International Journal of Scientific Research in Computer Science, Engineering and Information Technology
ISSN : 2456-3307 (www.ijsrcseit.com)
doi : https://doi.org/10.32628/IJSRCSEIT

# A Review on Heart Disease Detection Using Machine Learning

**Neha Madame[1], Prof. Mashhood Siddiqui[2]**
*[1]Computer Science, Bhabha Engineering Research Institute, Bhabha University, Bhopal, Madhya Pradesh, India
[2]Computer Science, Bhabha Engineering Research Institute, Bhabha University, Bhopal, Madhya Pradesh, India

**Article Info**

**Publication Issue :**

Volume 8, Issue 4
July-August-2022

**Page Number :** 206-214

**Article History**
Accepted: 10 July 2022
Published: 30 July 2022

## ABSTRACT

The successful application of data mining in highly visible fields like e-business, marketing, and retail has led to its application in other industries and sectors. Among these sectors just discovered is healthcare. The healthcare environment is still „information rich" but „knowledge poor". There is a wealth of data available within the healthcare systems. However, there is a lack of effective analysis tools to discover hidden relationships and trends in data. This research paper intends to provide a survey of current techniques of knowledge discovery in databases using data mining techniques that are in use in today"s medical research, particularly in Heart Disease Prediction. Number of experiments has been conducted to compare the performance of predictive data mining technique on the same dataset and the outcome reveals that the Decision Tree outperforms and sometime Bayesian classification is having similar accuracy as of decision tree but other predictive methods like KNN, Neural Networks, Classification based on clustering are not performing well. The second conclusion is that the accuracy of the Decision Tree and Bayesian Classification further improves after applying a genetic algorithm to reduce the actual data size to get the optimal subset of attribute sufficient for heart disease prediction.

**Keywords :** Data mining, Machine learning, Neural Network

## I. INTRODUCTION

It is difficult to identify heart disease because of several contributory risk factors such as diabetes, high blood pressure, high cholesterol, abnormal pulse rate and many other factors. Various techniques in data mining and neural networks have been employed to and out the severity of heart disease among humans. The severity of the disease is classi_ed based on various methods like K-Nearest Neighbor Algorithm (KNN), Decision Trees (DT), Genetic algorithm (GA), and Naïve Bayes (NB) [1], [2]. The nature of heart disease is complex and hence, the disease must be handled carefully. Not doing so may affect the heart or cause premature death. The perspective of medical science and data mining are used for discovering various sorts of metabolic syndromes. Data mining

with classifycation plays a significant role in the prediction of heart disease and data investigation.

Figure 1 depicts the parts of the human heart such as Left atrium, Right atrium, Right ventricle, Left ventricle, Aorta, pulmonary vein, Pulmonary valve, Pulmonary artery, Tricuspid valve, Aortic valve, Mitral valve, Superior vena cava and Interior vena cava.

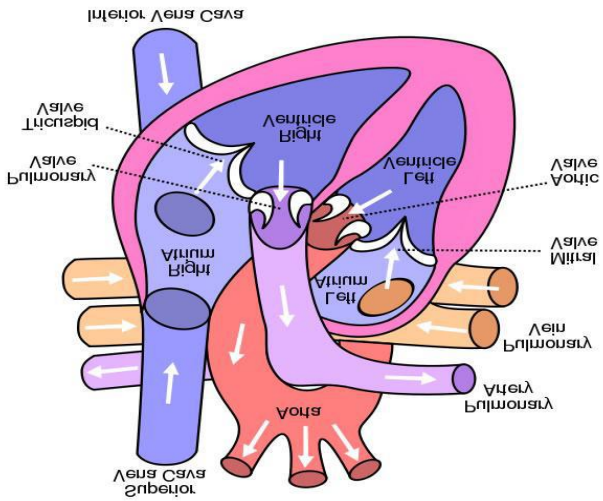


Figure 1 : Human heart[25],[26]

