Heart Disease Prediction based on Machine Learning and Deep Learning Techniques

Supriya Kaskar[1],
Department of Computer Engineering,
Wagholi, Pune

Sonali Gandewar[1],
Department of Computer Engineering,
Wagholi, Pune

Sapna Boinwad[1]
Department of Computer Engineering
Wagholi, Pune

Pradnya Bhusare[1]
Department of Computer Engineering
Wagholi, Pune

Dr. Vinod Wadne[1]
Department of Computer Engineering
Wagholi, Pune

Abstract: In recent times, heart condition prediction is one amongst the most sophisticated tasks in medical field. Within the epoch, approximately one person dies per minute because of heart condition. Data science plays an important role in process vast quantity of data within the field of health care. As heart condition prediction may be a complex task, there's a necessity to change the prediction method to avoid risks related to it and alert the patient well in advance. This paper makes use of heart condition dataset accessible in UCI machine learning repository. The planned work predicts the possibilities of heart condition and classifies patient's risk level by implementing totally different data processing techniques like Naive Bayes, call Tree, provision Regression and Random Forest, etc. Nowadays, health diseases are increasing day by day due to lifestyle, hereditary. Especially, heart disease has become more common these days, i.e. life of people is at risk. Each individual has different values for Blood pressure, cholesterol and pulse rate. But according to medically proven results the normal values of Blood pressure is 120/90, Cholesterol is 100-129 mg/dL ,Pulse rate is 72, Fasting Blood Sugar level is 100 mg/dL ,Heart rate is 60-100 bpm, ECG is normal, Width of major vessels is 25 mm (1 inch) in the aorta to only 8 μm in the capillaries. This project used different classification techniques for predicting the risk level of each person based on age, gender, Blood pressure, cholesterol, pulse rate. "Disease Prediction" system based on predictive modelling predicts the disease of the user on the basis of the symptoms that user provides as an input to the system
Keywords: Machine Learning, Supervised Learning, Heart disease, prediction, Imbalance data, etc.

INTRODUCTION

In today's world of stress Heart, being an essential organ in a human body which pumps blood through the body for the blood circulation is essential and its health is to be conserved for a healthy living. The health of a human heart is based on the experiences in a person's life and is completely dependent on professional and personal behaviours of a person. There may also be several genetic factors through which a type of heart disease is passed down from generations. According to the World Health Organization, every year more than 12 million deaths are occurring worldwide due to the various types of heart diseases which is also known by the term cardiovascular disease. The term Heart disease includes many diseases that are diverse and specifically affect the heart and the arteries of a human being. Even young aged people around their 20-30 years of lifespan are getting affected by heart diseases. The increase in the possibility of heart disease among young may be due to the bad eating habits, lack of sleep, restless nature, depression and numerous other

factors such as obesity, poor diet, family history, high blood pressure, high blood cholesterol, idle behavior, family history, smoking and hypertension. The diagnosis of the heart diseases is a very important and is itself the most complicated task in the medical field. All the mentioned factors are taken into consideration when analyzing and understanding the patients by the doctor through manual check-ups at regular intervals of time. The symptoms of heart disease greatly depend upon which of the discomfort felt by an individual. Some symptoms are not usually identified by the common people. However, common symptoms include chest pain, breathlessness, and heart palpitations. The chest pain common to many types of heart disease is known as angina, or angina pectoris, and occurs when a part of the heart does not receive enough oxygen. Angina may be triggered by stressful events or physical exertion and normally lasts under 10 minutes. Heart attacks can also occur as a result of different types of heart disease. The signs of a heart attack are like angina except that they can occur during rest and tend to be more severe. The symptoms of a heart attack can sometimes resemble indigestion. Heartburn and a stomach ache can occur, as well as a heavy feeling in the chest. Other symptoms of a heart attack include pain that travels through the body, for example from the chest to the arms, neck, back, abdomen, or jaw, light headedness and dizzy sensations, profuse sweating, nausea and vomiting. Heart failure is also an outcome of heart disease, and breathlessness can occur when the heart becomes too weak to circulate blood. Some heart conditions occur with no symptoms at all, especially in older adults and individuals with diabetes. The term 'congenital heart disease' covers a range of conditions, but the general symptoms include sweating, high levels of fatigue, fast heartbeat and breathing, breathlessness, chest pain. However, these symptoms might not develop until a person is older than 13 years. In these types of cases, the diagnosis becomes an intricate task requiring great experience and high skill. A risk of a heart attack or the possibility of the heart disease if identified early, can help the patients take precautions and take regulatory measures. Recently, the healthcare industry has been generating huge amounts of data about patients and their disease diagnosis reports are being especially taken for the prediction of heart attacks worldwide. When the data about heart disease is huge, the machine learning techniques can be implemented for the analysis.

