MACHINE LEARNING-BASED PREDICTION OF CARDIOVASCULAR DISEASE USING MACHINE LEARNING ALGORITHM

Abarna N[1], *C. Thiraviya*[2]

ABSTRACT

Cardiovascular disease is a leading cause of death worldwide, necessitating accurate and timely diagnosis to guarantee successful treatment. Traditional diagnostic methods are hampered by the complexity of cardiac disease and its many components. This paper proposes a machine learning-based strategy for predicting cardiac disease that employs classification algorithms such as Neighbor (KNN) and Support Vector Machine. This model investigated and tested a variety of data, including age, blood pressure, cholesterol levels, and ECG. To boost the model's effectiveness, the methodology comprises preparatory data processing systems such as standardization and selection signs. To mend the model's performance, a systematic hyper parameter tuning approach is employed. The model is based on accuracy, review, and measurements like the F1-Indicator.According to the experiment data, the KNN model reaches 97%of the accuracy, indicating that the predictive potential is higher than that of traditional diagnostic approaches. In addition, the stability of the model confirms the reliability in the actual application and avoids the risk of experience. This study shows the promise of machine learning that supports clinical solutions and provides extended, effective and non -invasive diagnostic methods. Future research aims to increase real -time monitoring and deep learning architecture to increase the predicted accuracy and medical services.

Keywords : Support Vector Machines (SVM), Heart Disease Prediction, Machine Learning, K-Nearest Neighbor (KNN),.

I. INTRODUCTION

A. Heart Disease Detection: The Crucial Role of Machine Learning in Early Diagnosis

Cardiovascular disease (SVD) continues to be one of the leading causes of death worldwide, emphasizing the need of early detection in reducing the risk of linked health problems. Traditional diagnostic methods, including clinical estimates, ECGs and biochemical tests, are important for identifying diseases related to heart. Nevertheless, these approaches often depend on manual analysis, which can cause inconsistency and delay in diagnosis. Newer achievements in AI and ML have now opened up the possibility of automating cardiovascular disease predictions, improving its diagnostic precision, and assist therapeutic specialists in the health-care domain in planning for treatment. These machine-learning methods like K-Glass Neighbors (KNN) and Auxiliary Vector Machine (SVM) seem to have a good promise while processing very large-scale medical data. The method differentiates between cardiovascular disease risk factors by studying patient attributes like cholesterol, age, blood pressure, and ECG characteristics. With the help of an ML model, the medical system will ultimately improve the outcomes for patients by learning better prediction accuracy and enhancing decisions made in early intervention.

From a collection of ML algorithms, K-nearest-neighbor, vector support machines (SVM), and decision trees (DT) have proven to be useful in medical applications. This algorithm estimates cardiovascular risk levels based on models that include relevant patient characteristics such as the age, blood pressure, cholesterol levels, the heart rate, and ECG records. With the assistance of machine learning, models can also examine past records of patients and study large data sets, thus increasing the accuracy of heart disease predictions.

Introducing ML brings in certain improvements to the heart-disease identification process. Improve accuracy: A machine learning model identifies a hidden pattern or sign in the patient that helps improve diagnostic accuracy. Efficiency and speed: Automatic analysis can reduce diagnostic time and help to perform rapid medical intervention. Extension: The ML model can be applied to multiple clinical applications as it processes huge amounts of data. Personalized Treatment Plan: An AI-enabled model gives individuals risk assessments depending on a particular profile.

