AI FOR HEART HEALTH

Atul Lanjewar

Date : 25th feb , 2024

1. *Abstract:*

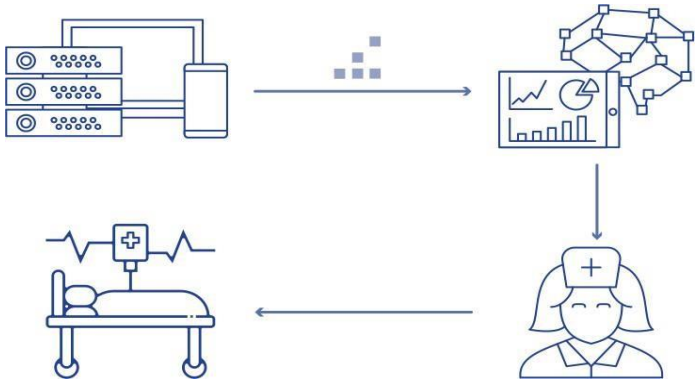
The ideal characteristic of artificial intelligence is its ability to rationalize and take actions that have the best chance of achieving a specific goal. A subset of artificial intelligence is machine learning, which refers to the concept that computer programs can automatically learn from and adapt to new data without being assisted by humans. Cardiologists are pretty good in diagnosing heart while they are not so good in the prognosis of heart disease. It takes more than 3 to 4 days to identify heart problems in an individual. To overcome this hazardous circumstance, Main objective is to use Machine Learning, which not only gives faster results but also demonstrates higher accuracy in the heart disease prediction process.

2. Problem statement :

Every year, Cardologists diagnose 1.5 million new patients with heart disease related problems around the world. Most of them have 96-98 % success rate in diagnosing it but have only 60% accuracy in prognoses. A prognosis is the part of a biopsy that predicts the development of the heart disease. Cardiologists take 3 or more days to evaluate the prognosis process, which is a huge time for heart patient as the risk of cardiovascular attack is rapid once they are produced. To overcome this hazardous situation and to help more and more people fight with heart disease, Machine Learning technology is used in cardiovascular disease prediction. Although predicting cardiovascular disease is neat, but predicting the recurrence of it is a way more complex task for humans. Machine learning not only performs complex task of recurrence, but also takes huge amount of data & gives accurate output in seconds faster than Cardiologists. Finally, something we are certain of is that ML is the next step of Cardiology.

3. Market / Customer/ Business Need Assessment:

Cardiologists diagnose 1.5 million new patients with cardiovascular disease around the world. That's millions of people who'll face years of uncertainty. Cardiologists have been performing cardiovascular diagnoses and prognoses for decades. Most Cardiologists have a 96–98% success rate for diagnosing it. They're pretty good at that part. The problem comes in the next part. According to the Oslo University Hospital, the accuracyof prognoses is only 60% for Cardiologists. A prognosis is the part of a biopsy that comes is predicting the development of the disease. It's time for the next step to be taken in cardiology.



It is not a secret that customers like great and fast service. Cardiologists take 3 or more days to evaluate the prognosis process, which is a huge time for heart patient as the spread of cardiovascular disease are rapid oncethey are produced. With the advent of the Internet of Things technology, there is so much data out in the world that humans can't possibly go through it all. That's where machines help us. They can do work faster than us and make accurate computations and find patterns in data. This will benefit the customer and brings greater profits to the cardiology firm by adapting this advanced technology.

In this article, I am going to emphasize about machine learning which is a branch of artificial intelligence helping cardiologists in prognosis and people suffering from cardiovascular disease. This is a system which takes in data, finds patterns, trains itself using the data and outputs an outcome. In this article we exploreapplications of artificial intelligence to provide business leaders with an understanding of current and emerging trends, presents representative examples of people facing cardiovascular problems and most importantly addressing the importance of cardiovascular prognosis process in eradicating heart disease.

So this information, depicts the need for usage of AI in cardiology.

