International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 10 Issue: 10 | Oct 2023 www.irjet.net p-ISSN: 2395-0072

A Comparative Analysis of Heart Disease Prediction System Using Machine Learning Based Techniques

Santosh P. Shrikhande [1]

[1]*School of Technology, S.R.T.M. University, Sub-Campus, Latur, Maharashtra, India*

Abstract - *Among different life-threatening diseases, heart disease is one of the most dangerous diseases in the world with high mortality rate since last two decades. Approximately 12 million people are dying due to the heart diseases every year. Therefore, diagnosis of heart diseases at early stage is very much important so that mortality rate of humans can be reduced. There are different methods to diagnose the heart diseases but early diagnosis is quite challenging task for the medical practitioners with better accuracy. Recently, machine learning based techniques are widely used in predicting the human heart diseases with great prediction accuracy. These techniques accurately predict the heart condition of patient so that further treatments can be made effective. This paper presents the comprehensive study and analysis of existing machine learning based techniques used in heart disease prediction. This comparative study also provides comprehensive analysis of existing methodologies that will help researcher for designing and developing novel and better heart disease prediction system using machine learning techniques.*

Key Words: Heart Diseases, Coronary Artery Diseases, Heart Attack Prediction, Machine Learning Techniques.

1. INTRODUCTION

Heart disease is one of the most dangerous disease that involves heart, blood vessels or both heart and blood vessels which is the leading cause of human deaths. The atherosclerosis and/or hypertension are main causes of the heart disease. Atherosclerosis is a condition that develops when a substance called plaque builds up in the walls of the arteries that narrows arteries. If a blood clot forms, it can stop the blood flow or slow down. This can cause the heart attack or stroke [1], [2]. The major risk factors for heart diseases are age, gender, obesity, high blood pressure, diabetes, tobacco smoking, processed meat consumption, excessive alcohol consumption, family history, lack of physical activity, psychosocial factors, and air pollution [3]. The diagnosis of heart diseases is usually done on the basis of symptoms and physical examination of the patient. Although heart diseases usually affects to older adults, the symptoms may begin in early life, making primary prevention efforts necessary from childhood [4]. Therefore risk factors may be modified by having healthy eating habits, exercising regularly and avoiding of smoking tobacco. If a heart disease person has not taken a proper treatment then it may cause deaths too. Therefore, early diagnosis of heart disease is necessary because many heart disease cases were severe due to late diagnosis and improper treatment [5], [6]. Recently, machine learning based techniques are widely used in early prediction of heart diseases. This research paper presents the comprehensive review of existing machine learning algorithms from literature for early prediction of heart diseases. This paper also illustrates comparative study of different machine learning techniques with their prediction accuracy, advantages and limitations. This will help researchers for understanding and finding new dimension to develop best heart disease prediction system using machine learning. This research paper is organized as follows. Section II, describes the design of heart disease prediction system using machine learning technique. Section III, presents literature review and comparison of different machine learning techniques used for heart disease prediction with comprehensive detail. The interpretations, discussion based on the comparative study and critical analysis of existing results is illustrated in the Section IV. The conclusions and suggestions are provided in Section V, those will help researcher in developing novel and better heart diseases prediction system in all aspect.

2. HEART DISEASE PREDICTION MODEL USING MACHINE LEARNING

The diagnosis or prediction accuracy in heart diseases is one of the major challenge for medical practioners. Therefore, mortality rate of heart disease patients is increasing worldwide due to inaccurate diagnosis and not in time treatment. Motivated by this, different machine learning based techniques have been proposed by several researchers for early prediction of the heart diseases [4]. The classification task of supervised machine learning is generally used for prediction of heart diseases using past data from the datasets. In the first phase of classification, pre-processed heart disease dataset is given to predictor for training purpose. Predictor learns from the training data and develops a model that will be further used for prediction of test data in second phase [6], [7], [8].

Following figure shows the framework of heart disease prediction system using supervised machine learning task.

