

A Survey on Early Heart Disease Detection using AI based Approach

Upasana Patil¹ and Prasath Nithiyanandam²

^{1,2}Dept. of Networking and Communication, SRM Institute of Science & Technology, SRM Nagar,
Kattankulathur, TamilNadu-603203, India

Email: up1715@srmist.edu.in, prasathn@srmist.edu.in

Abstract—With the development of machine intelligence and artificial intelligence, heart disease prediction models can be designed to avoid human errors, take precautionary measures and enhance the efficiency of detection in preliminary stages. Precise heart disease prediction can minimize mortality rate to a significant level and can provide clinical diagnosis support in the best possible manner. Certainly, these are the motivations to work in this domain and provide a detailed survey of heart disease prediction models. Several datasets and their attributes are analyzed to get the best possible insights related to heart disease. In this survey, several works of literature related to heart disease detection are taken based on machine learning, data mining and deep learning models from IEEE explore, Research Gate, PubMed, Google Scholar, Scopus, and Web of Science search engines. Most of the literature is recently published between 2015-2022. This literature survey identifies some important shortcomings of the state-of-art-heart disease prediction models considering various methodologies and datasets. Finally, a precise heart disease estimation system based on machine intelligence algorithms can become a genuine option for improving patient health.

Keywords— Machine Learning, Data Mining, Heart Disease Prediction, Classification, Deep Learning.

I. INTRODUCTION

The heart is one of the most important organs of the body and the function of the heart is to pump blood to other organs of the body. If the heart does not function properly, the other body parts also stopped functioning and ultimately can cause the death of suffered human being in just a few minutes. Thus, Heart Disease has become one of the most widespread diseases that significantly minimizes the lifespan of human beings. According to a report in [1], the mortality rate of human beings due to heart disease has risen up to 17.5 million per year across the world. According to World Health Organization (WHO) reports, over 60% of human mortalities caused due to heart disease across the globe [2]. Thus, heart failure has become a major and elaborated problem that needs to be handled appropriately and in the most effective manner. Therefore, one of the proper and genuine solutions put forward by many researchers and medical experts to handle heart disease problems is heart disease prediction in earlier stages. There are a few symptoms like chest pain, high blood pressure, hypertension, and cardiac arrest that can help to detect heart disease in the preliminary stage. Although many researchers and medical experts are working on a solution for the early detection of heart disease based on the various heart attributes and risk factors like cholesterol percentage, abnormal pulse rate, age, sex, hyper-tension, etc.

However, a heart failure estimation model can be built in an effective manner through deep regression analysis and longitudinal study on risk factors [3]. However, analysis of this amount of abundant data is a complex and challenging process. Thus, different automated and semi-automated methods have been presented for heart disease prediction by various researchers. Among them, Data Mining (DM) and Machine Learning (ML) techniques are the most dominant algorithms for the proper analysis of abundant medical data from different healthcare domains. These algorithms and techniques can generate massive insights, inferences, or conclusions from the analysis of medical datasets.

Datasets are built by various healthcare departments by gathering knowledge and data for varied health-related issues. Therefore, Machine-Learning (ML) techniques can be very useful for providing decent conclusions and insights to predict heart disease in a short time compared to any other algorithm or technique due to the fully automated process and capability of handling large datasets. The insights provided by ML algorithms are much more accurate and faster than DM techniques due to their automated process. DM techniques utilize data mining engines to analyze datasets and generate insights based on pattern estimation methods. However, ML algorithms work upon neural networks to form a predictive model and develop an automation process to generate insights and provide final conclusions. Thus, ML algorithms are more powerful and accurate than DM techniques [4].

Therefore, in this research survey paper, detailed information is gathered to design a proper and reliable heart disease prediction system. A comprehensive study is conducted by taking all possible dimensions regarding heart disease prediction systems namely research on available datasets, dataset attributes, previous heart disease estimation models, data mining techniques, tools, traditional machine learning methods, prediction classifiers, and performance metrics to measure system efficiency, and research challenges are taken into picture. The provided insights and knowledge can be advantageous for several research scholars of heart disease domain understanding and for future heart disease prediction systems and their relevant modules.

This paper is arranged in the following style. Section 2 discusses several available techniques for heart disease detection. Section 3 discusses the details of available datasets. Section 4, discusses the related work regarding the heart disease detection system and identifies issues related to heart disease detection. Section 5, describes the generalized methodology for the heart disease detection system. Section 6 discusses the simulation metrics and tools, and section 5 concludes the paper.

