



HEART DISEASE PREDICTION MODEL

Dr.Mahammad Shabana

HOD & Associate Professor of IT Department,
Aijaz Afzal, Syed Aqib, Mohd Irfan Ahmed,
UG Students
IT Department
ISL Engineering College, Telengana.

ABSTRACT

For healthcare professionals, the challenge of predicting and diagnosing cardiac disease has always been a difficult one. Treatments for heart illness are available via a variety of medical facilities, some of which charge exorbitant fees. People all across the globe will benefit from being able to detect the onset of heart disease early on so that they may take appropriate action before things worsen. Alcohol, smoke, and a lack of exercise are all factors that contribute to heart disease, which has become a major concern in recent years. Using a wide range of health care data, machine learning has shown itself to be an useful tool for generating predictions and judgments. It has become more commonplace to use machine learning in the medical field as a result of its fast progress. Computer-aided learning, or ML, is a kind of artificial intelligence (AI) that uses data to create better predictions. The purpose of this project is to use the UCI heart illness dataset to construct a web application for heart disease categorization. Deploying the model and creating a website are done using Flask API.

Introduction

Mortality from cardiovascular disease is on the rise globally [1]. It's a difficult condition to diagnose, and it affects millions of

individuals throughout the world [2]. This is a condition that affects the heart and blood vessels. Smoking, high blood pressure, high cholesterol, a poor diet, lack of exercise, and obesity all raise the risk of heart disease. Coronary artery disease (narrowing or blocking of the coronary arteries) is the most prevalent kind of heart disease and may result in chest discomfort, heart attacks, or a stroke. Congestive heart failure, irregular heartbeat, congenital heart disease, and endocarditis are among conditions that may affect the heart (inflamed inner layer of the heart). [3] Also referred to as heart disease [4]. Males and heart disease: What are the effects on men of heart disease? In the United States, heart disease is the top cause of mortality for males, accounting for 357,761 fatalities in 2019. Many ethnic and racial groups are affected by heart disease, including African-Americans, American Indians and Alaska Natives, Hispanics, and whites. Cardiovascular disease is the leading cause of death among Asian American and Pacific Islander males. 2

Coronary heart disease affects one in every 13 white men (7.7%) and one in every 14 black men (7.1%). Hispanic males have a higher rate of coronary heart disease than any other ethnic group (5.9%). There were no prior signs in half of the men who died abruptly from coronary heart disease. Even if you don't show any signs of heart disease, you may still be at risk. When it comes to heart disease, how do women fare? It is

estimated that in 2017, 299,578 women died as a result of heart disease, making it the top cause of mortality in this country for women.

In the United States, heart disease is the top cause of mortality for both black and white women. Every year, heart disease and cancer claim the lives of almost the same number of American Indian and Alaska Native women. Heart disease is the second leading cause of mortality among Hispanic and Asian/Pacific Islander women, behind cancer.³

coronary heart disease, the most prevalent kind of heart disease, affects around one in every sixteen women over the age of 20.61% of white women; 66% of black women; and 65% of Hispanic women are in this group (6 percent)o About one in thirty Asian women are affected (3.2 percent)

CAUSES OF HEART DISEASE

Heart disease causes depend on your specific type of heart disease. There are many different types of heart disease. To understand the causes of heart disease, it helps to understand how the heart works. How the heart works?

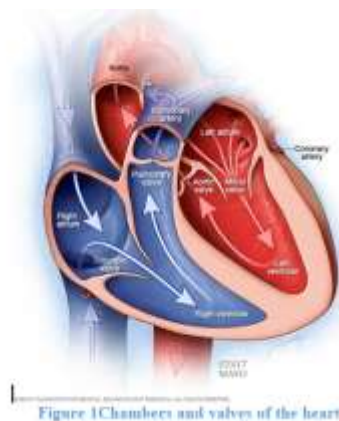


Figure 1 Chambers and valves of the heart

Heart is a pump. It's a muscular organ about the size of your fist, located slightly left of center in your chest. Your heart is divided into the right and the left sides. The right atrium and ventricle are located on the heart's right side. Through the pulmonary arteries, it gathers and pumps blood to the lungs. The lungs replenish the blood's oxygen supply. Carbon dioxide, a waste product, is also exhaled via the lungs.

Once in the left atrium and ventricle, oxygen-rich blood circulates throughout the body. The aorta, the body's biggest artery, is pumped by the left side of the heart, which carries oxygen and nourishment to every cell in the body.

Heart valves:

Four heart valves keep your blood moving the right way by opening only one way and only when they need to. To work properly, the valves must be formed properly, must open all the way and must close tightly so there's no leakage. The four valves are:

- Tricuspid
- Mitral
- Pulmonary
- Aortic

Heartbeats:

A beating heart squeezes (contracts) and relaxes in a continuous cycle. During contraction (systole), your ventricles squeeze tight, forcing blood into the vessels to your lungs and body. During relaxation (diastole), the ventricles are filled with blood coming from the upper chambers (left and right atria).

