INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 8, Issue 1 - V8I1-1473)
Available online at: https://www.ijariit.com

Transfer learning-based machine learning models for heart disease prediction in an earlier stage

Ganaga Muneeswari M.
1812073@nec.edu.in
National Engineering College, Kovilpatti, Tamil Nadu
Aishwaryalakshmi R. K.
aishwarya_cse@nec.edu.in
National Engineering College, Kovilpatti, Tamil Nadu
Soniya V.
1812085@nec.edu.in
National Engineering College, Kovilpatti, Tamil Nadu
Abisha D.
abisha_cse@nec.edu.in
National Engineering College, Kovilpatti, Tamil Nadu

ABSTRACT

Anticipating and identifying heart affliction has ordinarily been an intense and tedious task for specialists. To adapt to heart problems, medical clinics and explicit centers are giving steeply-evaluated rebuilding procedures and activities. Thus, hanging tight for heart disorder in its initial degrees is most likely valuable to individuals from one side of the planet to the other, allowing them to take required treatment ahead of time than it transforms into genuine. Heart ailment has been a main issue in state of the art years, with the essential intentions being unreasonable liquor use, tobacco use, and an absence of actual work. Machine gaining knowledge of has proven to be beneficial in making selections and predictions from a huge set of information created via way of means of the healthcare enterprise over time. Artificial neural networks (ANN), choice trees (DT), random forests (RF), and Naive Bayes) are a number of the supervised gadget gaining knowledge of strategies hired on this prediction of coronary heart ailment (NB).In addition, the results of various algorithms are summarized. This paper attempts to forecast cardiac disease at an early stage. We will compare the four algorithms with their accuracy score and will conclude which algorithm is best.

Keywords: *Heart Disease, ANN, Coronary Artery disease(CAD), Random Forest, Decision Tree, Naive Bayes.*

1. INTRODUCTION

After the brain, the coronary heart is one of the maximum vital additives of the human body. The coronary heart`s predominant responsibility is to pump blood to all elements of the body. Heart disorder refers to any situation which could impair the coronary heart's cappotential to operate. There are numerous styles of coronary heart contamination withinside the world; the maximum common coronary heart issues are coronary artery disorder (CAD) and coronary heart failure (HF). Heart attacks are prevalent, and heart disease has become one of the world's worst human diseases. This disease and issue do not appear anywhere. Scientists and doctors disclose that it is a continuous process that is the consequence of living a certain lifestyle for a long time, as well as the result of some fundamental and frequent symptoms appearing out of nowhere. Heart failure can be caused by a variety of circumstances. Medical scientists have divided these elements into two groups: risk factors that cannot be adjusted and risk factors that can be changed. Risk factors that cannot be changed include family history, sex, and age. High blood stress, excessive cholesterol, smoking, bodily inactivity, and excessive blood stress are all danger factors. The nature of the remedy to be supplied is decided via way of means of the prognosis, that's an essential a part of the affected person care cycle. Diagnosis is step one in nicely analyzing any sickness and also can offer the character of the remedy to be supplied. Even on this age of present day technology, 21st-century prognosis is some distance from perfect.

Because heart disease is such a serious problem, it is necessary to diagnose or forecast it. Heart disease can be diagnosed using a variety of approaches. Angiography is a popular procedure that is used by the majority of doctors around the world. However, there are several disadvantages to the angiography method. It is an expensive procedure, and clinicians must study a large number of parameters to diagnose a patient, making the job of a physician extremely difficult. These drawbacks inspire researchers to broaden a non-invasive approach for coronary heart sickness prediction. These traditional procedures deal with patient medical reports; also, these traditional methods are time-consuming and may produce incorrect results because they are carried out by people. We need an automated system to avoid these errors and obtain better and faster results. Researchers have shown that machine learning algorithms perform exceptionally well when assessing medical data sets in recent years. These data sets will be fed straight into machine learning algorithms, which will perform under their nature and produce results.

Problem Statement

Invasive-based procedures are typically used when patients present with certain symptoms, which are typically basic signs

that a layperson with less understanding might recognize as indicating that the patient is suffering from heart disease or stroke at the moment. Furthermore, the approaches are typically quite expensive and computationally difficult, and assessments take time. Meanwhile, we discovered during our research that we do not have a system in place that must assess specific aspects and symptoms related to patients, living habits, and parental history to provide patients with preventive information.