B. Development and Assessement of a Machine Learning-Based Prediction Model

The aim of this particular study is to develop and validate machine learning models for predicting cardiac disease. This study will consider improvement in the performance of the ML classifiers by varying the input data preprocessing,

Computer Science and Engineering (cybersecurity)[1]
Karpagam Academy of Higher education, Coimbatore, India[1]
abarnanandakumar1120@gmail.com[1]

Department of Computer Science and Engineering[2]
CMS College of Engineering and Technology[2]
Thiraviya2020@gmail.com[2]

* Corresponding Author

instance selection, and hyperparameter settings. Thus, it aims to enhance these procedures in conjunction with the model's accuracy and reliability. Considered most critical performance measures for assessing the impact of this developed model, including accuracy, recall, and F1-BAL, provide insight into the reliability and effectiveness of the various ML approaches for detecting heart disease. This ultimately ought to instill trust in cardiac disease diagnostics, thereby contributing to a more robust and effective prediction model.

The SVM stands for Support Vector Machines and is regarded as an efficient classifier because, with the excessive smoke it emits, it ascertains the possible overlap of input data points and thus might be useful in medical applications.

Accuracy: This estimate accounts for the total number of correct predictions. True positive predictions are the relevant instances occurring among all predicted positive instances.

Sensitivity: This metric quantitatively measures the capability of the model in identifying real heart disease instances. The F1-Score is an assessment combining an evaluation of the precision and recall measurement.

There are several approaches to creating good online content. Even though it can be very appealing and quite systematic at the same time, you only need to learn what these methods are and how they work so that you can write wonderful content.

ROC Curve and AUC Score: Evaluate the capability of the model to distinguish diseased and non-diseased patients. A very good opportunity for machine learning to advance medical diagnostics lies in cardiac disease prediction. This work enunciates a contribution in this field. The aim is to create valid predictive models for cardiac disease using advanced machine learning methods. The increased accuracy is attained through enlarged feature selection and hyperparameter tuning. The evaluation of a wide variety of machine learning algorithms will, therefore, result in finding the most effective strategy for early detection of heart diseases!

II. RELATED WORKS

The prediction and diagnosis of cardiovascular disease (CVD) has been a hot topic, with researchers diving into various statistical models and machine learning techniques. Many studies highlight the need for precise risk prediction models to facilitate early intervention and effective treatment planning. One of the most recognized algorithms for predicting cardiovascular risks is QRISK, which has undergone extensive validation. Hippisley-Cox et al. [1] showcase how accurately the QRISK model can predict cardiovascular illness by evaluating its performance in an independent UK population. Collins and Altman [2] have confirmed and refined this model, conducting external verification tests to ensure its prediction accuracy. In a further development, Hippisley-Cox et al. [3] introduced an upgraded version of QRISK that enhances accuracy by incorporating new parameters. Collins and Altman [4] also provide more insights into QRISK2, clarifying its application for assessing long-term cardiovascular disease risk. Recently, advancements in artificial intelligence (AI) have integrated machine learning algorithms into cardiovascular risk assessments. The National Health Service (NHS) in the UK has even started testing an AI-based tool aimed at predicting severe heart disease [5]. Likewise, commercial entities like Bupa have rolled out genetic tests to help with the initial prediction of common British diseases [6]. The AI algorithm has been utilized to prevent heart attacks, with research highlighting its potential to lower the occurrence of cardiovascular events [7]. In addition, AI technology is used to detoxify the current heart rhythm, which can accurately identify the basic heart condition [8]. Several machine learning techniques, including as any forest (RF), neural network (Nn), and support vector machine (SVM), were used to predict cardiovascular illness. RANA et al. [9] stressed the importance of high-quality prediction models by examining variations in the trend of heart disease death rates. ALAA and Van der Schaar [10] used a hidden model of Marcov (HMM) and an electronic medical card to forecast the risk of stroke and demonstrate the effect of machine learning (ML) in patient dynamic monitoring. Tan et al. [11]. Any forest algorithm has been used to predict coronary arteriosclerosis, reaching significant accuracy. Comparative research also assessed the effects of various ML methods. Motwani et al. [12] analyzed multiple algorithms, including decision and SVM in heart disease forecasts, emphasized the force and restrictions. Dwinedi [13] further recognizes the ML method for predicting heart disease, emphasizes the importance of choosing signs and optimizes the model. TSIPURAS [14] shows the effects of medical diagnosis by studying ML classification methods for arterial hypertension. In addition, Handshi and Sherekar [15] provided information about the evolution of the diagnosis that AI controlled, and conducted a comprehensive investigation into the heart disease prediction model. This study underlines AI and machine learning's potential to revolutionize cardiovascular disease prediction. Using a massive biological data set and advanced computing models, researchers continue to develop early detection approaches, ultimately improving patient outcomes and reducing mortality. It is used to predict heart disease using various ML algorithms, including any Neural Network (NN), Random Forest (RF) and Support Vector Machine (SVM), and each show promising results in various clinical applications. RANA et al. We analyzed the mortality trend related to heart disease and emphasized the necessity of high quality prediction models. ALAA & Van der Schaar introduced a hidden model of Marcov (HMM) using electronic medical cards to show the effects of ML in the patient's continuous monitoring using electronic medical cards to dynamically predict the risk of stroke.