Electrical system:

The electrical circuitry in your heart ensures that it continues to beat. A regular pulse keeps oxygen-rich blood flowing into your body, while oxygen-poor blood flows out. Your life depends on this trade.

This mechanism maintains your heart pumping in a coordinated and regular rhythm so that blood may circulate throughout your body. • Electrical signals originate high in the upper right chamber (right atrium) and travel down specific channels to the ventricles.

Causes of coronary artery disease

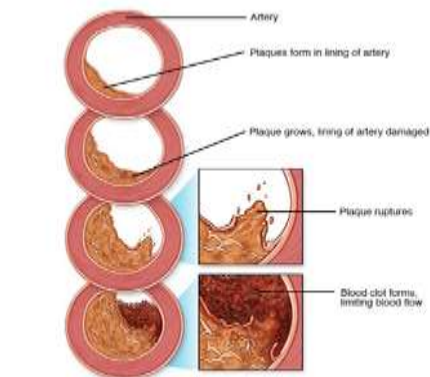


Figure 2 Development of atherosclerosis

Development of atherosclerosis:

The most prevalent cause of coronary artery disease is atherosclerosis, or the buildup of fatty plaques in the arteries (blockages). Atherosclerosis may be exacerbated by unhealthy behaviours such as smoking, being overweight, or not getting enough exercise.

Causes of heart arrhythmia:

Arrhythmias may be caused by coronary artery disease, diabetes, and many other illnesses.

- Abuse of illicit drugs
- Excessive use of alcohol or coffee might lead to health complications.
- Congenital heart defects (congenital heart defects)
- An elevated heart rate.
- Over-the-counter pharmaceuticals, prescription drugs and natural therapies
- Smoking
- Valvular

Heart disease

Without some external trigger, such as an electrical shock or the use of illicit substances, a lethal arrhythmia is unlikely to arise in a healthy individual with a normal, healthy heart. Heart illness or deformity may interfere with normal transmission of electrical impulses, increasing the risk of arrhythmias. During a baby's development in the womb, congenital cardiac abnormalities are most often discovered. A month after conception, the heart begins to acquire abnormalities, which may alter the flow of blood in the heart. Heart abnormalities may be caused by a variety of factors, including medical conditions, drugs, and even genes. Adults might acquire heart abnormalities as well. Alterations in the structure of your heart may result in an abnormality as you become older.

Restrictive cardiomyopathy

There is no known aetiology for this rare form of cardiomyopathy, in which the heart muscle stiffens and loses its elasticity. The development of aberrant proteins or



connective tissue problems may also be to blame (amyloidosis).

Causes of heart infection:

A heart infection, such as endocarditis, is caused when germs reach your heart muscle. The most common causes of heart infection include:

Bacteria

Viruses

Parasites

Causes of heart infection:
There are several different types of heart infections, including endocarditis. Heart infections may be caused by bacteria, viruses, and fungi.

- Viruses
- Parasites

Causes of valvular heart disease:

Heart valve disease may be caused by a variety of factors. For example, valves may be damaged by disorders like congenital valvular disease.

- Rheumatoid arthritis (infectious endocarditis)
- Disorders of the connective tissue.

RISK FACTORS AND COMPLICATIONS

Growing older is a risk factor for cardiovascular disease. Damaged and constricted coronary arteries and a weaker or thicker heart muscle may occur with ageing.

Sex: Heart disease is more common in men. After menopause, women's risk rises.

Genealogy: You're more likely to have coronary artery disease from a family history of heart disease, if a parent acquired it at a young age (before age 55 for a male

relative, such as your brother or father, and 65 for a female relative, such as your mother or sister).

Smoking: Carbon monoxide may damage the inner lining of your blood arteries, leaving them more vulnerable to atherosclerosis. Cigarette smokers have a higher risk of heart attacks than nonsmokers. A poor diet is a contributing factor. Heart disease may be exacerbated by a diet heavy in fat, salt, sugar, and cholesterol. There is a high risk of cardiovascular disease. Your arteries may stiffen and thicken as a consequence of uncontrolled high blood pressure, resulting in a constriction of the channels through which blood travels. High amounts of triglycerides in the blood. Plaque and atherosclerosis may be exacerbated by high cholesterol levels in the bloodstream.

Diabetes. Diabetes increases your risk of heart disease. Both conditions share similar risk factors, such as obesity and high blood pressure.

Obesity. Excess weight typically worsens other heart disease risk factors.

Physical inactivity. Lack of exercise also is associated with many forms of heart disease and some of its other risk factors as well.

Stress. Unrelieved stress may damage your arteries and worsen other risk factors for heart disease.