II. EASE OF USE HEART DISEASE PREDICTION TECHNIQUES:

A. Data Mining

Data Mining can be expressed as a knowledge-finding mechanism to analyze large data and convert their essential features into useful information. For a given dataset, finding the probability of disease presence is the main aim of this work, and data mining has been massively utilized over the years to achieve this objective. The main objectives of data mining are forecasting and descriptions. Data mining techniques require some attributes or variables to analyze the given dataset for precise prediction results. Here, the meaning of description is to find accurate patterns from the given dataset and interpret them into human understanding. Furthermore, DM is an advanced technique of information technology and is massively utilized in various domains like business, medicine, science, and technology. This technique generates meaningful information from the abundant information stored in the datasets and generates useful insights.

B. Machine Learning Algorithms

Machine Learning is a kind of automation process built on machine intelligence to link with the physical world. Artificial intelligence (AI) is subdivided into two important subsets namely Machine learning and Deep Learning algorithms. These algorithms are mainly focused on building a model for prediction and data analysis for generating important insights. Thus, both ML and DL algorithms are stronger and praiseworthy methods to generate meaningful information from raw data, especially in the case of medical data analytics. However, ML algorithms have an edge and are always considered an ideal method to predict heart-related disease. In addition, this technique can be useful for illness monitoring, prediction analysis, clinical management scenarios, etc. In another word, ML algorithms can be defined as the scientific discipline with the objective of learning data using an automation process through computers.

The main base of machine learning is a statistical analysis to discover insights from data using computing methods. This technique is capable of handling billions or trillions of data by designing a computational statistical model [5]. Furthermore, machine learning algorithms can be subdivided into three different categories namely, supervised learning, unsupervised learning, and Reinforcement Learning.

C. Deep Learning Algorithms

Deep learning algorithms are used to design some complex architectures and mathematical models. Heart disease prediction using DL algorithms mainly focuses on statistics and mathematical modeling to analyze large datasets. DL algorithms have gained high praise in the last few years in terms of designing complex deeper architectures and networks and have shown their worth in numerous applications like research, medicine, science, and technology.

III. HEART DISEASE DATASETS

A group of homogeneous data records builds a dataset. This section discusses various datasets that can be used to analyze heart-related attributes of patients and predict heart disease using machine learning algorithms. One of the most widely used and popular datasets related to heart disease analysis is the Cleveland heart disease data set [6]. This dataset is available at the University of California, Irvine (UCI) machine learning repository. This dataset consists of a total number of 303 data samples and each sample data has 76 attributes related to heart disease. The data values of 6 samples are missing from the available 303 samples.

Another widely used and popular dataset related to heart disease analysis is the Z-Alizadeh Sani dataset. This dataset consists of a total number of 303 data samples and each sample data has 55 input variables and a class variable related to heart disease. There are a total number of four groups that are associated with class variables namely normal, LCX, LAD, and RCA.

All these groups come under the coronary heart disease category. Some other datasets utilized by some of the researchers for the prediction of heart-related problems are Long Beach VA, Kaggle Framingham dataset, StatLog Heart, and Hungarian Heart Dataset.

StatLog Heart dataset consists of a total number of 270 data samples and each sample data has 13 attributes similar to the Cleveland heart disease data set. Hungarian Heart Dataset and Long Beach VA dataset are available at the University of California, Irvine (UCI) machine learning repository. Both datasets consist of a total number of 274 data samples and each sample data has 14 attributes related to heart disease same as the Cleveland dataset.

Whereas Kaggle Framingham dataset consists of a total number of 4240 data samples and each sample data has 16 attributes related to heart disease. This dataset is massive and contains demographic, medical, and behavioral risk factors

IV. LITERATURE SURVEY

Some of the widely used machine learning models are discussed below.

A. Naïve Bayes

Naïve Bayes is widely used to design a prediction model and it is part of machine learning algorithms and comes under supervised learning. Even further digging in the hierarchy, Naïve Bayes comes under the classification sub-category of supervised learning. By learning the dataset attributes and features a prediction model is built and the classifier provides the final output. It is one of the most adopted and fastest classification models which performs all the probabilities calculations based on the Bayes algorithm. This theorem works extremely well for large-volume databases such as the heart disease dataset. Each attribute of the heart disease data-base is treated as an independent attribute in this learning model. However, a strong relationship between all the attributes is maintained to get the desired output.