Neural networks are generally regarded as the best tool for prediction of diseases like heart disease and brain disease. The proposed method which we use has 13 attributes for heart disease prediction. The results show an enhanced level of performance compared to the existing methods in works like [3]. Carotid Artery Stenting (CAS) has also become a prevalent treatment mode in the medical _eld during these recent years. The CAS prompts the occurrence of major adverse cardiovascular events (MACE) of heart disease patients that areelderly. Their evaluation becomes very important. We generate results using a Arti_cial Neural Network ANN, which produces good performance in the prediction of heart disease[3], [4]. Neural network methods are introduced, which combine not only posterior probabilities but also predicted values from multiple predecessor techniques. This model achieves an accuracy level of up to 89:01% which is a strong result compared to previous works. For all experiments, the Cleveland heart dataset is used with a Neural Network NN to improve the performance of heart disease as we have seen previously in [5], [6].

Heart disease (HD) is the critical health issue and numerous people have been suffered by this disease around the world [7]. The HD occurs with common symptoms of breath shortness, physical body weakness and, feet are swollen [8]. Researchers try to come across an ef_cient technique for the detection of heart disease, as the current diagnosis techniques of heart disease are not much effective in early time identi_cation due to several reasons, such as accuracy and execution time [9]. The diagnosis and treatment of heart disease is extremely difficult when modern technology and medical experts are not available [10]. The effective diagnosis and proper treatment can save the lives of many people [11]. According to the European Society of Cardiology, 26 million approximately people of HD were diagnosed and diagnosed 3.6 million annually [12]. Most of people in the United States are suffering from heart disease [13]. Diagnosis of HD is traditionally done by the analysis of the medical history of the patient, physical examination report and analysis of concerned symptoms by a physician. But the results obtained from this diagnosis method are not accurate in identifying the patient of HD. Moreover, it is expensive and computationally difficult to analyze [14]. Thus, to develop a non invasive diagnosis system based on classifiers of machine learning (ML) to resolve these issues. Expert decision system based on machine learning classifiers and the application of artificial fuzzy logic is effectively diagnosis the HD as a result, the ratio of death decreases [15] and [16] and [26] . The Cleveland heart disease data set was used by various researchers for the identification problem of HD. The machine learning predictive models need proper data for training and testing. The performance of machine learning model can be increased if balanced dataset is use for training and testing of the model. Furthermore, the model predictive capabilities can improve by using proper and related features from the data. Therefore, data balancing and feature

selection is significantly important for model performance improvement. In literature various diagnosis techniques have been proposed by various researchers, however, these techniques are not effectively diagnosis HD. In order to improve the predictive capability of machine learning model data preprocessing is important for data standardization. Various Preprocessing techniques such removal of missing feature value instances from the dataset, Standard Scalar (SS), Min-Max Scalar etc. The feature extraction and selection techniques are also improve model performance. Various feature selection techniques= are mostly used for important feature selection such as, Least-absolute-shrinkage-selection-operator (LASSO), Relief, Minimal-Redundancy-Maximal-Relevance (MRMR), Local-learning-based-features-selection (LLBFS), Principle component Analysis (PCA), Greedy Algorithm (GA), and optimization methods, such as Anty Conley Optimization (ACO), fruit _y optimization (FFO), Bacterial Foraging Optimization (BFO) etc. Similarly Yun et al. [13] presented different techniques for different type of feature selection, such as feature selection for high-dimensional small sample size data, large-scale data, and secure feature selection. They also discussed some important topics for feature selection have emerged, such as stable feature selection, multi-view feature selection, distributed feature selection, multi-label feature selection, online feature selection, and adversarial feature selection. Jundong et al. [14] discussed the challenges of feature selection (FS) for big data. It is necessary to decrease the dimensionality of data for various learning tasks due to the curse of dimensionality. Feature selection has great influence in numerous applications such as building simpler, increasing learning performance, and creating clean and understandable data. The feature selection from big data is a challenging job creates the big problems because big data has many dimensions. Further, challenges of feature selection for structured, heterogeneous and streaming data as well as its scalability and stability issues. For big data analytics challenges of feature selection is very important to resolved. In [15] designed unsupervised hashing scheme, called topic hypergraphh hashing, to report the limitations. Topic hypergraph hashing effectively mitigates the semantic shortage of hashing codes by exploiting auxiliary texts around images. The proposed Topic hyper graph hashing can achieve superior performance equaled with numerous state-of-the art approaches, and it is more appropriate for mobile image retrieval. The feature selection algorithms are classified into three type such as alter based, wrapper based and embedded based. All these feature selection mechanisms have some advantages and limitations in certain cases. The alter based method measures the relevance of a feature by correlation with the dependent variable while the wrapper feature selection algorithm measure the usefulness of a subset of features by actually training the classifier on it. The allter method is less computationally complex than wrapper method. Thefeature set selected by the _lter is general and can be appliedto any model and it is independent of a speci_c model. In feature selection global relevance is of greater importance. On another hand suitable machine learning model is necessary for good results. Obviously, a good machine learning model is a model that not only performs well on data seen during training (else a machine learning model could simply learn the training data), but also on unseen data. To evaluate all classifiers on data and that they get, on average,50% of the cases right [16]. Furthermore, appropriate cross validation techniques and performance evaluation metrics arecritical necessary for a model when model is train and test on dataset.

