Smart Healthcare System

1st Vishnu Kumar, 2nd Yashwant Raj, 3rd Harsh Chahar

School of Computing Science & Engineering Galgotias University

Greater Noida, India

1st vishnukoli020@gmail.com, 2nd yashr0202@gmail.com, 3rd chaharharshh30@gmail.com

Abstract – According to reports, heart disease, a subtype of cardiovascular disease (CVD), is the leading cause of death worldwide at the moment, according to the World Health Organisation (WHO). Over the past few decades, heart disease has been rightly referred to as a 'quiet killer' since it silently takes its toll, often remaining unnoticed until it is too late. In that time, many researchers have applied several machine leraning and data mining techniques to help healthcare professionals team to diagnose heart-related diseases. This research paper is deals with an overview of heart disease prediction using Key words: - heart disease, machine learning, CVD, WHO, heart disease prediction

I.   INTRODUCTION

A vital aspect of living is receiving healthcare. Your heart (the main organ in your cardiovascular system) is a vital organ. The human heart is considered to be one of the most important organs, because it is responsible for maintaining life. The human heart is one of the most vital organ needed to keep life alive. The human heart is one of the strongest and most productive muscles in the body, and it works continuously throughout an individual's life. Globally, cardiovascular diseases (CVDs) are the leading cause of death due to heart attacks and stokes . 32% of all fatalities worldwide in 2019 were attributed to CVDs, with an estimated 17.9 million deaths [1] including Support vector machine (SVM), Logistic Regression, K-Nearest Neighbour (KNN), Naïve Bayes, Decision Tree (DT), Neural Network (NN), XGBoost, and Random Forest (RF), several machine learning classification techniques.

Out of various machine learning and algorithm-oriented techniques or methods to determine, specify, and diagnose multiple heart diseases with certainty and precision, Random Forest (RF) has stood out as a prominent method for detection, supported by a series of experiments and observations that resulted in an optimistically accurate way.

Eighty-five percent of the deaths were from heart attacks and strokes. Globally, the number of deaths from cardiovascular disease (CVD) increased from 12.1 million in 1990 to 20.5 million by 2021, according to a recent World Heart Federation (WHF) estimate [2]. Making accurate and realistic predictions for heart- related conditions is therefore essential. Cardiovascular medicine is seeing a growing use of machine learning, or ML, a branch of artificial intelligence (AI). Additionally, a new generation of information technology known as "smart healthcare" has evolved. Smart healthcare is a multifaceted, complete transformation rather than just a straightforward technological achievement. Here, showing you how machine learning can be used for risk prediction as far as cardiovascular disease is concerned and how this is helping to make a difference in how cardiovascular disease and other types of chronic illnesses are being helped by the practice of machine learning. So, cardiovascular disease still is the number one killer over 12 million people die from cardiovascular diseases every day and what we have learned so far is that some of these deaths are preventable so there is quite research that is going on both in the lab and in the field and as we will see in the area of data science that is helping to make a difference. We are living in an era that involves more digitization. There are numerous of machine learning algorithms have been implemented into day-to-day operations. For example, Algorithms in machine learning such as SVM can identify non-linear patterns that could be applied for facial recognition, handwriting analysis, or the spotting of fraudulent credit card transactions [3]. The remainder of the article is combined as follows. Research paper problem statement is covered in Section II. The literature review in Section can also be given by us.

Research GAP is covered in Section IV. Section V offers Goals of our Research Paper. Section VI provides the experimental setup generating Exploring Data and the algorithms employed. Section VII presents the dataset Statistics. Section VIII presents a discussion on the stated method outcomes and proposed model benchmarking. Flowchart is covered in Section IX. Section X outlines the approach of paper how the experiment was run and the outcomes reached. Section XI addresses the results of the stated approach. At last, Section XII finishes with a summary of present work and several comments on next

improvement.

II.   PROLEM STATEMENT

The World Health Organization (WHO) reports that heart disease is responsible for causing the deaths of 17.9 million people every year, making it one of the most significant causes of death across the globe [4]. Detecting heart disease early is crucial, for improving results. Still, how can medical healthcare professionals be accurate for effectively pinpointing this condition? In this case, we had been considering one approach that involves examining patient data sets that include variables such as gender, age, blood pressure, cholesterol levels [5], and other relevant variables. By utilizing this data, we can implement a machine learning model that can accurately estimate the likelihood of heart issues. We have a data which classified if patients have heart disease or not according to features in it. This progress will empower doctors to recognize high-risk patients and interface early to lower their chance of likelihood of developing heart disease.