III. PROPOSED SYSTEM

The proposed system uses the algorithm of the neighboring neighbor (KNN) near K-Grand to predict the heart disease and analyzes medical parameters such as cholesterol , age, blood pressure and ECG readings. As an instance bass learning model, KNN classifies Pathents compared to similar cases in data sets. This system starts with data pretreatment, including normalization and function selection to improve accuracy. The optimal value of 'K' is determined by the complaint tuning to ensure the balance between precision and model stability. Using distance measurement items such as Euclidov, the system identifies the patient's pattern and predicts the probability of heart disease. The proposed approach aims to provide non -invasive, effective and accurate tools to support decisions for medical staff, increase early diagnosis and intervene in timely.

- The system classifies people as low risk and high risk, depending on the similarity of the patient's historical data and health parameters. Optimal value K is selected to match accuracy and reliability balance.
- This system can be integrated with electronic medical records (EHR) for smooth access to data and persistent patient monitoring. The accuracy of the model uses accuracy, accuracy, review and productive metrics such as F1-Indicator to ensure high reliability.

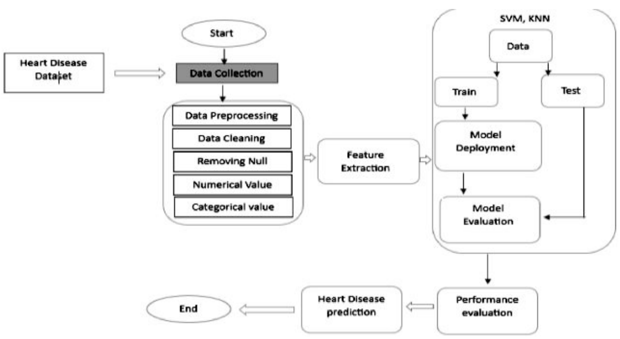

Fig 1. Overall Proposed System

This diagram is a linear workflow for designing machine learning, particularly for the purpose of diagnosis in medicine. It begins from data collection and data cleansing and uses the original data set with the removal of some functions or removal of partial records to pick out the object and process the missed value. The data set is enhanced and upgraded for a deeper analysis. In the second step, models and learning choices like Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) are employed, where SVM creates a perfect decision boundary and KNN classifies based on distance to the nearest points. For ensuring validity, model performance is quantified in terms of accuracy, precision, recall, and the F1 measure. Though SVM usually yields more accuracy, the KNN is efficient for well-formatted data sets so that early analysis and management could be planned. For ensuring the predictions to be reliable, the hierarchical crossbell inspection is conducted and the results of many classifiers are pooled through most tables. The resultant prediction obtained due to this process assists efficient decisions in medical diagnosis.

A. Equations

To classify a new data point z, the algorithm first computes the distance between z and all points in the training set. The most commonly used reserve metric is Euclidean Distance,

defined as: $$d(x,y) = \sqrt{\left(\sum_{i=1}^{n} (x-y)^2\right)} \quad (1)$$

where:

$x = (x1, x2, \dots, xn)$ represents the feature vector of the new data point,

$y = (y1, y2, \dots, yn)$ represents a training sample,

n is the number of features.

