ENGINEERING IN ADVANCED RESEARCH SCIENCE AND TECHNOLOGY

ISSN 2350-0174
Vol.03, Issue.04
June-2023
Pages: 813-828

Artificial Intelligence Based Solutions for Prediction of Cardiovascular Disease

[1]GUNDRA S G E SAI SREE, [2] SUNKARAPALLI SREENIVAS, [3]MOGILI SIVA
[1]M. Tech, Computers Science in Artificial Intelligence, Kakinada Institute of Engineering and Technology –II, KORANGI, A.P, sreesatya43@gmail.com.
[2]Assistant Professor, Computers Science in Artificial Intelligence, Kakinada Institute of Engineering and Technology –II, KORANGI, A.P, spsrinivas@gmail.com.
[3]Assistant Professor, Dept. of CSE (DATA SCIENCE), Kakinada Institute of Engineering and Technology –II, KORANGI, A.P, msiva438@gmail.com.

ABSTRACT: Cardiovascular diseases (CVD) are among the most common serious illnesses affecting human health. CVDs may be prevented or mitigated by early diagnosis, and this may reduce mortality rates. Identifying risk factors using machine learning models is a promising approach. We would like to propose a model that incorporates different methods to achieve effective prediction of heart disease. For our proposed model to be successful, we have used efficient Data Collection, Data Pre-processing and Data Transformation methods to create accurate information for the training model. We have used a combined dataset (Cleveland, Long Beach VA, Switzerland, Hungarian and Stat log). Suitable features are selected by using the Relief, and Least Absolute Shrinkage and Selection Operator (LASSO) techniques. New hybrid classifiers like Decision Tree Bagging Method (DTBM), Random Forest Bagging Method (RFBM), K-Nearest Neighbors Bagging Method (KNNBM), AdaBoost Boosting Method (ABBM), and Gradient Boosting Bagging Method (GBBM) are developed by integrating the traditional classifiers with bagging and boosting methods, which are used in the training process. We have also instrumented some machine learning algorithms to calculate the Accuracy (ACC), Sensitivity (SEN), Error Rate, Precision (PRE) and F1 Score (F1) of our model, along with the Negative Predictive Value (NPR), False Positive Rate (FPR), and False Negative Rate (FNR). The results are shown separately to provide comparisons. Data analysis is needed for this application, which is considered significant according to its approximately 90% accuracy rate over training data. We conclude the paper with objectives, limitations and research contributions.

Keywords: Least Absolute Shrinkage and Selection Operator, Cardiovascular diseases, Cardiovascular diseases, Random Forest Bagging Method, AdaBoost Boosting Method, Gradient Boosting Bagging Method.

INTRODUCTION: Machine Learning is a way of Manipulating and extraction of implicit, previously unknown/known and potential useful information about data". Machine Learning is a very vast and diverse field and its scope and implementation is increasing day by day. Machine learning incorporates various classifiers of Supervised, Unsupervised and Ensemble Learning which are used to predict and Find the Accuracy of the given dataset. We can use that knowledge in our project of HDPS as it will help a lot of people. Cardiovascular diseases are very common these days, they describe a range of conditions that could affect your heart. World health organization estimates that 17.9 million global deaths from (Cardiovascular diseases) CVDs. It is the primary reason of deaths in adults. Our project can help predict the people who are likely to diagnose with a heart disease by help of their medical history. It recognizes who all are having any symptoms of heart disease such as chest pain or high blood pressure and can help in diagnosing disease with less medical tests and effective treatments, so that they can be cured accordingly. This project focuses on mainly Five Machine learning techniques namely: Logistic regression, and Random Forest Classifier, naive byes classifier, SVM and Decision Tree Classifier. Machine learning proves to be effective in assisting in making decisions and predictions from the large quantity of data produced by the health care industry. This project aims to predict future Heart Disease by analyzing data of patients which classifies whether they have heart disease or not using machine-learning algorithm. Machine Learning techniques can be a boon in this regard. Even though heart disease can occur in different forms, there is a common set of core risk factors that influence whether someone will ultimately be at risk for heart disease or not. By collecting the data from various sources, classifying them under suitable headings & finally analyzing to extract the desired data we can say that this technique can be very well adapted to do the prediction of heart disease Cardiovascular disease has been regarded as the most severe and lethal disease in humans. The increased rate of cardiovascular diseases with a high mortality rate is causing significant risk and burden to the healthcare systems world-wide. Cardiovascular diseases are more seen in men than in women particularly in middle or old age [1], [2], although there are also children with similar health issues [3]. The associate editor coordinating the review of this manuscript and approving it for publication was Claudio Cusano . According to data provided by the WHO, one-third of the deaths globally are caused by the heart disease. CVDs cause the death of approximately 17.9 million people every year worldwide and have a higher prevalence in Asia [4], [5]. The European Cardiology Society (ESC) reported that 26 million adults worldwide have been diagnosed with heart disease, and 3.6 million are identi ed each year. Roughly half of all patients diagnosed with Heart Disease die within just 1-2 years and about 3% of the total budget for health care is deployed on treating heart disease [6]. To predict heart disease multiple tests are required. Lack of expertise of medical staff may results in false predictions [7]. Early diagnosis can be dif cult [8]. Surgical treatment of heart disease is challenging, particularly in developing countries which lack trained medical staff as well as testing equipment and other resources required for proper diagnosis and care of patients with heart problems [9]. An accurate evaluation of the risk of cardiac failure would help to prevent severe heart attacks and improve the safety of patients [10].

LITERATURE REVIEW: Artificial Intelligence in Cardiovascular Medicine, Karthik Seetharam, Sirish Shrestha and Partho P. Sengupta : The ripples of artificial intelligence are being felt in various sectors of human life. Machine learning, a subset of artificial intelligence, extracts information from large databases of information and is gaining traction in various fields of cardiology. In this review, we highlight noteworthy examples of machine learning utilization in echocardiography, nuclear cardiology, computed tomography, and magnetic resonance imaging over the past year. Recent findings In the past year, machine learning (ML) has expanded its boundaries in cardiology with several positive results. Some studies have integrated clinical and imaging information to further augment the accuracy of these ML algorithms. All the studies mentioned in this review have clearly demonstrated superior results of ML in relation to conventional approaches for identifying obstructions or predicting major adverse events in reference to conventional approaches. **Analytical Study of Heart Disease Prediction Comparing With Different Algorithms, Sana Bharti and Dr.Shaliendra Narayan Singh:** In data mining there are several ways, approaches to predict any disease and different researches are still going on. In this survey, we have studied several algorithms (like genetic algorithm, Particle Swarm Optimization, Artificial Neural Network) which play very essential role in determining or predicting heart disease. Here we firstly describe the basic concepts of these three algorithms ,and analyze how these algorithms helps in prediction of heart diseases. **The Current Research Landscape of the Application of Artificial Intelligence in Managing Cerebrovascular and Heart Diseases: A Bibliometric and Content Analysis, Bach Xuan Tran 1,2,* , Carl A. Latkin 2 , Giang Thu Vu 3 , Huong Lan Thi Nguyen 4 , Son Nghiem 5 , Ming-Xuan Tan 6 , Zhi-Kai Lim 6 , Cyrus S.H. Ho 7 and Roger C.M. Ho:** The applications of artificial intelligence (AI) in aiding clinical decision-making and management of stroke and heart diseases have become increasingly common in recent years, thanks in part to technological advancements and the heightened interest of the research and medical community. This study aims to provide a comprehensive picture of global trends and developments of AI applications relating to stroke and heart diseases, identifying research gaps and suggesting future directions for research and policy-making. A novel analysis approach that combined bibliometrics analysis with a more complex analysis of abstract content using exploratory factor analysis and Latent Dirichlet allocation, which uncovered emerging research domains and topics, was adopted.

Data were extracted from the Web of Science database. Results showed topics with the most compelling growth to be AI for big data analysis, robotic prosthesis, robotics-assisted stroke rehabilitation, and minimally invasive surgery.

EXISTING SYSTEM: The functional flow of this assessment is exemplified in Figure.. This research work focused on implementing a few classification algorithms and compares the outcomes. The dataset was divided into training and testing portions in the ratio of 70/30. Naive-Bayes, Decision Tree, Logistic-Regression, Random Forest, SVM and KNN classification models were used to predict CVD. The confusion matrix is used for identifying the mislabeling or error in prediction. It matches the actual and predicted values with four elements (True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN)). Type-I and Type-II errors are seeded by False Positive and False Negative values. The confusion matrix is very expedient to calculate Precision, Recall, F1-score and Accuracy.