2. LITERATURE SURVEY

Bo Jin, Chao Che, and colleagues (2018) **[1]** suggested a neural network-based "Predicting the Risk of Heart Failure With EHR Sequential Data Modeling" model. To conduct the experiment and forecast the heart disease before it occurred, this article used EHR data from real-world datasets relevant to congestive heart disease. To model the diagnostic events and predicted coronary failure events, we employed one-hot encryption and word vectors, which are the fundamental principles of an extended memory network model. We generally tend to illustrate the need of respecting the sequential person of medical statistics with the aid of using inspecting the findings.

"Heart Disease Prediction Using Evolutionary Rule Learning," **[2]** presented by Aakash Chauhan et al. (2018). This study eliminates the manual task, which also aids in the extraction of information (data) directly from electronic records. On the patient's dataset, we used frequent pattern growth association mining to generate strong association rules. This will assist (help) in reducing the number of services and demonstrating that the vast majority of the rules aid in the best prediction of coronary illness.

Senthilkumar Mohan, Chandrasegar Thirumalai, et al. (2019) **[3]** proposed an efficient technique using hybrid machine learning methodology in "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques." The hybrid approach combines the random forest and the linear method. For prediction, the dataset and attribute subsets were gathered. The pre-processed knowledge(data) set of cardiovascular disease was used to select a subset of some attributes. Following prep-processing, hybrid techniques were used to diagnose cardiovascular disease.

Dr. C.R.K.Reddy and K.Prasanna Lakshmi (2015) **[4]** created "Fast Rule-Based Heart Disease Prediction Using Associative Classification Mining." We used associative category mining over a landmark window of statistics streams withinside the proposed Stream Associative Classification Heart Disease Prediction (SACHDP). This paper is split into parts: the primary is set producing regulations from associative category mining, and the second one is set pruning the regulations with chi-rectangular trying out and arranging the regulations in an order to shape a classifier. These degrees may be used to without problems expect coronary heart disease.

M.Satish et al. (2015) [5] hired plenty of Data Mining strategies, which includes rule-based, Decision-tree, Naive Bayes, and synthetic neural networks. To are expecting Heart Disease, an green technique referred to as pruning category affiliation rule (PCAR) become used to generate affiliation guidelines from a cardiovascular sickness warehouse. The coronary heart assault statistics warehouse become used for mining pre-processing. All of the statistics mining strategies mentioned above had been described.

Lokanath Sarangi, Mihir Narayan Mohanty, and Srikanta Pattnaik (2015) [6] evolved a cost-powerful version the usage of the genetic set of rules optimizer method of their paper "An Intelligent Decision Support System for Cardiac Disease Detection." The weights have been optimized and fed into the given community as an input. The accuracy carried out with the hybrid method of GA and neural networks became 90%.Boshra Bahrami and Mirsaeid Hosseini Shirvani created **[7]** "Prediction and Diagnosis of Heart Disease Using Data Mining Techniques." This paper employs a variety of classification methodologies to diagnose cardiovascular disease. To divide the datasets, classifiers such as KNN, SVO classifiers, and Decision Trees are used. Following classification and performance evaluation, the Decision tree is evaluated as the best one from the dataset for cardiovascular disease prediction.

"Prediction and Diagnosis of Heart Disease Patients Using Data Mining Technique" **[8]** was created by Mamatha Alex P and Shaicy P Shaji (2019). This paper employs Artificial Neural networks, KNN, Random Forest, and Support Vector Machine techniques. When compared to the previously mentioned classification techniques in data mining, Artificial Neural Network predicts the highest accuracy for diagnosing heart disease.

Heart Disease Prediction System Using Data Mining Techniques [9] turned into proposed with the aid of using MeghaShahi et al. WEKA software program is utilized in healthcare facilities for computerized sickness analysis and to offer carrier quality. SVM, Nave Bayes, Association rule, KNN, ANN, and Decision Tree had been a number of the algorithms used withinside the paper. When as compared to different information mining algorithms, the paper recommends SVM due to the fact it's miles greater powerful and offers greater accuracy.