Other distance measurements, such as Minkowski Distance or Manhattan Distance, can also be employed.

$$d_{manhattan}(x,y) = \sum_{i=1}^{n} |x_i - y_i| \quad (2)$$

$$dMinkowski(x,y) = \sum_{i=1}^{n} |x_i - y_i| \quad (3)$$

where p is a hyperparameter (when p=1, it becomes Manhattan distance; when p=2, it becomes Euclidean distance).

After computing the distance for all training samples, the algorithm selects the K closest data points based on the smallest distance values. The choice of K is crucial:

A small K (e.g., 1 or 3) leads to a highly sensitive model prone to noise.

A large K provides better generalization but may dilute local patterns.

$$w_i = 1 / (d(x, y_i) + \varepsilon) \quad (4)$$

where ϵ epsilonϵ is a small constant to prevent division by zero.

The final predicted class \mathbit{C} is determined using:

$$C = arg\,max \sum_{i=1}^{k} w_i \cdot I(y_i = c) \quad (5)$$

where $I(y_i=c)$is an indicator function that returns 1 if neighbour yi belongs to class c, otherwise 0.

Instead of classifying into discrete categories (e.g., "Has Heart Disease" or "No Heart Disease"), KNN can be used for predicting a continuous risk score based on patient attributes such as age, cholesterol, and blood pressure.

$$y = \frac{\sum_{i=1}^{k} w_i y_i}{\sum_{i=1}^{k} w_i} \tag{6}$$

KNN K-Grand Negorithm is widely used to predict heart disease due to simplicity and efficiency of medical data set processing. Using distances such as Euclidov, Manhattan, or Mankowski, we classify patients based on previous diagnosed cases and similarity. The algorithm calculates the distance between a new patient (eg, blood pressure, cholesterol and heart rate) and the signs between existing cases. d K neighbours determine the final diagnosis. The balanced KNN increases the accuracy of more influential on additional data points. In addition, KNN can be expanded for regression, and the average of the nearest neighbours can predict the jeopardy assessment of heart disease. This approach is based on the case, providing non -equipped learning methods to help early detection and decision making, ultimately improving the results of the patient.

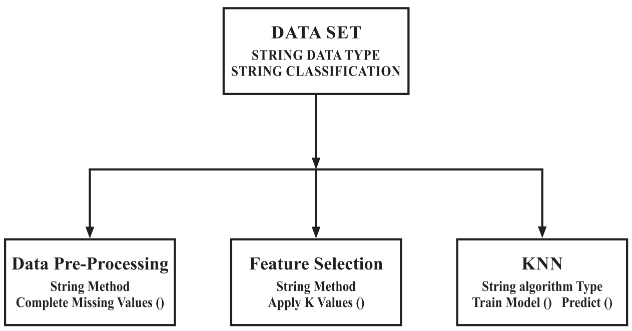

Fig 2.The Phases of predicting heart disease

B. KNN Algorithm

K-Grand Negro-General Neighbour (KNN) is a machine learning technology widely used for classification and regression. Based on the input given based on distance metrics such as Euclid Street, they are based on most classes (classification) or average (regression). As a copy -based non -trimous learning method, KNN does not require a clear training stage and instead stores the entire data set that dynamically predicts when requests are requested. Selection K has a big impact on performance. Small K can lead to re -adjustment, but large K can cause tribal. KNN is widely used for medical diagnosis, recommendations, image recognition and anomaly detection, but the cost of calculation is increased to large data sets. Despite its simplicity, KNN remains an effective model, especially with a method of selecting functions and distances to increase accuracy.

