Journal of Positive School Psychology http://journalppw.com
2022, Vol. 6, No. 6, 10551 – 10560

Hybrid Model of Machine Learning Algorithms for Prediction of Cardiovascular Disease

[1]Urja Desai, [2]Shamla Mantri

[1]School of Computer Engineering & Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India, Email: urjadesai1405@gmail.com
[2]School of Computer Engineering & Technology, Dr. Vishwanath Karad MIT World Peace University, Pune, India. Email: shamla.mantri@mitwpu.edu.in

Abstract

The utmost prevalent reason of death in worldwide is coronary artery disease, which is serious health issue for many people accounting for nearly 31% of all fatalities. Patients with common illnesses and symptoms are being reported in greater numbers. Many patients' lives can be saved by early identification of cardiovascular disorders such heart attacks and cardiovascular disease. It relates a slew of heart disease risk factors to the essential need for accurate, trustworthy, and practical approaches for early detection and management. In the healthcare, two popular approaches for evaluating huge datasets are mainly machine learning and deep learning algorithms. Many experts and researchers have practices a range of both deep learning and machine learning techniques for examine large quantities of complex medical data, assisting doctors in predicting heart problems. Accurately predicting, if a patient has cardiovascular diseases or not is the aim of this research. In this paper we have studied and implemented many traditional machine learning algorithms which are Naïve Bayes (NB), Logistic Regression (LR), Decision Tree (DT), Support Vector Machine (SVM), Random Forest (RF), Extreme Gradient Boost (XGBoost) and K-Nearest Neighbor algorithm (KNN). Based on output we have implemented hybrid model to archive more accuracy. ROC-curve, Accuracy, Error rate, Recall, Precision and F1 score performance evaluation metrics which are applied to compare the effectiveness and performance of the methodologies in study. We have obtained 93.4% accuracy using this proposed hybrid model using a stacking classifier technique.

Keywords: Heart disease prediction, Hybrid model, Machine learning, Deep learning.

1. Introduction

The One of fundamental concerns of human life is healthcare According to WHO recommendations, everyone have a right to decent health. A small percentage of fatalities from cardiac disorders are caused by natural clinical causes, while the majority one is caused by the diseases' delayed detection. According to the latest WHO study, heart disease is becoming more common. Every year, 17% people across the world die because consequence of this. According to statistics, 50 percent of Americans have at least one chronic illness. Ironically, this leads to the treatment of chronic diseases accounting for 80% of US healthcare costs [1]. Because the nature of heart disease is complicated, it should be handled carefully. According to a publication from the European Society of Cardiology (ESC), Heart Failure affects 26 million persons worldwide, yet 3.6 million adults receive new diagnoses each year [9]. Because of the late assessment of the heart attack's severity, developing nations, notably in Asia and Africa, have seen a significant number of failures to save human lives [15].

Heart disease refers to any condition that may impair the heart's ability to pump blood. Heart disease denotes to any kind of problem that can

compromises the heart's capacity of operating and working. As the population grows, diagnosing and initiating therapy at an early stage becomes more challenging [10]. Predicting cardiac illnesses using various methodologies has been a continuous endeavor over the last 2 centuries.

The human heart is susceptible to serious ailments as a result of today's unbalanced lifestyle. Diabetes, stress, and heavy smoking are all contributing to the emergence of important diseases and difficulties. All of these factors have a negative effect on the heart, resulting in a wide range of heart disorders. Every type of cardiac ailment has a unique set of symptoms and a different course of therapy. The earlier these signs are recognized, the better the odds of rapid treatment and preserving a human's lives. The use of Data mining approaches facilitates gathering of useful information from many perspectives. Predictive mining is being used in the current research project for developing a heart disease prediction and diagnosis system.

Today's people are fast-paced society want to live a lavish lifestyle, so they work like machines to make a lot of money and live comfortably. And Because of that, individuals neglect to take care for themselves, as well as changes in their food habits and entire way of life. As a result, they are more nervous, having a high blood pressure and diabetics from a very young age, don't take sufficient rest, and consume anything what they can. A number of factors can lead to heart failure. These aspects have been split into two groups by medical scientists: risk aspects that cannot be modified and risk aspects which can be altered. Family heredity, gender, and age etc. all are risk aspects which can't be modified. Risk factors include increased blood pressure, high cholesterol, alcoholism, insufficient exercise and high level of blood pleasure [8]. If detected early enough, cardiac illness may be treated; however that isn't always the scenario.