Heart disease is one of the prevalent disease that can lead to reduce the lifespan of human beings nowadays. Each year 17.5 million people are dying due to heart disease. Life is dependent on component functioning of heart, because heart is necessary part of our body. Heart disease is a disease that affects on the function of heart An estimate of a person's risk for coronary heart disease is important for many aspects of health

promotion and clinical medicine. A risk prediction model may be obtained through multivariate regression analysis of a longitudinal study [3]. Due to digital technologies are rapidly growing, healthcare centres store huge amount of data in their database that is very complex and challenging to analysis. Data mining techniques and machine learning algorithms play vital roles in analysis of different data in medical centres. The techniques and algorithms can be directly used on a dataset for creating some models or to draw vital conclusions, and inferences from the dataset. Common attributes used for heart disease are Age, Sex, Fasting Blood Pressure, Chest Pain type, Resting ECG(test that measures the electrical activity of the heart), Number of major vessels colored by fluoroscopy, Threst Blood Pressure (high blood pressure), Serum Cholestrol (determine the risk for developing heart disease), Thalach (maximum heart rate achieved), ST depression (finding on an electrocardiogram, trace in the ST segment is abnormally low below the baseline), painloc (chest pain location (substernal=1, otherwise=0)), Fasting blood sugar, Exang (exercise included angina), smoke, Hypertension, Food habits, weight, height and obesity[4]. Table 1 summarizes the most common types of the heart disease as follows.

Table 1: Different types of heart diseases

| | |
|---|---|
| Arrhythmia | The heart beat improper whether it may irregular, too slow or too fast |
| Cardiac arrest | An unexpected loss of heart function, consciousness and breathing occur suddenly |
| Congestive heart failure | The heart does not pump blood as well as it should, it is the condition of chronic. |
| Congenital heart disease | The heart abnormally which develops before birth |
| High blood pressure | It has a condition that the force of blood against the artery wall is too high |
| Peripheral artery diseases | The narrow blood vessels which reduce flow of blood in the limbs is the circulatory condition |
| Stroke | Interruption of blood supply occurs damage to the brain |

This paper is organized as follows. Section 2 gives an overall literature review of the existing work. Section provide challenges of prediction of heart diseases using machine learning Section 4 provides a conclusion and future work.

## II. LITERATURE REVIEW

Heart disease is one of the complex diseases and globally many people suffered from this disease. On time and efficient identification of heart disease plays a key role in healthcare, particularly in the field of cardiology. Ping et al [25] an efficient and accurate system to diagnosis heart disease and the system is based on machine learning techniques. The system is developed based on classification algorithms includes Support vector machine, Logistic regression, Artificial neural network, K-nearest neighbor, Naïve bays, and Decision tree while standard features selection algorithms have been used such as Relief, Minimal redundancy maximal relevance, Least absolute shrinkage selection operator and Local learning for removing irrelevant and redundant features. We also proposed novel fast conditional mutual information feature selection algorithm to solve feature selection problem. The features selection algorithms are used for features selection to increase the classification accuracy and reduce the execution time of classification system. Furthermore, the leave one subject out cross-validation method has been used for learning the best practices of model assessment and for hyperparameter tuning. The performance measuring metrics are used for assessment of the performances of the classifiers. The performances of the classifiers have been checked on the selected