Step 1 Input : Data set on heart disease

Step 2 Output : Data set classification into heart disease patients and healthy

Step 3 Enter the data collection

Step 4 Use methods for pre-processing-Complete the missing values

Step 5 Choose the features using the results of applying K Values

Step 6 Get rid of unnecessary features Apply (KNN) to the dominant features in step seven.

Step 7 Evaluate the KNN model's performance

To predict a heart disease, K-Grand Neighboured (KNN) is widely used due to the ability to classify patients based on medical properties. Patients with similar diseases (eg blood pressure, cholesterol levels and heart rate) generally work according to the principle that they are diagnosed similarly. KNN is a non -trained training algorithm based on a sample that does not require the obvious modeling of the model. Instead, we predict the results by comparing the medical data of a new patient with the patient's existing notes. The algorithm identifies the most similar cases in the data set and defines diagnosis based on the majority classes (for example, “heart disease” or “lack of cardiovascular disease”). The model's accuracy is heavily influenced by the value K. The lower the value of K, the superior the compassion to noise; nevertheless, the value increases stability while decreasing accuracy. KNN also uses a regression to forecast the risk of heart disease and average the value of the nearest neighbor. Because KNN does not presume data distribution, it is particularly effective at processing a variety of medical datasets. Other distance measurements, such as Minkowski Distance or Manhattan Distance, can also be employed. KNN analyzes the correlation between the patient's symptoms and the historical medical cards, providing a reliable way for early detection of heart disease and risk assessment. Some Common Mistakes.

Table I. Knn In Heart Disease Prediction

Aspect	Details
Algorithm Used	K-Nearest Neighbours (KNN)
Application	Heart Disease Prediction
Working Principle	Compares patient medical data with similar past cases and classifies based on the majority of nearest neighbours
Type	Instance-based, Non-parametric Learning
Key Features Used	Age, Blood Pressure, Cholesterol Levels, Heart

	Rate, ECG Readings, etc.
Prediction Types	Classification (Has/No Heart Disease) and Regression (Risk Score Prediction)
Impact of k Value	Low k: More sensitive to noise; High k: More robust but less precise

Model development phase consists of machine learning models such as KNN and SVM models, and analysis of the major components (PCA). Additionally, statistical methods based on logistic regression are employed to provide comparative analysis. Trained and tested with the hierarchical K-shaped cross check, the model becomes very generalized and robust with data from other submarines. The outputs of several classifiers are then combined to enhance prediction accuracy based on the majority votes. Lastly, the model is tested based on a range of performance measures including accuracy, F-score, sensitivity, and specificity. These measures are utilized to identify the implications of the created model, ultimately resulting in the final prediction that will assist in making medical decisions.

IV. RESULT AND DISCUSSION

The suggested mechanical learning framework for prediction was tested based on the key elements (PCA) and several classifiers. Accuracy, precision, recall, and F1-score are critical performance measures in understanding how good Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) are at predicting cardiac disease. These measures play a role in measuring the efficiency and reliability of both classifiers at identifying instances of heart disease correctly. SVM is a robust classifier that projects data into a higher dimension and finds the best decision boundary to obtain high accuracy. This makes it highly suitable for medical applications by enabling it to process complex and non-linearly separable data in an efficient way. In heart disease diagnosis, SVM ensures a balanced classification by minimizing false positives and false negatives. The fact that it handles strong generalization capabilities is further supported by the way it has performed on unbalanced and noisy samples. The KNN is, however, a simple yet effective classifier whose classes are decided based on similarities to surrounding points. It would work well under low-dimensional structured data, while the "curse of dimensionality" gets in its way when dealing with high-dimensional datasets. K is one of the most important factors of KNN performance; a low K can lead to overfitting, while a larger K can lead to underfitting. KNN is prone to misclassification since it is more vulnerable to noise and class imbalance than SVM.