The most important aspect of healthcare is illness diagnosis. An automated approach to predict and identify the heart illness, its cause and treatment plan in its initial stages is required. As a result, experts will be able to identify illnesses more quickly, and individuals will also be able to determine their health conditions with the use of this modern technology. For analyzing and interpreting enormous volumes of data, data mining techniques are quite useful. It's used to extract data and make judgments about what to do with it in the future. Clustering, association rule mining, and classifications are the most common data mining approaches. These data mining approaches can be implemented using a variety of algorithms. Python programming is gaining traction with these methods implemented using scikit learn packages. As a result, the implementation of data mining principles in real time is more trustworthy than ever before. In the healthcare-medical domain, improving categorization the accuracy of the models is the most significant determination of the data analytics. The classification technique is a supervised learning approach for classifying a dataset into various data classes. The dataset is then distributed into 2 sections: 80% of dataset is divided into training set, to train our proposed model and 20% of dataset is a test set to determine accuracy and correctness of model, which referred to as the train-test split.

Since a reliable, improved model is required for heart disease prediction, it can be presumed from the previous studies that uniqueness in the study is required. A speedy and effective detection system must be created due to the lack of professionals and the high rate of cases that are incorrectly diagnosed [23]. In order to enhance classification performance, researchers were inspired by the implementing the variety of the machine learning approaches to predict a likelihoods of occurrence of heart disease, in this research work we implemented hybrid model using stacking classifier technique instead of traditional machine learning algorithms. Researchers are always developing new methods to enhance the effectiveness of categorization. Researchers are always developing novel methods to enhance categorization performance. One of such novel method is heterogeneous ensemble or homogeneous ensemble learning approach [11].

The structure of remaining sections of the paper is as follow: The 2nd section represents literature survey and some existing works done on this. Section 3 introduces several machine learning algorithms and brief details of each algorithm. Our Proposed workflow is explained in to section 4. Description of the dataset has been given in depth in section number 5 along with proposed implementation result and

analysis. Section 6 describes future improvements and conclusion.

2. Literature Survey

Researchers are now doing extensive studies on the prediction and analysis of heart illness. Researchers used several data mining tools; like an association rules, regression, clustering, and classification to create framework to predict cardiac disease [4]. Several works on these are discussed farther down.

In [1] authors have used real-time patient data from hospital to test the prediction model. Structured, Semi-Structured, Unstructured data obtain to clean missing and noisy data. Then dataset classified into training and test dataset. Various ML algorithms are applied such as CNN, naïve bayes and KNN etc. Accuracy found using these algorithms are 85, 80 and 74 (in %) respectively. The created model achieves an accuracy of between 85 and 88 percent. The model accuracy is highest for the CNN designed model, followed by the Naive Bayes method, which is slightly less accurate than CNN, and finally the KNN algorithm, which has the least accuracy. The Convolutional Neural Network technique can identify early heart disease risk using structured data.

In [2] for the prediction of cardiac disease, a hybrid model was implemented. The model which combines two machine learning methods, random forest and decision tree methods were used on the Cleveland dataset for cardiac disease by authors. There are 303 cases in total in this collection, with 14 different features. This research attempts to use machine learning and an automated medical diagnosis system to anticipate cardiac disease. Accuracy obtains by DT and RF algorithms are 79 % and 81%. Accuracy obtain by of proposed hybrid model is 88%.In [3] authors have implemented KNN, Decision Tree and Random Forest ML algorithms and RF techniques were able to obtain 100% accuracy, specificity and sensitivity. They have observed that to create an extremely precise heart disease prediction model with a variety of applications the classic supervised machine learning algorithms are applicable.