Poor dental health. It's important to brush and floss your teeth and gums often, and have regular dental checkups. If your teeth and gums aren't healthy, germs can enter your bloodstream and travel to your heart,



causing endocarditis [6]. Among the heart disease-related complications are the following: cardiac failure. A frequent consequence of cardiac illness, heart failure happens when your heart is unable to pump enough blood to fulfil your body's requirements. There are several causes of heart failure, including congenital cardiac abnormalities, congenital or acquired heart disease, valvular heart disease, infections of the heart, and cardiomyopathy itself. Stroke of heart. One of the most common causes of heart attacks is the blockage of the heart's main supply of blood, which results in the heart muscle being damaged or destroyed. A heart attack may be caused by atherosclerosis.

Stroke: You may also have a stroke from the same risk factors that cause cardiovascular disease, such as smoking, high blood pressure, obesity, and high cholesterol. Within a few minutes following a stroke, brain tissue starts to die. This is a medical emergency

Aneurysm: A serious complication that can occur anywhere in your body, an aneurysm is a bulge in the wall of your artery. If an aneurysm bursts, you may face life-threatening internal bleeding.

Peripheral artery disease. When you develop peripheral artery disease, your extremities — usually your legs — don't receive enough blood flow. This causes symptoms, most notably leg pain when walking (claudication). Atherosclerosis also can lead to peripheral artery disease.

Sudden cardiac arrest. Sudden cardiac arrest is the sudden, unexpected loss of heart

function, breathing and consciousness, often caused by an arrhythmia. Sudden cardiac arrest is a medical emergency. If not treated immediately, it results in sudden cardiac death [6].

MAKING THE DIAGNOSIS

In addition to a physical examination, your doctor will inquire about your own and your family's medical history. If your doctor suspects you have heart disease, the tests you'll require may vary. Besides a blood test and a chest X-ray, there are more tests that may be used to identify heart disease:.

Electrocardiogram (ECG or EKG) During an ECG, your heart's electrical impulses are recorded quickly and painlessly. It has the ability to detect aberrant heartbeats. An ECG may be taken while you're lying down or while you're working out (stress electrocardiogram). **Holter monitoring.** With a Holter monitor, you can keep track of your heart rate and rhythm continuously for up to 72 hours. Holter monitoring is used to find irregularities in the heart's rhythm that aren't picked up by a standard electrocardiogram (ECG).

Echocardiogram. The heart's anatomy may be seen in great detail thanks to the use of sound waves in this noninvasive scan. This video demonstrates the heart's rhythm and blood flow.. **Stress test** Exercise or medication is used to raise your heart rate while cardiac tests and imaging are carried out to see how your heart functions. **Cardiac catheterization.** An artery or vein in your leg or arm is pierced with a small tube (sheath). The sheath is next opened and a hollow, flexible, and longer tube (the guide catheter)



is inserted. Your doctor uses X-ray pictures on a monitor to guide the catheter down the artery until it reaches your heart. Cardiac catheterization allows for pressure measurements and dye injections in the heart chambers. An X-ray shows the flow of blood through your heart, blood arteries, and valves, allowing your doctor to look for any abnormalities. Cardiac computerized tomography (CT) scan. You recline on a table in a doughnut-shaped machine for a heart CT scan. Images of your heart and lungs are captured by an X-ray tube that revolves around your body while you sit in the machine.

Cardiac magnetic resonance imaging (MRI):

Heart pictures are produced by cardiac MRI by using a magnetic field and computer-generated radio waves [7].

TREATMENT: The kind of cardiac condition you have determines the sort of therapy you will get. Typically, heart disease therapy requires a combination of medication and lifestyle changes.:

Lifestyle changes: Heart disease may be prevented by eating a diet low in fat and high in salt, exercising at least 30 minutes a day, stopping smoking, and reducing alcohol use.

Medications: In the event that lifestyle modifications alone are not adequate to manage your heart disease, your doctor may prescribe medication. Depending on the kind of heart condition you have, the medicine you take will be different.

Medical procedures or surgery: In certain cases, your doctor may propose a particular operation or surgery. Procedures and surgeries will be tailored to your heart condition and how much harm has been done to it [7].

Restrictive cardiomyopathy: This least common type of cardiomyopathy, which causes the heart muscle to become rigid and less elastic, can occur for no known reason. Or it may be caused by diseases, such as connective tissue disorders or the buildup of abnormal proteins (amyloidosis).

Causes of heart infection:

There are several different types of heart infections, including endocarditis. Infections of the heart may be caused by a variety of factors, including: Bacteria

- Viruses
- Parasites

Causes of valvular heart disease:

Heart valve disease may be caused by a variety of factors. For example, valves may be damaged by disorders like congenital valvular disease. Rheumatoid arthritis (infectious endocarditis) Disorders of the connective tissue [6]

RISK FACTORS AND COMPLICATIONS

Heart disease may be caused by a number of conditions, including::

Age. Damaged and constricted coronary arteries and a weaker or thicker heart muscle may occur with ageing.



Sex. Heart disease is more common in men. After menopause, women have an increased risk..

Family history. You're more likely to have coronary artery disease from a family history of heart disease, if a parent acquired it at a young age (before age 55 for a male relative, such as your brother or father, and 65 for a female relative, such as your mother or sister).