4. Target Specifications and Characterization:

A. To change traditional cardiology process to faster and accurate process.
B. Reducing frustration and death of patients due to delay in the prognosis process
C. Predetermined dataset of Cardiovascular patients and normal patients is taken and based on that prediction isperformed

Above, mentioned targets can be achieved by analyzing:

1. What the patient look for
2. How are present cardiology processes are being performed
3. Problems faced by people suffering from cardiovascular disease
4. How to identify and provide treatment in initial stage accurately.
5. How efficiently are the cardiologists performing prognosis process
6. When and where a patient likes to trust and spend on?
7. Analyzing the needs of the patients suffering from cardiovascular disease
8. To help patient fight it early stage
9. To send results to the patient within minutes and prescribing the next step to be taken by the patient if he's been found of suffering from cardiovascular disease.
10. To remind the patient about the latest changes in the cardiovascular prognosis process.

5. External Searches (Information searches):

5.1 Applications of Machine Learning in Cardiovascular Prediction and Prognosis

Machine learning is a branch of artificial intelligence that employs a variety of statistical, probabilistic and optimization techniques that allows computers to "learn" from past examples and to detect hard-to-discern patterns from large, noisy or complex data sets. This capability is particularly well-suited to medical applications. As a result, machine learning is frequently used in cardiovascular diagnosis and detection. More recently machine learning has been applied to cardiovascular prognosis and prediction. This latter approach is particularly interesting as it is part of a growing trend towards personalized, predictive medicine. In assembling this review we conducted a broad survey of the different types of machine learning methods being used, the types of data being integrated and the performance of these methods in cardiovascular prediction and prognosis.

The conclusions were made based on the data available on the internet.

The percentages were calculated roughly in a relative manner

5.2 Machine learning-based prediction of survival prognosis

Survival prediction after first diagnosis is important for both disease specialist and patients or their family members. First, as the survival ability of the cardiovascular patients largely depends on the structure of heart. On the other hand, patients and the families canset appropriate goals based on the accurate survival prediction. As the result, the timely prevention and treatment would be made and the worse treatment decision would be effectively avoided.

5.3 Cardiovascular dataset :

Features are computed from kaggle. They describe characteristics of a patient having cardiovascular problem .

Attribute Information:

1) Age.
2) Sex.
3) Chest pain type (4 values).
4) Resting blood pressure.
5) Serum Cholestoral in mg/dl.
6) Fasting blood sugar >120 mg/dl.
7) Resting Electrocardiograpic results(values 0,1,2)
8) Maximum heart rate achieved.
9) Exercise induced angina.
10) Oldpeak = ST depression induced by exercise relative to rest.
11) The slope of the peak exercise ST segment.
12) Number of major vessels(0-3) colored by flourosopy.
13) thal: 0 = normal; 1=fixed defect; 2=reversable defect

5.4 Machine learning is the future of cardiovascular problem prediction:

AI is set to change the medical industry in the coming decades — it wouldn't make sense for cardiology to not be disrupted too. Currently, ML models are still in the testing and experimentation phase for cardiovascular prognoses. As datasets are getting larger and of higher quality, researchers are building increasingly accurate models. Machine Learning is the next step forward for us to overcome this hurdle and create a high accuracy cardiology system

6 Benchmarking alternate products:

In a large study, patients with cardiovascular disease who chose alternative therapies had substantially worse survival than patients who received conventional heart treatments. After a median of 5 years, patients with cardiovascular disease were nearly five times as likely to dieif they had used an alternative therapy exit disclaimer as their initial treatment than if they had received conventional treatment. There's an increased risk of death with choosing alternative medicine, and that's something patients should consider when making their treatment decisions."

Overall, the patients who chose alternative treatments were more likely to be younger, female, healthier, and have higher incomes and education levels. Some of these traits, such as overall better health, would normally improve the odds of survival after a cardiovascular diagnosis.