The authors in [4] have described a supervised learning machine learning algorithm models such as decision trees, K-nearest neighbour, Naïve Bayes, and random forest algorithms, as well as numerous heart disease features in their research work. It is based on a dataset of the heart disease patients that already exists in the dataset from Irvine's University of California's Cleveland database. Following dataset consists of total 303 incidences and total 76 features. Despite their relevance in establishing the usefulness of alternative algorithms, only 14 of the 76 characteristics are observed. The aim of their study is to anticipate the chance of people developing cardiac attacks. According to the statistics, K-nearest neighbour has the highest accuracy score. In [5] authors reviewed and presented some recent work in the fields of machine learning and image fusion utilizing classification methods. They find that ANN has a decent effect in predicting heart disease in most models.

In [6] authors have developed a framework for understanding the fundamentals of anticipating the summary of patients' risk possibilities using patient information attributes throughout this research. The recommended model is built by using X2-statistical models and Deep Neural Networks (DNN). An ability of the model to correctly anticipate the existence or absence of heart disease was tested using DNN and ANN. In [10] The Cleveland dataset for cardiac illnesses was split into two halves using a present split for dividing the dataset into the training dataset and testing dataset by authors . We looked at 14 features and utilized four different approaches to assess the accuracy. They discovered that random forest has the highest level of accuracy, whereas decision tree has the lowest level of accuracy (85%).

3. Methodology

As machine learning algorithms are subset of an arising technology artificial intelligence (AI), the fundamental aim of this is to create systems which can learn from its past experiences and make predictions [22]. It creates a model by using a training dataset to train machine learning algorithms. Using the additional input data, the model predicts cardiac attack. Using machine learning, by discovering unseen patterns from the given dataset it creates models, and for the newer datasets, it can predict outcomes correctly. Noisy data has been cleaned from

dataset and Null values are handled using various methods. Using the updated input analysis to predict potential heart disease, the algorithm is then tested for accuracy [4]. For heart disease prediction model, we used seven standard machine learning approaches. The following is a detailed discussion of these machine learning algorithms that we have used for our research work.

3.1 Logistic Regression

In all supervised machine learning approaches, one of the utmost commonly used machine learning algorithms is the logistic regression model. In classification and predictive analytics, this statistical model which is also known as the logit model is commonly employed. Logistic regression determines the likelihood of an occurrence based on a set of independent factors. Based on a number of independent parameters, logistic regression calculates the probability of an event. The dependent variable can only vary from 0 to 1, as the result is a probability. There are three different forms of logistic regression.

i. A logistic regression algorithm using a binary variables

ii. A logistic regression algorithm with multinomial variables

iii. An ordinal logistic regression algorithm

For the prediction of cardiac related disease, binary logistic regression is the finest alternative which can be used. With discrete classes, it works nicely. The logit function may be defined using several concepts such as the odds ratio. It's the probability that anything will happen is defined by equation (1):

Probability Ratio = Pr/ (l — Pr) …….. (1)

Where, Pr is a probability of positive class.

It takes input in the [0, 1] range and converts it to values in the real-number range.

3.2 Naïve Bayes Classifier

It is supervised machine learning system which uses the Bayes' Theorem as its foundation. It's a probabilistic classifier that implies it can predict outcomes for many classes at the same time. Strong independence between characteristics is presumed. Probabilistic classifiers are those that predict several classes. The determinants don't have any connections to one another or a correlation with one another. The likelihood of maximization is affected by each feature separately [4].

The choice is made using conditional probability. Instead of a single algorithm, this paradigm utilizes a group of algorithms that all have the same assumption. Each attribute is believed to contribute an equal and unique amount to the outcome under this paradigm. This model has an advantage over others since it just takes a little amount of training data.

3.3 Random Forest Classifier

This algorithm consists of a number of classification-based trees that have not been pruned. Because it is unaffected by dataset noise and the risks of overfitting are limited, it works well in a range of real-world problems. It's faster than other tree-based methods and improves testing and validation data accuracy. Using this method, a forest is generated by a large number of trees. In random forest classifier algorithm every tree emits a class expectation, and most voted class becomes the prediction of model. For each set of data in the dataset, the cycle is repeated. The accuracy of the random forest classifier generally increases with the increment of number of trees.

This model is an ensemble classifier, that means it uses and combines a number of different decision tree classifiers [16]. The main purpose of utilizing a large number of trees is to train them sufficiently such that each tree's contribution is in the form of a model. The output is aggregated by the majority after the tree is created. It makes use of many decision trees, each of which is based on a single dataset with a similar distribution across the tree. This approach is capable of efficiently balancing errors in a class population of skewed data sets. It may be used to tackle classification and regression difficulties, among other things.

