International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-10 Issue-5, January 2022

Implementation of Machine Learning Model to Predict Heart Problem

Shruti Patil, Mrunal Annadate

Abstract: ***With the rapid growth of technology and data, the healthcare domain has emerged as one of the most important research areas in the modern period. Machine Learning is a novel method for disease prediction and diagnosis. This study demonstrates how machine learning can be used to forecast disease based on symptoms. Techniques of Machine learning such as Bayes, Random Forest, and SVM are used to forecast the disease on the supplied dataset. The research determines which algorithm is the best based on its accuracy. The accuracy of an algorithm is determined by its performance on a particular dataset. One of the most significant disorders is heart disease. We discovered machine learning models to predict heart problems in order to lower the incidence of death caused by heart disease. In this paper, we used a dataset from 1988 that included four databases: Cleveland, Hungary, Switzerland, and Long Beach V., and applied an algorithms to it to obtain the results. Previous studies had lower accuracy, therefore we focused on this research to enhance accuracy rate, precision, and recall which are very crucial parameters in medical field, in order to forecast heart problems and rescue patients. In this paper, we worked on different algorithms such as SVM, Random Forest, Naïve Bayes, Neural Network and Decision Tree. The model was implemented using the Python programming language. Analysis result indicates that SVM and Decision Tree algorithms have achieved highest accuracy which is 98.05%.***

Keywords: Heart Disease, Random Forest, Machine Learning, Naïve Bayes, Pre-processing.

I. INTRODUCTION

Heart disease (HD) is regarded as one of the most serious human diseases. In this disease, the heart is frequently unable to push an adequate amount of blood to other areas of the body in order to perform the body's regular functions, and heart failure develops as a result. Heart disease is a term used to describe a group of heart-related conditions. Coronary artery disease, heart rate issues (Arrhythmias), and congenital abnormalities are some of the examples of heart disease. The terms "cardiovascular disease" or "heart disease" are frequently used interchangeably. Cardiovascular disease refers to a condition in which blood vessels become narrow or sometimes blocked which can lead to a heart attack.

Manuscript received on January 16, 2022.
Revised Manuscript received on January 23, 2022.
Manuscript published on January 30, 2022.
* Correspondence Author
Shruti Gurudas Patil*, Department of Electronics and Telecommunication Engineering, Prof. Dr. Vishwanath Karad MIT World Peace University, Pune (MH), India. E-mail: shrutigpatil22@gmail.com
Dr. Mrunal Ninad Annadate, Department of Electronics and Telecommunication Engineering, Prof. Dr. Vishwanath Karad MIT World Peace University, Pune (MH), India. E-mail: mrunal.annadate@mitwpu.edu.in

Heart disease is one of the leading causes of death. One of the most crucial topics in clinical data analysis is heart disease prediction. A considerable amount of raw healthcare data is transformed into information that can be used to make better decisions and predictions using data mining. Machine learning algorithms such as SVM, DT, NB, RF, and ANN are used in the proposed system. In the proposed system, we have uploaded the Heart Disease Dataset and trained the model using a machine learning technique. We applied three separate data processing modules in this system: pre-processing, feature extraction, and classification, all of which utilizes different algorithms. Then, constructed a model and evaluated its effectiveness to predict the heart disease.

II. PROBLEM STATEMENT

In this study, we have used several Data Mining Techniques to develop a Heart Disease Prediction System. It also aims to increase the accuracy and precision with which heart disease risk can be predicted.

III. RELATED WORK

Artificial intelligence (AI) is revolutionary because it is a collection of computer science techniques that over the next several years and decades will hugely dominate the world. Artificial intelligence-based systems will be used to augment both defensive and offensive cyber operations. AI technology will enable new and innovative methods of conducting cyber-attacks, while also enhancing the security of AI technology. Finally, AI's requirement for massive amount of training data will emphasize the importance of data. This fundamentally alters the way we must think about data protection [4] [9]. Global prudence is essential to ensure broad safety and prosperity that this game-changing technology can bring.
The services above all use machine learning in one way or another. It powers self-driving vehicles and devices that can examine photos for the presence of medical issues. These days, businesses almost use machine learning so much that they are typically referred to as “artificial intelligence”, even if that is incorrect. In machine learning, computers can learn without being programmed. “Machine learning has been making steady inroads into AI in the last 5 to 10 years, and is becoming an even more critical method today,” says MIT Sloan professor Thomas W. Malone, who co-founded the MIT Center for Collective Intelligence. This suggests that a lot of AI development has focused on machine learning recently [7].