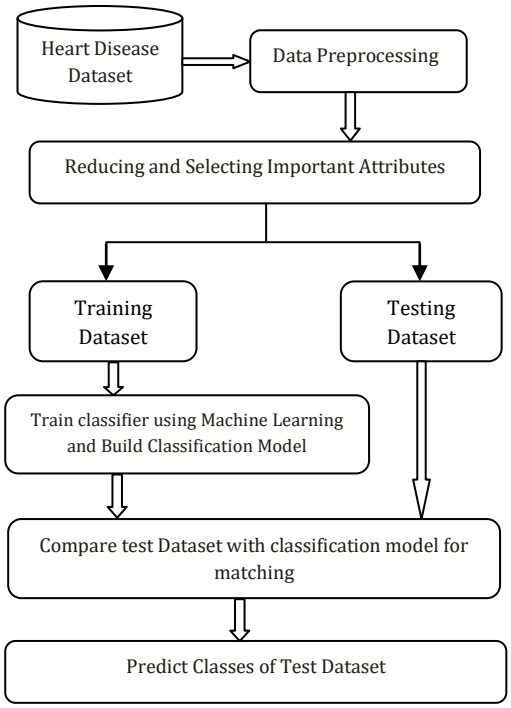

Fig-1: Heart Disease Prediction System Using Machine Learning

3. LITERATURE REVIEW

Now days, machine learning based techniques have been mostly used in the field of healthcare for heart diseases prediction and its severity analysis. So, several researchers have developed novel approaches using machine learning algorithms for early prediction of the heart diseases. The detailed literature review of existing machine learning techniques for early prediction of heart diseases is given below.

Nourah Alotaibi et al. [6], has done a comparative study and analysis of different feature extraction and feature selection techniques with different classifiers on the Cleveland heart disease dataset. From the study, it is found that: 1) The PES, EIA, CPT, MHR, THA, VCA, and OPK are the important features for predicting the heart diseases. 2) Naïve Bayes is the best classifier for heart disease prediction which has given 87.91% accuracy in 70-30 splitting of input data. 3) Then, Chi-squared feature selection is found the best method that reduced the number of features with improved prediction performance.

Pratiksha Shetgaonkar et al. [7], used Decision tree, Naïve Bayes and ANN based methods for heart disease prediction and evaluated all these methods with different parameters and optimization for best accuracy. Experiments conducted on Cleveland heart disease dataset concluded that Decision Tree has given 98.54%, Naïve Bayes has given 85.05% and Neural Network has given 81.83% prediction accuracy. Moreover, author has also concluded that chest pain and cholesterol are important attributes where as sex attribute does not play important role in the heart disease prediction.

Fadnavis et al. [8], implemented classification methods using Naïve Bayes and Decision Tree on Cleveland and Statlog datasets for heart disease prediction. After conduction of experiment, author concluded that Naïve Bayes has given better accuracy of 85.25% than the 81.97% accuracy of Decision Tree algorithm.

Devansh Shah et al. [9] presented supervised learning algorithms using Naïve Bayes, Decision Tree, KNN and Random Forest algorithm. Experiments conducted on Cleveland heart disease dataset has shown the highest accuracy of 90.78% using KNN, then 88.15%, 86.84%, 80.26% prediction accuracy has given by Naïve Bayes, Random Forest and Decision Tree classifiers respectively.

B. Santhi et al. [10] designed a machine learning model to predict heart diseases at early stage. Author used chi-squared statistical test for feature selection and correct tuning of hyper parameters to increase the prediction accuracy. Experiment carried out on Cleveland heart disease data set with 303 patients using Support Vector Machine (SVM), Random Forests, Ordinal Regression, Logistic Regression and Naïve Bayes algorithms. Author has improved accuracy of SVM classifiers with Gaussian kernel from 40% to 72% after feature selection and from 72% to 85% after correct hyper parameter tuning. Finally, the results are improved to 95% with the inclusion of rules based on observed statistical pattern.

Pranav Motarwar et al. [11] proposed a framework for predicting the heart disease using five machine learning algorithms such as Random Forest, Naïve Bayes, Support Vector Machine, Hoeffding Decision Tree (HDT), and Logistic Model Tree (LMT) on Cleveland dataset. Experimental results have shown that the Random Forest has given highest accuracy of 95.08% than the others algorithms. Gaussian Naïve Bayes has given 93.44%, SVM has given 90.16%, HDT has given 81.24%, and LMT has given 80.69% prediction accuracy.

Priya R. L. et al. [12] developed an efficient early prediction model that identify the best set of features for diagnosis using traditional machine learning algorithms along with modern gradient boosting approaches. Author used genetic algorithm for feature selection to optimize performance by reducing the number of parameters with 20% and keeping accuracy of the model intact. Moreover, hyper parameter optimization techniques were executed to improve performance of the predictive model. These hyper-parameters for Extra Tree Classifier have achieved a prediction accuracy of over 92%.