Smoking. Carbon monoxide may damage the inner lining of your blood arteries, leaving them more vulnerable to atherosclerosis. Smokers are more likely than nonsmokers to have a heart attack..

Poor dietHeart disease may be exacerbated by a diet heavy in fat, salt, sugar, and cholesterol.

High blood pressure

High blood cholesterol levels. High levels of cholesterol in your blood can• High amounts of triglycerides in the blood. It is possible to have high cholesterol levels in your blood. Hardening and thinning of the arteries may happen from uncontrolled high blood pressure. increase the risk of plaque formation and atherosclerosis.

Diabetes. Cardiovascular disease is more common in those with diabetes. Obesity and high blood pressure are common risk factors for both illnesses..

Obesity. Obesity has been shown to increase the likelihood of developing heart disease.

Physical inactivity. Many types of heart disease and other risk factors might be linked to a lack of exercise..

Stress. As a result of long-term stress, you may be more susceptible to heart disease..

Poor dental health. Brushing and flossing your teeth and gums on a regular basis is essential, as is seeing the dentist for regular examinations. Endocarditis (infection of the heart) may occur if bacteria from your mouth reach your circulation [6].

Heart failure:A frequent consequence of cardiac illness, heart failure happens when your heart is unable to pump enough blood to fulfil your body's requirements. There are several causes of heart failure, including congenital cardiac abnormalities, congenital or acquired heart disease, valvular heart disease, infections of the heart, and cardiomyopathy itself.Stroke of heartIt's possible that the heart muscle may be damaged or even destroyed by a clot impeding the flow of blood to the organ. A heart attack may be caused by atherosclerosis.

Stroke. You may also have a stroke from the same risk factors that cause cardiovascular disease, such as smoking, high blood pressure, obesity, and high cholesterol. Within a few minutes following a stroke, brain tissue starts to die. This is a medical emergency

Aneurysm: A serious complication that can occur anywhere in your body, an aneurysm is a bulge in the wall of your artery. If an aneurysm bursts, you may face life-threatening internal bleeding.

Peripheral artery disease: As a result of peripheral artery disease, your legs are deprived of enough blood supply. Leg discomfort while walking is the most notable symptom of this (claudication). Peripheral artery disease may also be caused by atherosclerosis..

Sudden cardiac arrest:Sudden cardiac arrest is the loss of heart function, respiration, and consciousness that occurs



suddenly and unexpectedly, usually as a result of an arrhythmic condition. A sudden cardiac arrest is a medical emergency. Sudden cardiac death may occur if it is not addressed quickly [6].

MAKING THE DIAGNOSIS

In addition to a physical examination, your doctor will inquire about your own and your family's medical history. If your doctor suspects you have heart disease, the tests you'll require may vary. Besides a blood test and a chest X-ray, there are more tests that may be used to identify heart disease:**Electrocardiogram:**(ECG or EKG). During an ECG, your heart's electrical impulses are recorded quickly and painlessly. It has the ability to detect aberrant heartbeats. An ECG may be taken while you're lying down or while you're working out (stress electrocardiogram).

Holter monitoring: With a Holter monitor, you can keep track of your heart rate and rhythm continuously for up to 72 hours. Holter monitoring is used to find irregularities in the heart's rhythm that aren't picked up by a standard electrocardiogram (ECG).

Echocardiogram.:The heart's anatomy may be seen in great detail thanks to the use of sound waves in this noninvasive scan. Your heart is seen beating and pumping blood using this device.

Stresstest: An exercise or medication-induced increase in heart rate is used in conjunction with a variety of cardiac tests and imaging modalities to determine how well your heart reacts to the stress.

Cardiac catheterizationA tiny tube is used to puncture an artery or vein in your leg or arm

(sheath). Afterwards, the sheath is opened, and a longer guide catheter (which is hollow, flexible, and longer in diameter) is inserted. To get the catheter as close to your heart as possible, your doctor will utilise X-ray images on a monitor. Cardiac catheterization allows for pressure measurements and dye injections in the heart chambers. An X-ray shows the flow of blood through your heart, blood arteries, and valves, allowing your doctor to look for any abnormalities.

Cardiac computerized tomography (CT) scan.:You recline on a table in a doughnut-shaped machine for a heart CT scan. Images of your heart and lungs are captured by an X-ray tube that revolves around your body while you sit in the machine.**Cardiac magnetic resonance imaging (MRI)** Heart pictures are produced by cardiac MRI by using a magnetic field and computer-generated radio waves [7].

TREATMENT

The kind of cardiac condition you have determines the sort of therapy you will get. Treatment for cardiac disease often comprises the following:**Lifestyle changes.** Heart disease may be prevented by eating a diet low in fat and high in salt, exercising at least 30 minutes a day, stopping smoking, and reducing alcohol use.**Medications.** If lifestyle changes alone aren't enough, your doctor may prescribe medications to control your heart disease. The type of medication you receive will depend on the type of heart disease.