features as selected by features selection algorithms. The experimental results show that the proposed feature selection algorithm (FCMIM) is feasible with classifier support vector machine for designing a high-level intelligent system to identify heart disease. The suggested diagnosis system (FCMIM-SVM) achieved good accuracy as compared to previously proposed methods. Additionally, the proposed system can easily be implemented in healthcare for the identification of heart disease.

S. Seema et al,[23] focuses on techniques that can predict chronic disease by mining the data containing in historical health records using Naïve Bayes, Decision tree, Support Vector Machine(SVM) and Artificial Neural Network(ANN). A comparative study is performed on classifiers to measure the better performance on an accurate rate. From this experiment, SVM gives highest accuracy rate, whereas for diabetes Naïve Bayes gives the highest accuracy.

Noura Ajam [22] recommended artificial neural network for heart disease diagnosis. Based on their ability, Feed forward Back propogation learning algorithms have used to test the model. By considering appropriate function, classification accuracy reached to 88% and 20 neurons in hidden layer. ANN shows result significantly for heart disease prediction.

Chala Beyene et al, [24] recommended Prediction and Analysis the occurrence of Heart Disease Using Data Mining Techniques. The main objective is to predict the occurrence of heart disease for early automatic diagnosis of the disease within result in short time. The proposed methodology is also critical in healthcare organisation with experts that have no more knowledge and skill. It uses different medical attributes such as blood sugar and heart rate, age, sex are some of the attributes are included to identify if the person has heart disease or not. Analyses of dataset are computed using WEKA software.

Abdullah et al [19] introduced an effective method based on the Synthetic Minority Oversampling Technique (SMOTE) to handle imbalance distribution issue, six different ML classifiers to detect the patient status, and Hyperparameter Optimization (HPO) to find the best hyperparameter for ML classifier together with SMOTE. Two public datasets were used to build and test the model using all features. The results show that SMOTE and Extra Trees (ET) optimized using hyperband achieved higher results than other models and outperformed the state-of-the-art works by achieving 99.2% and 98.52% in CVD detection, respectively. Also, the developed model converged to 95.73% severity classification using the Cleveland dataset.

Perk et al [20] propose a system that can predict and semantically interpret stroke prognostic symptoms based on machine learning using the multi-modal bio-signals of electrocardiogram (ECG) and photoplethysmography (PPG) measured in real-time for the elderly. To predict stroke disease in real-time while walking, we designed and implemented a stroke disease prediction system with an ensemble structure that combines CNN and LSTM. The proposed system considers the convenience of wearing the bio-signal sensors for the elderly, and the bio-signals were collected at a sampling rate of 1,000Hz per second from the three electrodes of the ECG and the index finger for PPG while walking. According to the experimental results, C4.5 decision tree showed a prediction accuracy of 91.56% while RandomForest showed a prediction accuracy of 97.51% during walking by the elderly.

Ahemad et al [21] proposes various Machine Learning algorithms such as LR, KNN, SVM, and GBC, together with the GridSearchCV, predict cardiac disease. The system uses a 5-fold cross-validation technique for verification. A comparative study is given for these four methodologies. The Datasets for both Cleveland, Hungary, Switzerland, and Long Beach V and UCI Kaggle are used to analyze the models' performance. It is found in the analysis that the Extreme Gradient Boosting Classifier with GridSearchCV gives the highest and nearly

comparable testing and training accuracies as 100% and 99.03% for both the datasets (Hungary, Switzerland &; Long Beach V and UCI Kaggle). Moreover, it is found in the analysis that XGBoost Classifier without GridSearchCV gives the highest and nearly comparable testing and training accuracies as 98.05% and 100% for both the datasets (Hungary, Switzerland & Long Beach V and UCI Kaggle).

