Heart Disease Prediction Using Supervised Machine Learning

Anil Behal
Department of CSE, Chandirarh university

DEEPANK YADAV
Department of CSE, Chandirarh university

THAKUR DIGVIJAY SINGH
Department of CSE, Chandirarh university

I. ABSTRACT

Heart disease occurrences are rising quickly, making it crucial to identify any potential illnesses beforehand. However, correct diagnosis is a challenging undertaking. The primary focus of the research paper is on whether a patient is more likely to have heart disease in light of various medical data.In this Research Paper, Using the patient's medical information, we have developed a technique to predict whether or not the patient will be diagnosed with a cardiac illness.

II. INTRODUCTION

A method of manipulating and extracting implicit, formerly unknown/known, and potentially relevant information from data is called machine learning . Machine learning is a highly broad and complex topic, and its application and breadth are expanding daily. Machine learning uses a variety of classifiers from supervised, unsupervised, and ensemble learning to predict outcomes and assess their accuracy.

provided dataset .Given that it will benefit many people, we can use that information to our HDPS project. These days, a wide variety of disorders that potentially harm your heart are referred to as cardiovascular diseases.

It is the primary cause of adult fatalities. Our project can determine who is most likely to be diagnosed with a heart issue by looking at a person's medical history . In order to properly treat patients, it can help in detecting disease with fewer medical tests and effective remedies. Anyone exhibiting any heart disease symptoms, such as chest pain or excessive blood pressure, can be identified by it.

Three data mining techniques, including logistic regression, KNN, and random forest classifier, are the major emphasis of this study. Our project's accuracy is 87.5

HDPS precision and effectiveness. The field of supervised learning includes logistic regression. Only

Logistic regression makes use of discrete values.

This project's goal is to determine, depending on the patient's medical characteristics—such as gender, age, chest discomfort, fasting blood sugar level, etc.—whether they are likely to be diagnosed with any cardiovascular heart illnesses. A dataset containing the characteristics and medical background of the patient is chosen from the UCI repository. By

Using this dataset, we make a prediction about the patient's potential for heart disease. This is predicted using 14 medical characteristics of a patient and determine whether he is likely to have a cardiac condition. These health Three algorithms—Logistic regression, KNN, and Random Forest Classifier—are used to train attributes. Most

KNN is the most effective of these algorithms, providing accuracy of 88.52

III. LITERATURE REVIEW

Around the world, heart disease is regarded as the condition that kills people the fastest. Particularly in this sort of illness, the heart is unable to provide the necessary amount of blood to the body's remaining organs to carry out their typical functions. [1]. Heart disease symptoms can include physical weakness, breathing problems, swelling feet, etc. The ability to identify complex heart disorders with a high risk of death is dependent on certain procedures. [2]. The lack of doctors and diagnostic tools that affect the treatment of cardiac patients makes diagnosis and treatment processes extremely difficult at the moment. [3]. Early detection of heart disease is crucial to reducing heart-related problems and safeguarding against serious hazards. [4]. Based on a patient's medical history, an expert's report on their study of their symptoms, and a physical laboratory report, invasive techniques are used to identify cardiac disorders. Additionally, it delays and results in an inaccurate diagnosis because of human interaction. . At the moment of assessment, it is

time-consuming, computationally demanding, and expensive.

[5]. HHeart disease can be predicted using a variety of symptoms, including age, gender, pulse rate, and others. Data analysis in healthcare helps in disease prediction, better diagnosis, symptom analysis, prescribing appropriate medications, enhancing treatment quality, lowering costs, extending life expectancy, and lowering the rate of heart attacks. patients . By placing an ECG (Electro Cardio Gramme) on a patient's chest and monitoring their heartbeat, ECG (Electro Cardio Gramme) aids in the early detection of irregular heartbeat and stroke. The precise clinical data that can help professionals make decisions is used to forecast heart disease. Blood's ability to operate properly is crucial for maintaining human life.

heart's blood vessels . Insufficient blood flow can also result in imminent death due to heart inactivity, kidney failure, and brain imbalance. Obesity, smoking, diabetes, blood pressure, cholesterol, inactivity, and poor diet are a few of the risk factors that might result in heart disease.

Acute Myocardial Infarction (AMI) is a cardiovascular condition that occurs when there is a disruption in the blood supply or circulation to the heart muscle, which results in necrotic (damaged or dead) heart muscle. [6]. The main cause of this condition is a blockage, which results in decreased or obstructed blood flow to the heart muscle. The ability of red blood cells to function and carry enough oxygen aids in restoring blood flow if it is restricted or clogged.