7 Applicable Regulations (Government and Environmental)

a. Patents on ML algorithms developed
b. Laws related to privacy for collecting data from users
c. Protection/ownership regulations
d. Creating an e-mail service to mail the report to the patient and doctor.
e. Being responsible by design.
f. Ensuring open-source, academic and research community for an audit of Algorithms.
h. Review of existing work authority regulations.

8 Applicable Constraints

Expertise:

A. Requires a lot of research to obtain universal dataset of cardiovascular disease patients in-order to provide moresophisticated and accurate results.
B. Establishing e-mail service in the product which have to send the report after the machine learning model is deployed in any server.
C. Confidential health data to be obtained to train the model.
D. Thorough understanding of dataset and verification of the results must be performed by the cardiologist from the machine learning model to provide a great health prescription and service to the user.

9 Business Opportunity:

Cardiologists are pretty good in diagnosing cardiovascular problem while they are not so good in the prognosis of it. It takes more than 3 to 4 days to identify cardiovascular problem in an individual. To overcome this hazardous circumstance, our main objective is to use Machine Learning, which not only gives faster results but also demonstrates higher accuracy in the Cardiovascular disease prediction process.

10 Concept Generation:

Cardiovascular problem occurs in different stages, as the stages increase the disease gets worse and takes patient towards the death. Treating it early might help the patient fight with heart attack.

1. Symptoms begin:Symptoms can vary but commonly include chest pain or discomfort,shortness of breath,nausea,sweating, and lightheadedness.

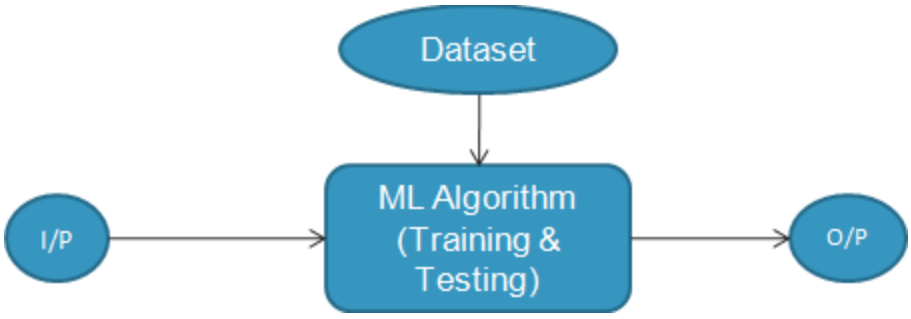
2. Initial damage: The blockage in a coronary artery restricts blood flow to heart muscle,causing damage.This can lead to the death of heart muscle cells if blood flow isn't restored promptly.

3. Heart muscle damage: Without intervention,the lack of oxygen-rich blood can cause significant damage to the heart muscle,leading to complications like arrhythmias(irregular heartbeats) or heart failure.

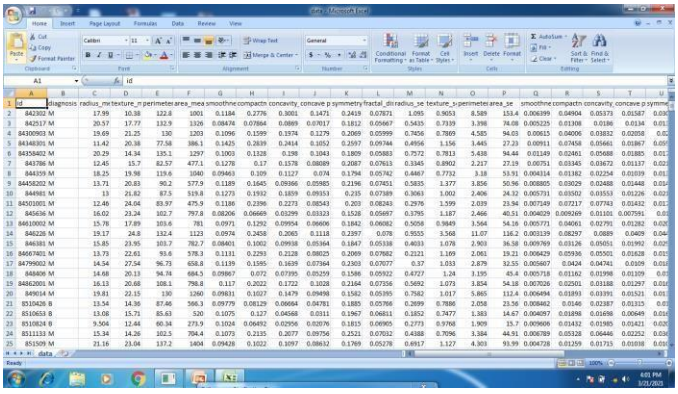
4. Treatment:Medical intervention,such as medication,angioplasty,or surgery,aims to restore bloodflow to the heart and minimize further damage.

5. Recovery:After trearment,the heart begins to heal.Rehabilitation and lifestyle changes are often recommended to improve heart health and prevent future heart attacks.