IV. LITERATURE SURVEY

Anjan Nikhil Repaka et al. [1] The authors' technique includes data collection, user registration and login (application-based), categorization using Naive Bayes model, prediction, and secure data transfer using AES (Advanced Encryption Standard). Following that, a result is generated. The research elaborates and offers various knowledge abstraction methodologies by utilizing data mining tools for heart disease prediction. The data suggests that the current diagnostic approach is capable of predicting risk factors for heart disease.

Pahulpreet Singh Kohli et al. [2] The authors' employed three separate classification methods, each with its own set of benefits, to predict disease in three different disease databases available in the UCI repository (Heart, Breast cancer, and Diabetes). The p-value test was used to choose features for each dataset using backward modeling. The findings of the study back up the notion of employing machine learning to detect diseases early. The prediction accuracy of this study technique is 87.1 percent for Heart Disease detection using Logistic Regression, 85.71 percent for Diabetes prediction using Support Vector Machine (linear kernel), and 98.57 percent for Breast Cancer detection using AdaBoost classifier.

Cincy Raju et al. [3] Data mining techniques have been suggested as a feasible solution for restorative circumstances. Data mining classification methods used to diagnose cardiac illnesses include decision trees, neural networks, Bayesian classifiers, Support vector machines, Association Rule, and K-nearest neighbor classification. According to this study, the Support Vector Machine (SVM) technique is a good tool for predicting cardiac disease. It achieves a high level of accuracy by analyzing a variety of research papers.

Rohit Binu Mathew et al. [4] They have developed a system to provide an alternative to the traditional way of visiting a hospital and scheduling an appointment with a doctor to receive a diagnosis. The goal of this study is to construct a chatbot application using natural language processing and machine learning technologies. People can communicate with the chatbot in the same way they would with another human, and the chatbot will identify the user's symptoms and, as a result, anticipate the disease and prescribe treatment through a series of questions.

This technique can be very useful in doing daily check-ups, making individuals aware of their health state, and encouraging them to take the necessary precautions to stay healthy. This study describes a medical chatbot that can be used to replace traditional disease diagnostic and therapy suggestion methods. A chatbot can perform the role of a doctor. The chatbot serves as a user interface. The user of this application can tell the chatbot about their symptoms, and the chatbot will tell them what health measures they should take.

Abderrahmane Ed-daoudy et al. [5] The authors' suggest a real-time heart disease prediction system based on Apache Spark, a strong large-scale distributed computing platform that can be used for streaming data events rather than machine learning via in-memory calculations. The system's two main components are streaming processing and data storage and display. To forecast heart illness, the first uses Spark MLlib and Spark streaming to apply a classification model to data events. The massive amount of data created is saved in Apache Cassandra in seconds.

Ashir Javeed et al. [6] The authors' of this study proposed that in diagnostic system RSA was used for feature selection, and a random forest model was used to predict heart failure. The grid search method was used to optimize this diagnostic system. To assess the method's precision, two types of experiments were carried out. In the first experiment, only a random forest model was constructed, however in the second experiment, the proposed RSA-based random forest model was created. System is more economical and less complicated than a classic random forest model, offering a 3.3 percent higher accuracy with only seven features.

Archana Singh et al. [11] A predictive system for disease awareness was presented as a solution to the problem. They investigated the accuracy of machine learning algorithms for predicting cardiac illness using the UCI repository dataset for training and testing, including k-nearest neighbor (87%), decision tree (79%), linear regression (78%), and support vector machine (SVM) (83 %). They used Anaconda (jupyter) notebook as it is the best tool for implementing Python programming since it comes with a variety of libraries and header files that help to make the job more accurate and efficient.

Md. Razu Ahmed et al. [8] In this paper, they described a cloud-based 4-tier architecture that may significantly improve patient health data prediction and monitoring. They used five common supervised learning-based machine learning approaches to diagnose heart disease early. The main purpose was to assess how well the chosen classification techniques function. Machine learning classifiers' prediction accuracy is as follows:

SVM (82%), ANN (84%), NB (82%), DT (77%), and RF (77%).