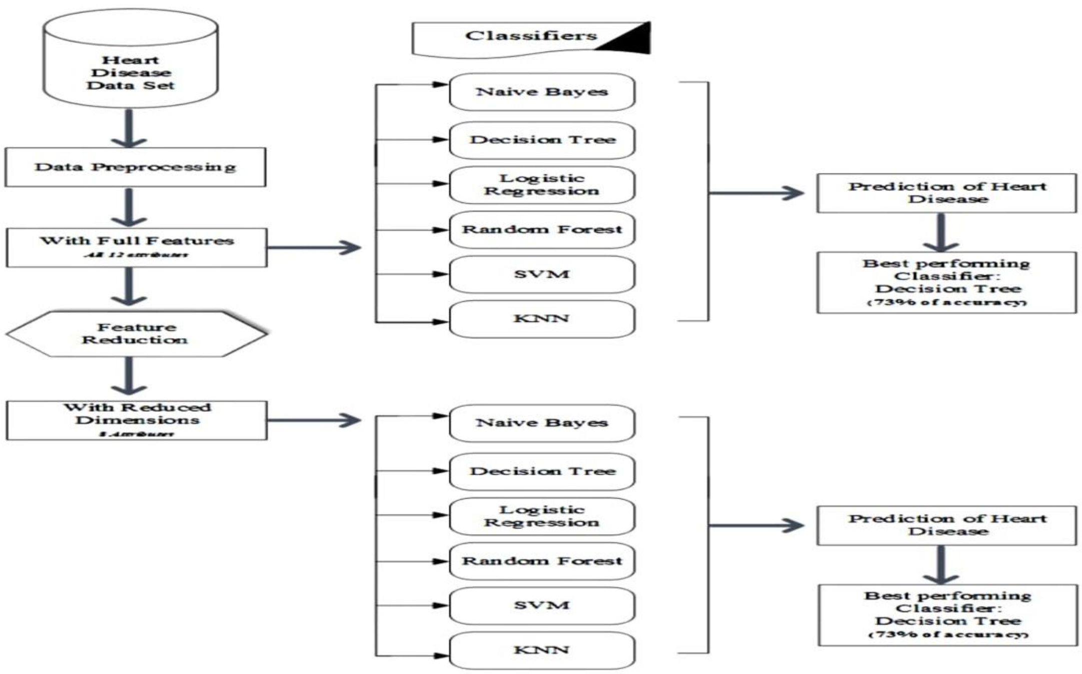


Fig:1 Existing architecture

The confusion matrix is used for identifying the mislabeling or error in prediction. It matches the actual and predicted values with four elements (True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN)). Type-I and Type-II errors are seeded by False Positive and False Negative values. The confusion matrix is very expedient to calculate Precision,

		Actual values	
		Positive	Negative
Predicted Values	Positive	True Positive(TP)	False Positive(FP)
	Negative	False Negative(FN)	True Negative(TN)

The accuracy denotes the properly predicted values. Fig. 3 represents the accuracy of each algorithm tested.

Accuracy = (True Positive+ True Negative)/ Total

The decision tree algorithm outperformed others by producing 73% accuracy. The logistic regression and SVM delivered 72% and Random forest made it with 71%. The KNN and Naive-Bayes algorithms delivered 66% and 60% of accuracy respectively. This study has been conducted on cardiovascular dataset by applying classification techniques. The Decision Tree algorithm delivered better prediction by providing 73% of accuracy. Since the dimension of dataset plays a major role in the performance of algorithms , the reduction of dimension affects the capability of Random Forest and KNN algorithms. The outcomes indicate that the dimension of dataset impacts the algorithms either positive or negative. In the next level we can apply bagging methods to each and every systems to get more accuracy.

PROPOSED SYSTEM ARCHITECTURE

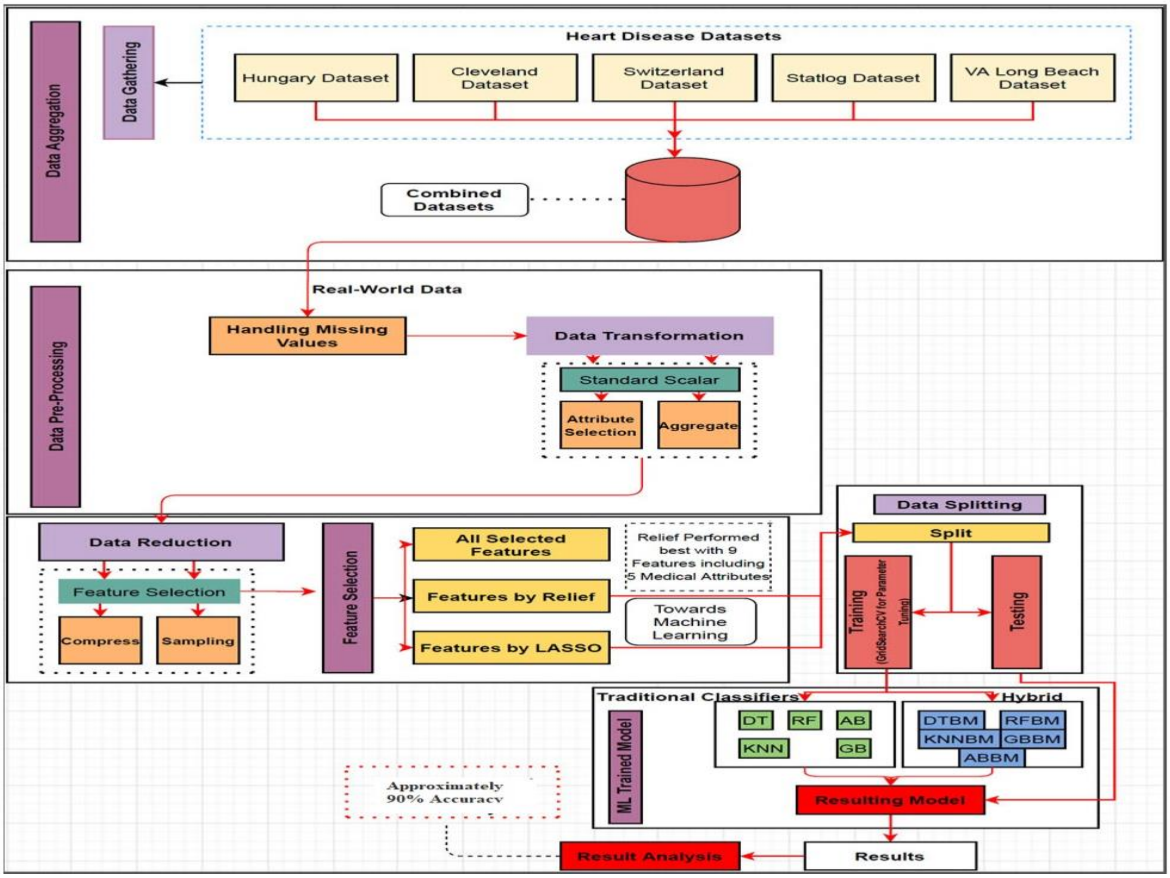


Fig2 Proposed Architecture diagram

PROPOSED SYSTEM:

AN OVERVIEW OF DATA PREPROCESSING AND CLEANING TECHNIQUES

There is a large amount of collected data in the modern world that can be gathered via the internet, surveys, and experiments, etc. Often the data to be used contain missing values, noise, and distortions, however. The combined dataset used for this research also contains missing or null values. There are some popular techniques, such as imputation and deletion that can be used to deal with missing values

FEATURE SELECTION TECHNIQUES:

Feature selection techniques are important for the machine learning procedure as the best attributes for classification need to be extracted. This also helps to reduce s+ to all the features in the dataset. These weights can then be modified gradually [64]. The aim is to ensure that the important features have a large and that the remaining features have low weights. Relief uses the similar techniques as in KNN to determine feature weights. This well known algorithm of feature selection approaches has been shown by Kira and Rendell [65]. *R* is for a randomly selected instance. Relief searches for its two nearest neighbors: one from the same class, called closest hit *H*, and one from the opposite class, called closest miss *M*. It adjusts the consistency calculation $W[A]$ for feature *A* according to the *R*, *M*, and *H* values. If there is a large difference between *R* and *H* occur this is not desirable, so the performance value $W[A]$ is reduced. On the other hand if there is a large difference between *R* and *M* for attribute *A* then *A* may be used to distinguish different classes, so the weight $W[A]$ is increased. This process will is continued form times where m is a parameter that can be adjusted.

2) LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR ALGORITHM (LASSO)

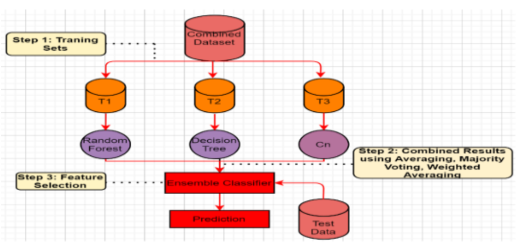
The minimum selection and shrinkage functionality of this operator depends on modifying the absolute value of the coefficient of functions. Some coefficient values of the features are zero, and features with negative coefficients can also be removed from the subset of features. The LASSO has a very good performance for feature values with small coefficients. Features which have large coefficient values will be available in the chosen subsets of features. Unnecessary features can be found with LASSO . Moreover, the reliability of this feature can be enhanced by repeating the above procedure many times eventually taking the most frequently found features in as the most important ones. This is called the randomized LASSO feature, which was introduced by Meinshausen and Buhlmann, in 2010 and Wang in 2011 [67]. It should be implemented on a powerful computer as it uses parallel programming.

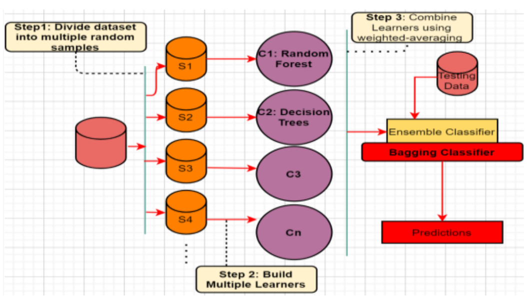
ENSEMBLE METHODS OF MACHINE LEARNING

Ensemble techniques mix multiple classifiers of a Decision Tree to achieve better classification results than only one Decision Tree classifier. The core idea behind the ensemble method is that a combination of weak learners can work together to *f* a strong learner, thus improving the model's accuracy and precision . Below Figure. depicts the ensemble process . When we seek to identify the target feature using any machine learning method, key reasons for in the difference real and identified outcomes are noise, uncertainty, and bias. Ensemble techniques assist in dealing some of these variables, particularly uncertainty and bias. In this study, we apply two ensemble techniques: Bagging and Boosting to obtain more accurate results. These techniques are explained below.