B. Decision Tree

A Decision Tree model is another widely used model to design a prediction mechanism and comes under supervised learning. The decision Tree model also comes under the classification sub-category of supervised learning. This classification model efficiently performs binary as well as multi-class classification using the available dataset attributes. This classifier model generates multiple similar subsets to learn the model and perform classification and takes varied combinations of attributes to understand attribute features and their significance, based on that decisions are given. Based on the data type of variables, the decision is made regarding binary classification or continuous variable classification. The decision Tree model is designed based on famous algorithms like ID3, C4.5, chi-square, and Gini Index. P Ghosh et al (2021) [7] studied machine learning algorithms and provided modifications to the existing Decision Tree model based on the bagging methods to get high efficiency for disease estimations. The main limitations of the Decision Tree model are over-fitting and weak continuous variable processing.

C. Neural Networks

Neural Networks are a popular machine-learning model and are used to understand the human brain or heart behavior by analyzing mul-ti-attribute large datasets. Neural Networks come under the regression sub-category of supervised learning and provide multiple facilities like back-propagation and forward propagation. Numerous ANN sub-networks are Hopfield Networks, Radi-al-Basis function networks, Single layer percep-tion networks, and multilayer perception net-works. In the case of heart disease prediction model designing, the most useful neural network is a multilayer perception network. The dataset attributes are used as input from the input layer and these layers are internally mapped with input layers. The output layer is used to obtain the final desired outputs. Some of the important recent literature are provided in the following para-graphs.

Lakshmana Rao et al (2019) [8] have present-ed a heart disease identification mechanism using machine learning algorithms and emphasized that external elements like diabetes, chain smokers, high cholesterol, and circulatory strain may en-hance the chances of heart failure and cause a difficult challenge to precisely predict the condi-tion of heart based on the given dataset. Several techniques and systems are designed in the last few years to study heart disease and the chances of heart failure using data mining and machine learning techniques and varied neural network architectures. The main idea is to make people aware of the seriousness of heart disease and the reasons which can elaborate the disease and ul-timately will take their lives. Determining solu-tions to deal with this issue. Thus, heart disease prediction in the initial stage is a priority. A ma-chine learning framework is introduced based on data analysis and past information from different models.

Mr. Santhana Krishnan.J and Dr. Geetha.S, (2019) [9] has introduced a data mining tech-nique to design a heart disease prediction system. Classification methods are used to estimate heart disease in male patients. Several attributes, fea-tures, facts, and types are studied related to heart disease analysis. The challenges and risk factors to designing an accurate heart disease prediction system are also studied. WEKA (Waikato Envi-ronment for Knowledge Analysis), a data mining tool for the application of Bioinformatics is used to analyze the given datasets and generate some important insights. There are three different kinds of interfaces are employed in WEKA namely, Artificial Neural Networks, Naive Bayes, and Decision Trees. Using these mentioned methods, heart disease prediction is performed. The decision Tree is the main technique used in this work for predicting heart disease which con-sists of some algorithms namely, C4.5, CHAID, J48, and ID3.

Avinash Golande et al (2019) [10] introduced a heart disease estimation technique based on machine learning algorithms and in conjunction, a few data mining methods are also explored to understand the insights of heart disease and sup-port the heart specialists. The adopted classifiers are KNN, Naïve Bayes, and Decision Tree. Some more technologies are explored in this work for a better understanding namely, packing calcula-tion, neural networks, Part thickness, repeated streamlining, and self-arranging kernel guide. High-performance results are obtained in this work.

Rishabh Magar et al. (2020) [11] presented a disease prediction model via a web application based on the machine learning model. The model is tested on the dataset from the UCI repository with a training-testing ratio of 75%-25%. The performance results are evaluated in terms of classification accuracy. Different models are tested on the given dataset and different perfor-mance results are observed. The logistic regres-sion model provides the best accuracy results with 82.89%, SVM provides 81.57% accuracy, and Naive Bayes provides the same accuracy as the Decision Tree model at 80.43%. This applica-tion can be widely utilized for checking the heart condition of patients as a preliminary test.

Apurb Rajdhan et al. (2020) [12] introduces a system to predict the condition of heart patient based on four different classification models namely, Decision Tree, Logistic Regression, Na-ive Bayes, and Random Forest. The dataset is separated into a training-testing ratio of 80%-20%. A confusion matrix is obtained for all four classification algorithms with true positive, true negative, false positive, and false negative val-ues. Based on these values, classification deci-sions are taken. The Random Forest classifier model obtain 90.16% classification accuracy.