Maintaining human life and consciousness has a significant impact. Without oxygen for six to eight minutes, Heart muscle arrest could occur, which would cause the patient to pass away. Plaque, which is a hard substance consisting of cholesterol (fat) that develops in the coronary arteries and is the major contributor to cardiovascular disease,

to restrict or limit blood flow. Atherosclerosis, which can occasionally develop in the arteries, has a chronic inflammatory component, according to research into its aetiology. the rise in white blood cell count

causes inflammation and other subsequent disorders such as stroke or reinfarction [7]. Monocytes and macrophages often go through two stages of wound healing: the inflammatory stage and the reparative stage. The two stages are necessary for optimal wound healing, though, and if the inflammation persists for too long, it might result in

to cardiac arrest.

Acute spasm or contraction in the coronary arteries is a rare form of cardiac disease. The cramps

suddenly become apparent in arteries without any indication of atherosclerosis [8].It prevents the blood flow that deprives the heart of oxygen. Males are more likely than females to have a heart attack. Additionally, whereas men typically experience discomfort for shorter time than an hour, women can endure agony for up to an hour.

1 hour. The entire physiological system, not only the heart, is affected by cardiovascular illness; alterations

occur anywhere, and even in distant organs like the spleen and bone marrow. [9] [?], [?], [?].

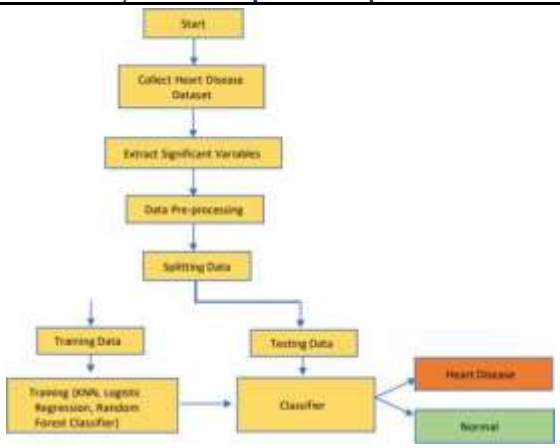

Fig. 1. FlowChart for Prediction of Heart Disease Through Supervised Machine Learning

```
[1] import pandas as pd
    import numpy as np
    from sklearn.model_selection import train_test_split
    from sklearn.preprocessing import StandardScaler
    from sklearn.ensemble import RandomForestClassifier
    from sklearn.metrics import accuracy_score, classification_report
```

Fig. 2. Importing Various Libraries of Python for Developing System

IV. Result and Discussion

We took the data set which contains the medical data of 304 patient of different age. This dataset gives us the information such as age, resting blood pressure, fasting sugar level etc. of the patient that helps us in detecting the patient that is diagnosed with any heart disease or not. This dataset includes 14 medical data of 304 patients that helps in detecting if the patient is at risk of a heart disease or not and it helps us classify patients which are at risk of a heart disease and that who are not at risk. This dataset is taken from the UCI repository. This medical dataset contains 304 rows and 14 columns

In this model we have used different libraries of python for training our machines like pandas,numpy,sci-kit learn. This paper shows the use of various machine learning algorithms, the algorithms that are used in this paper are Random Forest Classifiers,StandardScaler etc. which can be helpful for medical analysts for diagnosing Heart Disease. The methodology is a process which includes steps that transform given data into recognized data patterns for the knowledge of the users. Data preprocessing deals with the missing values, cleaning of data and normalization depending on algorithms used. So in conclusion, the proposed model is undertaken, where we assessed our model on the basis of accuracy and performance using various performance metrics. Here in this model, an effective Heart Disease Prediction System

V. Conclusion

AThree ML classification modelling techniques have been used to create a model for cardiovascular disease detection. By extracting the patient medical history that results in a fatal