1) BAGGING TECHNIQUE

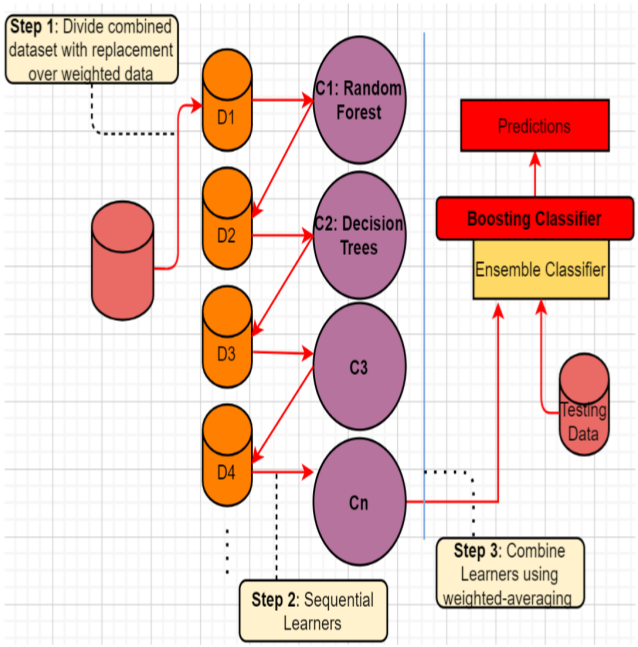
Bagging is used when the goal is to reduce the variance of Decision Tree classifiers. The objective is to create several subsets of data from the training samples. [68] Randomly chosen collections of subset data are used to train their Decision Tree. As a result, we get an ensemble of different models.





The average of all predictions from different trees is then used. This is more robust than a single Decision Tree classi er. It helps not only to reduce the over tting problem but also to handle higher dimensionality data properly. It resolves missing data issues and maintains accuracy. The process of the Bagging method is described in Pseudocode 1 and Fig..With the help of the Bagging technique, three ensemble hybrid models, based on DT, RF, and KNN, are constructed. The three hybrid models: DTBM, RFBM, and KNNBM are applied in both the training and the testing phase.

2) BOOSTING TECHNIQUE:



Boosting is a repetitive process which depends on the last prediction and changes the weight. Fig. 7 are added to better understand the workflow. If an instance is incorrectly classified its weight is increased. Usually, Boosting constructs good predictive models [6].

It generates different loss functions and works by combining the weak models to boost their performance. For this research, we have applied the Boosting technique on two classification algorithms: AB and GB to construct our hybrid models. The resulting ABBM and GBBM are applied in both the training and testing phases.

USE CASE DIAGRAM

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

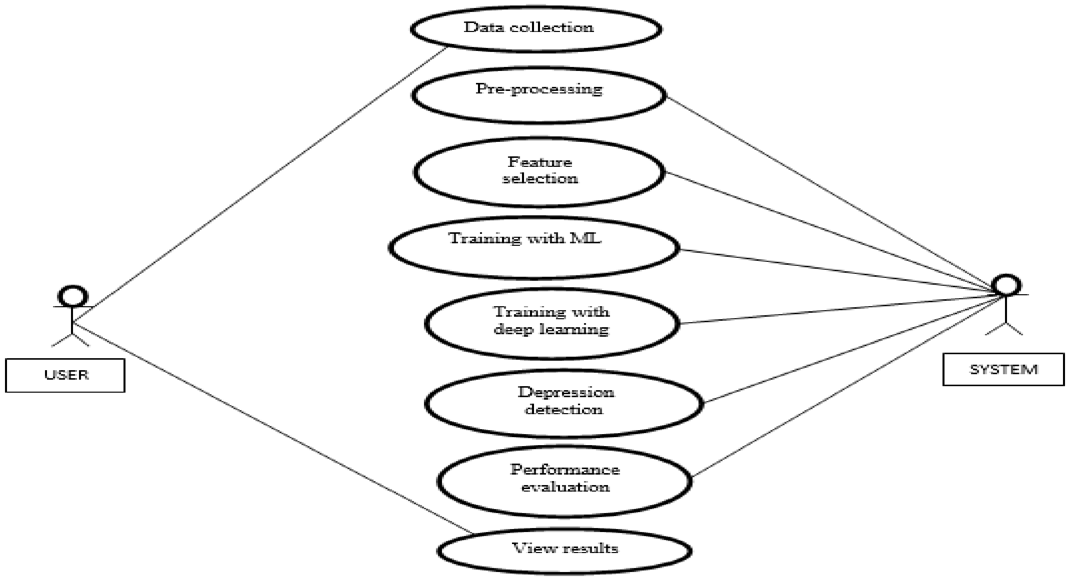


Fig.3 use case diagram

Sequence Diagram:

A sequence diagram is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. A sequence diagram shows object interactions arranged in time sequence. It depicts the objects and classes involved in the scenario and the sequence of messages exchanged between the objects needed to carry out the functionality of the scenario. Sequence diagrams are typically associated with use case realizations in the Logical View of the system under development. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

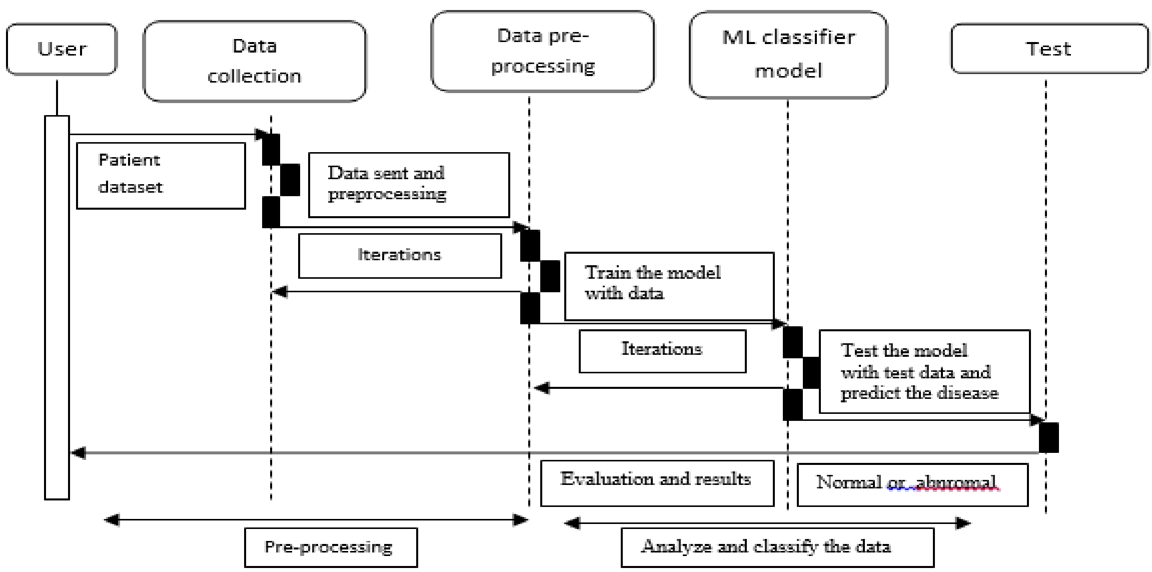


Fig.4Sequence diagram

Algorithm

Algorithm: Artificial intelligence based solutions for prediction of cardiovascular disease

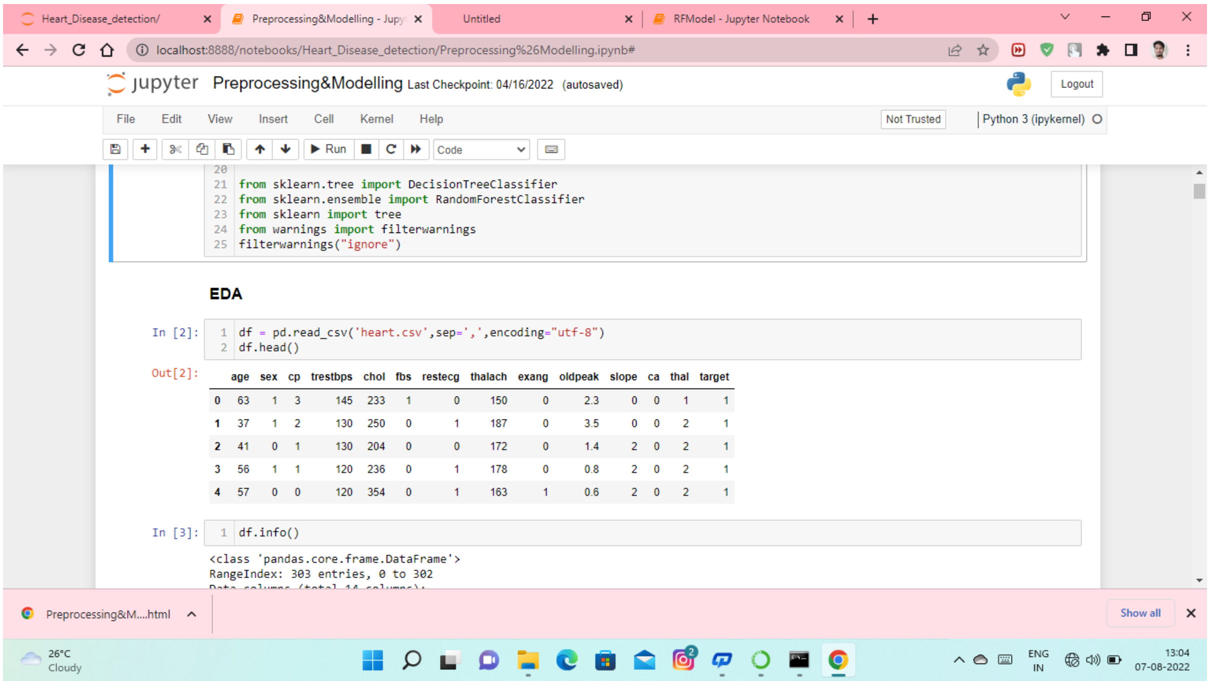
Inputs: patient dataset details as P, machine learning models as M