Devansh Shah et al. (2020) [13] present a su-pervised learning model using a WEKA tool. Four different classification models namely, De-cision Tree, Random Forest, Naive Bayes, and K-nearest neighbor are used to design a heart dis-ease prediction model. An exploratory data anal-ysis is performed on the given dataset such as cleaning, data transformation, smoothening, ag-gregation, and normalization is performed. KNN classifier model obtains the highest classification accuracy.

N. Saranya et al. (2020) [14] disease predic-tion model via a web application based on the machine learning model. The proposed model is a time and cost-efficient model. Two different classification models namely, the K-nearest neighbor and Random Forest. Deep data analysis is performed on Coimbatore's hospital dataset using KNN and Random Forest classifier. The obtained accuracies are 100% and 91.36% using RF and KNN,

respectively. A Logistic Regression model is also employed based on the ensemble methods to enhance classification accuracy. With logistic regression and without logistic regression, the accuracy of ensemble methods is 98.77% and 95.06%, respectively. V. S. Dehnavi et al (2020) [15], a classification model is presented based on Artificial Neural Networks (ANN) based on genetic algorithms. The non-linear attributes present in the heart dataset are optimized using fuzzy sets. The combination of different machine learning models can improve classification accuracy. Apurv Garg et al. (2021) [16] used machine learning algorithms to estimate and analyze heart disease based on two different classifiers KNN and Random Forest. The given data is analyzed and examined using both classification models to get correlations between different attributes. Their impact on human health is studied. The Kaggle dataset is used to get the performance of the presented machine learning model. The dataset is separated into a training-testing ratio of 80%-20%. A positive correlation is obtained between attributes like Chest Pain and Maximum heart rate against the target value. With KNN and random forest, the accuracy of machine learning models is 86.88% and 81.96%, respectively. Harshit Jindal et al. (2021) [17] provided a heart disease prediction system based on three different classification models namely, KNN, LR, and RF. The highest accuracy result achieved is 87.5%. In this system, hybrid classification models are also studied. The accuracy (88.52%) of Logistic Regression and KNN is much higher than RF with KNN. Overall it is estimated that the hybrid model provides better performance results. Aravind Akella et al. (2021) [18] provided a heart disease prediction system based on six different classification models considering the UCI dataset. The highest accuracy result achieved is 93.03% using neural network architecture. Recall using neural network architecture is obtained at 93.8%. Thus, neural networks provide more efficient results than compare to remaining classifiers with 80% accuracy results. A. Tyagi et al. (2021) [19] suggested a predictive model via web application based on data mining techniques to provide support to medical experts by estimating the solutions of pre-defined questions. These smart applications can be used by heart disease experts to get instant results. The base of this application is naïve Bayes with high accuracy. Ghulab Nabi Ahmad (2022) [20] utilizes different classifiers namely, RF, GBC, DT, SVM, KNN, and DT to estimate cardiac disease based on the feature selection concept. A K-fold cross-validation approach is also employed for validation. The performance is measured using multiple datasets Long Beach V, Cleveland, Switzerland, and Heart Statlog Cleveland Hungary. For all the datasets, the Decision Tree Classifier and Random Forest Classifier provide the best results in terms of classification accuracy at 99.40% and 99.76%, respectively. All the results are compared against various previous classification models. Table 1 demonstrates a description of previous heart disease prediction models.

TABLE I. BRIEF DESCRIPTION OF HEART DISEASES PREDICTION WORKS

Year	Author	Purpose	Techniques used	Accuracy
2016	Sonam Nikhar[21]	Heart Disease Estimation Using Machine Learning Methods	1)Decision Tree 2) NaiveBayes Classifier	DT has performed superior compared to NB.
2017	P.Sai Chandrashekhar Reddy et. Al [22]	Heart disease prediction using Data Mining techniques	ANN	JAVA programming language is used for accuracy
2017	Jayami Patel et al,[23]	Heart disease Prediction using Data Mining Techniques and Machine Learning Algorithms	LMT, UCI	UCI provides better accuracy results compared to LMT
2018	Aditi Gavhane[24]	Designing a Heart Disease Prediction model Using Machine Learning	1) Multi-layer perceptron Technique	The MLP Performs better than previous models with the best Accuracy results
2018	V.V. Ramalingam[25]	Estimation of Heart Disease using machine learning algorithms	1) NB 2) SVM 3) KNN 4) DT 5) RF 6) EM	SVM shows the best accuracy results.
2019	Abhay Kishore[26]	Heart Attack Estimation Using Deep Learning models	1) RNN	RNN accuracy is achieved as 92%.
2019	A. Lakshmanarao[18]	Machine Learning algorithm presented for Heart Disease estimation	1)Synthetic Minority (SM) Oversampling 2)Adaptive synthetic sampling (ASS) 3)Random Oversampling (ROS)	FOR ROS – SVM - 82.30%. For SM-RF-91.3% FOR ASS-RF-90.3%

