when referencing a self-reported history of persistent AF.

Nevertheless, despite significant advancements achieved, a full, human-like ECG interpretation package is still a long way off. Even in its most advanced form, the package doesn't have the precision required for execution without human supervision. The interpretation of an ECG generated by a machine also has the potential to affect human overreaders and, if erroneous, can be a cause of bias or systemic mistake. This worry is especially important if the algorithms were developed in populations that are different from the populations where they are used.

This flaw emphasises the need for a diverse derivation sample, rigorous external validation studies, phased implementation, and ongoing model performance and effectiveness evaluations (ideally including a diverse patient population and diverse means of data collection that reflect real-world practices).

**5.Deep Learning Methods applied to Phonocardiogram**

The first fundamental analytic technique used to assess the heart's functioning status is cardiac auscultation. If the heart sound from the Phonocardiogram (PCG) test indicates any abnormalities, an electrocardiography (ECG) test is required. The ability to hear and see the heart auscultation on the screen increases confidence in the precision of the first diagnosis. However, the diagnosis could not be as precise as expected due to noise and human error. If an artificial intelligence computer is utilized to generate potential diagnoses utilizing some distinguishing characteristics of the heart sound waves, diagnosis accuracy can be significantly improved. This method should lower death rates and healthcare expenses.

Numerous articles have put forth various methods for deriving characteristics from heart sounds and categorizing them using neural networks. Mohamed and Raafat created a mathematical model in the late 80s to use a limited set of parameters to characterize cardiac murmurs and noises. In this instance, features were retrieved using a fourth-order linear prediction of the cardiac cycle frames, and classification was done using the minimal difference between the features of the observed pattern and the reference patterns.

An intelligent heart attack prediction system based on data mining and artificial neural networks was suggested by Patil and Kumaraswamy. By applying the K-means clustering algorithm to the supplied data, this technique computes the parameters crucial to the heart attack. The Maximal Frequent Itemset Algorithm is used to extract these common patterns from the data (MAFIA). Following that, the designs are chosen based on the calculated significant weightage. Although the aforementioned study claimed that this technology may detect heart attacks using the MAFIA algorithm, the prediction accuracy for the work was not stated. Additionally, rather than using feature characteristics of the heart sound signal, this method makes use of features that correspond to the subject's behavioral patterns, such as drinking and smoking.

The GAL (Grow and Learn) algorithm is a revolutionary technique for segmenting heart sounds utilizing homomorphic filtering and feature extraction from wavelet coefficients. This method's accuracy was estimated to be 90.9%. In their study on the analysis of heart sounds for symptom identification, Reed et al. used wavelet decomposition to segment and alter the heart sounds. By eliminating levels with the shortest scales, the altered vectors were condensed into lower vector sizes. A three-layer neural network was used to classify each vector, and it provided 100% accuracy for all heart sounds. The drawback of this method is the requirement for using several hidden layer neurons—up to 50 layers.

Arrhythmia categorization based on heart rate variability (HRV) has been documented. Both the General Discriminant Analysis (GDA) and the Multi Layer Perceptron (MLP) methods are the foundations of this approach. The outcomes showed that this approach produced 100% accuracy for the data the authors obtained from the MIT-BIH database. But instead of PCG signals, this approach employs HRV signals based on the ECG. It should be highlighted that getting an ECG signal does not qualify as a regular test for primary care doctors since it necessitates laboratory setups, which takes time and is less efficient financially.

**6.Conclusion –** Cardiovascular Diseases are a threat that needs to be reduced. Traditional methods of identifying CVD need a lot of human interaction and are expensive hence there is a need for AI-augmented devices which can accurately predict the abnormalities of the heart. Applications for data mining are frequently utilised in the medical field to identify disorders and provide patients with a heart disease diagnosis based on their medical records. The most effective machine learning methods for categorizing cardiovascular illness have been established using patient data, and we have explored these methods in this work. Based on evaluation metrics including precision, recall, f1-score, accuracy, and training time, the various classification algorithms SVM, KNN, DT, LR, and NB have been contrasted.

We want to improve the effectiveness of these fundamental categorization algorithms in the future by creating a meta-model that will be applied to predicting cardiovascular disease in those at risk for heart disease.