Chaitanya Suvarna et al. [9] The project's purpose was to combine data mining and optimization approaches to create a prediction algorithm. Data mining is the process of identifying information or decision-making knowledge in a database and extracting it in such a way that it may be used in fields such as decision support, forecasts, forecasting, and estimation. CPSO (55.1%), PSO (53.1%), ID3 (51.4%), MLP with BP (48.5%) and CART (48.5%) were found to be the most common particle swarm data mining methods according to the results.

Rajesh N et al. [10] They have used a variety of factors connected to heart disease to develop a better prediction technique and algorithms. The Naive Bayes technique was applied to a dataset based on risk variables. They have used decision tree and a variety of algorithms to forecast heart disease based on the qualities provided. The Naive Bayes technique delivered accurate results when the dataset was small, while decision tree produced accurate results when the dataset was large.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

V. BLOCK DIAGRAM

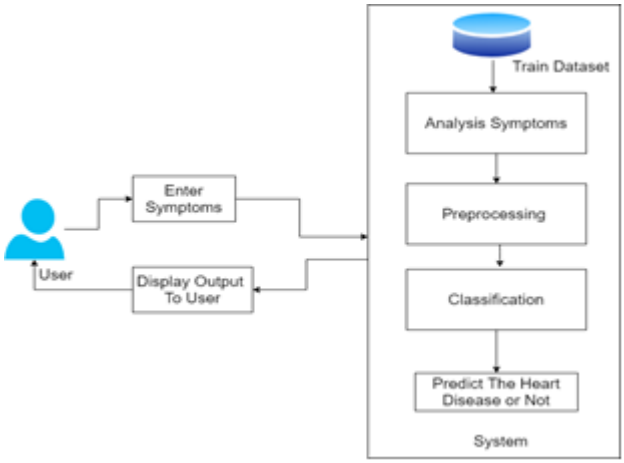


Fig 1. Block Diagram

A. Modules

- Enter Symptoms: Symptoms of a heart attack includes: chest pain (cp), cholesterol (chol), fasting blood sugar (fbs), resting electrocardiographic (restecg), The person's maximum heart rate achieved (thalach), Exercise induced angina (exang), ST depression induced by exercise relative to rest (oldpeak), slope of the peak exercise ST segment(slope), The number of major vessels (ca), thalassemia (thal) etc.
- Train Dataset: Identify your goal. The initial step is to pinpoint the set of objectives that you want to achieve through a machine learning application. Select suitable algorithms. Different algorithms are suitable for training artificial neural networks. Develop your dataset.
- Analyze Symptoms: In this module we analyze the parameters of the symptoms value like cp, chol etc.
- Pre-processing: Pre-processing is a term used to describe processes on datasets where the input and output are both intensity text at the lowest level of abstraction which was used to enhance text data before by suppressing undesired distortions or enhancing certain text features that are relevant for later processing.
- Feature Extraction: Using the feature extraction technique, we can create new features which are a linear mixture of current features. When compared to the original feature values, the new set of features will have different values. The main goal is to utilize fewer features to obtain the same quantity of data.
- Classification: Classification is a supervised machine learning technique in which an algorithm learns from the data provided and then applies what it has learnt to classify fresh observations. To put it another way, the training dataset is used to improve boundary conditions, which may subsequently be utilized to define each target class; after these boundary criteria is created, the next goal is to predict the target class. SVM, DT, RF, NB, and ANN are the most common classification methods.
- Result: Predict the heart disease detected or not.

VI. PROPOSED ALGORITHM

- Step 1: Choose a dataset.
- Step 2: Data Pre-processing - Overview of the data.
- Step 3: Detection and removal of outliers
- Step 4: Choose a model - Model selection is the third step in the machine learning process. There are different algorithms for different tasks; choose the right algorithm.
- Step 5: Train the model.
- Step 6: Evaluate the model- Make a confusion matrix and divide the data into 70/30 or 80/20.
- Step 7: Prediction- Check the result is accurate or not.