III.   LITERATURE REVIEW

[1]   Today in 2023On 2023 today Focussing on heart disease prediction via machine learning techniques, Mansi et al. released a paper titled "Heart Disease Prediction using Machine Learning". Based on the medical parameters supplied in the dataset, which clinicians can use for diagnosis [6], classification models predict heart disease using Logistic Regression, Support Vector Classifier (SVC), k-Nearest Neighbours (KNN), Neural Network (NN), and Random Forest (RF) Models.

[2]   [2]In their paper "Prediction of Heart Disease Based on Machine Learning Algorithms," published in 2023, Premkumar et al. discussed effective machine learning models for a heart disease prediction system that used classification techniques. Based on health factors, machine learning models— specifically a voting classifier that combines Random Forest and Logistic Regression— achieved the highest accuracy of 98.36% in predicting heart diseases. The highest accuracy, 98.36%, was achieved by the voting classifier. Voting ensemble technique overcomes model development challenges [7].

[3]   In 2023, Applied and Computational Engineering, published a paper titled "heart disease prediction based on machine learning algorithms" The article discusses the use of machine learning algorithms to predict heart disease. It compares different algorithms and evaluates their performance based on accuracy, precision, recall, and f1-score. Comparative study of Decision Tree, Logistic Regression, SVM, and Random Forest [8].

[4]   In 2024, Prof., Vinod et al. "Heart Disease

Prediction Using Machine Learning Algorithms" This study explores logistic regression, KNN, and random forest for heart disease prediction, enhancing risk assessment with machine learning algorithms and improving personalized treatment strategies in cardiovascular care. Investigates logistic regression, KNN, and random forest for heart disease prediction [9].

[5]   In 2024, Mohankumar et al. Published a paper titled "Precise heart: heart disease prediction using machine learning" The study demonstrates that the Support Vector Machine (SVM) technique effectively predicts heart disease, showcasing high accuracy and performance compared to other machine learning methods. Investigate accuracy levels of machine learning techniques (SVM, KNN, Naive Bayes, Decision Trees) for prediction [10].

[6]   In 2024, Syeda et al. Published a paper titled "A Hybrid Approach for Heart Disease Prediction using Genetic Algorithm and SVM" the study aimed that the SVM model, optimized through a genetic algorithm, achieved 98.0% accuracy in predicting heart disease, demonstrating its effectiveness in early detection and diagnosis. Utilizes genetic algorithm and SVM for heart disease prediction. Achieved 98.0% accuracy, outperforming other works in accuracy [11].

[7]   In 2023, Khandaker Mohammad et al. "Machine learning-based approach to the diagnosis of cardiovascular vascular disease using a combined dataset, Intelligence-Based Medicine, Volume 7" stated that The Decision Tree (DT) algorithm has been shown to have high accuracy, with one study reporting an accuracy of 99.16% [12].

[8]   In 2023, Ahmad AA et al. published online article "Prediction of Heart Disease Based on Machine Learning Using Jellyfish Optimization Algorithm. Diagnostics (Basel)" study aimed that The SVM classifier model has been shown to be the most accurate model in some studies. When combined with the Jellyfish optimization algorithm, the SVM- based Jellyfish approach has an accuracy of 98.47% [13].

[9]   Al-Alshaikh et al. Comprehensive evaluation and performance analysis of machine learning in heart disease prediction. *Sci Rep* **14**, 7819 (2024), described that the machine learning hybrid deep predictive model (ML-HDPM) has shown to be effective in predicting heart disease. It has a training accuracy of 95.5% and a testing accuracy of 89.1% [14]

M., S., Khan et al. (2024).With the use of Machine learning Algorithms we can easily pridict the Heart Diseases. Indian Scientific Journal Of Research In Engineering And Management [15]

- Similitudewith more different datasets.

- Research of additional machine learning algorithms.

Samer et al. (2024). 5.With the use of Machine learning Algorithms we can easily pridict the Heart Diseases. Science journal of University of Zakho [16]

- Enembles method hasn't been research much to predict heart diseases.

- Less Disscusion on the explianability of machine learning models.

Dina et al. (2024). 7. With the use of Machine learning Algorithms we can easily pridict the Heart Diseases. [17].