Medical procedures or surgery.In certain cases, your doctor may propose a particular operation or surgery. Procedures and surgeries will be tailored to your heart

condition and how much harm has been done to it [7].

Problem Definition

When it comes to making clinical judgments, physicians typically rely on their intuition and experience rather than the data's wealth of information. Errors and excessive expenditures on medical care are possible outcomes, but they are not guaranteed. Clinical judgments may be improved by the use of analytical techniques and data modelling. In other words, we're trying to create a tool that will make it easier for physicians to see signs of cardiac disease.

Scope of the project

Four distinct machine learning algorithms were used to make predictions regarding heart disease based on a dataset with many records. The prediction model was trained and tested on a limited number of variables. Another approach to improving model accuracy may be to include more techniques into the current model. – Further insights and improved prediction accuracy will be revealed by testing these models on big datasets with little or no missing attribute values.

Objectives

A dataset containing a large number of records was utilised to train four different machine learning algorithms to predict cardiac disease. A small number of variables were used to train and test the model. There may be a way to improve the model's accuracy by using other strategies. Testing these models on large datasets with few or no missing attribute values can provide more insights and enhance their prediction accuracy.

Motivation

Coronary artery disease is a severe and sometimes fatal condition. It has the potential to harm every organ in the body over time, including the kidneys, nerves, limbs, and eyes. Blood veins that feed the heart and other main organs of the body are affected by the condition, which is a significant cause of mortality globally. Early identification and treatment may lessen pain and avert serious health problems including heart failure, stroke, and renal disease.

3.1. Architecture Design

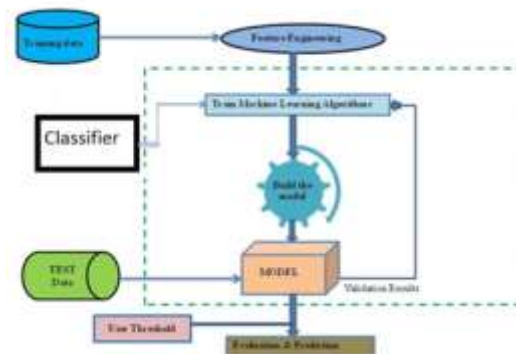


FIGURE 3. Block diagram Implementation

Implementation

Four different machine learning algorithms are included into the suggested study. These are the names of them: o Random Forest, a tree-based classification method The SVM algorithm is a functional classification method or Adaptive Enhancement

Amplification by Gradient5.1 Random Forest:

This method uses data samples to build decision trees, which it uses to generate predictions from each one before voting on the best one. There are many decision tree

classifiers in this ensemble, hence the collection of classifiers might be considered a "forest." An algorithm generates the decision trees. a random selection of attributes at The divide is determined by each node. Trees vote on the classes they want to be classified into.

Bagging and random attribute selection may be used to build random forests. d tuples are provided in the training set D . An example of an ensemble decision tree generation technique is shown below. A training set of d tuples, D_i , is sampled with replacement from D for each iteration, i ($i = 1, 2, \dots, k$). As a result, because each D_i is a bootstrap sample of D (Section 8.5.4), certain tuples may appear several times in D_i while others may be omitted. Each node is separated based on a limited number of characteristics (F), compared to the total amount of attributes. Randomly selecting F characteristics at each node as candidates for the split at the node is used to build a decision tree classifier, M_i . The trees are cultivated with the CART approach. The trees are allowed to reach their full potential and are not trimmed after that. When random input selection is used to create a forest, the result is known as Forest-RI.

Forest-RC is a kind of random forest that employs random linear combinations of the input characteristics. A linear combination of the existing characteristics is used instead of randomly picking a subset of the attributes. The number of original attributes to be merged, L , is used to construct a new attribute. All of the L characteristics at a particular node are randomly picked and combined with coefficients that are uniform random values on the range $[1,1]$. It is

necessary to build a large number of linear combinations before attempting to find the optimal split among them. When just a few variables are provided, this kind of random forest may be effective in reducing the correlation between individual classifiers [29].

If you're looking for a way to minimise overfitting, a random forest technique is the way to go. Even if a major amount of the data is missing, the Random Forest method still performs well.

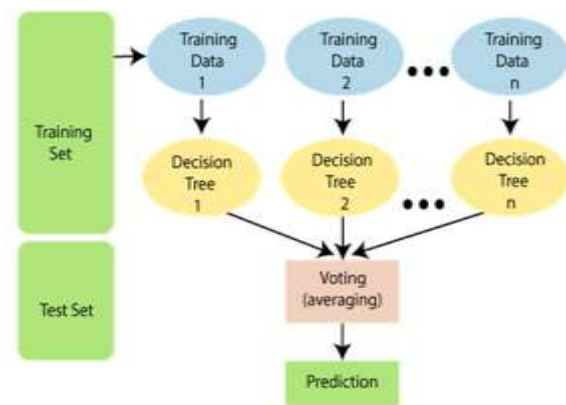


FIGURE 4. Working of Random Forest

SVM:

Linear and nonlinear data may be classified using the Support Vector Machine (SVM). This is how the algorithm works. The original training data is transformed into a higher dimension via a nonlinear mapping. A "decision border" that separates tuples of one class from those of another is sought inside this new dimension. Data from two classes may always be separated by a hyperplane if the nonlinear mapping to a sufficiently high dimension is adequate. With the help of "critical" training tuples

and margins, the SVM is able to discover this hyperplane (defined by the support vectors).