A. Support Vector Machine:

SVM (Support Vector Machine) is a controlled machine learning approach that can be used to solve classification and regression problems. It is employed largely in classification issues. It is most commonly utilize in classification issues. We represent each data item as a point in n-dimensional space (where n is the number of features you have) with the value of each feature being the value of a specific position in the SVM algorithm. The SVM is a boundary between two classes (hyper planes / rows). It is categorized by finding the hyper-plane that distinguishes the classes very well. Computing the (soft-margin) SVM classifier amounts to minimizing an expression of the form

$$\left[\frac{1}{n}\sum_{i=0}^{n}\max(0,1-y_i(w.x_i-b))\right]+\lambda\|w\|^2$$

We concentrate on the soft-margin classifier because with linearly classifiable input data, a sufficiently small value for lambda generates the hard-margin classifier.

B. Artificial Neural Network (ANN) Algorithm:

Artificial Neural Network (ANN) is fully connected multi-layer neural networks. An input layer, multiple hidden layers, and an output layer make up these layers. Every node in one layer is connected to every other node in the following layer. By increasing the number of stored layers, we can make the network deeper. [6]

1. The ANN algorithm's steps

- Step 1: Defining a Sequential model is the first step.
- Step 2: Apply a sigmoid activation function to a dense layer.
- Step 3: Use an optimizer and a loss function to compile the model.
- Step 4: Analyze the data and fit the model to it.

The weighted total of the inputs is computed by the artificial neural network, which also incorporates a bias. A transfer function is used to express this computation.

$$\sum_{i=0}^{n} wi * Xi + b$$

The weighted total is used as an input to an activation function to generate the output. Activation functions determine if a node should fire or not. The only ones who make it to the output layer are those who are fired. Depending on the work at hand, there are a variety of activation functions that can be used.

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

C. Random Forest Algorithm:

Unsupervised classification is accomplished using the random forest technique. In keeping with its name, this procedure results in a dense forest. The size of a forest appears to be inversely related to the number of trees in the forest. Similarly, the random forest classifier gets increasingly accurate as the number of trees in the forest grows. Knowing how a decision tree works should allow you to guess how accurate the results will be. Some decision tree algorithm rules will be developed as a result of the objective-feature training. The test dataset may be predicted using the same set of rules [8] [10]. The Random Forest level is the sum of all the trees in the forest. The total number of trees is divided by the sum of the features importance rating on each tree: fi sub(i) = the value of a feature I ni sub(j)= node j's importance

$$RFfi_i = \frac{\sum_{jealltrees} normfi_{ij}}{T}$$

- T denotes total number of trees
- RFfi sub(i) denotes the importance of feature I estimated from all trees in the Random Forest model
- normfi sub(ij) denotes the normalized feature importance for I in tree j

D. Naive Bayes Algorithm:

Classification problems are tackled using the Naive Bayes algorithm, which is a supervised learning method based on the Bayes' Theorem. Text classes with a lot of dynamic data in the training set often utilize this technique. In order to swiftly create prediction models, the Naive Bayes' Classification algorithm is one of the most easy and successful methods to date. The Naive Bayes' algorithm is widely used in spam filtering, sentiment analysis, and article classification [1]. As the name suggests, it's naive since it believes certain attributes exist regardless of other factors. For example, in addition to being called an apple, this reddish-orange, spherical, and soft fruit is also referred to be a fruit because of these characteristics: Without being mutually exclusive, each trait helps to identify the fruit as an apple. The Bayes' theorem allows you to calculate the posterior probability, P(c|x), from P(c), P(x), and P(x|c) using P(c), P(x), and P(x|c). The Naive Bayes classifier assumes that the impact of a predictor's value (x) on a given class (c) is independent of the values of other predictors. Class conditional independence is the term for this assumption.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$$P(c|x) = P(x_1|c) \times P(x_2|c) \times \ldots \times P(x_n|c) \times P(c)$$

- *P(c|x) denotes the posterior probability of class (target)*
- *given predictor (attribute).*
- *P(c) denotes the prior probability of class.*
- *P(x|c) denotes the likelihood which is the probability of predictor given class.*
- *P(x) denotes the prior probability of predictor.*