3.4 Extreme Gradient Boost

The extreme gradient boosting is a machine learning approach for tackling predictive analytic issues related to classification and regression. Decision tree models are employed in the creation of ensembles. The ensemble is then gradually incorporated to the trees, and they are then fitted to correct the prediction errors caused by the earlier models.

A machine learning ensemble model is comparable to the boosting strategy. It is the most generally used algorithm in contests for applied machine learning, and its popularity has increased as a result of effective data solutions in organized and tabular form.

3.5 K-Nearest Neighbor

KNN method is a supervised machine learning algorithm, which is one of the all fundamental techniques. The KNN model places the new instance in the class that resembles the current categories the most, presuming that the newest data/case and previous cases are equivalent. The K-NN model captures all of the data that is available and determines a set of points based on how similar it is to the old data .This suggests that employing the K-NN approach, new data might be efficiently grouped into a certain category.

Due to its slow rate of learning from the training set, this algorithm is sometimes referred as a lazy learner approach. It keeps the dataset and utilizes it when it's required to categories it. The KNN algorithm work by evaluating a distances using Manhattan or Euclidean distance among a query [12]. By selecting the nearest one instances of k to the query, and then vote for the most recurrent category type (for classification) or averaging the categories are how KNN works (for regression). Its analytical tractability and ease of implementation are the benefits of the method. Given that it only considers one instance while making predictions, the classifier is relatively efficient and performs well enough in diseases prediction.

3.6 Decision Tree

One of the most popular and commonly used machine learning techniques is the decision tree (DT), which used to build classification models. The structure of this classification model is tree-like structure. This is classified as supervised learning since the desired outcome is already known. The Decision tree technique may be used with both categorical and numerical data. Decision tree is made of three segments and they all are primarily employed as a node in the decision tree. The decision tree is made up of three parts: the leaf nodes, the root nodes and branches. Starting at the decision tree's root, a record bearing a class label will be predicted.

Root Node: It serves as the central node and effectively carries out all tasks.

Interior node: There are several restrictions and constraints that are checked at this node.

Leaf node: Only the leaf node stores the ultimate outcome.

The values are contrasted against the root attributes of the following record attributes. This comparison yields the corresponding value for the following node to depart [2]. This strategy can be used in the event of a multi-dimensional investigation with several classes. The past data, also known as the past vector, is used to create a model which can predicts the output result based on the input. There are multiple nodes in a tree, each of which corresponds to one of two vectors. The tree comes to a standstill at a leaf node, which represents a potential outcome or result.

3.7 Support Vector Machine

It is a supervised machine learning classification approach which used in regression and classification analysis to evaluate data and find trends. This method is also more successful in dealing with a wide range of practical issues. It's generally a basic approach that divides data into groups using a linear or hyperplane [17]. In this technique, the optimal hyper plane which divides all the data points of the one class from the other class is used to describe data. The better the model is, the greater the separation or edge between the two classes.

The hardest part of using SVM to construct a model is choosing the kernel and strategy to avoid underfit and overfit problems. Support vector machine is described as the discrete vector space with magnitudes of each characteristic of an item. Described as having dimensions for each object feature [5]. This method is preferable to a wide range of complicated and fundamental picture, text classification and segmentation, handwritten character recognition, and most of biology and related disciplines.

4. Proposed Workflow

The hybrid model is a revolutionary approach in which the output is obtained by entering one machine learning model result as a result into any other machine learning model. This hybrid

model offers us more highly optimized results based on the basic machine learning methods which are selected for implementations.

In order to address a particular machine learning issue, many learners are developed and merged in an ensemble learning process. The medical expert inputs the numbers from the patient's health report. The information is used to feed a model that anticipates the chance of cardiovascular disease.