2019	Avinash Golande[19]	Heart Disease Estimation Using Machine Learning Algorithms	1)DT 2)KNN, 3)KMC 4)Adaboost	Decision Tree- 86.60%
2019	Mr. Santhana Krishnan.J[27]	Estimation of Heart Disease Using Machine Learning Techniques	1)Decision Tree 2)Naïve Bayes	DT -91%
2019	Senthilkumar Mohan [28]	Hybrid Machine Learning Algorithms for Effective Heart Disease Prediction	1)DT 2)LM 3)SVM 4)RF 5)NB 6)NN 7)KNN 8)HRFLM	HRFLM - 88.4%
2019	Bashir, Saba [33]	Enhancing Heart Disease Estimation based on Feature Selection Methods	1)DT 2)LR 3)NB 4)RF 5)SVM	DT - 82.22% LR - 82.56%
2020	Shi et.al. [29]	Heart Disease Analysis	Statistical Analysis	cardiac problems estimated in 19.7% of patients
2020	Manohar Manur et.al [30]	Heart Disease prediction based on the recurrent neural network	BiLSTM – CRF and Data Analysis is performed and linear relationship between attributes is maintained.	Accuracy – 90.04% Limitations- 1) Number of attributes chosen for the estimation is not clear. 2.) Mean accuracy results are preferred over individual accuracies for varied datasets. 3.) No significance test is evaluated.
2020	Anna Karen Gárate Escamilla et.al., [31]	Heart Disease prediction based PCA and Feature Selection	Both feature selection and feature extraction methods are used to get high accuracy	Cleveland- 98.7% Hungry- 99% CH- 99.4% Limitations- 1.) Analyzed the system with the same type of features available in different datasets. 2.) Important parameters like age, RestECG, ST Depression (Slope), etc, are excluded in the model for the Cleveland dataset.
2020	Rishabh Magar et al. [11]	Heart disease estimation using machine learning	1)LR 2)NB 3)SVM 4)DT	LR - 82.89% SVM- 81.57% NB- 80.43% DT- 80.43
2020	Apurb Rajdhan et al. [12]	Heart disease estimation using machine learning algorithm	1)LR 2)NB 3)RF 4)DT	1)LR – 85.25% 2)NB – 85.25% 3)RF - 90.16% 4)DT – 81.97%
2020	Devansh Shah et al. [13]	Heart disease prediction using machine learning techniques	1)NB 2)KNN 3)RF 4)DT	1)KNN – 90.78% 2)NB – 88.15% 3)RF - 86.84% 4)DT – 80.26%
2021	Harshit Jindal et al. [17]	Heart disease prediction using machine learning Algorithms	1)LR 2)KNN 3)HYBRID (KNN & LR)	1)LR – 88.5% 2)KNN – 88.52% 3)HYBRID (KNN & LR) – 87.5%

2021	Aadar Pandita et al. [32]	Machine learning algorithms are used for Heart disease prediction	1)LR 2)KNN 3)SVM 4)NB 5)RF	1)LR – 84.38% 2)KNN – 89.06% 3)SVM – 87.5% 4)NB – 85.94% 5)RF – 87.50%
2021	Aravind Akella et al. [18]	Estimation of coronary artery disease using Machine learning algorithms	1)DT 2)KNN 3)SVM 4)NN 5)RF 6)GLM	1)DT – 79.78% 2)KNN – 84.27% 3)SVM – 86.52% 4)NN – 93.03% 5)RF – 87.64% 6)GLM – 87.64%
2021	Saba Bashir et al. [33]	Heart Disease Prediction using UCI Dataset	Ensemble-based voting scheme	EVC -83%
2022	D. Cenitta et al.[34]	Ischemic Heart Disease Prediction Using Optimized Feature Selection Technique	Optimized Squirrel Search Feature Selection Algorithm	IHDSSO with random forest classifier – 98% Improve feature selection and minimize the number of features while giving better decisions
2022	Abdallah Abdellatif et al.[35]	Heart Disease Detection using machine learning and classification model	Severity Level Classification Model based on machine learning algorithms and Hyperparameter Optimization Methods used	95.73% classification accuracy using the Cleveland dataset. six different ML classifiers are used to analyze the patient's status
2022	Wen-Hsien Ho et al.[36]	Heart Diseases Detection using kernel Ensembling method	Banded Kernel Ensemble Learning Method	Accuracy -81%
2022	Shishir Rao et al.[37]	Heart Failure Prediction based on deep learning techniques	Transformer-Based Deep Learning Model	Contribution analyses identified risk factors that are closely related to HF. The model achieved 0.93 area under the receiver operator curve.