Mohd Faisal Ansari et al. [13] proposed a modified machine learning algorithm using Logistic Regression with PCA for predicting the heart disease with more accuracy. Experiments conducted on Cleveland dataset has shown an

accuracy of 86%, 68% of recall, 69% of specificity, 77% of precision, and 72% of f1score using Logistic regression with PCA.

Archana Singh et al. [14] developed a heart disease prediction system using machine learning algorithms such as KNN, Decision Tree, Linear Regression and SVM on Cleveland heart dataset. Experimental results have shown the prediction accuracy of 87% using KNN, 83% using SVM, 79% using Decision Tree and 87% using Linear Regression algorithm.

Halima et al. [15] proposed a machine learning model for prediction of the heart diseases using Naïve Bayes, KNN, SVM, Random Forest, and Decision Tree. After conducting experiment, author concluded that Naïve Bayes outperforms using both cross-validation and train-test split techniques with an accuracy of 82.17%, 84.28%, respectively. The KNN has given 76.56%, 81.31%, SVM has given 79.20%, 81.42%, Random Forest has given 69.30%, 77.14%, and Decision Tree has given 75.57%, 82.28% of accuracy using cross-validation and train-test split techniques. Then, author concluded that accuracy of all algorithm decreases after applying the cross-validation technique.

Monther et al. [16] proposed novel heart disease prediction system that combines all techniques into one single algorithm called as hybridization. The experiments conducted on Cleveland heart dataset have shown that the result of hybridization method is 89.20% which is better than the other individual techniques.

Senthilkumar Mohan et al. [17] proposed a hybrid approach by combining the characteristics of Random Forest and Linear Method to find significant attribute combinations and improving accuracy in heart disease prediction. Experiment conducted on Cleveland heart dataset has shown the 88.7% of prediction accuracy with 90.01% of precision and 92.80% of sensitivity.

Amin Ul Haq et al. [18] developed heart disease prediction system using sequential backward selection algorithm for significant feature selection, improving the prediction accuracy and reducing computational time. The experimental results have shown the 90% of prediction accuracy using KNN algorithm with six reduced features.

Imran Mirza et al. [19] proposed a machine learning model for heart disease prediction using RBF SVM and Linear SVM along with KNN and Naïve Bayes classifiers. Author identified thirteen significant features from the Cleveland dataset and applied these algorithms on it. Experimental results have shown that the highest accuracy of 87.112% for linear SVM and 87.114% for RBF SVM. Then, accuracy of 84.70%, 80% is reported using KNN and Naïve Bayes algorithms respectively.

M. Ganesan et al. [20] developed an efficient framework for heart disease prediction using machine learning algorithms on Cleveland dataset. After conducting an experiment author concluded that J48 tree based classifier has given better performance of 91.48% and SVM, Logistic Regression and MLP have given 84.07%, 83.70%, 78.14% of prediction accuracy respectively.

Amin Ul Haq et al. [21] developed a machine learning based diagnosis system for heart disease prediction using Cleveland heart disease dataset. Author used Relief Feature Selection, Minimal-Redundancy-Maximal-Relevance (mRMR), Least Absolute Shrinkage and Selection Operator (LASSO) algorithms to select the significant features. Then, Logistic Regression, SVM, Naïve Bayes, ANN, Decision Tree and KNN algorithms used for classification. Experimental results have shown accuracy of 86% 84%, 83%, 83%, 76%, 74%, 74% using RBF SVM, Logistic Regression, Naïve Bayes, Random Forest, KNN, ANN and Decision Tree algorithms respectively.

Poornima Singh et al. [22] proposed an effective heart disease risk level prediction system using Artificial Neural Network (ANN). Author employed the multilayer perceptron neural network with back propagation as the training and testing algorithm in a Weka software. This proposed model has reported near about ~100% prediction accuracy on the Cleveland dataset.

Chaithra N. et al. [23] developed a model for heart disease prediction using echocardiography dataset of heart diseases. Author used J48 Decision Tree, Naive Bayes and ANN for the classification of heart disease. Experiment conducted on Echocardiography database have shown that the ANN performed better in prediction of heart disease with 97.91% of accuracy and 97.20 % of sensitivity whereas J48 and Naïve Bayes has given 92.55% and 74.40% accuracy respectively.