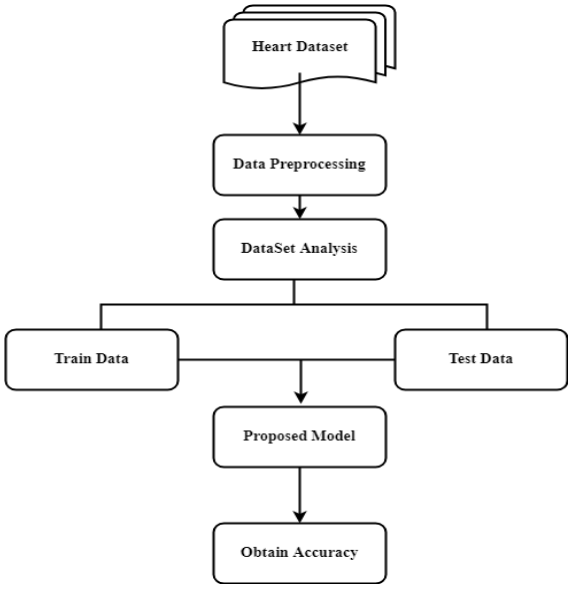

Figure 1: Proposed Workflow

Figure 1 illustrates the workflow of proposed system. First, we implemented seven machine learning algorithms using python libraries such as pandas, keras, numpy, matplotlib, sklearn libraries, and few other relevant libraries in the proposed work. Based on accuracy obtained from these ML algorithms, we build a hybrid model using algorithms which obtains highest accuracy.

5. Experimental Result and Analysis Discussion

This section describes in brief about the implementation phase and result analysis of research work carried out by our proposed model. A description of dataset, implementation and performance evaluation measure is discussed.

5.1 Description of Dataset

We collected dataset from the UCI repository for this study. The dataset is consists of 303 instances and 14 attributes. From these 14 attributes 13 are predictor attributes and 1 is a target attribute. Description of these attributes is as followed:

Table 1: Details of Features in Dataset

Sr. No.	Feature	Possible Values
1	Age- Age of Patient	Between 29-71
2	Sex- Gender of Patient	0=Female, 1=male
3	CP- Type of Chest pain	4 Possible values (Type 0 or 1 or 2 or 3)
4	TRESTBPS- Resting Blood Pressure Rate	Between 94-200
5	CHOL- Serum cholesterol	Between 126-564
6	FPS- Fasting Blood Sugar	FPS>120=1; FPS<120=0.
7	RESTECH-resting electro-cardiograph	0 or 1 or 2
8	THALACH- Maximum heart rate	Between 71-202
9	EXANG- Exercise-induced angina	0 = no, 1 = yes;
10	OLDPEAK-Level of depression	Between 0-6.2
11	SLOPE- During exercise slope of peak.	1 or 2 or 3
12	CA- fluoroscopy result	0 or 1 or 2 or 3
13	THAL- thalassemia	0 or 1 or 2 or 3
14	TARGET- Indicator attribute of heart attack prediction.	0 = No heart attack, 1 = heart attack

There are eight category features and six numerical features in the dataset. Table 1 covers all of the dataset's details. In the dataset, range of age feature is Minimum age 29 to maximum age 77. Target feature is indicator of the heart disease. Value 0 indicates No heart attack and 1 indicates the heart attack.

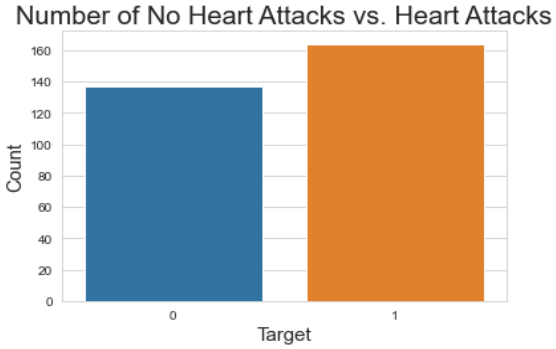

Figure 2: Heart Disease Data Visualization in a Dataset

Figure 2 represents the visualization of dataset to find comparison of No Heart Disease Vs. Heart Disease. So in dataset out of 303 instances total numbers of heart attack are 165 and total numbers of no heart disease are 138.

5.2 Data Preprocessing and Implementation

Several features in the original dataset have missing values; it might impair the accuracy of the model and produce erroneous results [18]. The required dataset preparation must be placed in a manner that the algorithm can comprehend during the preprocessing stage, which is of utmost importance. The mean mode approach was used to find and replace these missing values. After data preprocessing we performed Exploratory Data Analysis. We divided dataset into 2 parts: Training set and test set. For our model's training set, we divided the dataset by 80%. Then we have implemented 7 different machine learning algorithms. We obtain performance evaluation based on confusion matrix and accuracy for each machine learning algorithm.

