JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

A Review of Different Approaches Used for Cardiovascular Disease Using Machine Learning Techniques

[1]Ms. Sneha Kumbhar, [2]Prof. Kishor Mane

[1]MTech Student, [2]Assistant Professor

[1]Department of Computer Science and Engineering,

[1]D. Y. Patil College of Engineering & Technology, Kolhapur, India

[2]Department of Computer Science and Engineering,

[2]D. Y. Patil College of Engineering & Technology, Kolhapur, India

Abstract: An early detection of cardiovascular disease is crucial for saving lives, as it is a leading cause of death worldwide. Healthcare organizations generate vast amounts of data that can be utilized by researchers to improve disease diagnosis accuracy. Researcher aims to explore various algorithm combinations and find efficient techniques for disease diagnosis by using various optimization algorithms. Accurate prognosis and effective management of cardiovascular disease necessitate the expertise of medical professionals specialized in this field. Researchers have offered a plethora of algorithms and learning strategies to aid doctors in this technological age. Due to a lack of trained medical personnel and other essential resources, accurate prognosis and treatment of heart disease patients is particularly difficult in poor nations. The primary causes are inadequate preventative measures and the presence of incompetent or unqualified medical providers. Even though many cases of heart disease are avoidable, the epidemic is spreading largely due to insufficient preventative efforts. The health care region generates huge amounts of data daily, particularly in regard to individuals and illnesses. Although academics and practitioners have access to this information, they are not making effective use of it. The modern healthcare sector has an abundance of data but a scarcity of insight. It is possible to improve diagnostic and decision making with the use of accessible data mining as well as machine learning tools and methodologies. In this paper various Machine learning classification approaches for diagnosis of cardiovascular disease are reviewed and finding has been represented in assisting healthcare professionals in accurately diagnosing these diseases. Based on the review different drawbacks and challenges has been identified for providing more accurate system than the existing systems.

***Index Terms* - Machine Learning, Cardiovascular disease, Diagnosis, Treatment, Multilayer perceptron, Logistic Regression,Naïve Bayes, Random Forest, Artificial Intelligence, K-Nearest Neighbor, J48, bagging, Neural Network, Support Vector machine.**

I. INTRODUCTION

The healthcare sector is among the domains where machine learning finds widespread application. Machine learning, a branch of computer science, empowers computers to exhibit increased creativity. Its intelligent nature makes it possible for the healthcare sector to analyze data efficiently. Although data has been collected and used for research purposes for decades, it has not been used precisely. It is challenging for humans to extract convenient information from vast amounts of data [26]. This is where machine learning comes in handy, as it is extensively used to analyze such data and detect complications in the healthcare organization. Machine learning algorithms learn from past detected patient's cases.

Within the healthcare sector, ensuring prompt and accurate identification of cardiovascular disease plays a vital role in preserving lives. To achieve this, researchers have been exploring various machine learning approaches for disease prediction, which have been implemented in different cardiovascular disease detection systems. Nevertheless, the patient must undergo all necessary tests, and healthcare professionals must carefully examine the results. This research paper seeks to shed light on the machine learning classification techniques employed by previous researchers for detecting cardiovascular diseases. It underscores the significant role played by machine learning in the healthcare industry, which can lead to more accurate predictions and assist healthcare professionals [11]. Previous surveys have mainly relied on machine learning techniques with feature extraction algorithms, while optimization algorithms, which could improve the accuracy of the model, have not been sufficiently used.

The utilization of machine learning has gained significant traction in the medical domain, specifically for the purpose of disease identification. By utilizing machine learning approaches, researchers aim to reduce diagnosis time while increasing accuracy and efficiency. While supervised machine learning techniques can be applied to detect any type of disease, this paper will focus on heart disease (also known as cardiovascular disease) diagnosis [34]. Ensuring accurate and timely identification of cardiac conditions is of utmost importance in safeguarding lives, as it remains a significant global contributor to mortality rates. According to World Health Organization, cardiovascular diseases result in approximately 12 million fatalities annually, rendering it the foremost cause of death in numerous developing and developed nations. In the United States, a single life succumbs to heart disease every 34 seconds. Similarly, heart diseases are also a significant threat to adult lives in India, where it is considered the most dangerous disease. Hence, the precise identification of heart disease has emerged as an urgent requirement both in India and worldwide.