**Table 2: A cooperative Literature Review of various algorithm for heart disease prediction**

| Year | Author | Mehtod | Technique | Accuracy |
|---|---|---|---|---|
| 2015 | Noura Ajam et al, [22 ] | Heart Disease Diagnoses using Artificial Neural Network. | ANN | 88% |
| 2016 | S. Seema et al, [23] | Chronic Disease Prediction by mining the data. | Naïve Bayes | Highest accuracy achieved by SVM, in the case of heart disease 95.56% |
| 2018 | Chala Bayen et al,[24] | Prediction and Analysis the occurrence of Heart Disease using data mining techniques. | Naïve Bayes, Support Vector Machine. | J48, It gives short time result which helps to give quality of services and reduce co |
| 2020 | Ping et al[25] | Heart Disease Identification Method Using Machine Learning Classificatio n in E-Healthcare | Support vector machine, Logistic regression, Artificial neural network, K-nearest neighbor, Naïve bays, and Decision tree | |
| 2022 | Abdallah et al[19] | An Effective Heart Disease Detection and Severity Level Classification Model Using Machine Learning and Hyperparameter Optimization Methods | Synthetic Minority Oversampling Technique (SMOTE) | 99.2% |
| 2022 | Park et al [20] | AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals | modal bio-signals of electrocardiogram (ECG) and photoplethysmography (PPG) | 4.5 decision tree showed a prediction accuracy of 91.56% while RandomForest showed a prediction accuracy of 97.51% during walking by the elderly |
| 2022 | Ahmead et al.[21] | G. N. Ahmad, H. Fatima, S. Ullah, A. Salah Saidi and | Extreme Gradient Boosting Classifier with | gives the highest and nearly comparab |

| | | | | |
|---|---|---|---|---|
| | | Imdadullah, "Efficient Medical Diagnosis of Human Heart Diseases Using Machine Learning Techniques With and Without GridSearch CV," in IEEE Access, vol. 10, pp. 80151-80173, 2022, doi: 10.1109/ACCESS.2022.3165792. | GridSearch CV | le testing and training accuracies as 98.05% and 100% |

## III. ISSUE AND CHALLENGES

Medical diagnosis is considered as a significant yet intricate task that needs to be carried out precisely and efficiently. The automation of the same would be highly beneficial. Clinical decisions are often made based on doctor‟s intuition and experience rather than on the knowledge rich data hidden in the database. This practice leads to unwanted biases, errors and excessive medical costs which affects the quality of service provided to patients. Data mining have the potential to generate a knowledge-rich environment which can help to significantly improve the quality of clinical decisions.

A major challenge faced by health care organizations, such as hospitals and medical centre's, is the provision of quality services at affordable costs.1 The quality service implies diagnosing patients properly and administering effective treatments. The available heart disease database consists of both numerical and categorical data. Before further processing, cleaning and filtering are applied on these records in order to filter the irrelevant data from the database.

Issue faced in data set

1. Prediction of cardiovascular disease results is not accurate. 2. Data mining techniques does not help to provide effective decision making.

3. Cannot handle enormous datasets for patient records.

.

## IV. CONCLUSION AND FUTURE WORK

It has been concluded that various data mining and machine learning methods can be used to forecast the occurrence of cardiac disease. Find out how well each algorithm predicts, then put the recommended system to work where it's needed. Better the accuracy of algorithms by using better feature selection techniques. Once a patient has been diagnosed with a certain form of heart disease, they can choose from a variety of treatment options. Mining this kind of data for insights can be really useful.

Based on the results of the literature review, it appears that relatively limited success has been made in developing a prediction model for people with heart disease improve the accuracy of early cardiac disease prediction, we need combinational and more complicated models. The more information that is added to the database, the smarter it will become. In order to make this prediction system more scalable and accurate, there are a wide variety of avenues that could be investigated. The following studies/tasks need to be done in the future, however time constraints prevent us from doing so now. We are interested in putting a number of discretization methods to the test, as well as the multiple classifiers voting technique and two varieties of decision trees (information gain and gain ratio). eager to test the waters with new approaches, such as the association rule, logistic regression, and clustering techniques.