- Novel fast experiment for feature selection

- Lack of exploration with other current models

HIMANSHI et al. (2024). 10. With the use of Machine learning Algorithms we can easily pridict the Heart Diseases.and Deep learning Models [18].

- Technical nuances of model construction

- Ethical issues, privacy concerns, data scientist-medical expert relationship

By addressing following gaps, future research can contribute to more robust and effective heart disease prediction systems that are better suited to meet the needs of healthcare professionals and patients alike. Disquition of additional machine learning algorithms.

V. OBJECTIVES

To create a web application that can determine if a person has heart disease.

- Assess the accuracy of machine learning models in predicting heart disease.

- To Predict and monitoring all symptoms and illeness of cardiac disease using machine learning algorithms.

- Analyse medical data for early disease prediction.

- Improve precision and effectiveness of cardiovascular risk identification.

VI. EXPLORING DATA

We get an insight into the trends and the relationships that may exist in our data there we can do some hypothesis testing whether that's between two of the predictor variables or between a predictor variable and the outcome. This is where we're going to generate the charts, graphs, and other types of visualizations that would help you to understand what is the information contained in our data because looking at the numbers and data frame that is generally not going to give you all of the insights that is contained in there.

Figure 1. Heart disease prediction using various classification algorithms.
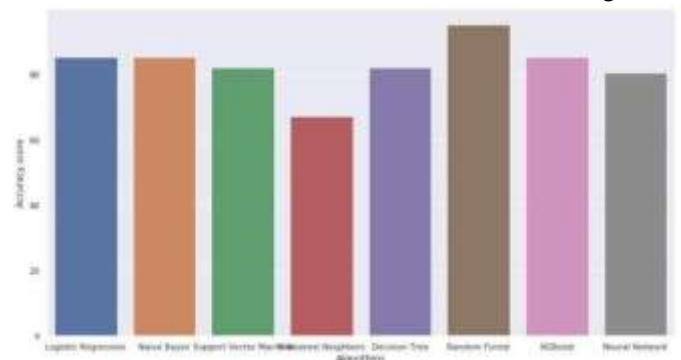

Figure 2. Random Forest algorithm achieved highest accuracy level of 95.08%

The accuracy score achieved using Logistic Regression is: 85.25 %
The accuracy score achieved using Naive Bayes is: 85.25 %
The accuracy score achieved using Support Vector Machine is: 81.97 %
The accuracy score achieved using K-Nearest Neighbors is: 67.21 %
The accuracy score achieved using Decision Tree is: 81.97 %
The accuracy score achieved using Random Forest is: 95.08 %
The accuracy score achieved using XGBoost is: 85.25 %
The accuracy score achieved using Neural Network is: 80.33 %

Figure 3. Dataset

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1

dataset.sample(5)

VI. STATISTICS

1. Global Health Concern – The paper highlight that Cardiovascular disease continues to be the leading cause of death, with approximately 12 million deaths each day.

2.   Machine Learning Algorithms used – The research implements various classification algorithms that applied in heart disease prediction like:

- Decision Tree,
- Naïve Bayes,
- Support Vector Machine (SVM),
- K-Nearest Neighbour (KNN),
- Random Forest,
- Logistic Regression,
- Neural Network, and
- XGBoost

Random Forest achieving the highest accuracy of 95.08% in predicting heart disease.

3.   Featured Analyzed – The study utilizes standard feature to enhance performance of the model that are critical in predicting heart disease. These features include:

- Age
- Sex
- Chest Pain (cp)
- Cholesterol Levels
- Blood Pressure (bp)
- Resting Electrocardiographic Result
- Fasting Blood Sugar (fbs), etc these features are important for establishing a predictive system.

5. The outcome of the prediction – the final result of using the method applied that is the key result of machine learning used in healthcare.

5.   ◦ Heart disease prediction using various classification algorithms.

6.   ◦ Random Forest algorithm achieved highest accuracy level of 95.08%.

7.   ◦ Systematic prediction can help patients manage health risks.

V I. PROPOSED SYSTEM

The suggested model improves performance and highlights the accuracy comparison between the various algorithms. The results show that, with an accuracy level of 95.08%, the Random Forest algorithm performed better than the others. The Random Forest is a crucial component of our suggested model since its high accuracy shows that it is an excellent predictor of heart diseases. The main objective of the suggested model is to develop a methodical prediction system that can help patients determine their risk of heart disease. Through the use of machine learning techniques, the model seeks to offer practical insights that can assist people in taking health-related preventive actions. Improving patient outcomes and

healthcare strategies requires this methodical approach.