II. LITERATURE REVIEW

AH Chen et al. [18] introduced a heart disease prognosis method to assist the clinicians determine if patient having heart disease or not dependent on their medical details. Classification of cardiac disease was performed using an artificial neural network technique. Information was obtained from UCI's machine learning repository. The artificial neural network architecture comprised separate sets of neurons dedicated to input and output layers, incorporating multiple layers for both input and output functionalities. There are some potential drawbacks and challenges associated with this approach: The accuracy and reliability of any machine learning model heavily depend on the quality and representativeness of the data used for training. Generalization to real-world scenarios: The performance of a neural network model trained on a specific dataset might not generalize well to real-world scenarios. The UCI dataset used in this study may not fully capture the complexity and heterogeneity of heart disease cases encountered in clinical practice. The utility of artificial intelligence & machine learning in healthcare raises ethical and legal considerations. Privacy and data security are paramount when handling sensitive patient information.

Mohan, S., Thirumalai, C. and Srivastava, G. [27] presented Clinical data analysis for predicting cardiovascular disease accurately. The healthcare sector produces an extensive volume of data, emphasizing the importance of employing efficient decision-making tools like machine learning. Machine learning has demonstrated its efficacy in numerous Internet of Things (IoT) advancements, and recent research has unveiled its potential in predicting cardiovascular disease. The author introduces an innovative methodology in this paper, which utilizes machine learning techniques to improve the precision of cardiovascular disease prediction. The presented prediction model incorporates different combinations of features and multiple established classifying methods. Notably, the hybrid random forest with a linear model (HRFLM) demonstrates exceptional performance in accurately predicting heart disease. The primary goal of this study is to enhance prediction accuracy for cardiovascular disease by leveraging machine learning methods and identifying crucial features. The paper showcases a prediction model that combines various features and employs different classification approaches, with the HRFLM algorithm showcasing superior performance in accurately predicting heart disease. While the authors mention identifying significant features, the process of feature selection is not explicitly described. Choosing the right set of features is challenging, as it requires domain expertise and understanding of the underlying relationships between features and the target variable. Additionally, handling high-dimensional data can pose challenges, leading to potential overfitting or increased computational complexity. ML models, such as random forests, are known for their black box nature, making it challenging to interpret the underlying reasons behind their predictions. While the authors propose a hybrid model with a linear model, which generally offers more interpretability, the interpretability of the overall HRFLM model is not discussed.

Mrudula Gudadhe et al. [16] introduced a decision support system designed to assist in the diagnosis of CVD. The system utilized support vector machine (SVM) and artificial neural network (ANN) techniques. A three-layered multilayer perceptron neural network (MLPNN) was employed, and training was carried out using the computationally efficient back-propagation technique. The findings indicate that utilizing an MLPNN with back-propagation proves effective in diagnosing heart disease. The Author does not specify the interpretability of the MLPNN and SVM models. Neural networks, including MLPNNs, are known to be complex and often lack interpretability. Understanding the underlying features and factors contributing to the model's predictions is important in healthcare to gain trust from healthcare professionals and facilitate better understanding of the diagnostic decisions made by the system.

Shaikh Abdul Hannan et al. [21] introduced a Radial Basis Function (RBF) for prediction of prescriptions related to heart disease. The RBFNN, known as the Radial Basis Function-Neural Network, comprises three layers: input, hidden, and output layers. Within the hidden layer (bk), RBF units (nh) and a bias value (bias) are incorporated. Within the hidden layer, every neuron employs a non-linear processing technique, employing a radial basis function, with the Gaussian function being the widely preferred choice. The design of an RBFNN involves selecting parameters such as the number of latent surface units, weights and width. The results obtained demonstrate the effectiveness of using the radial basis function in predicting heart disease medication prescriptions. The performance of an RBFNN heavily relies on the appropriate selection of parameters, including the number of latent surface units, weights, and width of the radial basis functions. Determining the optimal values for these parameters can be challenging and may require extensive experimentation and fine-tuning. An inappropriate choice of parameters could lead to suboptimal performance and less accurate predictions. RBFNN models, similar to other neural networks, tend to be complex and lack interpretability. The underlying reasoning behind the network's predictions may not be easily understandable or explainable. This lack of interpretability could limit the acceptance and adoption of the model by clinicians and other healthcare professionals who require transparency and clear explanations for the prescriptions provided