Cite this article as :

## V. REFERENCES


[1] M. Durairaj and V. Revathi, ``Prediction of heart disease using back propagationMLPalgorithm,'' Int. J. Sci. Technol. Res., vol. 4, no. 8, pp. 235_239, 2015.

[2] A. Gavhane, G. Kokkula, I. Pandya, and K. Devadkar, ``Prediction of heart disease using machine learning,'' in Proc. 2nd Int. Conf. Electron.Commun. Aerosp. Technol. (ICECA), Mar. 2018, pp. 1275_1278.

[3] L. Baccour, ``Amended fused TOPSIS-VIKOR for classi_cation (ATOVIC) applied to some UCI data sets,'' Expert Syst. Appl., vol. 99,pp. 115_125, Jun. 2018. doi: 10.1016/j.eswa.2018.01.025.

[4] P. S. Kumar, D. Anand, V. U. Kumar, D. Bhattacharyya, and T.-H. Kim, ``A computational intelligence method for effective diagnosis of heart

disease using genetic algorithm,'' Int. J. Bio-Sci. BioATOMICol., vol. 8, no. 2, pp. 363_372, 2016.

[5] F. Dammak, L. Baccour, and A. M. Alimi, ``The impact of criterion weights techniques in TOPSIS method of multi-criteria decision making in crisp and intuitionistic fuzzy domains,'' in Proc. IEEE Int. Conf. Fuzzy Syst.(FUZZ-IEEE), vol. 9, Aug. 2015, pp. 1_8.

[6] M. J. Liberatore and R. L. Nydick, ``The analytic hierarchy process in medical and health care decision making:Aliterature review,'' Eur. J. Oper.Res., vol. 189, no. 1, pp. 194_207, 2008.

[7] A. L. Bui, T. B. Horwich, and G. C. Fonarow, ``Epidemiology and risk pro_le of heart failure,'' Nature Rev. Cardiol., vol. 8, no. 1, p. 30, 2011.

[8] M. Durairaj and N. Ramasamy, ``A comparison of the perceptive approaches for preprocessing the data set for predicting fertility successrate,'' Int. J. Control Theory Appl., vol. 9, no. 27, pp. 255_260, 2016.

[9] L. A. Allen, L.W. Stevenson, K. L. Grady, N. E. Goldstein, D. D. Matlock, R. M. Arnold, N. R. Cook, G. M. Felker, G. S. Francis, P. J. Hauptman, E. P. Havranek, H. M. Krumholz, D. Mancini, B. Riegel, and J. A. Spertus, ``Decision making in advanced heart failure: A scienti_c statement from the American heart association,'' Circulation, vol. 125, no. 15,pp. 1928_1952, 2012.

[10] S. Ghwanmeh, A. Mohammad, and A. Al-Ibrahim, ``Innovative arti_cial neural networks-based decision support system for heart diseases diagnosis,'' J. Intell. Learn. Syst. Appl., vol. 5, no. 3, 2013, Art. no. 35396.

[11] Q. K. Al-Shayea, ``Arti_cial neural networks in medical diagnosis,'' Int. J. Comput. Sci. Issues, vol. 8, no. 2, pp. 150_154, 2011. [6] J. Lopez-Sendon, ``The heart failure epidemic,'' Medicographia, vol. 33, no. 4, pp. 363_369, 2011.

[12] P. A. Heidenreich, J. G. Trogdon, O. A. Khavjou, J. Butler, K. Dracup, M. D. Ezekowitz, E. A. Finkelstein, Y. Hong, S. C. Johnston, A. Khera,D. M. Lloyd-Jones, S. A. Nelson, G. Nichol, D. Orenstein, P.W. F.Wilson, and Y. J. Woo, ``Forecasting the future of cardiovascular disease in theunited states: A policy statement from the American heart association,'' Circulation, vol. 123, no. 8, pp. 933_944, 2011.