So in order to generate the model based on the problem stated above, we need to use Machine learning. Machine learning (ML) is the study of computer algorithms that improve automatically through experience and by the use of data. It is seen as a part of artificial intelligence. Machine learning algorithms build a model based on sample data, known as "training data", in order to make predictions or decisions without being explicitly programmed to do so.



Machine learning algorithms are trained and executed in the Google Colab(IDE) which uses python as its programming language. Firstly cardiovascular dataset is prepared, in which data is extracted from the results of the patients who have undergone test for cardiovascular disease. This dataset includes Age, Sex, Chest pain type (4 values), Resting blood pressure, Serum Cholestoral in mg/dl, Fasting blood sugar >120 mg/dl, Resting Electrocardiograpic results(values 0,1,2), Maximum heart rate achieved, Exercise induced angina, Oldpeak = ST depression induced by exercise relative to rest, The slope of the peak exercise ST segment, Number of major vessels(0-3) colored by flourosopy, thal: 0 = normal; 1=fixed defect; 2=reversable defect.

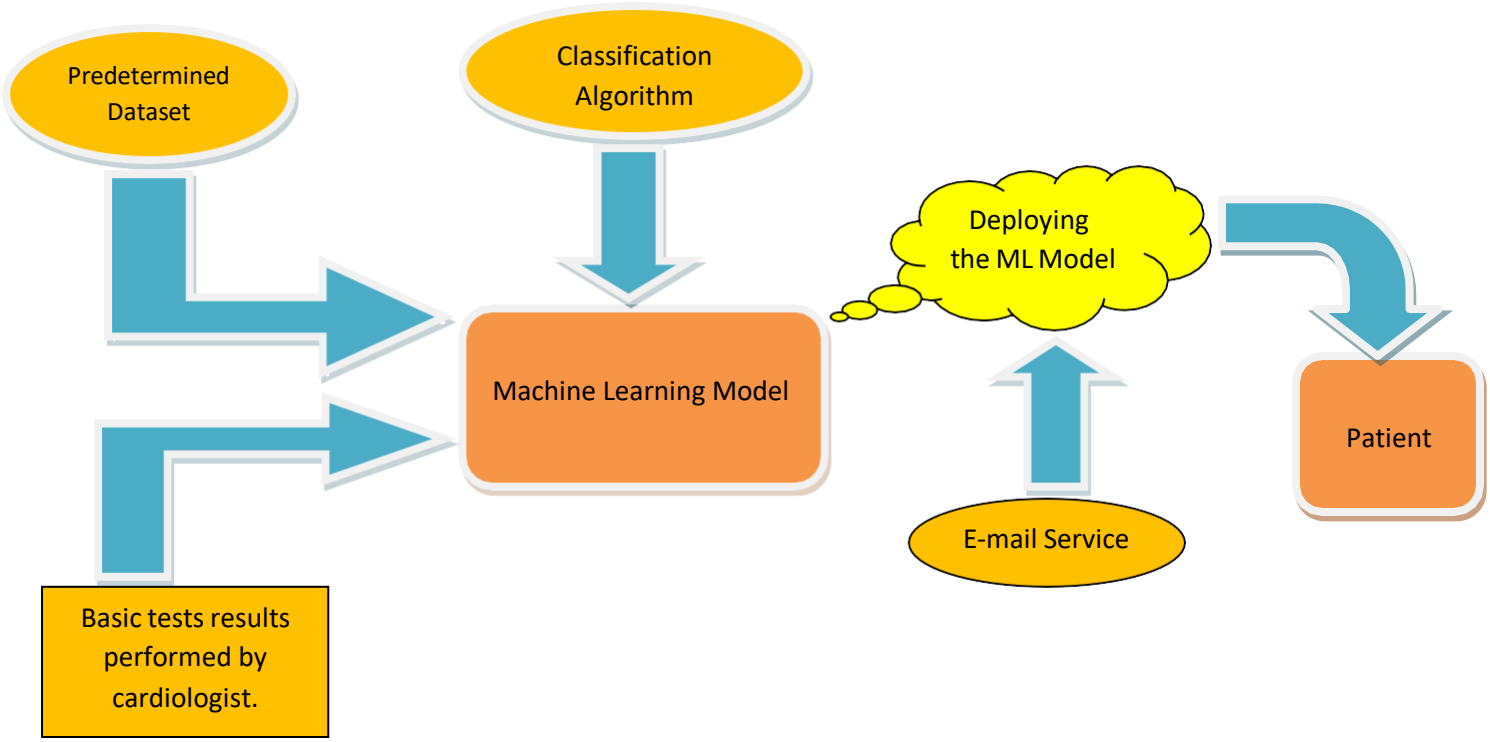


This dataset is imported into the colab environment. Python contains some in-built libraries such as pandas, scikit learn, Seaborn which is exclusively used for training machine learning algorithms.

Clinical tests are performed either at a clinic or at home. Data is inputted into a cardiological ML system.A few minutes later, you receive an email which is programmed in the deployment software with a detailed report that has an accurate prediction about the development of your cardiovascular disease.