Mathur, P., Srivastava [25] proposed the concept of Artificial intelligence-based applications have been widely used across scientific, technological, and medical fields. The concept of utilizing machine computing power in clinical medicine and diagnostics dates back to the 1960s. However, recent progress in computing and algorithms, notably deep learning networks, has revitalized the interest in employing these systems in clinical medicine, with a specific focus on cardiac medication. Artificial intelligence (AI) systems have been implemented in diverse domains, showcasing their utility and versatility, including the prediction of cardiovascular risk. This comprehensive review of diverse AI applications, encompassing machine learning and deep learning, sheds light on their profound impact on cardiovascular medicine, fostering a deeper comprehension of heart failure and congenital heart disease phenotype. Data quality and representativeness: AI-based applications heavily rely on the quality and representativeness of the data used for training. Obtaining large, diverse, and high-quality datasets in cardiovascular medicine can be challenging. Biases, missing data, and data variability can affect the performance and generalizability of the AI models. Ensuring data accuracy, completeness, and proper representativeness across different patient populations is crucial to avoid biased or inaccurate predictions. Clinical integration and adoption: Integrating AI-based applications into clinical practice requires collaboration between data scientists, clinicians, and healthcare administrators. Implementation challenges, such as resistance to change, lack of technical expertise, and concerns about liability and accountability, can hinder the widespread adoption of AI systems in cardiovascular medicine. Collaboration between different stakeholders is crucial to address these challenges and ensure successful integration into routine clinical workflows.

Asif, M., Nishat, M.M., and Faisal [29] conducted a research study was conducted to assess the effectiveness of various machine learning algorithms in predicting cardiovascular disease. Timely detection of life-threatening conditions is vital for ensuring appropriate treatment and mitigating severity in many instances. To train and test the model, researchers utilized the University of California, Irvine (UCI) data repository. They compared the performance of multiple machine learning algorithms using default hyperparameters (DHP), grid search cross-validation (GSCV), and random search cross-validation (RSCV) techniques. The computational time required for GSCV and RSCV was also recorded. The study revealed that both hard and soft voting ensemble classifiers (EVCH and EVCS) demonstrated commendable accuracy. However, the Adaboost algorithm outperformed EVCH and EVCS in terms of correctness and attentiveness. It is worth noting that hyperparameter tuning and algorithm selection are not universally applicable approaches. Optimal hyperparameters and algorithm choices may differ depending on the specific dataset, problem domain, and evaluation metrics. Therefore, the author's findings should not be generalized to all datasets or clinical scenarios. Computational time and scalability: The author mention the calculation of computational time for hyperparameter tuning methods (GSCV and RSCV). However, the computational time required by different algorithms, especially for larger datasets or more complex models, is an important consideration. Scalability and efficiency in real-world app.

Haq, A.U., Li, J.P., Memon [31] discussed that the traditional medical history for diagnosing heart disease is not considered reliable in many cases. Thus, machine learning techniques have proven to be effective and dependable in accurately classifying individuals as healthy or having cardiac disease. The scientists devised a system for predicting heart disease. that employed machine learning algorithms along with a heart disease dataset. The system incorporated various feature selection algorithms, a cross-validation method, and several performance evaluation standards. By utilizing machine learning-based decision support system, medical professionals can receive valuable assistance in diagnosing patients with heart conditions. However, it is important to acknowledge that the performance and reliability of machine learning models are contingent upon the quality and representativeness of the training dataset. The heart disease dataset utilized in the study may possess certain limitations, such as a small sample size, data imbalance, or missing information. These limitations can affect the generalizability and accuracy of the developed prediction system, particularly when applied to diverse patient populations or different healthcare settings. Validation and generalization: Author mention the use of cross-validation methods for performance evaluation. While cross-validation helps in assessing the model's performance on the available dataset, it is crucial to validate the system's performance on independent and unseen datasets to evaluate its generalization capability. External validation is necessary to determine the system's accuracy and reliability in real-world clinical scenarios and different patient populations.