A key objective of the SVM method is to find the optimal decision boundary or line that can divide n-dimensional space into classes, allowing us to quickly classify fresh data points in the future. A hyperplane denotes the border of the optimal choice. To create the hyperplane, the SVM uses the most extreme points/vectors in the dataset. As a result, the technique is referred recognised as a Support Vector Machine [30]. [30] Even the quickest SVMs may take a long time to train, but their ability to simulate complicated nonlinear decision limits makes them exceptionally accurate. In comparison to other approaches, they are less prone to overfitting. The trained model may be concisely described using the support vectors that were discovered. Both classification and numeric prediction may be done using SVMs.

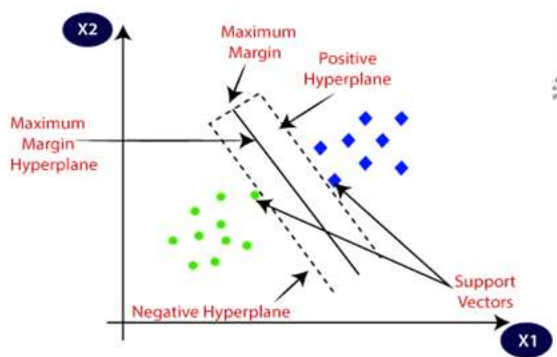


FIGURE 5. Classification using hyperplane SVM can be of 2 types:

Linear SVM:

For data that can be divided into two distinct categories by a single straight line, it is known as linearly separable data, or LSD data. There is no limit to the number of lines that may be drawn to separate two objects. Ideally, we're looking for the "best" one, which will have the least amount of classification mistake on previously unclassified tuples.

The optimal separation plane may be found if the data are 3-D (i.e., with three qualities). We discover the optimum hyperplane when we extrapolate to dimensions n. Searching for the hyperplane with the biggest margin, or the maximum marginal hyperplane, may be done via an SVM (MMH). Classes are most clearly separated by the related boundary. If the sides of the margin are parallel to the hyperplane, the shortest distance from one side of the margin to the other is equal to the smallest distance from the hyperplane to the other side. This is the smallest distance between the MMH and the nearest training tuple in either class when dealing with the MMH. Linear SVM is seen in the figure.5.2.2 Non-linear SVM:

Classification of non-linear data requires a classifier known as Non-linear SVM classifier [10], since a straight line cannot be utilised to divide the data.

We may get a nonlinear SVM by extending the linear SVM technique in the following way. Basically, there are two stages: Using a nonlinear mapping, we first shift the original input data into a higher-dimensional space. As we'll see in the following section, a variety of nonlinear mappings may be applied here. The second stage is to look for a linear separation hyperplane in the new

higher space once the data have been changed. We have a quadratic optimization issue that can be handled using the linear SVM formulation once again.. New space has a non-linear separating hyper surface that corresponds to the new space's greatest margina hyper plane.

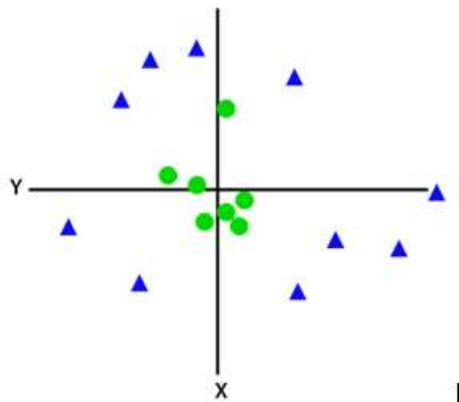
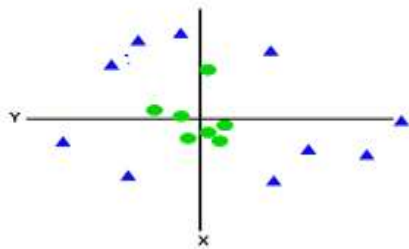


Figure 6 Non Seperable data



Adaptive Boosting:(gfg)

When it comes to binary classification, AdaBoost is the first effective boosting method. "Weak classifiers" may be combined into a single "strong classifier" using AdaBoost, a common boosting approach known as "adaptive boosting".

The algorithm is as follows: 1. Create a dataset and give each data point equal

weight. Input this into the model and find the data points that were incorrectly categorised. Increase the importance of data items that were incorrectly categorised.