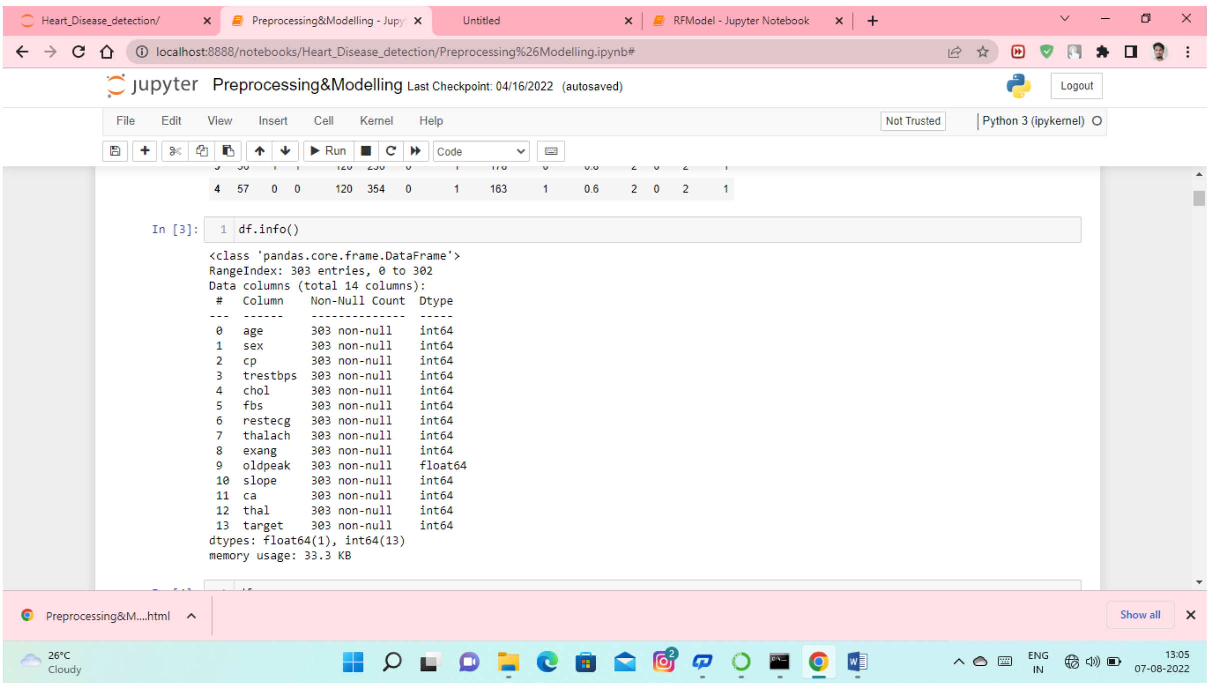
Output: results as R

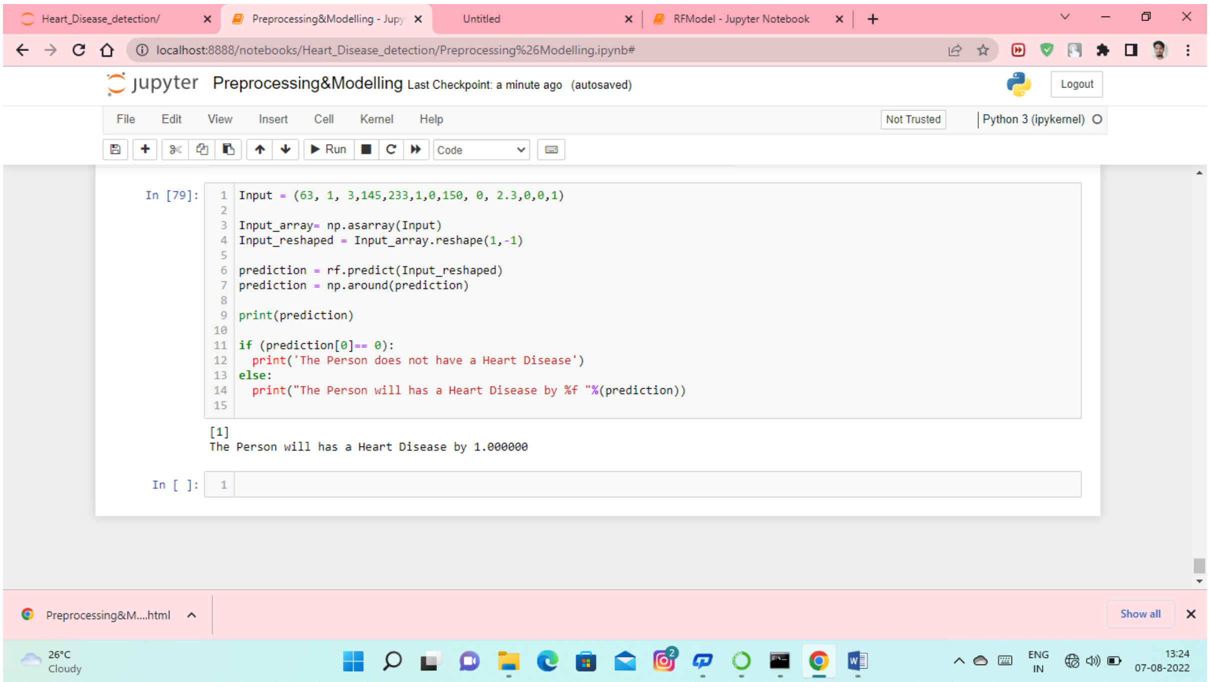
1. Start
2. Input patient dataset , P
3. Pre-processing
4. Splitting data
5. Extract features from training set()
6. For each model m in M
7.   Train the model m
8. End For
9. For each model m in M
10.   Use model for testing
11.   Evaluate
12.   Display results
13. End For
14. Save the model()
15. Predict the disease
16. Return R

RESULT

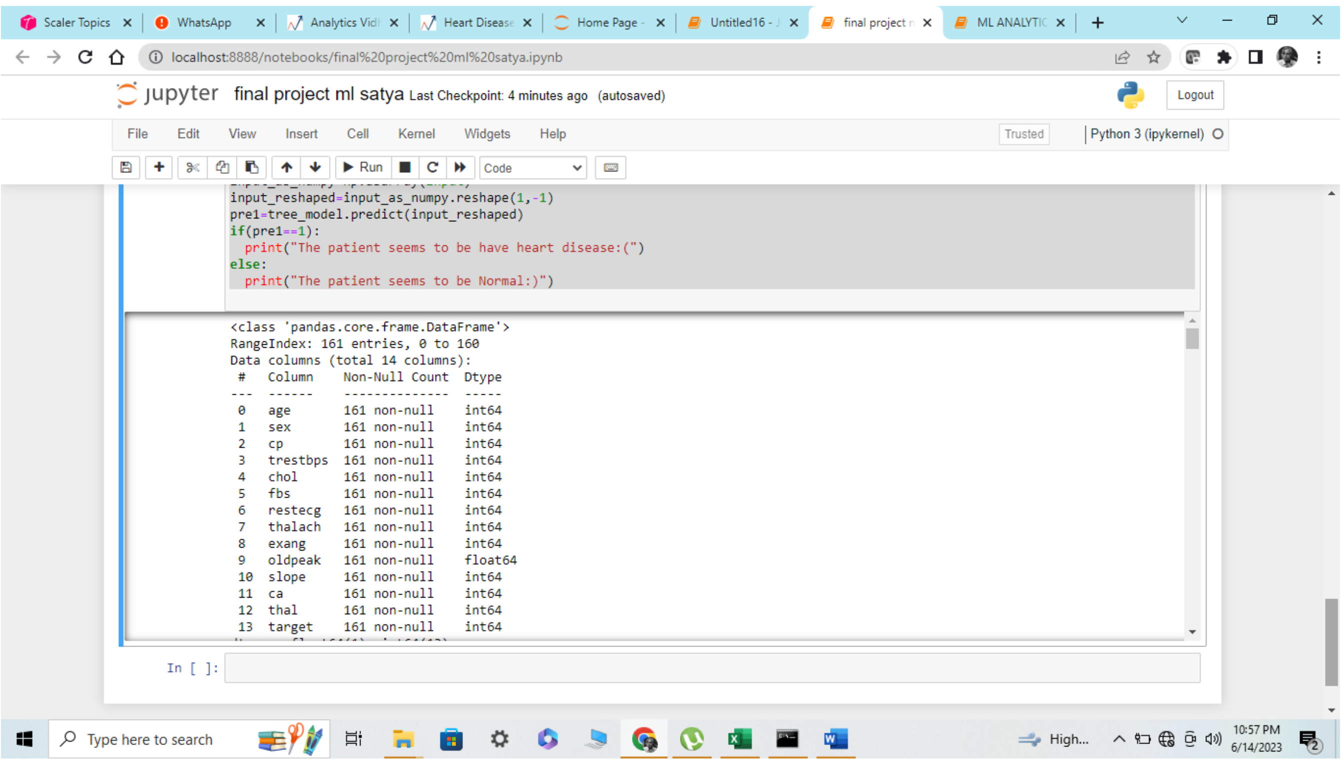
EXISTIN SYSTEMS Cardio vascular disease prediction results