Dinesh, K.G., Arumugaraj, K., Santosh, K.D., and Mareeswari, V. [17] introduced a system that incorporates data pre-processing techniques involving the removal of noisy and missing data, imputation of relevant default values, and attribute categorization to facilitate prediction and decision-making across various levels. The identification model was developed using methods such as categorization, correctness, responsiveness, and particularity analysis. The author put forth a prediction model that demonstrates high accuracy in determining the presence or absence of heart disease, providing valuable insights into the condition. Author does not mention the use of external validation or the evaluation of the model's performance on independent datasets. To assess the generalizability of the identification model, it is important to validate its performance on diverse and unseen datasets. External validation is necessary to determine the model's reliability and effectiveness in real-world clinical settings.

Ghumbre, S.U. and Ghatol, A.A. [19] announced the increasing use of machine learning algorithms for predicting various diseases and how assembling utilizes a model can enhance reporting consistency and accuracy by facilitating disease visualization and analysis. The researchers carried out an investigation to explore the detection of heart disease using various machine learning algorithms, including k-nearest neighbor, logistic regression, random forest, and gradient boosting. An evaluation of these algorithms was carried out using a confusion matrix. The authors employed a two-step approach, where the heart disease dataset was first prepared in the required format for machine learning algorithms and subsequently utilized to determine the presence or absence of heart disease in patients. Findings indicated that Logistic Regression algorithm exhibited a superior level of accuracy when compared to other methods. Author does not explicitly mention the process of feature selection or engineering. Selecting the most relevant features and transforming them into meaningful representations are critical steps in developing an effective prediction model. Inadequate feature selection or inappropriate feature engineering techniques may lead to suboptimal model performance or introduce biases.

Asha Rajkumar et al. [15] employed a supervised machine learning technique to classify and diagnose heart disease. They utilized a freely available academic and research data mining software called Tanagra, which offers diverse data mining techniques.

The researchers trained their model on 80% of the data and performed testing on the remaining 20% using 10-fold cross-validation. The findings revealed that the Naive Bayes technique exhibited the lowest error rate and demanded the least amount of time. The author mentions the use of Naive Bayes technique with the lowest error rate and time requirement. While Naive Bayes is a popular classification algorithm, its performance may vary depending on the dataset and the problem at hand. To ensure optimal heart disease diagnosis, it is crucial to assess and compare multiple classification algorithms. Evaluating the performance of the models should extend beyond error rate and encompass other metrics like reactivity, accuracy, correctness, and the area under the receiver operating characteristic curve (AUC-ROC). Incorporating these additional metrics provides a comprehensive evaluation of the model's effectiveness.

Gunsai Pooja Dineshgar et al. [26] conducted a survey on knowledge discovery techniques in databases that utilize data mining methods. They developed an intelligent heart disease prediction system prototype that uses a historical heart database for diagnosis. The research study explored clustering techniques, specifically the k-means and k-medoid algorithms, aiming to attaining global optimality in clustering based on partition requires specific optimization techniques. PAM (Partitioning Around Medoids) utilizes iterative optimization by relocating points between clusters and considering them as potential medoids. On the other hand, CLARA (Clustering Large Applications) adopts a random search approach, generating neighbors through the random selection of a starting node and subsequent neighbor checks. Both techniques represent variations of the k-medoid algorithm. It is essential to note that the accuracy and reliability of the cardiac disease diagnosing system heavily rely on a quality and representativeness of the historical heart database utilized for diagnosis. It is crucial to ensure that the database encompasses a diverse range of patient characteristics, including various demographics, medical histories, and disease profiles. Inadequate representation of certain patient populations or limited availability of certain features may impact the generalizability and performance of the system. The author does not discuss the interpretability of the clustering results obtained from the k-means, k-medoid, PAM, and CLARA algorithms. Interpreting and understanding the clusters formed by these algorithms can provide valuable insights into different patient groups and their risk profiles for heart disease. The interpretability of the clustering results can facilitate better decision-making and patient management.

Jyoti Soni et al. [26] observed that the Weighted Associative Classifier (WAC) exhibited superior performance contrasted to other Associative Classifiers when diagnosing heart disease using patient test results. The researchers developed a user-friendly graphical interface on the Java platform, which utilizes CAR rules generated by WAC. WAC employs Weighted ARM to extract association rules, incorporating a weighted support and confidence framework to eliminate irrelevant associations and prioritize the most significant ones. Author mentions the use of Weighted ARM to mine association rules with a weighted support and confidence framework. The effectiveness of the Weighted ARM algorithm can be affected by factors like the selection of weight assignment strategy and the determination of threshold values for support and confidence. It is essential to carefully choose and optimize these parameters to attain optimal performance and ensure the effectiveness of the WAC classifier. Appropriate selection and optimization of these parameters play a critical role in achieving optimal performance and ensuring the effectiveness of the WAC classifier.