The accuracy of the datasets, the techniques used to pre-process the data, and the algorithms applied, in addition to the algorithms utilized, have a significant impact on the efficiency and accuracy of the categorization and prediction model. Models with fewer parameters have poor capacity, which causes them to be underfit problem. The model should be constructed in such a manner that it produces a hypothesis with the greatest possible capacity since a model with more parameters than necessary results to high capacity and over fitting. After implementing all seven machine learning algorithms we have built hybrid model using stacking technique. We employ ensembling to increase the model's accuracy (hybrid by stacking). Stacking, or stacked generalization, is an ensemble machine learning approach. Using a meta-learning strategy, it figured out way to combine predictions using two or more machine learning algorithms. Stacking ensembles are regularly varied due to the utilization of several learning approaches at the base level.

So using stacking method we build a hybrid model using three base machine learning algorithms which obtain high accuracy out of all seven approaches. Result of base algorithms implemented in this study is discussed in section 5.4.

5.3 Performance Evaluation Matrix

The confusion matrix and ROC curve are used after the classification to evaluate the performance of machine learning algorithms.

Recall: It is a metric for assessing a classification model's completeness. The higher the recall, the less false negatives are produced, and the findings are more accurate. The lesser the recall, the higher the false negatives.

Equation to calculate recall is defined by eq. (2)

$$\text{Recall} = \frac{TP}{TP+FN} \quad \ldots\ldots\ldots\ldots\ldots\ldots (2)$$

Precision: It is a measure of exactness used for evaluating the performance of a classification model. There will be fewer false positives if the precision is high. False positives are more likely in a model with lower precision means. Precision can be calculated using equation (3):

$$\text{Precision} = \frac{TP}{TP+FP} \quad \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

F1 Score: In machine learning algorithms, the F1-score is also one of the most essential assessment parameter. To properly depict a model's prediction ability, it combines two contradicting criteria: accuracy and recall. Using equation 4 we have obtained F1 Score.

$$\text{F1 Score} = \frac{\text{Recall}*\text{Precision}}{\text{Recall}+\text{Precision}} \quad \ldots\ldots\ldots\ldots (4)$$

ROC Curve: The ROC curve (Receiver Operating Characteristic Curve is a graph that depicts the accomplishment of a classification models over all categorization criteria. On this graph, two parameters are displayed: True Positive Rate and False Positive Rate.

5.4 Evaluation Results and Discussion

We have implemented proposed prediction model in python language using several in built libraries like pandas, numpy, matplotlib, keras, tensorflow, sklearn libraries and some other related libraries. Accuracy for prediction of the heart disease is being calculated using different 13 attributes in dataset. In table 2 the experiment result has been displayed.

Table 2: Performance evaluation of models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1Score (%)
Logistic Regression	85.2	83.8	91.2	87.3
Naïve Bayes	85.2	83.8	91.2	87.3
Random Forest	86.9	86.1	91.2	88.6
Extreme Gradient Boost	91.8	93.9	9.12	92.5
K-Nearest Neighbour	88.5	90.9	88.2	89.6
Decision Tree	82.0	87.1	79.4	83.1
Support Vector Machine	88.5	88.6	9.12	89.9
Proposed Hybrid Model	93.4	94.1	94.1	94.1

As we can see that after implementation of machine learning algorithms we have obtain accuracy as shown in Table 2. In which out of 7 models K-Nearest Neighbour (KNN), Extreme Gradient Booster (XGBoost) and Support Vector Machine (SVM) are the three base machine learning algorithms which have obtained highest accuracy among all. So for our hybrid model using stacking technique we selected these 3 algorithms and after building hybrid model using these algorithms we have obtain 93.4 % accuracy.