Hidayet et al. [24] used machine learning and feature selection algorithms together to predict the heart attacks. Author used best machine learning methods with optimum parameters and several feature selection methods and evaluated its performance using Statlog heart disease dataset. Experimental results have shown that SVM with linear kernel and reliefF method pair is best feature selection method. This pair of algorithm has given highest accuracy of 84.81%.

Ritika Chadha et al. [25] implemented a heart disease prediction system using ANN, Decision Tree and Naive Bayes on Cleveland dataset with only eight attributes out of fourteen. Experimental results have shown that the ANN has given near about ~100% accuracy whereas 88.025%, and 85.86% prediction accuracy has given by Decision tree and Naïve Bayes respectively.

4. COMPARATIVE STUDY AND ANALYSIS

This section presents the heart disease prediction accuracy achieved by different authors in their research work from the literature. In addition, this section has also presented a complete analysis of what type of methodologies and data sets were used by authors in their research. It also explores

the pre-processing techniques of significant attributes reduction and selection with different classification algorithms used by various authors for conducting their experiments. The prediction accuracy of different existing machine learning techniques is shown in the following table.

Table -1: Heart Diseases Prediction Accuracy of Existing Technique

Ref. No.	Year	Technique Used	Prediction Results	Dataset Used
Nourah Alotaibi et al. [6]	2022	Chi-squared feature selection and PCA for feature reduction with Naïve Bayes(NB), Support Vector Machine (SVM), AdaBoost, Random Forest (RF), Logistic Regression algorithms for classification	Naïve Bayes with all 13 attributes=87.91%, Naïve Bayes with Chi-squared 7 attributes = 87.91%, Naïve Bayes with PCA used 2 attributes = 87.91%, SVM = 85.15%, RF = 80.72%, PCA with LR = 81.82%, Chi-squared AdaBoost = 85.45% accuracy.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Pratiksha Shetgaonkar et al.[7]	2021	Decision Tree, Naïve Bayes, and ANN algorithms with keeping significant attributes and removing insignificant attributes	Decision Tree = 98.54%, Naïve Bayes = 85.01%, ANN = 81.83%.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
R. Fadnavis et. al.[8]	2021	Naïve Bayes and Decision Tree classification algorithms	Naïve Bayes = 85.25%, Decision Tree = 81.97% prediction accuracy.	Cleveland and Statlog both Heart Disease Dataset
Devansh Shah et al. [9]	2020	Naïve Bayes, Decision tree, KNN and Random forest algorithms	Prediction accuracy using KNN=90.78%, Naïve Bayes= 88.15%, Random Forest=86.84% and Decision Tree=80.26%	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
B. Santhi et al. [10]	2020	Chi-squared statistical test for feature selection and SVM with Gaussian kernel used for classification.	Accuracy of SVM with Gaussian kernel improved from 40% to 72% after feature selection and from 72% to 85% after correct hyper parameter tuning. Final improved result is 95% with inclusion of rules based on observed statistical pattern.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Pranav Motarwar et al. [11]	2020	Random Forest, Naïve Bayes, Support Vector Machine, Hoeffding Decision Tree, and Logistic Model Tree	Random Forest=95.08%, Gaussian Naïve Bayes = 93.44%, SVM=90.16%, Hoeffding Decision Tree=81.24%, Logistic Model Tree=80.69% .	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Priya R. L. et. al. [12]	2020	Machine learning algorithms with modern gradient boosting approaches. Genetic algorithm for feature selection and Hyper parameter optimization technique for Extra Tree Classifier to improve performance.	Decision Tree=88.7%, Random Forest=90.7%, AdaBoost=85.5%, Gaussian Naive Bayes=62.7%, Logistic Regression = 70.4%, KNN =84.5%, XGBClassifier=85.2%, Gradient Boosting=86.8% prediction accuracy.	Framingham's Heart Study Dataset