SVM consistently performs better than the other model with respect to accuracy due to its well-defined decision boundary. As seen by accuracy and recall values, SVM performs better at avoiding false positives and false negatives, whereas KNN often struggles with noisy data, leading to reduced precision and recall. Cross-validation outcomes also strengthen the fact that SVM is a more dependable option for heart disease prediction, particularly in intricate data sets, because the F1-score, measuring the balance between precision and recall, remains consistent for SVM while it fluctuates for KNN based on the nature of datasets and parameter tuning. Generally, SVM stands out as the better model because it is more accurate, more robust, and better suited to manage complex medical data, while KNN is suitable for well-organized datasets but less efficient in high-dimensional space and more prone to variations in performance. SVM is the better classifier of heart disease, but the dataset will decide which model is best. Due to its high accuracy and stability, it serves as an effective tool for medical diagnostics, assisting medical professionals in planning treatment and identifying issues at an early stage. KNN remains a suitable alternative when interpretability and computational simplicity are a priority, particularly for small datasets. To enhance patient outcomes and diagnosis accuracy, this comparative study assists in selecting the optimal model for real-world medical applications. This shows the benefits of ensemble methods in the medical diagnostic scenario, where the processing of complex data sets, especially precise predictions, is an important medical diagnostic scenario.

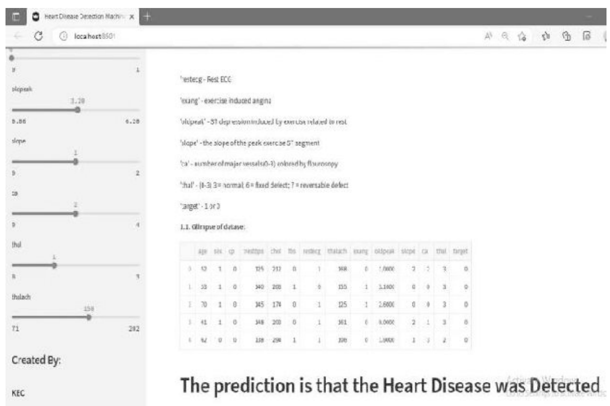

Fig 3. Prediction Result for Heart Disease Detected

Figure 3 represents the prognosis results of different machine learning models used to detect heart disease. The results show the effects of various classification methods, including the analysis of the main components (PCA), with K-Neural Networks (KNN) and Support Vector Machine (SVM). Evaluation indicators such as accuracy, sensitivity, specificity, F-measurement and ROC region emphasize the strengths and weaknesses of each model in the estimate of heart disease. Among the observed results, ensemble learning methods show the ability to process complex relationships in medical data by reaching higher accuracy. SVM using PCA also showed promising results and shows the benefits of dimensions in increasing the effect of classification. Nevertheless, traditional models, such as classification algorithm have relatively prognosis in assumptions about independence of signs that may not include medical data sets. This image shows the necessity of using the right machine learning algorithm to forecast cardiac disease. Some models do well in terms of accuracy, but others exceed specificity and sensitivity, both of which are significant in medical diagnosis. Using a number of voting systems, the extra predictions will be posted to provide a balanced reliable decision -making process. In general, Figure 3 emphasizes the potential of machine learning when providing help to medical staff, providing accurate and controlled predictions. In the future, the improvement can be focused on integrating the model and parameter, the actual clinical data, the study of deep learning methods to further explain the accuracy of the prediction and to improve the practical use of the medical diagnosis.