E. Decision Tree Algorithm:

In general, decision tree analysis is a flexible technique that may be used in a wide range of predictive modeling scenarios. An algorithmic strategy that splits the data set in a variety of ways according to specified conditions can be used to create decision trees. Decision settings are the most powerful algorithms in the supervised algorithm category. They can be used to tackle problems with regression as well as classification. The two main parts of the tree are decision-making nodes, which divide the data and leave the outcomes. For a person who is fit or unfit, a binary tree is depicted to provide information on their age, eating habits, and practices [3]. "Information Gain" is a term that refers to the process of dividing data using entropy. The drop in entropy after splitting the dataset on an attribute is used to compute it:

$$Gain\ (T, X) = Entropy\ (T) — Entropy\ (T, X)$$

- T stands for target variable.
- Entropy (T, X) denotes the entropy estimated after splitting the data on feature X.
- X = Feature to be split on

VII. RESULTS

To get the desired result, we used the Python programming language. For well-balanced datasets, accuracy is a suitable metric to use. Precision is an important factor in reducing false positives. In a field like medicine, where we must reduce the probability of missing positive cases, recall is crucial. Precision and recall aren't always sufficient on their own. The F1-score is a combination of precision and recall.[13]

SVM and Decision Tree Classifier have more accuracy than the other machine learning models, according to the results. The comparison of the various parameters of the machine learning models is shown in the diagram below.

Here in Fig.2 and Fig.3, we can see that the accuracy of SVM and DT are the same, 98.05. As a result, we evaluated the execution time to see which classifiers produce the fastest results. We can see from the results that the decision tree produces the best results.

```
Classification Report :
              precision    recall  f1-score   support

           0       1.00      0.96      0.98       152
           1       0.96      1.00      0.98       156

    accuracy                           0.98       308
   macro avg       0.98      0.98      0.98       308
weighted avg       0.98      0.98      0.98       308

Confusion Matrix :
[[146   6]
 [  0 156]]


Execution Time: 2.797 seconds

Support Vector Machine :
Accuracy =  98.05194805194806
```

Fig 2. SVM Accuracy

```
Classification Report :
              precision    recall  f1-score   support

           0       0.96      1.00      0.98       135
           1       1.00      0.97      0.98       173

    accuracy                           0.98       308
   macro avg       0.98      0.98      0.98       308
weighted avg       0.98      0.98      0.98       308

Confusion Matrix :
[[135   0]
 [  6 167]]


Execution Time: 0.1795 seconds

Decision Tree Classifier :
Accuracy =  98.05194805194806
```

Fig 3. Decision Tree Classifier Accuracy

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

The accuracy rate of the Naive Bayes is 91.88, as shown in the Fig.4 which is lowest among these five algorithms.

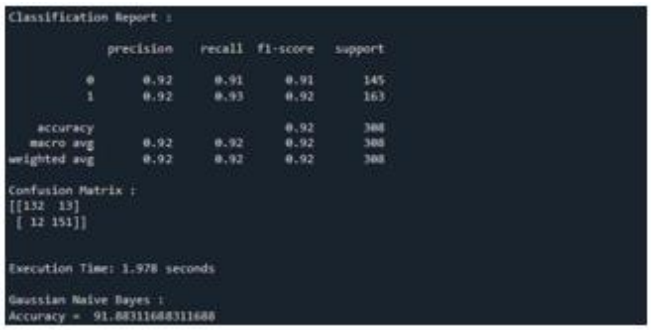


Fig 4. Naive Bayes Accuracy

The accuracy rate of the RF is 97.40, as shown in the Fig.5.The accuracy rate of the ANN is 97.56, as shown in the Fig. 6 with highest execution time.