Mohd Faisal Ansari et al. [13]	2020	Logistic Regression with PCA for feature reduction, selection and classification	Accuracy of 86%, Recall=68%, Specificity=69%, Precision=77%, and F1score=72% using Logistic Regression with PCA.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Archana Singh et al. [14]	2020	KNN, Decision Tree, SVM, and Linear Regression algorithms for heart disease prediction	Prediction accuracy of KNN=87%, SVM=83%, Decision Tree=79% and Linear Regression=87%	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Halima et. al. [15]	2020	Machine learning model for prediction of the heart disease using Naïve Bayes, KNN, SVM, Random Forest, and Decision Tree with cross-validation and train-test split method.	Accuracy of 82.17% and 84.28%, 76.56%, and 81.31%, 79.20%, and 81.42%, 69.30% and 77.14%, 75.57% and 82.28% using Naïve Bayes, KNN, SVM, Random Forest and Decision Tree with cross-validation and train-test split respectively.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Monther Tarawne et al. [16]	2019	Information gain, gain ratio, reliefF, symmetrical uncertainty, and oneR for feature selection and Hybridization using combination of Naïve Bayes, SVM, KNN, ANN, J48 Tree algorithms for classification	89.20% accuracy using hybridization technique with only 12 reduced features out of 14 features from dataset.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
S. Mohan et al. [17]	2019	Hybrid approach is developed by combining the characteristics of Random Forest (RF) and Linear Method (LM)	Prediction accuracy using Hybrid Random Forest with a Linear Model (HRFLM) is 88.70%.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Amin Ul Haq et al. [18]	2019	Sequential backward selection algorithm is used for feature selection and KNN classifier is used for heart disease prediction.	90.0% of accuracy is reported on only six numbers of reduced features using sequential backward selection with KNN algorithm.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Imran Mirza et al. [19]	2019	Author identified 13 significant features from dataset and applied RBF SVM and Linear SVM along with KNN and Naïve Bayes classifiers for classification	The highest accuracy of 87.112% using linear SVM and 87.114% using RBF SVM. The accuracy of 84.70% using KNN and 80% using Naïve Bayes algorithms.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
M. Ganesan et. al. [20]	2019	Developed an efficient technique using machine learning algorithms such as J48, SVM, Logistic Regression, and Multilayer Perceptron	J48 tree based classifiers given better performance of 91.48%, SVM = 84.07%, Logistic Regression = 83.70% and MLP = 78.14% accuracy	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Amin Ul Haq et. al. [21]	2018	Hybrid intelligent model using Logistic Regression, KNN, ANN, SVM, Naïve Bayes, Decision Tree, Random forest with Relief, mRMR, and LASSO for important features selection	Logistic regression = 89% with Relief, Naive Bayes = 84% with mRMR, RBF SVM = 88% with LASSO Features Selection algorithm.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository

Poornima Singh et. al. [22]	2018	Multilayer Perceptron Neural Network (MLP) with Back Propagation model for classes predictions	Multilayer Perceptron Neural Network with Back Propagation model has given ~100% of prediction accuracy	Cleveland Heart Disease Dataset from UCI Machine Learning Repository
Chaithra et. al. [23]	2018	Author used J48 Decision Tree, Naive Bayes and Artificial Neural Network algorithms for classification of heart disease.	ANN performed better in predicting accuracy with 97.91% and 97.20% of sensitivity whereas J48 Decision Tree and Naïve Bayes have given 92.55% and 74.40% accuracy respectively.	336 records with 24 attributes were obtained from the Echocardiography heart disease database
Hidayet TAKCI et. al. [24]	2017	Stepwise regression models, Fisher filtering (FF) and reliefF algorithms are used for feature selection and Regression, SVM, Decision trees and MLP and MLR used for classification	The highest accuracy without feature selection is 82.59%. The Linear kernel SVM with reliefF feature selection method reported 84.81% accuracy and KNN= 80.00%, Naïve Bayes = 83.70%, MLP = 83.33%, MLR = 83.70% accuracy	Statlog Heart Diseases Dataset
Ritika Chadha et. al.[25]	2016	Implemented heart disease prediction system using ANN, Decision Tree and Naive Bayes algorithm with only eight attributes out of fourteen.	ANN has given near about 100% accuracy whereas Decision Tree has given 88.025%, and Naïve Bayes has given 85.86% prediction accuracy.	Cleveland Heart Disease Dataset from UCI Machine Learning Repository