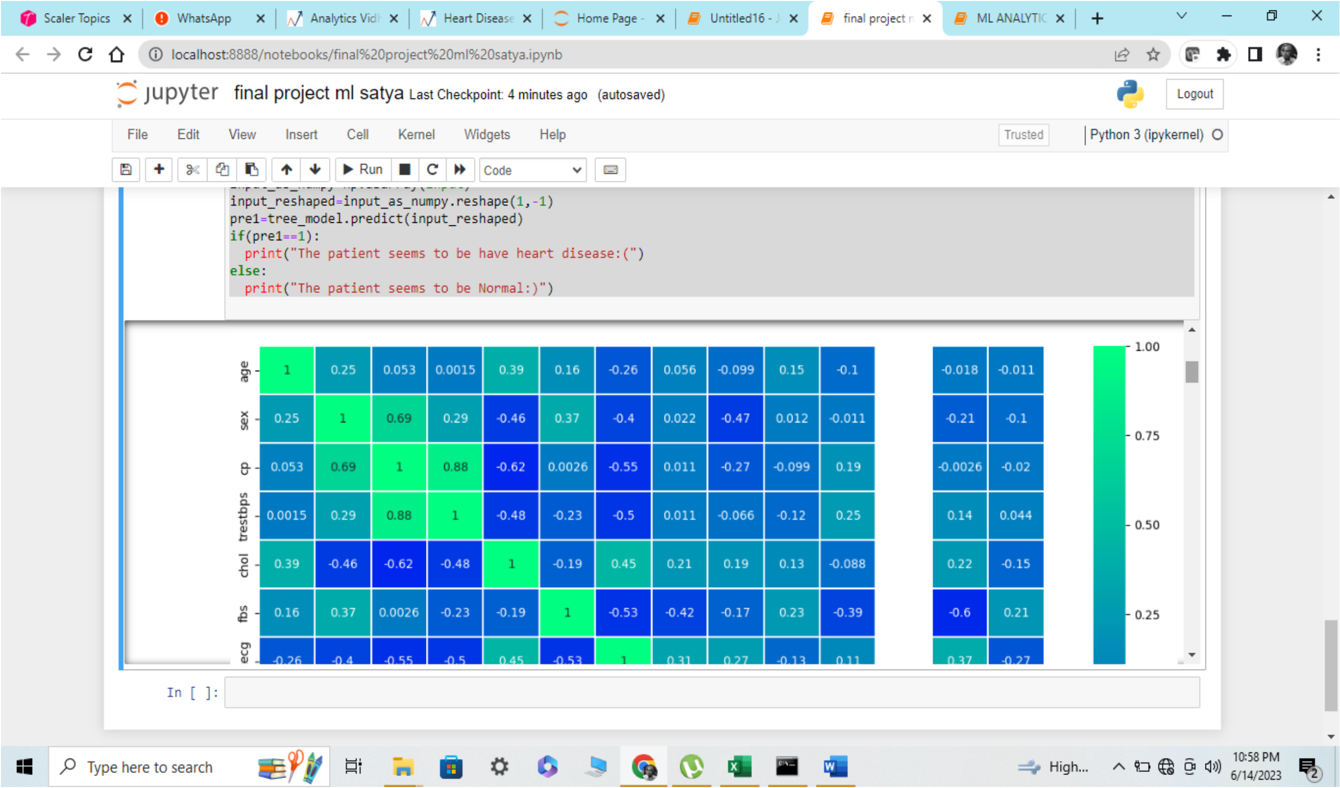


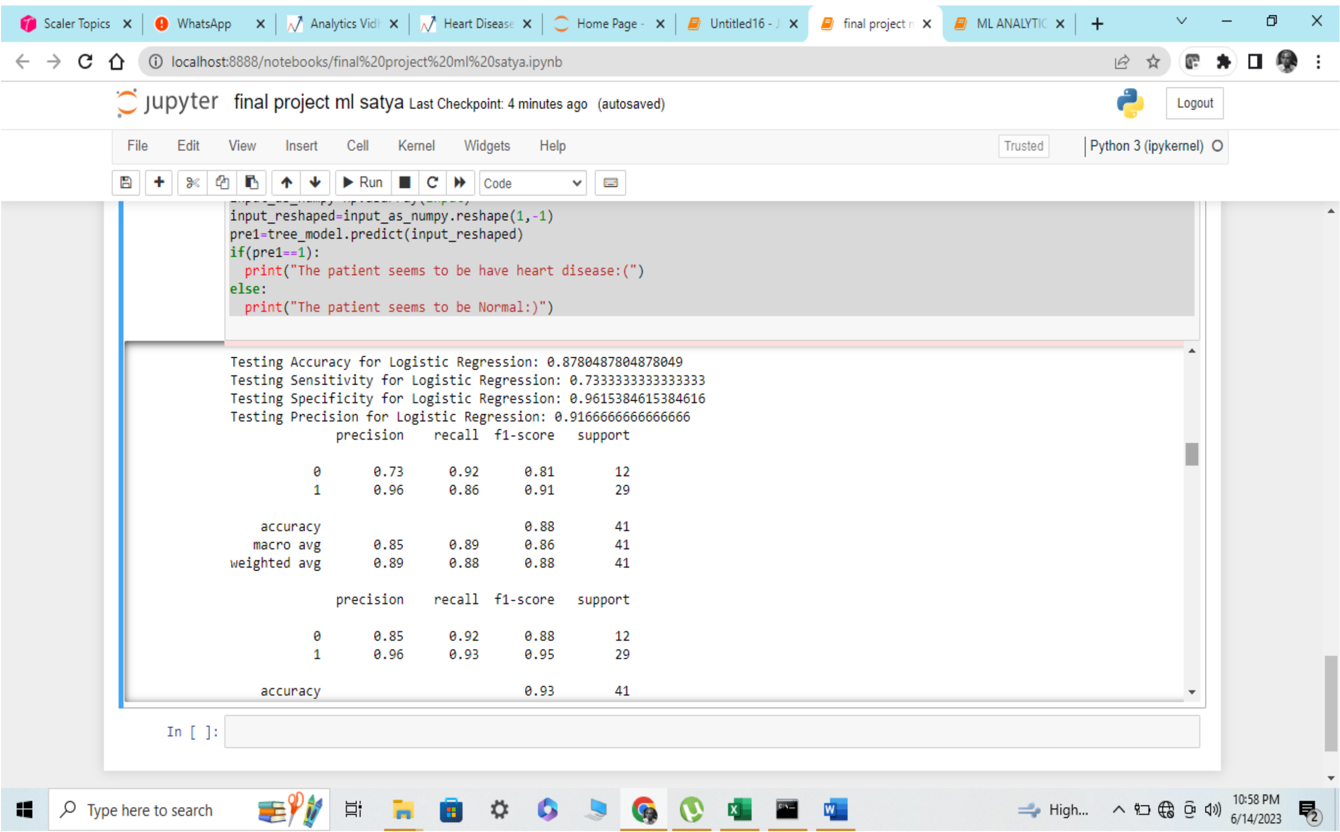


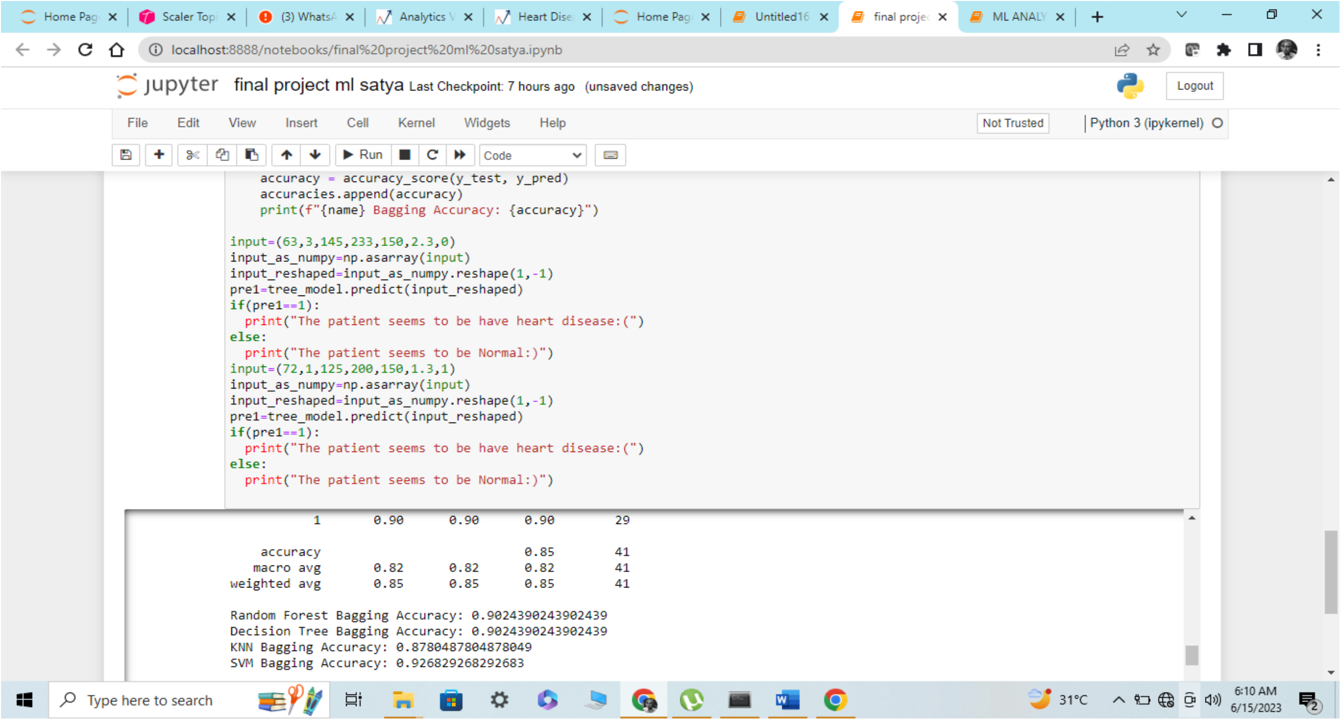


PROPOSED SYSTEM RESULTS:




localhost:8888/notebooks/final%20project%20ml%20satya.ipynb
jupyter final project ml satya Last Checkpoint: 4 minutes ago (autosaved)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=tree_model.predict(input_reshaped)
if(pre1==1):
print("The patient seems to be have heart disease:(")
else:
print("The patient seems to be Normal:)")
age
sex
cp
trestbps
chol
fbs
ecg
1 0.25 0.053 0.0015 0.39 0.16 -0.26 0.056 -0.099 0.15 -0.1 -0.018 -0.011
0.25 1 0.69 0.29 -0.46 0.37 -0.4 0.022 -0.47 0.012 -0.011 -0.21 -0.1
0.053 0.69 1 0.88 -0.62 0.0026 -0.55 0.011 -0.27 -0.099 0.19 -0.0026 -0.02
0.0015 0.29 0.88 1 -0.48 -0.23 -0.5 0.011 -0.066 -0.12 0.25 0.14 0.044
0.39 -0.46 -0.62 -0.48 1 -0.19 0.45 0.21 0.19 0.13 -0.088 0.22 -0.15
0.16 0.37 0.0026 -0.23 -0.19 1 -0.53 -0.42 -0.17 0.23 -0.39 -0.6 0.21
1.00
0.75
0.50
0.25
In [ ]:


localhost:8888/notebooks/final%20project%20ml%20satya.ipynb
jupyter final project ml satya Last Checkpoint: 4 minutes ago (autosaved)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=tree_model.predict(input_reshaped)
if(pre1==1):
print("The patient seems to be have heart disease:(")
else:
print("The patient seems to be Normal:)")
Testing Accuracy for Logistic Regression: 0.8780487804878049
Testing Sensitivity for Logistic Regression: 0.7333333333333333
Testing Specificity for Logistic Regression: 0.9615384615384616
Testing Precision for Logistic Regression: 0.9166666666666666
precision recall f1-score support
0 0.73 0.92 0.81 12
1 0.96 0.86 0.91 29
accuracy 0.88 41
macro avg 0.85 0.89 0.86 41
weighted avg 0.89 0.88 0.88 41
precision recall f1-score support