IX. METHODOLOGY

This study is about computerized heart disease prediction that can be useful in healthcare of patients, healthcare professions and doctors. We develop a user-friendly web application that makes use of medical data to make predictions on the probability of heart disease to empower doctors and medical personnel. In this study, To achieve the goals,we use variety of Machine learning algorithm to a heart disease dataset. Our approch of heart disease pridiction in this paper consists of following steps – Data processing, Data collection, Traning and Testing, Model selection , Evaluation of performance and creation of a pricing system. The goal of each step is to increase the model's prediction accuracy and dependability.
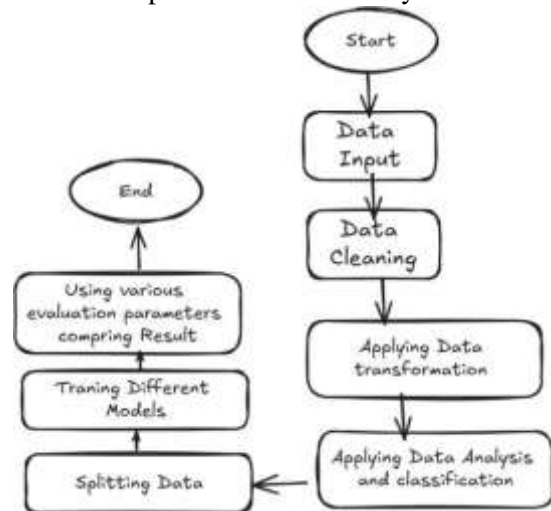

Figure 4: Model of flow diagram

9.   Data Sources

The dataset exploit in this study,as described in [19], Fixture of 70,000 patient records used.

In table 2 , 12 specific features used in this study , that includes systolic blood pressure , diastolic blood pressure , Age , sex. The target variable label as "cardio"thai shows that the patient has cardio vascular disease (marked as 1) or if healthy (marked as 0).
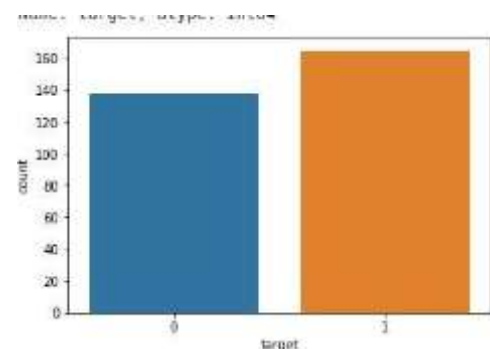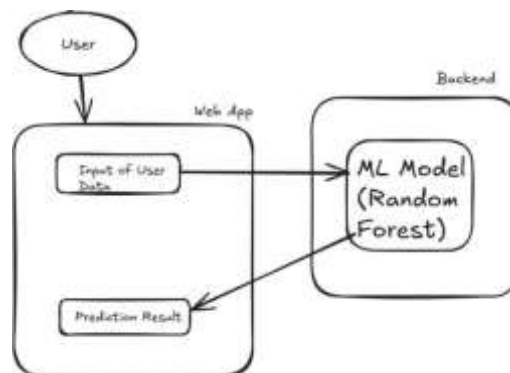
Figure 5. Data Set (Number of Patient)

Table. Datasets attributes.

Feature	Variable
Age	Age
Sex	Sex
Chest Pain	cp
The person's resting blood pressure	trestbps
Cholesterol	chol
Fasting Blood Sugar	fbs
Resting electrocardiogram (ECG)	restecgs
Maximum heart rate achieved	thalach
Exchange Transfusion	exang
Cardiac Arrest	ca

9.1 Removing Outliers

The dataset is manifest that are present in outliers.

Outlier detection – reduction of the number of outliers in data entry, which may (hopefully) improve our predictive model.

Issue: We removed all outliers (weight, and height) that did not fall within the 2. 5% – 95% threshold. That outlier identification and outlier elimination was done by hand.