[13] A. Tsanas, M. A. Little, P. E. McSharry, and L. O. Ramig, ``Nonlinear speech analysis algorithms mapped to a standard metric achieve clinicallyuseful quanti_cation of average Parkinson's disease symptom severity,'' J. Roy. Soc. Interface, vol. 8, no. 59, pp. 842_855, 2011.107580 VOLUME 8, 2020 J. P. Li et al.: HD Identification Method Using ML Classification in E-Healthcare

[14] S. I. Ansarullah and P. Kumar, ``A systematic literature review on cardiovascular disorder identi_cation using knowledge mining andmachine learning method,'' Int. J. Recent Technol. Eng., vol. 7, no. 6S, pp. 1009_1015, 2019.

[15] S. Nazir, S. Shahzad, S. Mahfooz, and M. Nazir, ``Fuzzy logic based decision support system for component security evaluation,'' Int. Arab J. Inf. Technol., vol. 15, no. 2, pp. 224_231, 2018.

[16] R. Detrano, A. Janosi,W. Steinbrunn, M. P_sterer, J.-J. Schmid, S. Sandhu, K. H. Guppy, S. Lee, and V. Froelicher, ``International application of a new probability algorithm for the diagnosis of coronary artery disease,'' Amer. J. Cardiol., vol. 64, no. 5, pp. 304_310, Aug. 1989.

[17] S. Mohan, C. Thirumalai and G. Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques," in IEEE Access, vol. 7, pp. 81542-81554, 2019, doi: 10.1109/ACCESS.2019.2923707.

[18] J. P. Li, A. U. Haq, S. U. Din, J. Khan, A. Khan and A. Saboor, "Heart Disease Identification Method Using Machine Learning Classification in E-Healthcare," in IEEE Access, vol. 8, pp. 107562-107582, 2020, doi: 10.1109/ACCESS.2020.3001149.

[19] A. Abdellatif, H. Abdellatef, J. Kanesan, C. -O. Chow, J. H. Chuah and H. M. Gheni, "An Effective Heart Disease Detection and Severity Level Classification Model Using Machine Learning and Hyperparameter Optimization Methods," in *IEEE Access*, vol. 10, pp. 79974-79985, 2022, doi: 10.1109/ACCESS.2022.3191669.

[20] J. Yu, S. Park, S. -H. Kwon, K. -H. Cho and H. Lee, "AI-Based Stroke Disease Prediction System Using ECG and PPG Bio-Signals," in IEEE Access, vol. 10, pp. 43623-43638, 2022, doi: 10.1109/ACCESS.2022.3169284.

[21] G. N. Ahmad, H. Fatima, S. Ullah, A. Salah Saidi and Imdadullah, "Efficient Medical Diagnosis of Human Heart Diseases Using Machine Learning Techniques With and Without GridSearchCV," in IEEE Access, vol. 10, pp. 80151-80173, 2022, doi: 10.1109/ACCESS.2022.3165792.

[22] Noura Ajam, "Heart Disease Diagnoses using Artificial Neural Network", The International Insitute of Science, Technology and Education, vol.5, No.4, 2015, pp.7-11

[23] Dr.S.Seema Shedole, Kumari Deepika, "Predictive analytics to prevent and control chronic disease", https://www.researchgate.net/punlication/316530782, January 2016.

[24] Mr. Chala Beyene, Prof. Pooja Kamat, "Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques", International Journal of Pure and Applied Mathematics, 2018.

[25] 25Nimai Chand Das Adhikari, Arpana Alka, and rajat Garg, "HPPS: Heart Problem Prediction System using Machine Learning".

[26] Muthuvel, Marimuthu & Abinaya, M & Hariesh, K & Madhankumar, K & Pavithra, V. (2018). A Review on Heart Disease Prediction using Machine Learning and Data Analytics Approach. International Journal of Computer Applications. 181. 975-8887. 10.5120/ijca2018917863.

**Cite this article as :**

Neha Madame, Prof. Mashhood Siddiqui, "A Review on Heart Disease Detection Using Machine Learning", International Journal of Scientific Research in Computer Science, Engineering and Information Technology (IJSRCSEIT), ISSN : 2456-3307, Volume 8 Issue 4, pp. 206-214, July-August 2022.
Journal URL : https://ijsrcseit.com/CSEIT228438