V. METHODOLOGY FOR HEART DISEASE PREDICTION

Machine learning algorithms mainly utilize 7 different steps to design a predictive model. A brief discussion and a block diagram are given in the following paragraph. Here, Figure 1 demonstrates the basic working of machine learning algorithms.

VI. RESULT AND DISCUSSION

The tools, programming languages, and performance metrics required to get high-performance results using machine learning algorithms to design a disease prediction model are discussed in the next section.

Tools that are massively used in machine learning model implementation with python and MATLAB programming are.

1.) WEKA, 2.) Rapid Miner, 3.) Apache Mahout, 4.) MATLAB

The performance of the designed machine learning models using these mentioned tools is obtained in a statistical manner. For example, the performance of the heart disease prediction model is evaluated in terms of classification accuracy, precision, recall, F1-measure, recall, and sensitivity based on the obtained confusion matrix (true positive, true negative, false positive, false negative values).

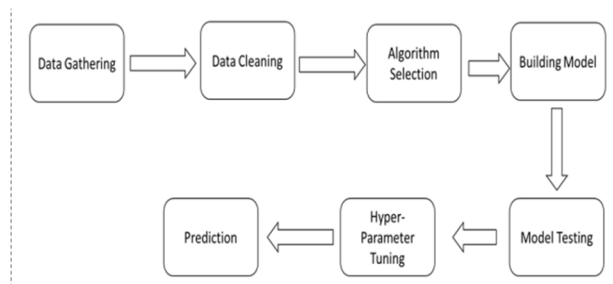


Fig. 1. Block Diagram for working process of Machine Learning Algorithm

VII. RESEARCH CHALLENGES

Some of the research challenges are discussed in the next section which is analyzed from the literature review. It is observed that most HDPS systems, generate weights using the attributes for classification. Further, it is well-known that all the attributes are linked to each other and possess certain correlations among them. Therefore, some hidden attributes like high blood pressure, smoking, and alcohol consumption also need to be considered for proper analysis and classification because their attributes can affect patient health. Some literature has used hybrid prediction models for heart disease estimation and obtained high classification performance results in comparison with standalone models. However, designing an accurate hybrid HDPS system need to look at varied constraints like interoperability, flexibility, and compatibility. Other limitations in designing a proper HDPS system are proper model selection, relevant data gathering, data processing, designing issues, sample size, result optimization, and adequate model training.

VIII. CONCLUSION AND FUTURE WORK

In the current survey paper, a comprehensive discussion and review are conducted on the state-of-art heart disease prediction systems. In this survey report, different heart disease-related datasets, different machine learning algorithms, data mining techniques, their working process, implementation tools, performance metrics, several prediction methods, and varied research limitations are discussed. Numerous authors have emphasized that understanding of research limitations and benefits is useful for analyzing any prediction model. The most used machine learning implementation interfaces are Mahout, MATLAB, Rapid Miner, and WEKA. The ultimate goal of this research survey is to provide details of current heart disease prediction methods, and their working model and highlight their limitations so that an efficient heart disease prediction system can be built in near future. This survey can definitely help researchers to design a reliable and accurate heart disease prediction system in the early stages. It can be concluded that to design an enhanced heart disease prediction system, proper pre-processing of data, exploratory data analysis, data cleansing, proper model selection, feature selection, and efficient classification model selection are the mandatory requirements. In future work, designing an effective predictive mechanism related to heart disease detection is the cardinal interest in minimizing most of the limitations of existing works.

ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

- [1] Animesh Hazra, Arkomita Mukherjee, Amit Gupta, Asmita Mukherjee, “Heart Disease Diagnosis and Prediction Using Machine Learning and Data Mining Techniques: A Review”, Research Gate Publications, July 2017, pp.2137-2159.
- [2] World Health Organization, Department of Health Statistics and Information Systems. Global Health Estimates: key figures and tables [Internet] Geneva: World Health Organization; 2016. http://www.who.int/healthinfo/global_burden_disease/en
- [3] Guizhou Hu, Martin M. Root, “Building Prediction Models for Coronary Heart Disease by Synthesizing Multiple Longitudinal Research Findings”, European Science of Cardiology, 10 May 2005
- [4] Jiawei, Micheline Kamber and Jian Pei, “Data Mining Concepts and Techniques”, Third Edition, Elsevier
- [5] Mirpouya Mirmozaffari, Alireza Alinezhad and Azadeh Gilanpour, “Data Mining Classification Algorithms for Heart Disease Prediction”, International Journal of Computing, Communications & Instrumentation Engg (IJCCIE), Vol. 4, No. 1, 2017.
- [6] Blake, C.L., Mertz, C.J.: “UCI Machine Learning Databases”, <http://mllearn.ics.uci.edu/databases/heart-disease/>, 2004.
- [7] P. Ghosh et al., “Efficient Prediction of Cardiovascular Disease Using Machine Learning Algorithms With Relief and LASSO Feature Selection Techniques,” in IEEE Access, vol. 9, pp. 19304-19326, 2021
- [8] A.Lakshmanarao, Y.Swathi, P.Sri Sai Sundareswar,” Machine Learning Techniques For Heart Disease Prediction”, International Journal Of Scientific & Technology Research Volume 8, Issue 11, November 2019.
- [9] Mr.Santhana Krishnan.J, Dr.Geetha.S,” Prediction of Heart Disease Using Machine Learning Algorithms”,2019 1st International Conference on Innovations in Information and Communication Technology(ICIICT),doi:10.1109/ICIICT1.2019.8741465.
- [10] Avinash Golande, Pavan Kumar T,” Heart Disease Prediction Using Effective Machine Learning Techniques”, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019.