Sharan Monica L et al. [30] used data mining approaches in the WEKA tool to forecast heart disease using fewer variables. They explored different ways of data mining, including J48, Naive Bayes, and CART, to generate predictive models. J48 uses the C4.5 decision-making algorithm, the Java-based implementation of this solution is open-source, allowing for easy access and customization and Naive Bayes works well with continuous data. CART provides clear visualization of data linkages. All three decision tree methods were used in WEKA, with CART providing the best accuracy and J48 being the fastest to construct. The study aimed to refine the exactness of heart disease diagnosis while minimizing the number of variables used in a process. The author explores different data mining techniques, including J48, Naive Bayes, and CART. The choice of these techniques and their configuration parameters can significantly affect the performance and interpretability of the predictive models. It is important to carefully select and evaluate the appropriate data mining techniques for heart disease forecasting, considering their strengths and limitations in handling different types of data and capturing complex relationships.

Buettner, R. and Schunter, M. [32] described a methodology utilizing the Random Forests algorithm to detect potential heart disease. On a global scale, cardiovascular diseases continue to be a prominent cause of mortality, resulting in an estimated 17.9

million deaths attributed to these ailments. This machine learning approach contributes to the healthcare field by utilizing clinical data and test results from diverse patients to identify the existence of cardiac infraction. The outcome & significance of this research lie in its ability to decide whether a patient has cardiac disease dependent on specific medical data and experimental results, providing valuable support to doctors in making knowledgeable resolution regarding patient treatments. The author should discuss the integration of the Random Forests algorithm into the clinical workflow and the challenges associated with its adoption by healthcare professionals. Considerations such as usability, interpretability, and acceptance by medical practitioners are crucial for successful implementation and utilization of the algorithm in real-world clinical settings. Also address the potential ethical considerations and bias associated with the use of the Random Forests algorithm in healthcare. Machine learning algorithms can inadvertently perpetuate biases present in the training data, leading to unequal or unfair treatment of certain patient groups. It is essential to carefully consider and mitigate these biases to ensure the algorithm's fairness and equitable application in diagnosing heart disease. The summary of above survey has been alleged in Table 1.

Table 1: Summarization of Distinctive techniques used for Diagnosis of Cardiovascular Disease

Techniques	References	Approach Used	Drawbacks/ Limitations
Naive Bayes	[15], [30]	• Data mining techniques explored for forecasting heart disease	1. Lack of feature interactions 2. Limited expressiveness
J48	[30]	• Builds a decision tree by recursively partitioning the dataset	1. Overfitting 2. Difficulty in handling continuous or numeric attributes. 3. Lack of interpretability
Random Forest	[19], [20], [22], [32], [35]	• This approach utilizes ensemble learning, which involves the combination of multiple decision trees, to identify and make predictions for heart disease.	1. Computational complexity 2. Imbalanced data 3. Sensitivity to noisy or irrelevant features
SVM (Support Vector Machine)	[16], [19], [22]	• It is one of the techniques for diagnosing cardiovascular disease. • It is widely utilized for classification tasks in supervised machine learning.	1. Complexity with large datasets. 2. Sensitivity to parameter tuning.

This review was conducted based on multiple research articles, and the following are the drawbacks of several popular machine learning algorithms:

1. Naive Bayes can assume features to be independent, which is not always the case in real-world problems.
2. It cannot handle missing data, and it may suffer from zero probabilities. Decision Trees can overfit, and even small changes in data can lead to a completely different tree.
3. They may not work well on imbalanced datasets or with continuous values.
4. A considerable quantity of data is necessary to adequately train the Multilayer Perceptron, as insufficient data can result in overfitting.
5. It is sensitive to feature scaling, and training time may be lengthy for large datasets.
6. Random Forests are difficult to interpret, require more memory and computational resources than decision trees, and may not work well on imbalanced datasets or with continuous values. SVMs may be sensitive to input feature scaling and require time to choose optimal kernel functions and parameters.
7. imbalanced datasets, ANNs might not perform effectively. They demand a substantial volume of data for training, which can lead to potential overfitting. Furthermore, ANNs are challenging to interpret and computationally expensive, often requiring substantial memory resources.
8. In comparison, Logistic Regression and K Nearest Neighbor have fewer drawbacks, making them more favorable in some situations.