After completion of comprehensive literature review of heart disease prediction system using machine learning approaches, it has been observed that most of techniques from literature were used Cleveland heart disease dataset with only fourteen attributes and reported their prediction results. It is found that very few researchers have used clinically generated or own collected data set from the real patients. The above mentioned machine learning techniques have reported their prediction results using fourteen attributes of the Cleveland heart dataset. The heart disease prediction accuracy using Cleveland dataset is found quite well than the other primary datasets. The prediction accuracy is found varied depending on the type of heart disease data set and number of attributes were used for the prediction. It means, appropriate dataset and significant attributes selection plays an important role in improving the prediction accuracy of the proposed heart disease system. The attributes those are not playing much significant role in improving prediction accuracy need to be excluded as they are increasing computational time and slowing down the speed of system. But, there will be a problem of over fitting, if too many attributes from the dataset are excluded or reduced and it will be specific for a one technique only. Some of the researchers from literature have designed their approaches using one single machine algorithm technique and some researchers have used ensemble and hybrid techniques to verify the performance of heart diseases prediction system. It is also observed from comparison table that single algorithm technique using Naïve Bayes, Random Forest and KNN reported better prediction accuracy than other algorithms. Artificial neural network with different hidden layers and multilayer perceptron technique has given good and acceptable prediction results. It is also observed form the study that the performance of single machine learning algorithm can be improved by ensemble and/or hybridization of more than one algorithm. From the comparative analysis it is found that hybridization techniques using combination of Logistic Regression, KNN, ANN, SVM, Naïve Bayes, Decision Tree, and Random forest with PCA for features reduction and Relief, mRMR, LASSO and Chi-squared techniques for important features selection has reported improvement in prediction accuracy than individual classification algorithms. It means appropriate or significant attributes selection techniques have direct impact on the prediction accuracy of the heart disease prediction system. Moreover, ensemble techniques using combination of different classification algorithms with significant features selection also has impact on the prediction accuracy improvement.

Therefore, in order to design an effective and accurate heart disease prediction system, the methodology using ensemble or hybridization technique should be adopted so that significant attributes will be selected using one algorithm and classification will be done using other algorithms. This will improve the prediction performance of the designed technique for heart disease prediction. The appropriate preprocessing technique for attributes reduction and

selection is also essential so that over-fitting and under-fitting issues should be avoided and significant feature are selected. Moreover, the real performance of any machine learning algorithm is verified using different heart disease datasets from different countries with distinguished attributes rather than only fourteen attributes from the Cleveland heart disease dataset. If designed system performs well for all these datasets then this developed system will be considered suitable for the heart disease prediction.

5. CONCLUSIONS AND FUTURE WORK

This research paper has presented a comprehensive literature review and comparative analysis of existing machine learning techniques used in the heart diseases prediction. After exploring the different existing machine learning techniques used in heart diseases prediction along with their prediction results, following interpretations are highlighted. The Cleveland heart disease dataset with only fourteen attributes from UCI machine learning repository has been mostly used to verify the performance of heart disease prediction and found acceptable prediction accuracy. It is also observed that the real prediction performance of system is varied depending on the datasets, number of attributes, instances, preprocessing and classification algorithms. Some of the classification algorithms from the literature such as Naïve Bayes, Random Forest, K-Nearest Neighbour and Artificial Neural Network have reported the acceptable prediction accuracy using Cleveland heart disease dataset. But, in order to verify the real performance of the system different datasets along with distinguished attributes should be used and tested using these machine learning techniques.

Therefore, it is concluded from this study that the various heart disease datasets of real patients with different attributes from different geographical area should be used for verifying the real performance of the heart disease prediction system.

The appropriate pre-processing technique should be adopted for significant attributes selection and reduction so that prediction accuracy will be increased and computational time will be reduced. Then, proper classification methodology using ensemble or hybridization technique should be utilized for accurate classes prediction and minimizes the misclassification rate. Moreover, some other classification techniques can be adopted using Artificial Neural Network (ANN), Deep Learning algorithms for better prediction accuracy. Our future work is to give emphasis on the above addressed aspects and develop a new methodology for early prediction of heart diseases.

REFERENCES

[1]. Sujata Joshi, Mydhili K. Nair, "Prediction of Heart Disease Using Classification Based Data Mining Techniques", Computational Intelligence in Data Mining - Volume 2, Smart Innovation, Systems and Technologies 32, DOI 10.1007/978-81-322-2208-8_46.

[2]. The top 10 causes of death. World Health Organization, http://www.who.int/mediacentre/factsheets/fs310/en/. Accessed 20 June 2023).

[3]. Baban Rindhe, Nikita Ahire et. al., "Heart Disease Prediction Using Machine Learning", International Journal of Advanced Research in Science, Communication and Technology (IJARSCT), Vol. 5, Issue 1, pp.267-273, 2021.