Machine learning not only predicts cardiovascular prognosis results faster but also gives higher accuracy which is around 70 to 80 % which is greater when compared with the cardiologists. Heart attack one such kind of disease which when treated early would save many lives. So machine learning is most sort aftertechnique which is very useful in replacing present cardiovascular prognosis process.

11 Final Product prototype:



12 Product details:

It's a system which takes in data, finds patterns, trains itself using the data and outputs an outcome. ML has key advantages over cardiologists. Firstly, machines can work much faster than humans. A biopsy usually takes a cardiologist 10 days. A computer can do thousands of biopsies in a matter of seconds. Machines can do something which humans aren't that good at. They can repeat themselves thousands of times without getting exhausted. After every iteration, the machine repeats the process to do it better. Humans do it too, we call it practice. While practice may make perfect, no amount of practice can put a human even close to the computational speed of a computer.

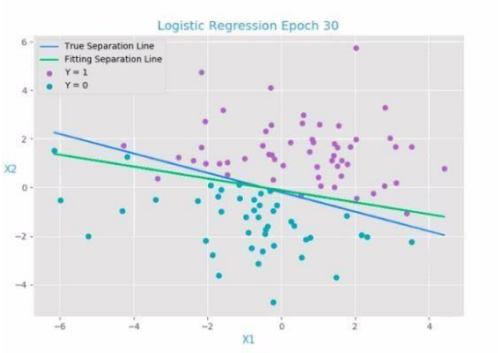
Machines have greater accuracy. With the advent of the Internet of Things technology, there is so much data out in the world that humans can't possibly go through it all. That's where machines help us. They

can do work faster than us and make accurate computations and find patterns in data. That's why they're called computers

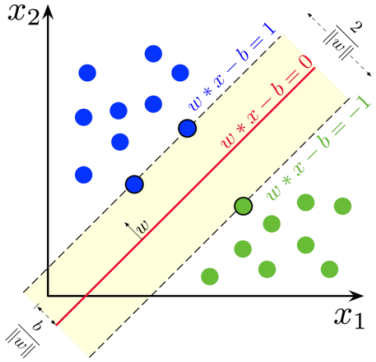
12.1 Algorithm:

Classification model contains different algorithms which can be used to classify weather a person is suffering from cardiovascular disease or not.