0 0.85 0.92 0.88 12
1 0.96 0.93 0.95 29
accuracy 0.93 41
In [ ]:


localhost:8888/notebooks/final%20project%20ml%20satya.ipynb
jupyter final project ml satya Last Checkpoint: 7 hours ago (unsaved changes)
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
accuracy = accuracy_score(y_test, y_pred)
accuracies.append(accuracy)
print(f"{name} Bagging Accuracy: {accuracy}")
input=(63,3,145,233,150,2.3,0)
input_as_numpy=np.asarray(input)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=tree_model.predict(input_reshaped)
if(pre1==1):
print("The patient seems to be have heart disease:(")
else:
print("The patient seems to be Normal:)")
input=(72,1,125,200,150,1.3,1)
input_as_numpy=np.asarray(input)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=tree_model.predict(input_reshaped)
if(pre1==1):
print("The patient seems to be have heart disease:(")
else:
print("The patient seems to be Normal:)")
1 0.90 0.90 0.90 29
accuracy 0.85 41
macro avg 0.82 0.82 0.82 41
weighted avg 0.85 0.85 0.85 41
Random Forest Bagging Accuracy: 0.9024390243902439
Decision Tree Bagging Accuracy: 0.9024390243902439
KNN Bagging Accuracy: 0.8780487804878049
SVM Bagging Accuracy: 0.926829268292683
Type here to search
31°C
6:10 AM
6/15/2023

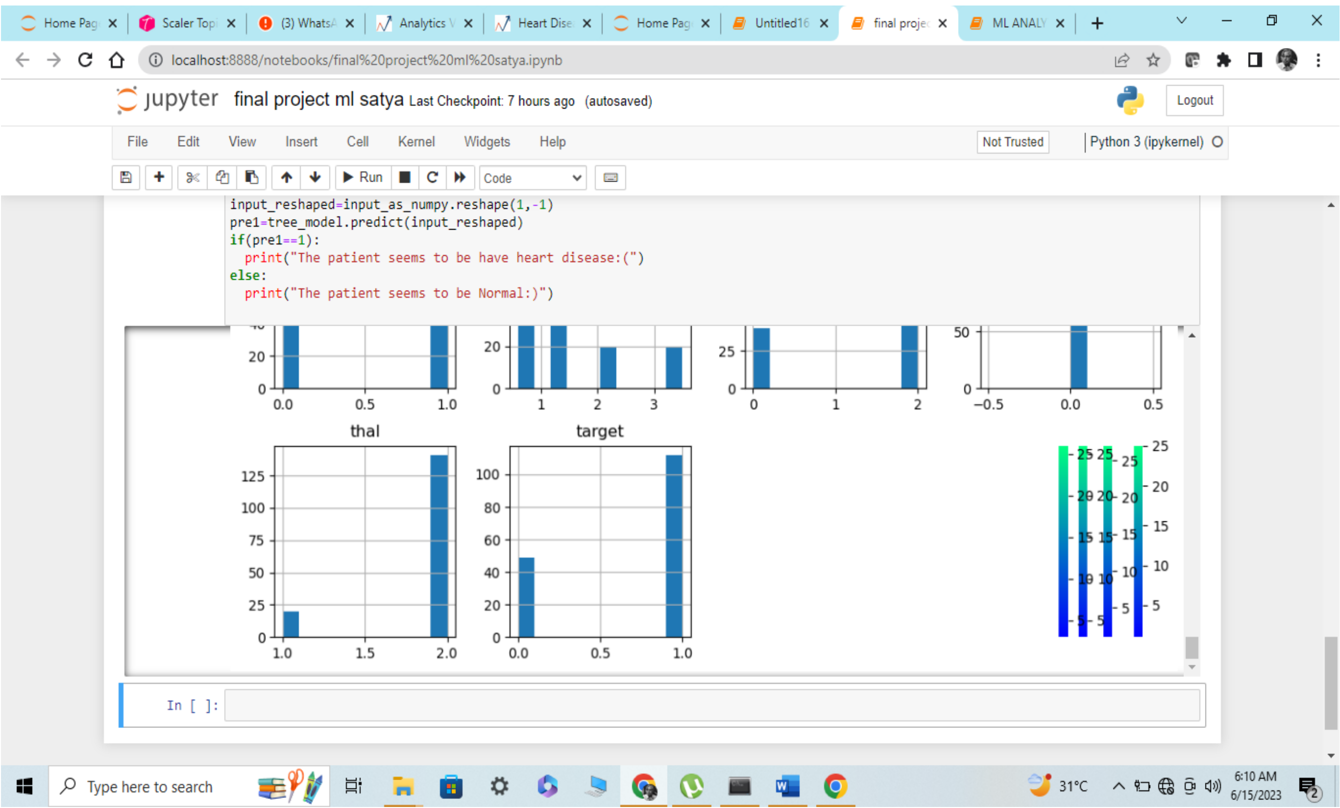

localhost:8888/notebooks/final%20project%20ml%20satya.ipynb
jupyter final project ml satya Last Checkpoint: 7 hours ago (autosaved)
File Edit View Insert Cell Kernel Widgets Help
Not Trusted
Python 3 (ipykernel)
input_reshaped=input_as_numpy.reshape(1,-1)
pre1=tree_model.predict(input_reshaped)
if(pre1==1):
print("The patient seems to be have heart disease:(")
else:
print("The patient seems to be Normal:)")
thal
target
In [ ]:
Type here to search
31°C
6:10 AM
6/15/2023