III. Discussion

This review paper highlights the challenges faced by various methods and algorithms used in cardiovascular disease diagnosis systems. These challenges include the difficulty in selecting appropriate kernel solutions for large datasets, long training processes, and challenges in interpreting and comprehending the results due to variable weights not being fixed.

Logistic regression and K-nearest neighbor (KNN) are effective machine learning algorithms for classification problems. Logistic regression is suitable for problems with few features and can handle different types of data, providing insight into the importance of each feature. KNN is a non-parametric method that can handle both continuous and categorical data without making assumptions about data distribution. Both algorithms have exhibited promising capabilities in the diagnosis and management of heart disease, encompassing risk prediction and patient classification based on factors like blood pressure, and cholesterol levels, age.

To overcome the drawbacks of other methods, Logistic regression and KNN can be used together. These algorithms are easy to implement and interpret, and their training processes are highly efficient. They can quickly classify unknown data records and are robust to noise in training data.

Evaluation of the suggested approach is crucial and should compare its effectiveness to tried-and-true approaches using suitable evaluation criteria. Both training and testing accuracy must be evaluated. In conclusion, using Logistic regression and KNN together can improve the accuracy and effectiveness of heart disease diagnosis and treatment. When selecting an algorithm, it is crucial to carefully evaluate and acknowledge the advantages and constraints of each one, ultimately opting for the most suitable option that aligns with the specific problem at hand.

IV. Conclusion

Heart disease is a remarkable aspect of deaths in a healthcare domain, and timely detection and treatment can enhance patient outcomes and reduce costs. Traditional methods of diagnosis and treatments are time-consuming, expensive, and prone to errors. Automating the heart disease diagnosis system can overcome these limitations and provide more accurate and timely diagnosis and treatment Many existing methods and algorithms for heart disease diagnosis have limitations, such as difficulty in selecting the appropriate kernel solution function and lengthy training processes for large datasets. Additionally, interpreting the results can be challenging due to variable importance and individual considerations. Logistic regression and K-nearest neighbor (KNN) machine learning algorithms offer viable solutions to overcome the limitations mentioned above and can be effectively utilized in the diagnosis of heart disease. Logistic regression, well-suited for problems with a limited number of features, demonstrates its versatility by handling diverse types of data while providing valuable insights into the importance of each feature. On the other hand, KNN, a non-parametric approach, is capable of accommodating both continuous and categorical data without imposing assumptions about data distribution. These algorithms possess the potential to forecast the probability of heart disease and classify patients into different risk categories based on significant risk factors. To assess the algorithms' effectiveness, appropriate evaluation criteria should be employed, and the model's accuracy in both training and testing phases must be thoroughly evaluated . Furthermore, by integrating these algorithms with complementary techniques, it is possible to further enhance accuracy and effectiveness in a prediction of heart disease.

In summary, automating cardiovascular disease diagnosis can improve patient outcomes and reduce healthcare costs. Machine learning algorithms such as logistic regression and KNN can be effective tools for heart disease diagnosis, and combining them with other methods can improve accuracy and effectiveness.

V. REFERENCES

[1] Nikam, A., Bhandari, S., Mhaske, A., & Mantri, S. (2020). *Cardiovascular Disease Prediction Using Machine Learning Models. 2020 IEEE Pune Section International Conference (PuneCon).* doi:10.1109/punecon50868.2020.9362367

[2] M. Abdar, U. R. Acharya, N. Sarrafzadegan and V. Makarenkov, "NE-nu-SVC: A New Nested Ensemble Clinical Decision Support System for Effective Diagnosis of Coronary Artery Disease," in IEEE Access, vol. 7, pp. 167605-167620, 2019, doi: 10.1109/ACCESS.2019.2953920.

[3] Mohan, S., Thirumalai, C., & Srivastava, G. (2019). *Effective Heart Disease Prediction using Hybrid Machine Learning Techniques. IEEE Access, 1–1.* doi:10.1109/access.2019.2923707

[4] Alalawi, H.H. and Manal, S.A., 2021. Detection of Cardiovascular Disease using Machine Learning Classification Models. International Journal of Engineering Research & Technology, ISSN, pp.2278-0181.