LITERATURE SURVEY

[1]According to Ordonez the heart disease can be predicted with some basic attributes taken from the patient and in their work have introduced a system that includes the characteristics of an individual human being based on totally 13 basic attributes like sex, blood pressure, cholesterol and others to predict the likelihood of a patient getting affected by heart disease. They have added two more attributes i.e. fat and smoking behavior and extended the research dataset, and not handled large dataset[2] Yilmaz, have proposed a method that uses least squares support vector machine (LS-SVM) utilizing a binary decision tree for classification of cardiotocogram to find out the patient condition. Duff, et al. have done a research work involving five hundred and thirty-three patients who had suffered from cardiac arrest and they were integrated in the analysis of heart disease probabilities, major disadvantages of this paper has Underfitting problem and not handle noise and outliers efficiently Frawley, et al. have performed a work on prediction of survival of Coronary heart disease (CHD) which is a challenging research problem for medical society. They also used 10-fold cross-validation methods to determine the impartial estimate of the three prediction models for performance comparison purposes. The existing system having number of disadvantages accuracy of the algorithm are very low and not meet the performance. In the existing system mining techniques can-not handle efficient decision making result. To small size data set are used for prediction and its affect to the performance of the system and model cannot meet the scalability of the algorithm.

BACKGROUND

The most important background of machine learning algorithms their technique and mathematical formulation are outlined in this section. Analysing the Healthcare data

1. Machine Learning

Machine learning algorithm can be group into two main categories, they include

1. **Supervised Learning**: supervised learning algorithm main feature is target variable and outcome variable to predict. Supervised learning technique is achieved using regression and classification problem.

2. **Unsupervised learning**: in unsupervised learning algorithm no target or outcome variable to predict. It is used for clustering entities into an different groups.

2. Classification Algorithms:

Classification algorithms work by predicting the simplest cluster to that a knowledge purpose belongs to by learning from labelled observations; it uses a group of input options for the educational method. Classification algorithms square measure sensible for grouping knowledge that square measure ne'er seen before into their numerous groupings and square measure thus extensively employed in machine learning tasks.

3. Evaluation Matrix:

1. Accuracy:

it is measured how many true positive and true negative cases is correct. Mathematically it is defined as

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN}$$

2. Sensitivity or Recall:

1. Recall: tells us how many of the actual positive cases we were able to predict correctly with our model. Mathematically it is defined as

$$Recall = \frac{TP}{TP+FN}$$

2. Specificity: tell us how many times classifier gets true negative correct value, mathematically is defined as

$$Specificity = \frac{TN}{TN+FP}$$

3. Precision:

Precision tells us how many of the correctly predicted cases actually turned out to be positive. Mathematically it is defined as,

$$Precision = \frac{TP}{TP+FP}$$

PROPOSED SYSTEM

Heart Diseases affect a large population in today's world, where the lifestyle is moved from active to comfort-oriented. We live in era of fast foods. Which build up cholesterol, diabetes and many more factors which in turn affects the heart in some way or the other. According to the World Health Organization Cardiovascular Diseases (CVD) or Heart Diseases cause more death than any other diseases globally. The amount of data in medical sectors is quite large and computerized as well. They are not utilized or put to any use. This data if studied and analyzed could be put to good use like prediction of diseases or even prevent them Diseases such as cancer can be detected, and the stage can also be predicted by training dataset with pictures of cancer cells. Similarly, heart disease can be predicted based on aspects like cholesterol, diabetes,

heart rate etc. The prediction of heart diseases is a challenge and very risky. We observed that in some cases solutions of problems does not rely on a single method. It varies from situation to situation. It is also a challenge as most of the data are sparse or missing as they were not stored in the motive of analysing. We therefore set out goal to finding which method would be best for predicting the diseases using data of four different hospitals from four different places. This is a comparative study on the efficiency of different machine learning techniques such as Logical Regression, Random forest, K-Nearest Neighbors, Decision Tree in predicting heart diseases. The Machine learning techniques are analyzed, and the accuracy of prediction is noted for each method used. The result showed that heart diseases can be predicted with accuracy of above 96%.In this project we have used 5 algorithms to find out the reasons of heart disease and create a model to get the maximum accuracy possible