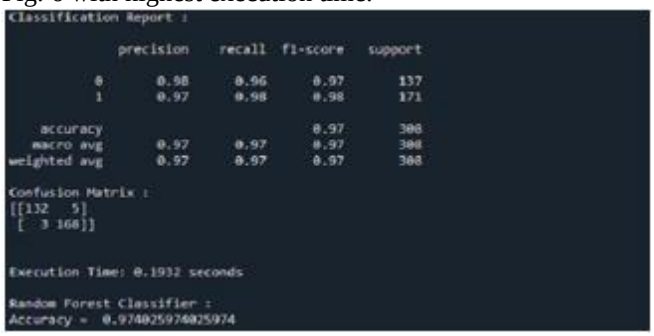


Fig 5. Random Forest Accuracy

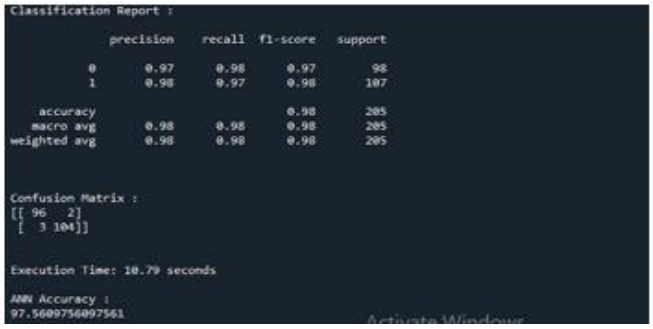


Fig 6. ANN Accuracy

The graph between True Positive Rate (or sensitivity) and False Positive Rate (specificity) is depicted by the ROC curve. Classifiers with curves that are closer to the top-left corner perform better. A random classifier is expected to give points that are diagonal (FPR = TPR) as a baseline. The test becomes less accurate when the curve approaches the ROC space's 45-degree diagonal.

Here, in Fig.7 and Fig.8, TPR is displayed on y-axis, and the FPR is displayed here on x-axis. From graph it is observed that ROC curve of SVM and DT is almost to the perfect classifier. In contrast, we can see in Fig.9 that the ROC curve is slightly off from the ideal classifier.