Further, we discuss the advancement of AI and ML on the ECG and PCG devices in the literature review and discuss the contribution of various researchers in the advancement of automated detection of CVD.

**7. References –**

1.P. Keikhosrokiani, Perspectives in the development of mobile medical information systems: Life cycle, management, methodological approach and application. 2019. 2. WHO, “Cardiovascular diseases (CVDs): key facts,” World Health Organization, 2017.

3. P. Keikhosrokiani, N. Mustaffa, and N. Zakaria, “Success factors in developing iHeart as a patient-centric healthcare system: A multigroup analysis,” Telematics and Informatics, vol. 35, no. 4, 2018, doi: 10.1016/j.tele.2017.11.006.

4. P. Keikhosrokiani, “Chapter 5 - Success factors of mobile medical information system (mMIS),” P. B. T.-P. in the D. of M. M. I. S. Keikhosrokiani, Ed. Academic Press, 2020, pp. 75–99.

5. P. Keikhosrokiani, N. Mustaffa, N. Zakaria, and R. Abdullah, “Assessment of a medical information system: the mediating role of use and user satisfaction on the success of human interaction with the mobile healthcare system (iHeart),” Cognition, Technology & Work, vol. 22, no. 2, pp. 281–305, 2020, doi: 10.1007/s10111-019-00565-4.

6. A. D’Souza, “Heart disease prediction using data mining techniques,” International Journal of Research in Engineering and Science (IJRES) ISSN (Online), pp. 2320–9364, 2015.

7. J. Patel, D. TejalUpadhyay, and S. Patel, “Heart disease prediction using machine learning and data mining technique,” Heart Disease, vol. 7, no. 1, pp. 129–137, 2015.

8. I. Ayon, M. M. Islam, and M. R. Hossain, “Coronary artery heart disease prediction: a comparative study of computational intelligence techniques,” IETE Journal of Research, pp. 1–20, 2020.

9. S. S. Tripathy. “System for diagnosing valvular heart disease using heart sounds”, Master’s Thesis, India, 2005

10. A. S. A. Mohamed and H. M. Raafat. “Recognition of heart sounds and murmurs for cardiac diagnosis”. In Proceedings of 9th International Conference on Pattern Recognition, 1988.

11. W. M. Jinjri, P. Keikhosrokiani and N. L. Abdullah, "Machine Learning Algorithms for The Classification of Cardiovascular Disease- A Comparative Study," 2021 International Conference on Information Technology (ICIT), 2021, pp. 132-138, doi: 10.1109/ICIT52682.2021.9491677.

12. I. Abdel-Motaleb and R. Akula, "Artificial intelligence algorithm for heart disease diagnosis using Phonocardiogram signals," 2012 IEEE International Conference on Electro/Information Technology, 2012, pp. 1-6, doi: 10.1109/EIT.2012.6220714.

13. Zaibunnisa L. H. Malik, Momin Fatema, Nikam Pooja, Gawandar Ankita, 2021, Heart Disease Prediction using Artificial Intelligence, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NREST – 2021 (Volume 09 – Issue 04),

14. Javed Azmi, Muhammad Arif, Md Tabrez Nafis, M. Afshar Alam, Safdar Tanweer, Guojun Wang, A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data, Medical Engineering & Physics, Volume 105, 2022, 103825, ISSN 1350-4533,

15. Safial Islam Ayon, Md. Milon Islam & Md. Rahat Hossain (2022) Coronary Artery Heart Disease Prediction: A Comparative Study of Computational Intelligence Techniques, IETE Journal of Research, 68:4, 2488-2507,

16. Siontis KC, Noseworthy PA, Attia ZI, Friedman PA. Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. Nat Rev Cardiol. 2021 Jul;18(7):465-478. doi: 10.1038/s41569-020-00503-2. Epub 2021 Feb 1. PMID: 33526938; PMCID: PMC7848866.

17. Siontis, K.C., Noseworthy, P.A., Attia, Z.I. *et al.* Artificial intelligence-enhanced electrocardiography in cardiovascular disease management. *Nat Rev Cardiol* **18**, 465–478 (2021).

18. Zaibunnisa L. H. Malik, Momin Fatema, Nikam Pooja, Gawandar Ankita, 2021, Heart Disease Prediction using Artificial Intelligence, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) NREST – 2021 (Volume 09 – Issue 04),