[4]. Manoj Diwakar, Amrendra Tripathi. Kapil Joshi, Minakshi Memoria, Prabhishek Singh, Neeraj kumar, "Latest trends on heart disease prediction using machine learning and image fusion", Materials Today: Proceedings By Eleveir, https://doi.org/10.1016/j.matpr.2020.09.078. Nancy Masih, Sachin Ahuja, "Prediction of Heart Diseases Using Data Mining Techniques: Application on Framingham Heart Study", International Journal of Big Data and Analytics in Healthcare Volume 3, Issue 2, pp. 1-9, 2018.

[5]. Nourah Alotaibi, Mona Alzahrani, "Comparative Analysis of Machine Learning Algorithms and Data Mining Techniques for Predicting the Existence of Heart Disease", International Journal of Advanced Computer Science and Applications, Vol. 13, No. 7, pp. 810-818, 2022.

[6]. Pratikha Shetgaonkar, Shailendra Aswale, "Heart Disease Prediction Using Data Mining Techniques", International Journal of Engineering Research and Technology (IJERT), Vol. 10, Issue 02, pp. 281-286, February-2021.

[7]. R Fadnavis, K Dhore, D Gupta, J Waghmare, D Kosankar, "Heart disease prediction using data mining", International Conference on Research Frontiers in Sciences (ICRFS 2021), Journal of Physics: Conference Series 1913, doi:10.1088/1742-6596/1913/1/012099, pp.1-6, 2021.

[8]. Devansh Shah, Samir Patel, Santosh Kumar Bharti, "Heart Disease Prediction using Machine Learning Techniques", SN Computer Science, Springer Nature Journal , https://doi.org/10.1007/s42979-020-00365-y, pp.1-6, 2020.

[9]. B. Santhi, K. Renuka, "Study and Analysis of Prediction Model for Heart Disease Data Using Machine Learning Techniques", Journal of Computer Science 16 (3), DOI: 10.3844/jcssp.2020.344.354, pp.344.354, 2020.

[10]. Pranav Motarwar, Ankita Duraphe, G Suganya, M Premalatha, "Cognitive Approach for Heart Disease

Prediction using Machine Learning", International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), 978-1-7281-4141-1/$31.00, 978-1-7281-4142-8, pp.1-5, 2020.

[11]. Priya R. L, S. Vinila Jinny, Yash Vijay Mate, "Early prediction model for coronary heart disease using genetic algorithms, hyper-parameter optimization and machine learning techniques", Health and Technology, IUPESM and Springer-Verlag GmbH Germany, part of Springer Nature, https://doi.org/10.1007/s12553-020-00508-4, 2020.

[12]. Mohd Faisal Ansari, Bhavya AlankarKaur, and Harleen Kaur, "A Prediction of Heart Disease Using Machine Learning Algorithms", Springer Nature Switzerland, https://doi.org/10.1007/978-3-030-51859-2_45, pp. 497–504, 2021.

[13]. Archana Singh, Rakesh Kumar, "Heart Disease Prediction Using Machine Learning Algorithms", IEEE, International Conference on Electrical and Electronics Engineering (ICE3-2020), 978-1-7281-5846-4/20/$31.00 ©2020 IEEE, 2020.

[14]. Halima Hamdaoui, Saïd Boujraf, Nour Houda Chaoui, Mustapha Maaroufi, "A Clinical support system for Prediction of Heart Disease using Machine Learning Techniques", 5th International Conference on Advanced Technologies MIA-51 for Signal and Image Processing, Sfax, Tunisia, September, 2020.

[15]. Monther Tarawneh, Ossama Embarak, "Hybrid Approach for Heart Disease Prediction Using Data Mining Techniques", Springer Nature Switzerland AG 2019, https://doi.org/10.1007/978-3-030-12839-5_41, pp. 447–454, 2019.

[16]. Senthilkumar Mohan, Chandrasegar Thirumalai, Gautam Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", Special Section On Smart Caching, Communications, Computing And Cyber security For Information-Centric Internet Of Things, DOI. 10.1109/ACCESS.2019.2923707, pp.81542-81554, 2019.