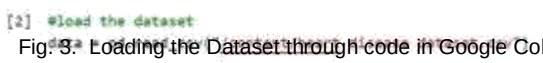

Fig. 3. Loading the Dataset through code in Google Colab

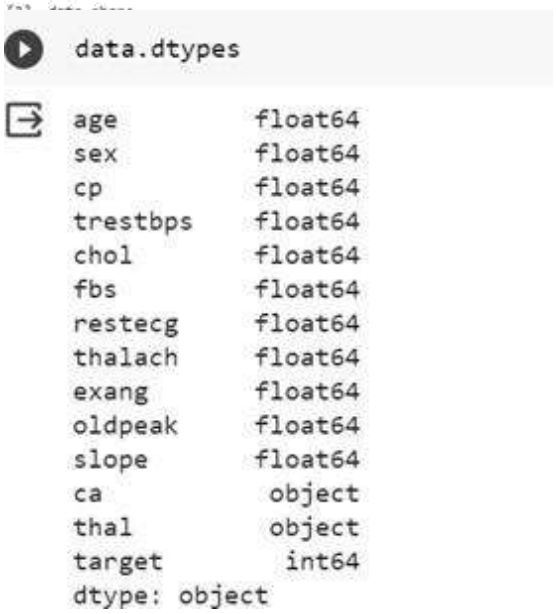

Fig. 4. Using Commands of Pandas Library to Know the Data Types of Different Columns in our Dataset

Fig. 5. Using Algorithm to Train the Machine

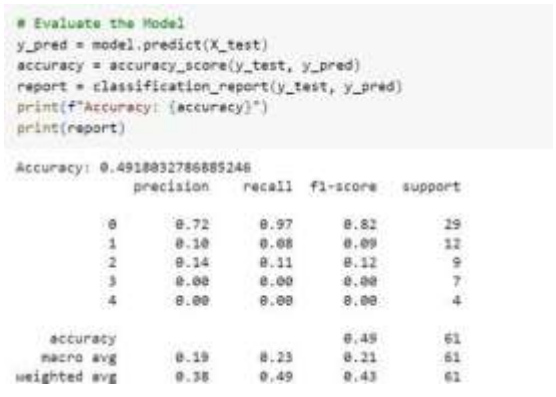

Fig. 6. Accuracy Shown by Algorithm After building the Prediction model through Machine Learning

Fig. 7. Daily Expenses Shown by Pie Chart

Fig. 8. Daily Expenses Shown by Pie Chart

heart illness from a dataset that contains patients' medical histories such as chest pain, this method predicts persons with cardiovascular disease,blood pressure, sugar level, etc. The patient is helped by this technique for detecting heart disease depending on his or her clinical details on whether they have a history of heart disease. The formulas for building the proposed model includes KNN, Random Forest Classifier, and Logistic Regression. The precision of our 87.5

REFERENCES

[1] S. H. Jee, Y. Jang, D. J. Oh, B.-H. Oh, S. H. Lee, S.-W. Park, K.-B. Seung, Y. Mok, K. J. Jung, H. Kimm et al., "A coronary heart disease prediction model: the korean heart study," BMJ open, vol. 4, no. 5, p. e005025, 2014.

[2] A. Ganna, P. K. Magnusson, N. L. Pedersen, U. de Faire, M. Reilly, J. Ärnlöv, J. Sundström, A. Hamsten, and E. Ingelsson, "Multilocus genetic risk scores for coronary heart disease prediction," Arteriosclerosis, thrombosis, and vascular biology, vol. 33, no. 9, pp. 2267–2272, 2013.

[3] M. A. Jabbar, B. L. Deekshatulu, and P. Chandra, "Heart disease prediction using lazy associative classification," in 2013 International Mutli-Conference on Automation, Computing, Communication, Control and Compressed Sensing (iMac4s). IEEE, 2013, pp. 40–46.

[4] C. S. Dangare and S. S. Apte, "Improved study of heart disease prediction system using data mining classification techniques," International Journal of Computer Applications, vol. 47, no. 10, pp. 44–48, 2012.

[5] J. Soni, U. Ansari, D. Sharma, S. Soni et al., "Predictive data mining for medical diagnosis: An overview of heart disease prediction," International Journal of Computer Applications, vol. 17, no. 8, pp. 43–48, 2011.

[6] A. H. Chen, S.-Y. Huang, P.-S. Hong, C.-H. Cheng, and E.-J. Lin, "Hdps: Heart disease prediction system," in 2011 computing in Cardiology. IEEE, 2011, pp. 557–560.

[7] L. Parthiban and R. Subramanian, "Intelligent heart disease prediction system using canfis and genetic algorithm," International Journal of Biological, Biomedical and Medical Sciences, vol. 3, no. 3, 2008.

[8] G. Wolgast, C. Ehrenborg, A. Israelsson, J. Helander, E. Johansson, and H. Manefjord, "Wireless body area network for heart attack detection [education corner]," IEEE antennas and propagation magazine, vol. 58, no. 5, pp. 84–92, 2016.

[9] S. Patel and Y. Chauhan, "Heart attack detection and medical attention using motion sensing device-kinect," International Journal of Scientific and Research Publications, vol. 4, no. 1, pp. 1–4, 2014.