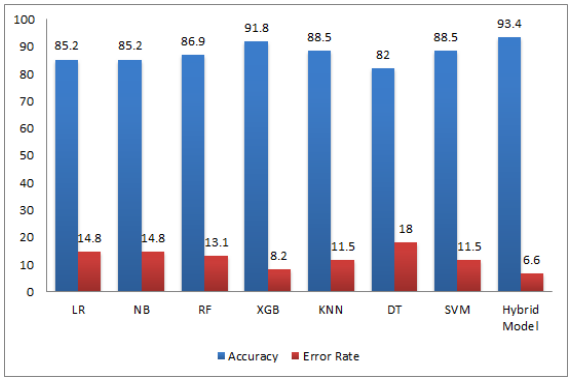

Figure 3: Accuracy and Error Rate of All Models

Accuracy and error rate of all machine learning models has been shown in figure 3 in a histogram. Another measure for evaluating the performance of algorithms is ROC curve.

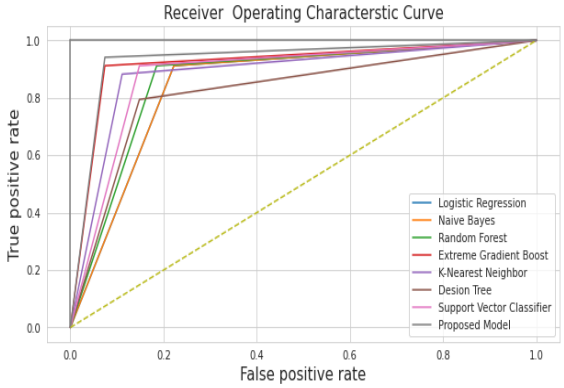

Figure 4: ROC curve

Figure 4 shows the ROC curve for base machine learning algorithms.

6. Conclusion and Future Scope

The most important organ of the human body, after the brain, is the heart. Heart disease is by far the most prevalent cause of death among the various causes of mortality. Due to medical practitioners' lack of knowledge and expertise about warning indications of heart failure, detecting heart disease can be difficult. We used seven different machine learning algorithms in our proposed research work and then created a hybrid model from Extreme Gradient Booster, KNN and Support Vector Machine using stacking Technique. And we have obtained 93.4% accuracy using our proposed Model.

In the medical field, Machine Learning (ML) and Data Mining (DM) both the approaches are proven to be useful and essential [7]. In the future work, innovative ways for anticipating heart disease with help of a hybrid model of machine learning and deep learning might be explored in order to accomplish a higher accuracy. In addition to that we might use real time data collected from hospital and doctor instead of readymade dataset available on internet. We believe that our research will aid future researchers in making the best decisions possible.

References

[1] Shankar, VirenViraj, et al. "Heart disease prediction using CNN algorithm." SN Computer Science 1 (2020): 1-8.

https://doi.org/10.1007/s42979-020-0097-6

[2] Kavitha, M., et al. "Heart Disease Prediction using Hybrid machine Learning Model." 2021 6th International Conference on Inventive Computation Technologies (ICICT). IEEE, 2021. https://doi.org/10.1109/ICICT50816.2021.9358597

[3] Ali, Md Mamun, et al. "Heart disease prediction using supervised machine learning algorithms: performance analysis and comparison." Computers in Biology and Medicine 136 (2021): 104672. https://doi.org/10.1016/j.compbiomed.2021.104672

[4] Shah, Devansh, Samir Patel, and Santosh Kumar Bharti. "Heart disease prediction using machine learning techniques." SN Computer Science 1.6 (2020): 1-6. http://dx.doi.org/10.48175/IJARSCT-1131

[5] Diwakar, Manoj, et al. "Latest trends on heart disease prediction using machine learning and image fusion." Materials Today: Proceedings 37 (2021): 3213-3218. https://doi.org/10.1016/j.matpr.2020.09.078

[6] Ramprakash, P., et al. "Heart Disease Prediction Using Deep Neural Network." 2020 International Conference on Inventive Computation Technologies (ICICT). IEEE, 2020. https://doi.org/10.1109/ICICT48043.2020.9112443

[7] Franklin, Ramya G., and B. Muthukumar. "Survey of Heart Disease Prediction and Identification using Machine Learning Approaches." 2020 3rd International Conference on Intelligent Sustainable Systems (ICISS). IEEE, 2020. https://doi.org/10.1109/ICISS49785.2020.9316119