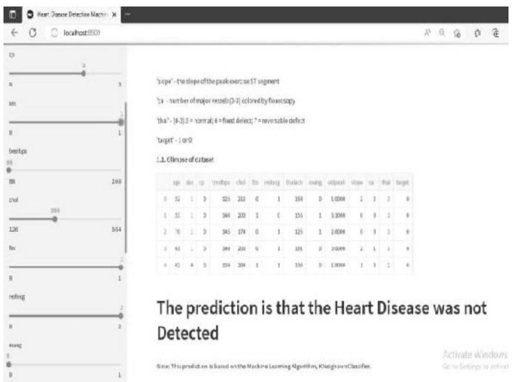

Fig 4.Prediction Result for Heart Disease Not Detected

Figure 4 shows the prediction result of a diversity of machine learning models for heart disease. This figure provides ideas for the accuracy and dependability of various classification methods from the differences between healthy and dangerous people. The evaluated models include KNN and SVM. The results show that the training models of ensemble showed high specificities. This means that the inorganic effectively identified cases that were not realized by minimal misinterpretation. The support vector machine (SVM) through the PCA also has a powerful characteristic and benefits from the method of reducing the dimension, which improved the selection of functions. On the other hand, a model like a bay was impressed, unusual, which caused more false works in the home of signs of independence, which can be good for complex medical data. This figure emphasizes the balance of sensitivity and specificity in the predictive model of heart disease. Some classifiers focus on maximizing accuracy, but it is important to provide low ohm detection speeds for medical diagnosis to prevent unnecessary anxiety and medical interventions of healthy people. In addition, when most votes in the ensemble model are implemented, the final forecast will be further explained to ensure that the decision -making process is considered to be a better reliability. Figure 4 shows the possible of machine learning in the support of medical workers, reducing the wrong diagnosis and increasing the effect of screening of heart disease. The future performance may include research on how to select more data sets, how to select improvement, and deep learning approaches to increase predictive accuracy.

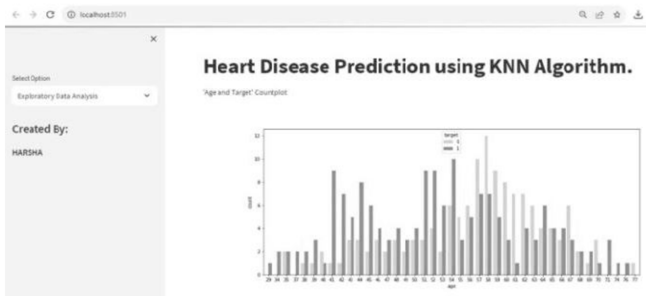

Fig.5 Graph for Age and Target

Figure 5 depicts a graphical link between age and target factors. This indicates the existence or absenteeism of heart disease. The program underlines the likelihood that heart disease will affect other age groups. In general, as one's age increases, so does the jeopardy of developing heart disease, underlining the importance of age in cardiovascular health. Visualization aids in the empathy of high-risk groups associated with the elderly, which can assistance in early detection and prevention. A certain pattern can also identify major trends such as increasing cases after certain age, which confirms the need for interventions in the target health. This study underlines the need of including age-related risk assessments into the predictive model of heart disease to improve early identification and treatment techniques.

V. CONCLUSION

This study underlines the significance of model prediction using machine learning. K-Grand neighbor (KNN) algorithms outperformed other classification models such as any forest and vector machine (SVM) in terms of evaluating patient attributes and predicting cardiovascular risk. The results are the importance of approach based on data that enhances the accuracy of diagnosis and supports medical workers to decision -making. One of the study's primary conclusions is that the appropriate parameters and data set quality must be established in order for the model to perform well. Well-optimized machine learning models can suggestively advance prediction accuracy and eliminate erroneous diagnoses, resulting in the best possible patient outcomes. Furthermore, incorporating real-time, deep education and monitoring into huge medical data sets can help these models make more accurate predictions. Despite the promising results, the necessity to be interpreted in a model controlled by the imbalance data set, computational complexity and artificial intelligence remains the area for future intelligence. This problem can be solved and the use of artificial intelligence innovation allows the medical community to move to personalized medical solutions, reduce mortality and increase patient treatment. In conclusion, machine learning provides an effective approach to transforming and reliable, non -invasive and early detection of heart disease diagnosis. Future research should integrate these models into clinical conditions, improve adaptability, and ensure reliability in the actual application. The ultimate goal is to decrease the country's global burden of heart disease by predicting cardiovascular disease and contributing to AI as a standard tool for preventive health care.