TABLE. A comparison of accuracy between proposed system and existed outcomes

Models	Accuracy of all features (13)	Accuracy of 11 features (LASSO)	Accuracy of 10 features (Relief)	Existing Systems (Accuracy)	Percentage calculator for ALL features	Percentage calculator for 11 features	Percentage calculator for 10 features
DT	86.97%	88.6%	89.12%	75.55% [40]	15.11% increase	17.27% increase	17.97% increase
RF	88.65%	86.97%	97.89%	80.89% [41]	9.60% increase	7.51% increase	21.01% increase
KNN	83.61%	93%	94.11%	90.16% [43]	7.26% decrease	3.15% increase	4.38% increase
						increase	101.85% increase
						increase	14.18% increase
						increase	13.42% increase
						% increase	12.04% increase
						% increase	16.62% increase
						increase	7.17% increase
						increase	19.17% increase

CONCLUSION & FUTURE SCOPE: Heart diseases are a major killer in India and throughout the world, application of promising technology like machine learning to the initial prediction of heart diseases will have a profound impact on society. The early prognosis of heart disease can aid in making decisions on lifestyle changes in high-risk patients and in turn reduce the complications, which can be a great milestone in the field of medicine. The number of people facing heart diseases is on a raise each year. This prompts for its early diagnosis and treatment. The utilization of suitable technology support in this regard can prove to be highly beneficial to the medical fraternity and patients. In this paper, the five different machine learning algorithms used to measure the performance are SVM, Decision Tree, Random Forest, Naïve Bayes, Logistic Regression, and KNN are applied on the dataset. If all the features taken into the consideration then the efficiency of the system the author gets is less. To increase efficiency, attribute selection is done. In this n features have to be selected for evaluating the model which gives more accuracy. The correlation of some features in the dataset is almost equal and so they are removed. If all the attributes present in the dataset are taken into account then the efficiency decreases considerably. Hence, the aim is to use various evaluation metrics like confusion matrix, accuracy, precision, recall, and f1-score which predicts the disease efficiently. Additionally, various hybrid approaches, including Bagging and Boosting, are implemented to improve the testing rate and reduce the execution time. the highest accuracy of 90%. The future scope of this system aims at giving more sophisticated prediction models, risk calculation tools and feature extraction tools for other clinical risks. In the future we aim to generalize the model even further so that it can work with other feature selection algorithms and be robust against datasets where the level of missing data is high. The application of Deep Learning algorithms is another future approach. The primary aim of this research was to improve upon the existing work with an innovative and novel way of building the model, as well as to make the model useful and easily implementable to practical settings

REFERENCES

[1] Seetharam, Karthik; Shrestha, Sirish; Sengupta, Partho P. (2019). Artificial Intelligence in Cardiovascular Medicine. Current Treatment Options in Cardiovascular Medicine, 21(6), 25, p1-14.

[2] Bharti, Sana; Singh, Shaliendra Narayan (2015). [IEEE 2015 International Conference on Computing, Communication & Automation (ICCCA) - Greater Noida, India (2015.5.15-2015.5.16)] International Conference on Computing, Communication & Automation - Analytical study of heart disease prediction comparing with different algorithms. , p78–82.

[3] Rajamhoana, SP; Devi, C. Akalya; Umamaheswari, K.; Kiruba, R.; Karunya, K.; Deepika, R. (2018). [IEEE 2018 11th International Conference on Human System Interaction (HSI) - Gdansk, Poland (2018.7.4-2018.7.6)] 2018 11th International Conference on Human System Interaction (HSI) - Analysis of Neural Networks Based Heart Disease Prediction System. , p233–239.

[4] Tran, Bach Xuan; Latkin, Carl A.; Vu, Giang Thu; Nguyen, Huong Lan Thi; Nghiem, Son; Tan, Ming-Xuan; Lim, Zhi-Kai; Ho, Cyrus S.H.; Ho, Roger C.M. (2019). The Current Research Landscape of the Application of Artificial Intelligence in Managing Cerebrovascular and Heart Diseases: A Bibliometric and Content Analysis. International Journal of Environmental Research and Public Health, 16(15), 2699, p1-14.

[5] Konstantinos C. Siontis;Peter A. Noseworthy;Zachi I. Attia;Paul A. Friedman; (2021). Artificial intelligence-enhanced electrocardiography in cardiovascular disease management . Nature Reviews Cardiology, p1-14.

[6] Amin, Syed Umar; Agarwal, Kavita; Beg, Rizwan (2013). [IEEE 2013 IEEE Conference on Information & Communication Technologies (ICT) - Genetic neural network based data mining in prediction of heart disease using risk factors. , p1227–1231.

[7] Eleni I. Georga, Nikolaos S. Tachos, Antonis I. Sakellarios, Vassiliki I. Kigka,. (2019). Artificial Intelligence and Data Mining Methods for Cardiovascular Risk Prediction. Springer, pp.279-301.

[8] Benjamins, J. W.; Hendriks, T.; Knuuti, J.; Juarez-Orozco, L. E.; van der Harst, P. (2019). A primer in artificial intelligence in cardiovascular medicine. Netherlands Heart Journal, p1-9.

[9] M G Dorrer;S E Golovenkin;S Yu Nikulina;Yu V Orlova;E Yu Pelipeckaya; (2020). Using artificial intelligence to predict the human body's response to cardiovascular disease . Journal of Physics: Conference Series, p1-10.

[10] Krittanawong, Chayakrit; Zhang, HongJu; Wang, Zhen; Aydar, Mehmet; Kitai, Takeshi (2017). Artificial Intelligence in Precision Cardiovascular Medicine. Journal of the American College of Cardiology, 69(21), p2657–2664.

[11] Amma, N. G. Bhuvaneswari (2012). [IEEE 2012 International Conference on Computing, Communication and Applications (ICCCA) - Cardiovascular disease prediction system using genetic algorithm and neural network. , p1–5.

[12] Lopez-Jimenez, Francisco; Attia, Zachi; Arruda-Olson, Adelaide M.; Carter, Rickey; Chareonthaitawee, Panithaya; Jouni, Hayan; Kapa, Suraj; Lerman, Amir; Luong, Christina; Medina-Inojosa, Jose R.; Noseworthy, Peter A.; Pellikka, Patricia A.; Redfield, Margaret M.; Roger, Veronique L.; Sandhu, Gurpreet S.; Senecal, Conor; Friedman, Paul A. (2020). Artificial Intelligence in Cardiology: Present and Future. Mayo Clinic Proceedings, 95(5), p1015–1039.

[13] Swathy, M., & Saruladha, K. (2021). A comparative study of classification and prediction of Cardio-Vascular Diseases (CVD) using Machine Learning and Deep Learning techniques. ICT Express. p1-12.

[14] Zhanpeng Jin, ; Yuwen Sun, ; Cheng, A.C. (2009). [IEEE 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society - Predicting cardiovascular disease from real-time electrocardiographic monitoring: An adaptive machine learning approach on a cell phone, p6889–6892.

[15] Mathur, Pankaj; Srivastava, Shweta; Xu, Xiaowei; Mehta, Jawahar L (2020). Artificial Intelligence, Machine Learning, and Cardiovascular Disease. Clinical Medicine Insights: Cardiology, 14, p1-9.

[16] Dinesh, Kumar G; Arumugaraj, K; Santhosh, Kumar D; Mareeswari, V (2018). [IEEE 2018 International Conference on Current Trends towards Converging Technologies (ICCTCT) - Prediction of Cardiovascular Disease Using Machine Learning Algorithms, p1–7.

[17] Muhammad Saqib Nawaz;Bilal Shoaib;Muhammad Adeel Ashraf; (2021). Intelligent Cardiovascular Disease Prediction Empowered with Gradient Descent Optimization . Heliyon, p1-8.

[18] I. D. Mienye, Y. Sun, and Z. Wang, ''An improved ensemble learn-ing approach for the prediction of heart disease risk,'' *Informat. Med. Unlocked*, vol. 20, Jan. 2020, Art. no. 100402.

[19] H. Wang, Z. Huang, D. Zhang, J. Arief, T. Lyu, and J. Tian, ‘‘Integrat-ing co-clustering and interpretable machine learning for the prediction of intravenous immunoglobulin resistance in kawasaki disease,’’ *IEEE Access*, vol. 8, pp. 97064 97071, 2020.

[20] B. A. Tama, S. Im, and S. Lee, ‘‘Improving an intelligent detection system for coronary heart disease using a two-tier classi er ensemble,’’ *BioMed Res. Int.*, vol. 2020, Apr. 2020, Art. no. 9816142.

[21] J. Mishra and S. Tarar, *Chronic Disease Prediction Using Deep Learning*. Singapore: Springer, 2020, pp. 201 211.

[22] F. Z. Abdeldjouad, M. Brahami, and N. Matta, *A Hybrid Approach for Heart Disease Diagnosis and Prediction Using Machine Learning Techniques*. Cham, Switzerland: Springer, 2020, pp. 299 306.

[23] M. Tarawneh and O. Embarak, ‘‘Hybrid approach for heart disease predic-tion using data mining techniques,’’ *Acta Sci. Nutritional Health*, vol. 3, no. 7, pp. 147 151, Jul. 2019.

[24] C. B. C. Latha and S. C. Jeeva, ‘‘Improving the accuracy of prediction of heart disease risk based on ensemble classi cation techniques,’’ *Informat. Med. Unlocked*, vol. 16, Jan. 2019, Art. no. 100203.