9.2 Building Web Application

We develop a web application where users can input clinical attributes such as age, gender, cholesterol levels, and blood pressure. The web app sends this input to a backend server, which processes the data and communicates with a pre-trained Random Forest model. This model, which has been trained on heart disease data, provides a prediction based on the user's input. With an accuracy of 95%, the Random Forest model offers reliable predictions. The backend processes the input, runs it through the model, and returns the prediction, which is then displayed to the user in real-time.

CONCLUSION

To sum up, the Smart Healthcare System described in this study offers a novel way to improve patient care, monitoring, and general healthcare administration. This system improves patient outcomes and lowers hospital readmissions by combining IoT, AI, and data analytics to provide real-time health monitoring, predictive analysis, and personalized care recommendations. In addition to giving individuals more control over their health, this technology helps medical professionals make wise judgments. Expanding data privacy regulations, enhancing interoperability across different healthcare platforms, and upgrading algorithms for even more precise diagnosis could be the main areas of future research. In line with the changing demands of contemporary medical services, this smart healthcare framework is a major step toward effective, patient-centered healthcare solutions.

REFERENCE

[1]   Hon, Lye-Quen, et al. "Vascular closure devices: a comparative overview." *Current problems in diagnostic radiology* 38.1 (2009): 33-43.

[2]   Punetha, Mayank, et al. "Industrial Scale Production, Commercialization, and Global Market of Functionalized Carbon Nanostructures." *Handbook of Functionalized Carbon Nanostructures: From Synthesis Methods to Applications*. Cham: Springer International Publishing, 2024. 1-58.

[3]   Stamler, Jeremiah. "Epidemiology of coronary heart disease." *Medical Clinics that are present in North America* 57, no. 1 (1973): 5-46

[4]   Satapathy, Santosh Kumar, et al. "Heart Disease: Automatic Prediction from the Numerical and Categorical Features by Machine Learning Methods."

[5]   Hon, Lye-Quen, et al. "Vascular closure devices: a comparative overview." *Current problems in diagnostic radiology* 38.1 (2009): 33-43.

[6]   Punetha, Mayank, et al. "Industrial Scale Production, Commercialization, and Global Market of Functionalized Carbon

Nanostructures." *Handbook of Functionalized Carbon Nanostructures: From Synthesis Methods to Applications.* Cham: Springer International Publishing, 2024. 1-58.

[7] Stamler, Jeremiah. "Epidemiology of coronary heart disease." *Medical Clinics that are present in North America* 57, no. 1 (1973): 5-46

[8] Satapathy, Santosh Kumar, et al. "Heart Disease: Automatic Prediction from the Numerical and Categorical Features by Machine Learning Methods." *2023 IEEE 8th International Conference for convergence in Technology (I2CT)*. IEEE, 2023.

[9] With the use of Machine learning Algorithms we can easily pridict the Heart Diseases. | IEEE Conference Publication | IEEE Xplore

[10] Prediction of Heart Disease Based on Machine Learning Algorithms | IEEE Conference Publication | IEEE Xplore

[11] Heart disease prediction based on machine learning algorithms | Applied and Computational Engineering (ewadirect.com)

[12] ijraset.com/best-journal/heart- disease-prediction-using-machine- learning-629.

[13] ijraset.com/best-journal/heart- disease-prediction-using-machine- learning-629

[14] ijarcce.com/wp-content/uploads/2024/05/IJARCCE.20 24.134183.pdf

[15] A Hybrid Approach for Heart Disease Prediction using Genetic Algorithm and SVM | IEEE Conference Publication | IEEE Xplore

[16] https://www.sciencedirect.com /science/article/pii/S26665212230001 45

[17] Ahmad, Ahmad Ayid, and Huseyin Polat. "Prediction of heart disease based on machine learning using jellyfish optimization algorithm." *Diagnostics* 13.14 (2023): 2392.

[18] https://www.nature.com/articl es/s41598-024-58489-7#citeas

[19] Heart Disease Prediction Using Machine Learning Algorithms – IJSREM

[20] THE PREDICTION OF HEART DISEASE USING MACHINE LEARNING ALGORITHMS | Science Journal of University of Zakho (uoz.edu.krd)

[21] Heart Disease Prediction Using Machine Learning Algorithms | IEEE Conference Publication | IEEE Xplore

[22] Heart Diseases Prediction Using machine learning and Deep learning Models | IEEE Conference Publication | IEEE Xplore

[23] Cardiovascular Disease dataset (kaggle.com)

[24] Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques | IEEE Journals & Magazine | IEEE Xplore