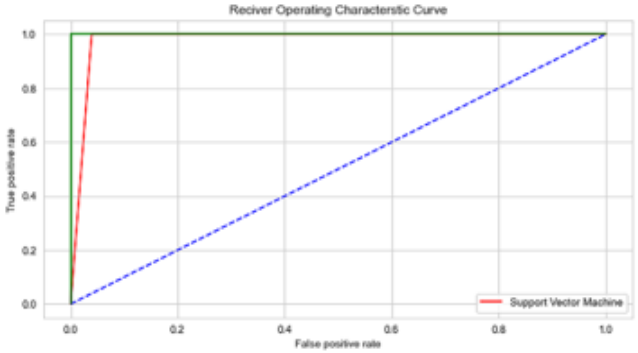


Fig 7. SVM Graph

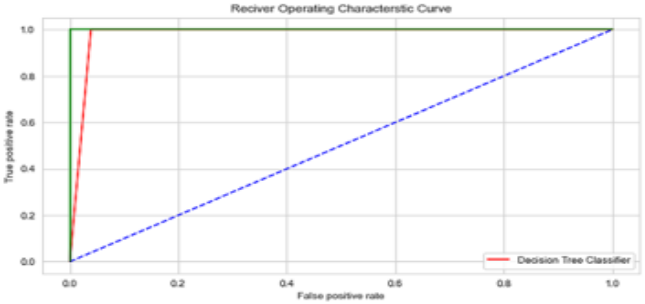


Fig 8. Decision Tree Classifier Graph

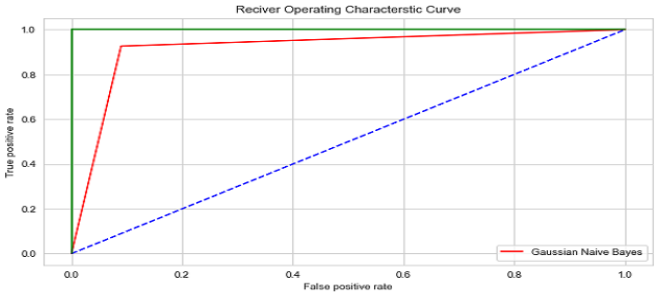


Fig 9. Naive Bayes Graph

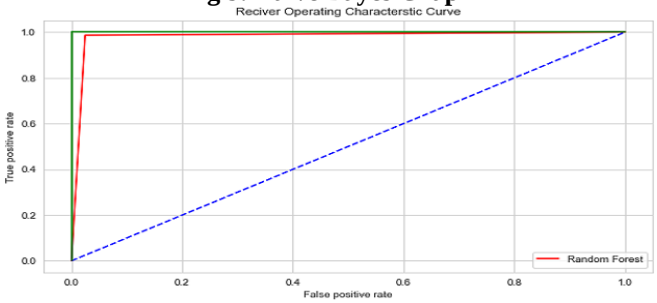


Fig 10. Random Forest Graph

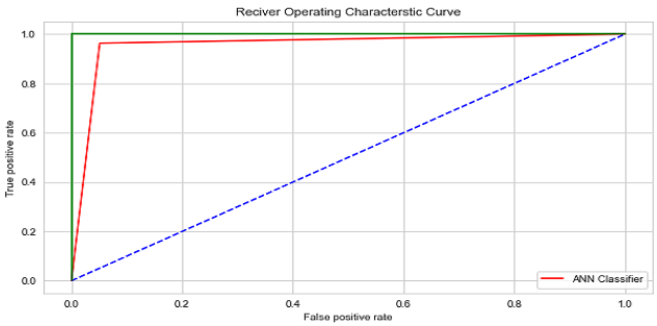


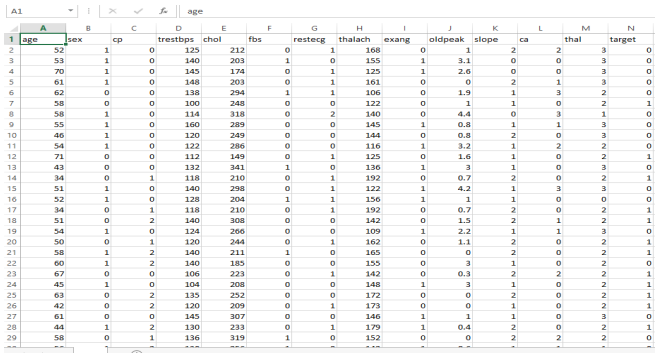
Fig 11. Artificial Neural Network Graph

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 1. Model Accuracies

Sr. No.	Models	Accuracy	Precisions	Recall	f1 score	Execution Time
1	SVM	98.05%	0.98	0.98	0.98	2.683 sec
2	Random forest	96.42%	0.96	0.97	0.96	0.1822 sec
3	Naïve Bayes	91.88%	0.92	0.92	0.92	2.099 sec
4	Decision Tree	98.05%	0.98	0.98	0.98	0.179 sec
5	ANN	96.10%	0.96	0.96	0.96	11.62 sec

VIII. DATASET



age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
52	1	0	125	212	0	1	168	0	1	2	2	3	0
53	1	0	140	203	1	0	155	1	3.1	0	0	3	0
70	1	0	145	174	0	1	125	1	2.6	0	0	3	0
61	1	0	148	203	0	1	161	0	0	2	1	3	0
62	0	0	138	294	1	1	106	0	1.9	1	3	2	0
58	0	0	100	248	0	0	122	0	1	1	0	2	1
58	1	0	114	318	0	2	140	0	4.4	0	3	1	0
55	1	0	160	289	0	0	145	1	0.8	1	1	3	0
46	1	0	120	249	0	0	144	0	0.8	2	0	3	0
54	1	0	122	286	0	0	116	1	3.2	1	2	2	0
71	0	0	112	149	0	1	125	0	1.6	1	0	2	1
43	0	0	132	341	1	0	136	1	3	1	0	3	0
34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
51	1	0	140	298	0	1	122	1	4.2	1	3	3	0
52	1	0	128	204	1	1	156	1	1	1	0	0	0
34	0	1	118	210	0	1	192	0	0.7	2	0	2	1
51	0	2	140	308	0	0	142	0	1.5	2	1	2	1
54	1	0	124	266	0	0	109	1	2.2	1	1	3	0
50	0	1	120	244	0	1	162	0	1.1	2	0	2	1
58	1	2	140	211	1	0	165	0	0	2	0	2	1
60	1	2	140	185	0	0	155	0	3	1	0	2	0
67	0	0	106	223	0	1	142	0	0.3	2	2	2	1
45	1	0	104	208	0	0	148	1	3	1	0	2	1
63	0	2	135	252	0	0	172	0	0	2	0	2	1
42	0	2	120	209	0	1	173	0	0	1	0	2	1
61	0	0	145	307	0	0	146	1	1	1	0	3	0
44	1	2	130	233	0	1	179	1	0.4	2	0	2	1
58	0	1	136	319	1	0	152	0	0	2	2	2	0

Fig 12. Dataset

People of all ages are at risk of developing heart disease. Every year, a large number of people die as a result of heart disease. If a person with heart disease follows all of the laws and regulations, he/she may be able to live a normal life. So, the most crucial thing is to figure out whether or not someone has heart disease. A dataset with 14 columns comprising various information about a person was provided. Using the other 13 columns, we must forecast the target column. This is an issue of supervised classification. We have included 14 columns in this dataset, including age, sex, cp, and trestbps, among others. In addition, the dataset has a total of 1024 rows. As a result, we complete the assignment over 1024 rows and 14 columns. [12]