[25] I. Javid, A. Khalaf, and R. Ghazali, ‘‘Enhanced accuracy of heart disease prediction using machine learning and recurrent neural networks ensem-ble majority voting method,’’ *Int. J. Adv. Comput. Sci. Appl.*, vol. 11, no. 3, 2020.

[26] N. Kumar and K. Sikamani, ‘‘Prediction of chronic and infectious dis-eases using machine learning classi ers A systematic approach,’’ *Int.*
J. Intell. Eng. Syst., vol. 13, no. 4, pp. 11 20, 2020.

[27] S. M. Saqlain, M. Sher, F. A. Shah, I. Khan, M. U. Ashraf, M. Awais, and A. Ghani, ‘‘Fisher score and matthews correlation coef cient-based feature subset selection for heart disease diagnosis using support vector machines,’’ *Knowl. Inf. Syst.*, vol. 58, no. 1, pp. 139 167, Jan. 2019.

[28] S. Mohan, C. Thirumalai, and G. Srivastava, ‘‘Effective heart disease prediction using hybrid machine learning techniques,’’ *IEEE Access*, vol. 7, pp. 81542 81554, 2019.

[29] F. Miao, Y.-P. Cai, Y.-X. Zhang, X.-M. Fan, and Y. Li, ‘‘Predictive modeling of hospital mortality for patients with heart failure by using an improved random survival forest,’’ *IEEE Access*, vol. 6, pp. 7244 7253, 2018.

[30] C. Raju, E. Philipsy, S. Chacko, L. P. Suresh, and S. D. Rajan, ‘‘A survey on predicting heart disease using data mining techniques,’’ in *Proc. Conf. Emerg. Devices Smart Syst. (ICEDSS)*, 2018, pp. 253-255.

[31] D. Chicco and G. Jurman, ‘‘Machine learning can predict survival of patients with heart failure from serum creatinine and ejection fraction alone,’’ *BMC Med. Informat. Decis. Making*, vol. 20, no. 1, p. 16, Dec. 2020.

[32] E. Ahmad, A. Tiwari, and A. Kumar, ‘‘Cardiovascular Diseases (CVDs) Detection using Machine Learning Algorithms,’’

[33] L. Wang, W. Zhou, Q. Chang, J. Chen, and X. Zhou, ‘‘Deep ensemble detection of congestive heart failure using short-term RR intervals,’’ *IEEE Access*, vol. 7, pp. 69559 69574, 2019.

[34] A. Gupta, R. Kumar, H. S. Arora, and B. Raman, ‘‘MIFH: A machine intelligence framework for heart disease diagnosis,’’ *IEEE Access*, vol. 8, pp. 14659 14674, 2020.

[35] C. B. C. Latha and S. C. Jeeva, ‘‘Improving the accuracy of prediction of heart disease risk based on ensemble classi cation techniques,’’ *Informat. Med. Unlocked*, vol. 16, no. 2, 2019, Art. no. 100203.
[36] G. O. Rashmi and U. M. A. kumar, ‘‘Machine learning methods for heart disease prediction,’’ *Int. J. Eng. Adv. Technol.*, vol. 8, no. 5S, pp. 220 223, May 2019.
[37] K. G. Dinesh, K. Arumugaraj, K. D. Santhosh, and V. Mareeswari, ‘‘Pre-diction of cardiovascular disease using machine learning algorithms,’’ in *Proc. Int. Conf. Current Trends Towards Converging Technol. (ICCTCT)*, Coimbatore, India, Mar. 2018, pp. 1 7.
[38] S. Mohan, C. Thirumalai, and G. Srivastava, ‘‘Effective heart disease prediction using hybrid machine learning techniques,’’ *IEEE Access*, vol. 7, pp. 81542 81554, 2019.
[39] S. Sharma and M. Parmar, ‘‘Heart diseases prediction using deep learning neural network model,’’ *Int. J. Innov. Technol. Exploring Eng.*, vol. 9, no. 3, pp. 1 5, Jan. 2020.
[40] R. Alizadehsani, J. Habibi, Z. A. Sani, H. Mashayekhi, R. Boghrati, A. Ghandeharioun, F. Khozeimeh, and F. Alizadeh-Sani, ‘‘Diagnosing coronary artery disease via data mining algorithms by considering labo-ratory and echocardiography features,’’ *Res. Cardiovascular Med.*, vol. 2, no. 3, pp. 133 139, Aug. 2013.
[41] A. A. Shetty and C. Naik, ‘‘Different data mining approaches for predict-ing heart disease,’’ *Int. J. Innov. Sci. Eng. Technol.*, vol. 5, pp. 277 281, May 2016.
[42] C. A. Cheng and H. W. Chiu, ‘‘An arti cial neural network model for the evaluation of carotid artery stenting prognosis using a national-wide database,’’ in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 2566 2569.
[43] K. C. Tan, E. J. Teoh, Q. Yu, and K. C. Goh, ‘‘A hybrid evolutionary algorithm for attribute selection in data mining,’’ *Expert Syst. Appl.*, vol. 36, no. 4, pp. 8616 8630, May 2009.
[44] I. K. A. Enriko, ‘‘Comparative study of heart disease diagnosis using top ten data mining classi cation algorithms,’’ in *Proc. 5th Int. Conf. Frontiers Educ. Technol.*, 2019, pp. 159-164.
[45] M. Saqlain, W. Hussain, N. A. Saqib, and M. A. Khan, ‘‘Identi ca-tion of heart failure by using unstructured data of cardiac patients,’’ in *Proc. 45th Int. Conf. Parallel Process. Workshops (ICPPW)*, Aug. 2016, pp. 426 431.
[46] A. K. Dwivedi, ‘‘Evaluate the performance of different machine learn-ing techniques for prediction of heart disease using ten-fold cross-validation,’’ *Neural Comput. Appl.*, vol. 29, pp. 685 693, Sep. 2016.
[47] A. Kaur, ‘‘A comprehensive approach to predict heart diseases using data mining,’’ *Int. J. Innov. Eng. Technol.*, vol. 8, no. 2, pp. 1 5, Apr. 2017.
[48] V. Chaurasia and S. Pal, ‘‘Data mining approach to detect heart diseases,’’ *Int. J. Adv. Comput. Sci. Inf. Technol.*, vol. 2, no. 4, pp. 56 66, 2014.
[49] R. Bhuvaneeswari, P. Sudhakar, and G. Prabakaran, ‘‘Heart disease pre-diction model based on gradient boosting tree (GBT) classi cation algo-rithm,’’ *Int. J. Recent Technol. Eng.*, vol. 8, no. 2, pp. 41 51, Sep. 2019.
[50] F. M. J. M. Shamrat, P. Ghosh, M. H. Sadek, A. Kazi, and S. Shultana, ‘‘Implementation of machine learning algorithms to detect the progno-sis rate of kidney disease,’’ in *Proc. IEEE Int. Conf. Innov. Technol.*, Nov. 2020, pp. 1 7.
[51] S. Shultana, M. S. Moharram, and N. Neehal, ‘‘Olympic sports events classi cation using convolutional neural networks,’’ in *Proc. Int. Joint Conf. Comput. Intell. (IJCCI)*, Dhaka, Bangladesh, 2018, pp. 507 518.
[52] S. V. J. Jaikrishnan, O. Chantarakasemchit, and P. Meesad, ‘‘A breakup machine learning approach for breast cancer prediction,’’ in *Proc. 11th Int. Conf. Inf. Technol. Electr. Eng. (ICITEE)*, Pattaya, Thailand, Oct. 2019, pp. 1 6.
[53] A. Gavhane, G. Kokkula, I. Pandya, and K. Devadkar, ‘‘Prediction of heart disease using machine learning,’’ in *Proc. 2nd Int. Conf. Elec-tron., Commun. Aerosp. Technol. (ICECA)*, Coimbatore, India, Mar. 2018, pp. 1275 1278.
[54] G. Singh, ‘‘Breast cancer prediction using machine learning,’’ *Int. J. Sci. Res. Comput. Sci., Eng. Inf. Technol.*, vol. 8, no. 4, pp. 278 284, Jul. 2020.

[55] *Heart Disease Datasets From UCI Machine Learning Repository*. Accessed: May 31, 2020. [Online]. Available: https://archive.ics.uci. edu/ml/datasets/Heart+Disease

[56] *Heart Disease Statlog Dataset of UCI Machine Learning Repos-itory*. Accessed: May 31, 2020. [Online]. Available: http://archive. ics.uci.edu/ml/datasets/statlog+(heart)

[57] S. Ralston, I. Penman, M. Strachan, and R. Hobson, *Davidson's Prin-ciples and Practice of Medicine*, 23rd ed. U.K.: Elsevier, Apr. 2018, pp. 219 225.

[58] A. Rairikar, V. Kulkarni, V. Sabale, H. Kale, and A. Lamgunde, ''Heart disease prediction using data mining techniques,'' in *Proc. Int. Conf. Intell. Comput. Control (IC)*, Jun. 2017, pp. 1 8.

[59] A. Acharya, ''Comparative study of machine learning algorithms for heart disease prediction,'' M.S. thesis, Helsinki Metropolia Univ. Appl. Sci., Helsinki, Finland, Apr. 2017. [Online]. Available: https:// www.theseus. /bitstream/handle/10024/124622/Final%20Thesis.pdf? sequence=1&isAllowed=y