- They are highly scalable.
- It runs efficiently on large databases
- It is very robust and can be simply implemented on classification datasets
- require little effort for their preparation
- It works well with data that has highly dependent attributes

BLOCK DIAGRAM

The proposed module can be divided in to different sections, machine learning, Flask, HTML, CSS, Anaconda-Jupiter notebook. Architecture used in proposed system are given below
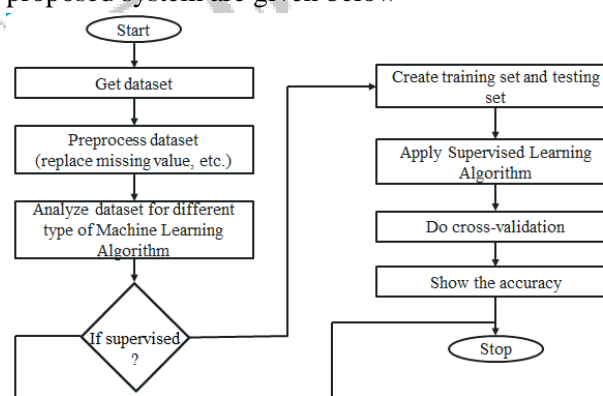

Figure: Architecture of System Design

CONCLUSION

prediction techniques help to detect Heart Disease , before going to advance stage here we are use different machine learning and deep Learning algorithm on Heart disease. After implementation and comparison of all algorithm, multiple based models gives more prediction Accuracy result Hence we are use machine learning and Deep Learning model for Heart disease

prediction and save lives. In future the accuracy can improve the data size.

BIBLIOGRAPHY

[1] "Prediction of Heart Disease using Artificial Neural Network," VFAST Trans. Softw. Eng., pp. 102–112, 2019

[2] P. Selvakumar and S. P. Rajagopalan, "A survey on neural network models for heart disease prediction," J. Theor. Appl. Inf. Technol., vol. 67, no. 2, pp. 485–497, 2014

[3] A. Methaila, P. Kansal, H. Arya, and P. Kumar, "Early Heart Disease Prediction Using Data Mining Techniques," vol. 6956, no. October, pp. 53–59, 2014

[4] I. S. F. Dessai, "Intelligent Heart Disease Prediction System Using Probabilistic Neural Network," no. 5, pp. 38–44, 2013.

[5] T. Karayilan and Ö. Kiliç, "Prediction of Heart disease using neural network," in 2nd International Conference on Computer Science and Engineering, UBMK 2017, 2017, pp. 719–723.

[6] M. Mardiyono, R. Suryanita, and A. Adnan, "Intelligent Monitoring System on Prediction of Building Damage Index using Neural-Network," TELKOMNIKA (Telecommunication Comput. Electron Control, vol. 10, no. 1, p. 155, 2015

[7] N. Guru, A. Dahiya, and N. Rajpal, "Decision support system for heart disease diagnosis using Neural Network," 2007.

[8] J. S. Singh Navdeep, "No Title," Intell. Hear. Dis. Predict. Syst. using CANFIS Genet. Algorithm, Int. J. Biol. Med. Sci., no. International Journal of Advance Research, Ideas and Innovations in Technology© 2018, www.IJARIIT.comAll Rights, p. Reserved Page | 987, 2008

[9] R. G. S. Rajkumar Asha, "No Title," Diagnosis Hear. Dis. Using Datamining Algorithm, vol. Issue 10 V, no. Global Journal of Computer Science and Technology, p. Page 38, 2010

[10] Samruddhi Kaware and Dr. V. S. Wadn, "EnhanceCancer and Diabetes Detection by using Machine Learning Techniques" IJRAR March 2020, volume 7, Issue 1.