[5] Chang, V., Bhavani, V.R., Xu, A.Q. and Hossain, M.A., 2022. An artificial intelligence model for heart disease detection using machine learning algorithms. Healthcare Analytics, 2, p.100016.

[6] Mohan, S., Thirumalai, C. and Srivastava, G., 2019. Effective heart disease prediction using hybrid machine learning techniques. IEEE access, 7, pp.81542-81554.

[7] Mrudula Gudadhe, Kapil Wankhade, and Snehlata Dongre, Sept 2020, "Decision Support System for Heart Disease Based on Support Vector Machine and Artificial Neural Network", International Conference on Computer and Communication Technology, DOI:10.1109/ICCCT.2010.5640377, 17-19, *IEEE access*.

[8] Nikhar, S. and Karandikar, A.M., 2016. Prediction of heart disease using machine learning algorithms. International Journal of Advanced Engineering, Management and Science, 2(6), p.239484.

[9] Mohan, S., Thirumalai, C. and Srivastava, G., 2019. Effective heart disease prediction using hybrid machine learning techniques. *IEEE access*, *7,* pp.81542-81554.

[10] Chen, A.H., Huang, S.Y., Hong, P.S., Cheng, C.H. and Lin, E.J., 2011, "HDPS: Heart Disease Prediction System", Computing in Cardiology, ISSN: 0276-6574, pp.557- 560.

[11] Meshref, H., 2019. Cardiovascular Disease Diagnosis: A Machine Learning Interpretation Approach. International Journal of Advanced Computer Science and Applications, 10(12).

[12] Yan, Y., Zhang, J.W., Zang, G.Y. and Pu, J., 2019. The primary use of artificial intelligence in cardiovascular diseases: what kind of potential role does artificial intelligence play in future medicine? journal of geriatric cardiology: JGC, 16(8), p.585.

[13] Shaikh Abdul Hannan, A.V. Mane, R. R. Manza, and R. J. Ramteke, Dec 2021, "Prediction of Heart Disease Medical Prescription using Radial Basis Function", IEEE International Conference on Computational Intelligence and Computing Research, DOI: 10.1109/ICCIC.2010.5705900 ,28-29.

[14] Asha Rajkumar, and Mrs. G. Sophia Reena, 2020, "Diagnosis of Heart Disease using Data Mining Algorithms", Global Journal of Computer Science and Technology, Vol. 10, Issue 10, pp.38-43, September.

[15] Gunsai Pooja Dineshgar, and Mrs. Lolita Singh, February 2021, "A Review on Data Mining for Heart Disease Prediction", International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE), Vol. 5, Issue 2, pp. 462-466. Jyoti Soni, Uzma Ansari, Dipesh Sharma, and Sunit

[16] Harimoorthy, K. and Thangavelu, M., 2021. Multi-disease prediction model using improved SVM-radial bias technique in healthcare monitoring system. Journal of Ambient Intelligence and Humanized Computing, 12(3), pp.3715-3723.

[17] Mathur, P., Srivastava, S., Xu, X. and Mehta, J.L., 2020. Artificial intelligence, machine learning, and cardiovascular disease. *Clinical Medicine Insights: Cardiology, 14, p.1179546820927404, IEEE access*.

[18] Asha Rajkumar, and Mrs. G. Sophia Reena, 2020, "Diagnosis of Heart Disease using Data Mining Algorithms", Global Journal of Computer Science and Technology, Vol. 10, Issue 10, pp.38-43, September.

[19] Ghumbre, S.U. and Ghatol, A.A., 2012. Heart disease diagnosis using machine learning algorithm. In Proceedings of the International Conference on Information Systems Design and Intelligent Applications 2021 (INDIA 2021) held in Vishakhapatnam, India, January 2021 (pp. 217-225). Springer, Berlin, Heidelberg.

[20] AH Chen, SY Huang, PS Hong, CH Cheng, and EJ Lin, 2021, "HDPS: Heart Disease Prediction System", Computing in Cardiology, ISSN: 0276-6574, pp.557- 560.

[21] Dinesh, K.G., Arumugaraj, K., Santosh, K.D. and Mareeswari, V., 2018, March. Prediction of cardiovascular disease using Machine learning algorithms. In 2018 International Conference on Current Trends towards Converging Technologies (pp. 1- 7), IEEE access.