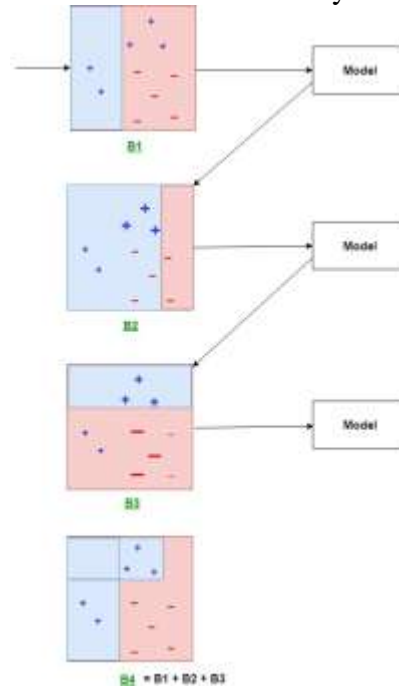


Figure 7 Adding third dimension to obtain Seperabledata

Figure 11 AdaBoost algorithm

The graphic above shows the AdaBoost algorithm in an easy-to-understand manner. Let's take it one step at a time, starting with this:

There are 10 data points in B1 which are divided into two categories, plus (+) and minus (-), with five plus (+) and five negative (-) data points apiece. Despite its best efforts, the first model incorrectly labels three pluses as minuses and draws a vertical divider line (-). Using the previous model's 10 data points, B2 includes the three incorrectly categorised pluses (+) in order to make the present model more likely to properly identify these pluses.. A vertical



separator line is generated by this model, which properly identifies the previously misclassified pluses(+), but in this effort, it incorrectly classifies three minuses (-).B3 is made up of the previous model's 10 data points, with the 3 incorrectly categorised minuses(-) being given a heavier weighting so that the current model can do a better job of classifying them. For each of the previously mislabeled negatives, this model creates a horizontal divider line (-).As a result, B4 is a considerably more accurate prediction model than any of the individual models utilised.

Proposed work:

It is the goal of this study to identify people who are at risk of developing heart disease in the future. Using a random sample of 271 patients, the predictive analytics is performed. Fourteen distinct parameters regulate the dataset used in this research, including age, sex, cp, trestbps, chol, fbs, restecg, thalach, exang, oldpeak, slope. ca. thal. and the study's aim.

Conclusion

A potential technique to early illness detection is machine learning, which may aid doctors in making diagnostic decisions. Many different approaches are available to help solve these problems. A difficult endeavour, predictive analytics in healthcare may assist practitioners make timely judgments regarding patients' health and treatment based on massive data. In the healthcare industry, analytics and human-centered design may guarantee that healthcare professionals eliminate inefficiencies along the patient journey and

personalise services to match the specific requirements of the patient group.

In this study, four different machine learning algorithms were utilised to predict heart disease using a dataset of 271 records. Predicting the development of cardiac disease with varying accuracy and classification report parameters is possible using the support vector machine, random forest, adaptive boost and gradient boosting classifiers. Patients can be prescribed the appropriate medication to prevent heart disease in the future by using a proposed prediction methodology that helps determine which algorithm among the four in consideration is most accurate and efficient so that it can be used as a part of normal/routine check-ups at hospitals. Studying the accuracies of several classifiers, researchers found that SVM was the most accurate in predicting whether or not a patient would develop heart disease in the future. The project's future scope may include adjustments and/or expansions that improve research productivity. Other types of machine learning methods might be included into the study in the future, thereby broadening the scope of this endeavour. Predictive analytics may also be used to other illnesses, such as cancer, using the same factors.

REFERENCES

- [1] Mohan, S., Thirumalai, C. and Srivastava, G., 2019. Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques. *IEEE Access*, 7, pp.81542-81554.
- [2] Li, J., Haq, A., Din, S., Khan, J., Khan, A. and Saboor, A., 2020. Heart Disease Identification Method Using Machine



Learning Classification in E-Healthcare. IEEE Access, 8, pp.107562-107582.

[3]"NCI Dictionary of Cancer Terms", National Cancer Institute, 2022. [Online]. Available: <https://www.cancer.gov/publications/dictionaries/cancer-terms/def/heart-disease>.

[Accessed: 11- May- 2022].

[4]"Men and heart disease", Centers for disease control and prevention, 2022. [Online].

Available: <https://www.cdc.gov/heartdisease/men.htm>.

[Accessed: 11- May- 2022].

[5] "Women and heart disease", Centers for disease control and prevention, 2022. [Online].

Available: <https://www.cdc.gov/heartdisease/women.htm>.

[Accessed: 11- May- 2022].

[6]"Heart disease - Symptoms and causes", Mayo Clinic, 2022. [Online]. Available:

<https://www.mayoclinic.org/diseases-conditions/heart-disease/symptoms-causes/syc-20353118>. [Accessed: 11- May- 2022].

[7]"Heart disease - Diagnosis and treatment - Mayo Clinic", MayoClinic.org, 2022. [Online].