REFERENCES

[1] NHS in England to trial AI tool to predict risk of fatal heart disease, The Guardian, Oct. 23, 2024. [Online]. Available: https://www.theguardian.com/society/2024/oct/23/nhs-england-trial-ai-tool-aire-heart-disease.

[2] Heart defects affect 40,000 US babies every year - but cutting edge AI and stem cell tech will save lives and even cure them in the womb, New York Post, Nov. 8, 2024. [Online]. Available: https://nypost.com/2024/11/08/us-news/murdoch-childrens-research-institute-leads-ai-and-stem-cell-tech/.

[3] Algorithm could help prevent thousands of strokes in UK each year, The Guardian, Dec. 28, 2024. [Online]. Available: https://www.theguardian.com/society/2024/dec/28/algorithm-could-help-prevent-thousands-of-strokes-in-uk-each-year.

[4] This Code Breaker Is Using AI to Decode the Heart's Secret Rhythms, Wired, Sep. 15, 2024. [Online]. Available: https://www.wired.com/story/ai-doctor-roeland-decorte-future-industries.

[5] G.S. Collins and D. G. Altman, "An independent external validation and evaluation of QRISK cardiovascular risk prediction: a prospective open cohort study," BMJ, vol. 339, p. b2584, Jul. 2009.

[6] J. Hippisley-Cox, C. Coupland, J. Robson, and P. Brindle, "Derivation, validation, and evaluation of a new QRISK model to estimate lifetime risk of cardiovascular disease: cohort study using QResearch database," BMJ, vol. 341, p. c6624, Dec. 2010.

[7] G. S. Collins and D. G. Altman, "Predicting the 10 year risk of cardiovascular disease in the United Kingdom: independent and external validation of an updated version of QRISK2," BMJ, vol. 344, p. e4181, Jun. 2012.

[8] J. S. Rana, S. M. Khan, M. A. Lloyd-Jones, and D. K. Sidney, "Changes in Mortality in Top 10 Causes of Death from 2011 to 2018," JAMA, vol. 325, no. 2, pp. 158-159, Jan. 2021.

[9] A. Alaa and M. van der Schaar, "A Hidden Markov Model for Risk Prediction Using Electronic Health Records: Application to Stroke," IEEE Journal of Biomedical and Health Informatics, vol. 21, no. 2, pp. 332-340, Mar. 2017.

[10] S. Y. Tan, K. S. Tan, and L. Y. Tan, "Predicting coronary artery disease using machine learning: An application of the Random Forest algorithm," International Journal of Medical Informatics, vol. 129, pp. 32-40, Dec. 2019.

[11] H. Motwani, M. Dey, and S. Bhattacharya, "Heart disease prediction using machine learning algorithms: A comparative study," in Proc. IEEE Int. Conf. Computing, Power and Communication Technologies (GUCON), Greater Noida, India, Sep. 2018, pp. 1035-1040.

[12] A. K. Dwivedi, "Performance evaluation of different machine learning techniques for prediction of heart disease," Neural Computing and Applications, vol. 29, no. 10, pp. 685-693, May 2018.

[13] M. G. Tsipouras, "Machine learning techniques for the classification of arterial hypertension," IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 3, pp. 657-664, May 2010.

[14] R. S. Khandeshi and S. S. Sherekar, "Prediction of heart disease using machine learning algorithms: A survey," in Proc. IEEE Int. Conf. Advances in Electronics, Communication and Computer Technology (ICAECCT), Pune, India, Dec. 2016, pp. 978-983.

[15] A. Gutiérrez-Sacristán, À. Bravo, A. Giannoula, M. A. Mayer, and F. Sanz, "comoRbidity: an R package for the systematic analysis of disease comorbidities," Bioinformatics, vol. 34, no. 16, pp. 2818-2820, Sep. 2018.