[22] Buettner, R. and Schunter, M., 2019, October. Efficient machine learning based detection of heart disease. In 2019 IEEE international conference on E-health networking, application & services (HealthCom) (pp. 1-6).

[23] Haq, A.U., Li, J.P., Memon, M.H., Nazir, S. and Sun, R., 2018. A hybrid intelligent system framework for the prediction of heart disease using machine learning algorithms. Mobile Information Systems, 2019.

[24] Chowdary, G.J., Suganya, G. and Premalatha, M., 2020. Effective prediction of cardiovascular disease using cluster of machine learning algorithms. Journal of Critical Reviews, 7(19), pp.1865-1875.

[25] Gunsai Pooja Dineshgar, and Mrs. Lolita Singh, February 2019, "A Review on Data Mining for Heart Disease Prediction", International Journal of Advanced Research in Electronics and Communication Engineering , Vol. 5, Issue 2, pp. 462-466.

[26] Chen, H., Zhang, L., & Yan, F. (2020). A review on machine learning methods for cardiovascular disease diagnosis and prediction. BioMed Research International, 2020.

[27] Shah, D., Patel, S. and Bharti, S.K., 2020. Heart disease prediction using machine learning techniques. SN Computer Science, 1(6), pp.1-6.

[28] Asif, M., Nishat, M.M., Faisal, F., Dip, R.R., Udoy, M.H., Shikder, M. and Ahsan, R., 2021. Performance Evaluation and Comparative Analysis of Different Machine Learning Algorithms in Predicting Cardiovascular Disease. Engineering Letters, 29(2).

[29] Jyoti Soni, Uzma Ansari, Dipesh Sharma, and Sunita Soni, June 2021, "Intelligent and Effective Heart Disease Prediction System using Weighted Associative Classifiers", International Journal on Computer Science and Engineering (IJCSE), Vol. 3, No. 6, pp. 2385-2392.

[30] kkus, Z., Galimzianova, A., Hoogi, A., Rubin, D. L., & Erickson, B. J. (2017). Deep learning for cardiovascular imaging: a review. Journal of the American College of Radiology, 14(5), 652-662.

[31] Oskouei, A. E., Bahrami, S., & Shiri, I. (2021). Machine learning-based approaches for cardiovascular disease diagnosis: a comprehensive review. Artificial Intelligence in Medicine, 113, 102044.

[32] Motlagh, M. A. Z., Farahani, A. V., & Langarizadeh, M. (2020). Application of machine learning techniques in the diagnosis of cardiovascular disease: A comprehensive review. Journal of Medical Systems, 44(5), 86.

[33] Sharan Monica L, and Sathees Kumar B, February 2022, "Analysis of Cardiovascular Heart Disease Prediction Using Data Mining Techniques", International Journal of Modern Computer Science, ISSN: 2320- 7868(Online), Vol. 4, Issue 1, pp. 55-58.

[34] Zhang, Q., & Wang, L. (2021). A systematic review of machine learning methods for cardiovascular disease prediction.

Computers in Biology and Medicine, 130, 104236.

[35] Alomari, R. S., Al-Madi, N. M., & Al-Tahat, H. (2019). Application of Machine Learning Techniques in Cardiovascular Disease Diagnosis: A Systematic Review. Journal of Healthcare Engineering, 2019, 1-15.

[36] Alizadehsani, R., Habibi, J., Hosseini, M. J., & Ebrahimzadeh, A. (2020). Cardiovascular disease diagnosis using machine learning algorithms: A review. Journal of Medical Signals and Sensors, 10(2), 67-80.

[37] Zabihi, M., & Lashkari, A. H. (2020). A comprehensive review of machine learning techniques for cardiovascular disease diagnosis. Computer Methods and Programs in Biomedicine, 188, 10525

[38] Géron, A. (2019). Hands-on machine learning with Scikit-Learn, Keras, and TensorFlow: Concepts, tools, and techniques to build intelligent systems. O'Reilly Media, Inc.

[39] K. Pahwa and R. Kumar, "Prediction of heart disease using hybrid technique for selecting features", *2019 4th IEEE Uttar Pradesh Sect. Int. Conf. Electr. Comput. Electron. UPCON 2019*, pp. 500-504, Jun. 2019.