- [11] "HEART DISEASE PREDICTION USING MACHINE LEARNING", International Journal of Emerging Technologies and Innovative Research (www.jetir.org), ISSN:2349-5162, Vol.7, Issue 6, page no.2081-2085, June-2020, Available :<http://www.jetir.org/papers/JETIR2006301.pdf>
- [12] Apurb Rajdhan , Avi Agarwal , Milan Sai , Dundigalla Ravi, Dr. Poonam Ghuli, 2020, Heart Disease Prediction using Machine Learning, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 09, Issue 04 (April 2020)
- [13] Shah, D., Patel, S. & Bharti, S.K. Heart Disease Prediction using Machine Learning Techniques. SN COMPUT. SCI. 1, 345 (2020).
- [14] N. Saranya, P. Kaviyarasu, A. Keerthana, C. Oveya. Heart Disease Prediction using Machine Learning International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-9 Issue-1, May 2020, Page No: 700-70
- [15] V. S. Dehnavi and M. Shafiee, "The risk predict ion of heart disease by using neuro-fuzzy and improved GOA," 2020 11th Internat ional Conference on Information and Knowledge T echnology (IKT), T ehran, Iran, 2020, pp. 127-131, doi: 10.1109/IKT51791.2020.9345630.
- [16] Garg, Apurv & Sharma, Bhartendu & Khan, Rizwan. (2021). Heart disease prediction using machine learning techniques. IOP Conference Series: Materials Science and Engineering. 1022. 012046. 10.1088/1757-899X/1022/1/012046.
- [17] Harshit Jindal et al 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1022 012072
- [18] Akella, Aravind and Akella, Sudheer. Machine learning algorithms for predicting coronary artery disease: efforts toward an open source solution. Future Science OA Volume 7, Number 6, Pages FSO698, 2021
- [19] A. Tyagi and R. Mehra, "Intellectual heartbeats classification model for diagnosis of heart disease from ECG signal using hybrid convolutional neural network with Goa," SN Applied Sciences, vol. 3, no. 2, p. 265, 2021.
- [20] G. N. Ahmad, S. Ullah, A. Algethami, H. Fatima and S. M. H. Akhter, "Comparative Study of Optimum Medical Diagnosis of Human Heart Disease Using Machine Learning Technique With and Without Sequential Feature Selection," in IEEE Access, vol. 10, pp. 23808-23828, 2022
- [21] Sonam Nikhar, A.M. Karandikar" Prediction of Heart Disease Using Machine Learning Algorithms" International Journal of Advanced Engineering, Management and Science (IJAEMS) Infogain Publication,[Vol-2, Issue-6, June-2016].
- [22] Mr.P.Sai Chandrasekhar Reddy, Mr.Puneet Palagi, S.Jaya, "Heart Disease Prediction using ANN Algorithm in Data Mining", International Journal of Computer Science and Mobile Computing, April 2017, pp.168-172.
- [23] Jayami Patel, Prof. Tejal Upadhyay, Dr. Samir Patel, "Heart disease Prediction using Machine Learning and Data mining Technique", March 2017.
- [24] Aditi Gavhane, Gouthami Kokkula, Isha Pandya, Prof. Kailas Devadkar (PhD)," Prediction of Heart Disease Using Machine Learning", Proceedings of the 2nd International conference on Electronics, Communication and Aerospace Technology (ICECA 2018). IEEE Conference Record # 42487; IEEE Xplore ISBN:978-1- 5386-0965-1
- [25] V.V. Ramalingam, Ayantan Dandapath, M Karthik Raja," Heart disease prediction using machine learning techniques: a survey", International Journal of Engineering & Technology, 7 (2.8) (2018) 684-687.
- [26] Abhay Kishore1, Ajay Kumar2, Karan Singh3, Maninder Punia4, Yogita Hambir5," Heart Attack Prediction Using Deep Learning", International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 04 Apr-2018.
- [27] X.-Y. Gao, A. A. Ali, H. S. Hassan, and E. M. Anwar, "Improving the accuracy for analyzing heart diseases prediction based on the ensemble method," Complexity, vol. 2021, Article ID 6663455, 10 pages, 2021.
- [28] Senthilkumar Mohan, Chandrasegar Thirumalai, Gautam Srivastava —Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques!, Digital Object Identifier 10.1109/ACCESS.2019.2923707, IEEE Access, VOLUME 7, 2019
- [29] Shi S, Qin M and Shen B, "Association of cardiac injury with mortality in hospitalized patients with COVID-19 in Wuhan," JAMA Cardiol, Vol. 5, No. 7, pp. 802-810, 2020
- [30] Manohar Manur et.al., "A Prediction Technique for Heart Disease Based on Long Short-Term Memory Recurrent Neural Network", International Journal of Intelligent Engineering and Systems, 2020, Vol. 13, No.2, pp.31-39
- [31] Anna Karen Gárate Escamilla et.al., "Classification models for heart disease prediction using feature selection and PCA", Informatics in Medicine Unlocked, 2020, Vol.19. 100330, pp.1-11.
- [32] Aadar Pandita, Siddharth Vashisht, Aryan Tyagi, Prof. Sarita Yadav."Prediction of Heart Disease using Machine Learning Algorithms", Volume 9, Issue V,International Journal for Research in Applied Science and Engineering Technology (IJRASET) Page No: 2422-2429, ISSN : 2321-9653
- [33] S. Bashir, A. A. Almazroi, S. Ashfaq, A. A. Almazroi and F. H. Khan, "A Knowledge-Based Clinical Decision Support System Utilizing an Intelligent Ensemble Voting Scheme for Improved Cardiovascular Disease Prediction," in IEEE Access, vol. 9, pp. 130805-130822, 2021, doi: 10.1109/ACCESS.2021.3110604.
- [34] D. Cenitta, R. Vijaya Arjunan and K. V. Prema, "Ischemic Heart Disease Prediction Using Optimized Squirrel Search Feature Selection Algorithm," in IEEE Access, vol. 10, pp. 122995-123006, 2022, doi: 10.1109/ACCESS.2022.3223429.

- [35] A. Abdellatif, H. Abdellatef, J. Kanesan, C. -O. Chow, J. H. Chuah and H. M. Ghenni, "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.
- [36] W. -H. Ho, C. -T. Liao, Y. J. Chen, K. -S. Hwang and Y. Tao, "Quickly Convert Photoplethysmography to Electrocardiogram Signals by a Banded Kernel Ensemble Learning Method for Heart Diseases Detection," in *IEEE Access*, vol. 10, pp. 51079-51092, 2022, doi: 10.1109/ACCESS.2022.3173176.
- [37] S. Rao et al., "An Explainable Transformer-Based Deep Learning Model for the Prediction of Incident Heart Failure," in *IEEE Journal of Biomedical and Health Informatics*, vol. 26, no. 7, pp. 3362-3372, July 2022, doi: 10.1109/JBHI.2022.3148820.