Available: <https://www.mayoclinic.org/diseases-conditions/heart-disease/diagnosis-treatment/drc-20353124>.

[Accessed: 11- May- 2022].

[8] "The Python Tutorial — Python 3.8.2 documentation", Docs.python.org, 2020. [Online].

Available: <https://docs.python.org/3/tutorial/>.

[Accessed: 19- Apr- 2020].

[9] "Python: 7 Important Reasons Why You Should Use Python", Medium, 2020. [Online].

Available: <https://medium.com/@mindfiresolutions.usa>

[/python-7-important-reasons-why-you-should-use-python-5801a98a0d0b](https://medium.com/@mindfiresolutions.usa/python-7-important-reasons-why-you-should-use-python-5801a98a0d0b).

[Accessed: 19- Apr- 2020]

[10]"Python Pandas Tutorial: Dataframe, Date Range, Slice", Guru99.com, 2020. [Online].

Available: <https://www.guru99.com/python-pandas-tutorial.html>.

[Accessed: 19- Apr- 2020].

[11]Python | Data analysis using Pandas - GeeksforGeeks", GeeksforGeeks, 2020. [Online].

Available: <https://www.geeksforgeeks.org/python-data-analysis-using-pandas/>.

[Accessed: 19- Apr- 2020].

[12] "NumPy in Python | Set 1 (Introduction) - GeeksforGeeks", GeeksforGeeks, 2020. [Online].

Available: <https://www.geeksforgeeks.org/numpy-in-python-set-1-introduction/>.

[Accessed: 19- Apr- 2020].

[13]. S. year! and K. Jain, "Scikit-Learn In Python - Important Machine Learning Tool", Analytics Vidhya, 2020. [Online].

Available: <https://www.analyticsvidhya.com/blog/2015/01/scikit-learn-python-machine-learning-tool/>.

[Accessed: 19- Apr- 2020].

[14]. J. Brownlee, "A Gentle Introduction to Scikit-Learn", Machine Learning Mastery, 2020. [Online].

Available: <https://machinelearningmastery.com/a-gentle-introduction-to-scikit-learn-a-python-machine-learning-library/>.

[Accessed: 19- Apr- 2020].

[15]"What is StarUML - StarUML Documentation", Documentation.help, 2020. [Online].

Available: https://documentation.help/StarUML/what_is_staruml.htm.

[Accessed: 19- Apr- 2020].

[16] "UML - Architecture - Tutorialspoint", Tutorialspoint.com, 2020.



- [Online]. Available: https://www.tutorialspoint.com/uml/uml_architecture.htm. [Accessed: 19- Apr- 2020].
- [17] R. Detrano, A. Janosi, W. Steinbrunn, M. P. Sterer, J.-J. Schmid, S. Sandhu, K. H. Guppy, S. Lee, and V. Froelicher, "International application of a new probability algorithm for the diagnosis of coronary artery disease," *Amer. J. Cardiol.*, vol. 64, no. 5, pp. 304-310, Aug. 1989.
- [18] M. Gudadhe, K. Wankhade, and S. Dongre, "Decision support system for heart disease based on support vector machine and artificial neural network," in *Proc. Int. Conf. Comput. Commun. Technol. (ICCCCT)*, Sep. 2010, pp. 741-745.
- [19] H. Kahramanli and N. Allahverdi, "Design of a hybrid system for the diabetes and heart diseases," *Expert Syst. Appl.*, vol. 35, nos. 1-2, pp. 82-89, Jul. 2008.
- [20] R. Das, I. Turkoglu, and A. Sengur, "Effectiveness of diagnosis of heart disease through neural network ensemble," *Expert Syst. Appl.*, vol. 36, no. 4, pp. 7675-7680, May 2009.
- [21] S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," in *Proc. IEEE/ACS Int. Conf. Comput. Syst. Appl.*, Mar. 2008, pp. 108-115.
- [22] E. O. Olaniyi, O. K. Oyedotun, and K. Adnan, "Heart disease diagnosis using neural network arbitration," *Int. J. Intell. Syst. Appl.*, vol. 7, no. 12, p. 72, 2015.
- [23] O. W. Samuel, G. M. Asogbon, A. K. Sangai, P. Fang, and G. Li, "An integrated decision support system based on ANN and Fuzzy_AHP for heart failure risk prediction," *Expert Syst. Appl.*, vol. 68, pp. 163-172, Feb. 2017.
- [24] X. Liu, X. Wang, Q. Su, M. Zhang, Y. Zhu, Q. Wang, and Q. Wang, "A hybrid classification system for heart disease diagnosis based on the RFRS method," *Comput. Math. Methods Med.*, vol. 2017, pp. 1-11, Jan. 2017.
- [25] S. Mohan, C. Thirumalai, and G. Srivastava, "Effective heart disease prediction using hybrid machine learning techniques," *IEEE Access*, vol. 7, pp. 81542-81554, 2019.