[8] Katarya, Rahul, and Polipireddy Srinivas. "Predicting heart disease at early stages using machine learning: A survey." 2020 International Conference on Electronics and Sustainable Communication Systems (ICESC). IEEE, 2020. https://doi.org/10.1109/ICESC48915.2020.9155586

[9] Selvakumari, S. Jeya, et al. "An Extensive Survey on Heart Disease Prediction." Annals of the Romanian Society for Cell Biology (2021): 13013-13020. https://www.annalsofrscb.ro/index.php/journal/article/view/4244

[10] Sharma, Vijeta, Shrinkhala Yadav, and Manjari Gupta. "Heart Disease Prediction using Machine Learning Techniques." 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN). IEEE, 2020. https://doi.org/10.1109/ICACCCN51052.2020.9362842

[11] Mienye, Ibomoiye Domor, Yanxia Sun, and Zenghui Wang. "An improved ensemble learning approach for the prediction of heart disease risk." Informatics in Medicine Unlocked 20 (2020): 100402. https://doi.org/10.1016/j.imu.2020.100402

[12] Jothi, K. Arul, et al. "Heart disease prediction system using machine learning." Materials Today: Proceedings (2021). https://doi.org/10.1016/j.matpr.2020.12.901

[13] Tougui, Ilias, Abdelilah Jilbab, and Jamal El Mhamdi. "Heart disease classification using data mining tools and machine learning techniques." Health and Technology 10 (2020): 1137-1144. http://dx.doi.org/10.1007/s12553-020-00438-1

[14] Jeyaranjani, J., T. Dhiliphan Rajkumar, and T. Ananth Kumar. "Coronary heart disease diagnosis using the efficient ANN model." Materials Today: Proceedings (2021). https://doi.org/10.1016/j.matpr.2021.01.257

[15] Mehmood, Awais, et al. "Prediction of Heart Disease Using Deep Convolutional Neural Networks." Arabian Journal for Science and Engineering 46.4 (2021): 3409-3422. http://dx.doi.org/10.1007/s13369-020-05105-1

[16] Mohan, Senthilkumar, Chandrasegar Thirumalai, and Gautam Srivastava. "Effective heart disease prediction using hybrid machine learning techniques." IEEE access 7 (2019): 81542-81554. https://doi.org/10.1109/ACCESS.2019.2923707

[17] Maru, Ajay, Ajay Kumar Sharma, and Mayank Patel. "Hybrid Machine Learning Classification Technique for Improve Accuracy of Heart Disease." 2021 6th International Conference on Inventive

Computation Technologies (ICICT). IEEE, 2021. https://doi.org/10.1109/ICICT50816.2021.9358616

[18] Bharti, Rohit, et al. "Prediction of heart disease using a combination of machine learning and deep learning." Computational intelligence and neuroscience 2021. https://doi.org/10.1155/2021/8387680

[19] Singh, Poornima, Sanjay Singh, and Gayatri S. Pandi-Jain. "Effective heart disease prediction system using data mining techniques." International journal of nanomedicine 13.T-NANO 2014 Abstracts (2018): 121. https://doi.org/10.2147%2FIJN.S124998

[20] Gárate-Escamila, Anna Karen, Amir Hajjam El Hassani, and Emmanuel Andrès. "Classification models for heart disease prediction using feature selection and PCA." Informatics in Medicine Unlocked 19 (2020): 100330. https://doi.org/10.1016/j.imu.2020.100330

[21] Li, Jian Ping, et al. "Heart disease identification method using machine learning classification in e-healthcare." IEEE Access 8 (2020): 107562-107582. https://doi.org/10.1109/ACCESS.2020.3001149

[22] Hossen, M. D., Tazin, T., Khan, S., Alam, E., Sojib, H. A., Monirujjaman Khan, M., & Alsufyani, A. (2021). "Supervised machine learning-based cardiovascular disease analysis and prediction" Mathematical Problems in Engineering, 2021. https://doi.org/10.1155/2021/1792201

[23] Jagtap, Ankita, Kamthe, Pratiksha, bebal, Arbaaz et al. "Blockchain-based Secure Healthcare for Cardio Disease Prediction of Arrhythmia", International Journal for Research in Applied Science & Engineering Technology (2022). https://doi.org/10.22214/ijraset.2022.43299