IX. CONCLUSION

In this paper, we applied different machine learning algorithms like SVM, ANN, Random Forest, Decision tree, Naïve Bayes for diagnosis of heart disease. The goal of this study is to predict heart disease based on symptoms. The project is set up in such a way that the system takes the user's symptoms as input and outputs a prediction of heart disease. Proposed model recorded highest accuracy (98.05%) with SVM and Decision tree classification model. As a performance results of all classifier are promising models for medical data sets like heart disease with dependent attributes for diagnosis of disease.

REFERENCES

1. Anjan Nikhil Repaka, Sai Deepak Ravikanti," Design And Implementing Heart Disease Prediction Using Naives Bayesian" Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019) IEEE Xplore Part Number: CFP19J32-ART; ISBN: 978-1-5386-9439-8.
2. Pahulpreet Singh Kohli, Shriya Arora," Application of Machine Learning in Disease Prediction", 2018 4th International Conference on Computing Communication and Automation (ICCCA).
3. Cincy Raju et al." A Survey on Predicting Heart Disease using Data Mining Techniques" Proc. IEEE Conference on Emerging Devices and Smart Systems (ICEDSS 2018) 2-3 March 2018, Mahendra Engineering College, Tamilnadu, India.
4. Rohit Binu Mathew et al." Chatbot for Disease Prediction and Treatment Recommendation using Machine Learning", Proceedings of the Third International Conference on Trends in Electronics and Informatics (ICOEI 2019) IEEE Xplore Part Number: CFP19J32-ART; ISBN: 978-1-5386-9439-8.
5. Abderrahmane Ed-daoudy,Khalil Maalmi," Real-time machine learning for early detection of heart disease using big data approach", 978-1-5386-7850-3/19/$31.00 ©2019 IEEE
6. ASHIR JAVEED et al." An Intelligent Learning System based on Random Search Algorithm and Optimized Random Forest Model for Improved Heart Disease Detection", Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000. Digital Object Identifier 10.1109/ACCESS.2017.DOI.
7. Amin Ul et al." A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms", Hindawi Mobile Information Systems Volume 2018, Article ID 3860146, 21 pages https://doi.org/10.1155/2018/3860146.
8. Md. Razu Ahmed et al." A Cloud Based Four-Tier Architecture for Early Detection of Heart Disease with Machine Learning Algorithms", 2018 IEEE 4th International Conference on Computer and Communications.
9. Chaitanya Suvarna, Abhishek Sali," Efficient Heart Disease Prediction system using Optimization Technique", Proceedings of the IEEE 2017 International Conference on Computing Methodologies and Communication (ICCMC).
10. Rajesh N et al." Prediction of Heart Disease Using Machine Learning Algorithms", International Journal of Engineering & Technology, 7 (2.32) (2018) 363-366.
11. Archana Singh, Rakesh Kumar, "Heart Disease Prediction Using Machine Learning Algorithms", 2020 International Conference on Electrical and Electronics Engineering (ICE3-2020).
12. https://www.kaggle.com/bakar31/heart-disease-analysis-and-prediction
13. https://towardsdatascience.com/a-look-at-precision-recall-and-f1-score-36b5fd0dd3ec

AUTHORS PROFILE

Shruti Gurudas Patil received the BE degree in Electronics and Telecommunication from MKSSS's Cummins college for Women, Pune, India, in 2016. Currently a student perusing MTech degree in Communication Network and Software from Dr. Vishwanath Karad, MIT World Peace University, Pune, India.

Dr. Mrunal N. Annadate, is currently working as Assistant Professor in MIT World Peace University, Pune, India. in the School of Electronics and Communication Engineering. She is having teaching experience of more than 23 years. She has published more than 20 research papers in various reputed journals and conferences. She is an author of one book chapter based on her research work also she has published one patent. Her areas of interest include Image processing, Signal/Video processing, Programming and Embedded systems. She was awarded Ph.D. degree by University of Pune, India in 2021 in Electronics and Telecommunication Engineering in the domain of Video Processing in biomedical Engineering.