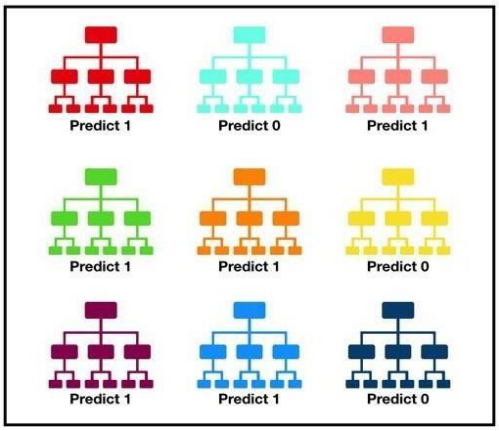
1. Logistic Regression: Logistic regression is a classification algorithm. It is used to predict a binary outcome based on a set of independent variables. A binary outcome is one where there are only two possible scenarios—either the event happens (1) or it does not happen (0). Independent variables are those variables or factors which may influence the outcome (or dependent variable). So: Logistic regression is the correct type of analysis to use when you're working with binary data. You know you're dealing with binary data when the output or dependent variable is dichotomous or categorical in nature; in other words, if it fits into one of two categories (such as "yes" or "no", "pass" or "fail", and so on).



2. Support Vector Machine: The goal of an SVM algorithm is to classify data by creating a boundary with the widest possible margin between itself and the data. The advantages of support vector machines are: Effective in high dimensional spaces. Still effective in cases where number of dimensions is greater than the number of samples. Uses a subset of training points in the decision function (called support vectors), so it is also memory efficient Versatile: different Kernel functions can be specified for the decision function.



3. Random Forest: The fundamental concept behind random forest is a simple but powerful one — the wisdom of crowds. In data science speak, the reason that the random forest model works so well is: A large number of 16 relatively uncorrelated models (trees) operating as a committee will outperform any of the individual constituent models. The low correlation between models is the key. Just like how investments with low correlations (like stocks and bonds) come together to form a portfolio that is greater than the sum of its parts, uncorrelated models can produce ensemble predictions that are more accurate than any of the individual predictions. The reason for this wonderful effect is that the trees protect each other from their individual errors. While some trees may be wrong, many other trees will be right, so as a group the trees are able to move in the correct direction.



Tally: Six 1s and Three 0s
Prediction: 1

12.2 Python-libraries for Cardiovasculer Prognosis:

- Pandas: Pandas is a Python library used for working with data sets. It has functions for analyzing, cleaning, exploring, and manipulating data. Pandas allows us to analyze big data and make conclusions based on statistical theories. Pandas can clean messy data sets, and make them readable and relevant.

  Syntax: import pandas as pd
  df = pd.read_csv('heartdata.csv')

- Scikit Learn: Scikit-learn (formerly scikits.learn and also known as sklearn) is a free software machine learning library used based on python programming language in colab IDE. Itfeatures various classification, regression and clustering algorithms including (SVM) support vector machines , random forests, gradient boosting, k-means and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy.

Syntax for Logistic Regression:
From sklearn.linear_model import LogisticRegression

Syntax for SVM:
From sklearn.svm import SVC

Syntax for Random Forest:
From sklearn.ensemble import RandomForestClassifier

- Seaborn: Seaborn is a library for making statistical graphics in Python. It is built on top of matplotlib and closely integrated with pandas data structures. This is a visualization tool .

Syntax: Import seaborn as sns

Algorithm with highest accuracy among classification algorithms is chosen as the best algorithm for Cardiovascular disease prediction.

12.3 Team required to develop:

1. Machine learning engineering
2. Business analyst
3. Software developer
4. Cloud engineer
5. Data Researcher

13 Conclusion:

AI is set to change the medical industry in the coming decades — it wouldn't make sense for cardiology to not be disrupted too. Currently, ML models are still in the testing and experimentation phase for cardiovasculer prognoses. As datasets are getting larger and of higher quality, researchers are building increasingly accurate models.

While we might not see AI doing the job of a cardiologist today, we can expect ML to replace our local cardiologist used for prognosis in the coming decades, and it's pretty exciting! ML models still have a long way to go, most models still lack sufficient data and suffer from bias. Machine learning can train just as well as doctor prognosis, it doesn't require extra pay for prognosis. Manual cardiovascular treatment take long time to show the result ,while machine learning gives output in seconds .To save people's life and allow doctor to fully concentrate in diagnosis, Yet, something we are certain of is that ML is the next step of cardiology, and it will disrupt the industry.