[17]. Amin Ul Haq, Jianping Li et. al.. "Heart Disease Prediction System Using Model Of Machine Learning and Sequential Backward Selection Algorithm for Features Selection", 5th International Conference for Convergence in Technology (I2CT) Pune, India. Mar 29-31, 2019, 978-1-5386-8075-9/19/$31.00 ©2019 IEEE.

[18]. Imran Mirza Arnav Mahapatra, "Human Heart Disease Prediction Using Data Mining Techniques", IEEE International Conference on Advances in Computing, Communication and Control (ICAC3)-Mumbai, India, pp. 1-5, 2019.

[19]. M.Ganesan, Sivakumar, "IoT based heart disease prediction and diagnosis model for healthcare using machine learning models", Procedings of IEEE International Conference on System Computation Automation and Networking, 2019.

[20]. Amin Ul Haq, Jian Ping Li, Muhammad Hammad Memon, Shah Nazir, "A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms", Hindawi, Mobile Information Systems, Volume 2018, Article ID 3860146, https://doi.org/10.1155/2018/3860146, pp.1-21. 2018.

[21]. Poornima Singh, Sanjay Singh, Gayatri S Pandi-Jain, "Effective Heart Disease Prediction System Using Data Mining Techniques", International Journal of Nanomedicine, Dovepress Journal, pp. 121-124, 2018.

[22]. Chaithra N and Madhu B, "Classification Models on Cardiovascular Disease Prediction using Data Mining Techniques", Journal of Cardiovascular Diseases and Diagnosis, Vol.-6, Issue-6, DOI: 10.4172/2329-9517.1000348, pp.1-4, 2018.

[23]. Hidayet TAKCI, "Improvement of heart attack prediction by the feature selection methods", Turkish Journal of Electrical Engineering and Computer Science, doi:10.3906/elk-1611-235, pp.1-10, 2018.

[24]. Ritika Chadha, Shubhankar Mayank, "Prediction of heart disease using data mining techniques", Special Issue REDSET of CSIT Springer, DOI 10.1007/s40012-016-0121-01, pp.1-6, 2016.

[25]. Ibrahim M. El-Hasnony, Omar M. Elzeki, Ali Alshehri Hanaa Salem, "Multi-Label Active Learning-Based Machine Learning Model for Heart Disease Prediction", Sensors 2022, 22, 1184. Pp.1-18, https://doi.org/10.3390/s22031184, 2022.

[26]. Santosh P. Shrikhande, Prashant Agnihotri, "Analysis and Prediction of Customer's Purchasing Pattern Using Market Basket Analysis with Association Rule Mining", International Journal of Computer Science and Information Technologies, Vol. 13 (4) , 2022, 97-102, pp. 97-102, 2022.

[27]. Montu Saw, Tarun Saxena, Sanjana Kaithwas, Rahul Yadav, Nidhi Lal, "Estimation of Prediction for Getting Heart Disease Using Logistic Regression Model of Machine Learning", IEEE International Conference on Computer Communication and Informatics (ICCCI - 2020) Coimbatore, pp. 1-6, 2020.

[28]. B. Keerthi Samhitha, Sarika Priya.M.R, Sanjana.C, Suja Cherukullapurath Mana and Jithina Jose,

"Improving the Accuracy in Prediction of Heart Disease using Machine Learning Algorithms", International Conference on Communication and Signal Processing, July 28 - 30, pp. 1326-1330, 2020.

[29]. 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) pp. 684-687, 2018.

[30]. Santosh P. Shrikhande, Prashant P. Agnihotri, "Comparative Study of Various Data Mining Techniques for Early Prediction of Diabetes Disease", International Journal of Scientific Research in Computer Science, Engineering and Information Technology, Volume 8, Issue 1, DOI: https://doi.org/10.32628/CSEIT228139, pp. 287-295, 2022.

[31]. Garima Choudhary, Dr. Shailendra Narayan Singh, "Prediction of Heart Disease using Machine Learning Algorithms", International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE 2020), pp.197-202, 2020,

[32]. Salma Banu N.K, Suma Swamy, "Prediction of Heart Disease at early stage using Data Mining and Big Data Analytics: A Survey", International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), pp. 256-261, 2016.

[33]. S. Kiruthika Devi, S. Krishnapriya, Dristipona Kalita, "Prediction of Heart Disease using Data Mining Techniques", Indian Journal of Science and Technology, Vol. 9(39), DOI: 10.17485/ijst/2016/v9i39/102078, pp. 1-5, 2016.