Chala Beyene et al. [10] counseled the use of statistics mining strategies to are expecting and examine the prevalence of coronary heart sickness. The fundamental intention is to are expecting the prevalence of coronary heart sickness to carry out an early automated analysis of the sickness and attain outcomes in a brief period. The proposed method is likewise essential in healthcare companies with professionals who lack know-how and skills. It makes use of numerous scientific attributes together with blood sugar and coronary heart rate, age, and gender to decide whether or not or now no longer someone has coronary heart sickness. WEKA software program is used to compute dataset analyses.

R. Sharmila et al. [11] proposed the use of a non-linear type set of rules to are expecting coronary heart ailment. It is proposed to apply massive records gear together with Hadoop Distributed File System (HDFS), MapReduce, and SVM for coronary heart ailment prediction with an optimized characteristic set. This examine regarded into using numerous records mining strategies for predicting coronary heart ailment. It shows the use of HDFS to shop massive quantities of records throughout a couple of nodes and strolling the SVM prediction set of rules on a couple of nodes on the identical time. SVM is utilized in parallel, which leads to a quicker computation time than sequential SVM.

S.Prabhavathi et al. [12] proposed a choice tree-primarily based totally Neural Fuzzy System (DNFS) method for reading and forecasting diverse coronary heart diseases. This paper examines the studies at the analysis of coronary heart sickness. DNFS is an abbreviation for Decision Tree-Based Neural Fuzzy System. The purpose of this studies is to broaden an sensible and cost-powerful device even as additionally enhancing the overall performance of the present device. Data mining strategies are

used particularly on this paper to enhance coronary heart sickness prediction. The findings of this look at display that SVM and neural networks carry out extraordinarily nicely in predicting coronary heart sickness. Still, statistics mining strategies aren't promising for predicting coronary heart sickness.

Sarabi H.Mujawar et al. [13] anticipated coronary heart sickness the usage of k-way and naive Bayes. The aim of this paper is to create a machine that gives analysis via way of means of making use of a historic coronary heart database. The machine turned into constructed with thirteen attributes in mind. Data mining strategies together with clustering and category techniques may be used to extract expertise from databases. The Cleveland Heart Database turned into used for thirteen attributes with a complete of three hundred records. Based at the values of thirteen attributes, this version predicts whether or not the affected person has coronary heart sickness or not.

Sharan Monica. L et al [14] proposed a cardiovascular disorder analysis. This paper proposed records mining strategies for disorder prediction. Its purpose is to offer a survey of present day strategies for extracting records from datasets so that it will be beneficial to healthcare practitioners. The time it takes to construct the system`s selection tree may be used to calculate performance. The number one purpose is to expect the disorder the use of a small variety of attributes.
P.Sai Chandrasekhar Reddy et al. [15] proposed the usage of an ANN set of rules in records mining to are expecting coronary heart disorder. Due to the growing expenses of coronary heart disorder diagnosis, there has been a want to broaden a brand new gadget that might are expecting coronary heart disorder. The prediction version is used to are expecting the patient`s situation after comparing diverse parameters including coronary heart rate, blood pressure, cholesterol, and so on. The gadget's accuracy is verified in Java.

3. PROPOSED SYSTEM

The current system is being developed to detect heart disease before it becomes a chronic occurrence in which people's lives are jeopardized. This study will attempt to predict heart disease using the data science classifiers Naive Bayes NB, Decision Tree DT, and other useful algorithms. The model's goal is to predict heart disease and then provide useful suggestions for maintaining a healthy heart. The issue with heart diseases is that they require immediate attention in the event of a heart attack or stroke. The number of cases encountered worldwide is "death due to heart attack occurs due to a lack of knowledge and first aid given to them." As people's lifestyles around the world have changed, which is the primary cause of various heart complications, there has been clear research done to predict heart disease well in advance, with nearly 90 percent or higher prediction accuracy. The issue here is that the prediction alone cannot completely rule out the disease from the body. Three basic things must be done to cure it. 1. Medication 2. Safety precautions 3. Changing patients' lifestyles by recommending physical activity based on their unique characteristics. As a result, our model will predict the level of heart disease, and we will compare the four algorithms based on their accuracy score to determine which algorithm is the best.

A. Block Diagram

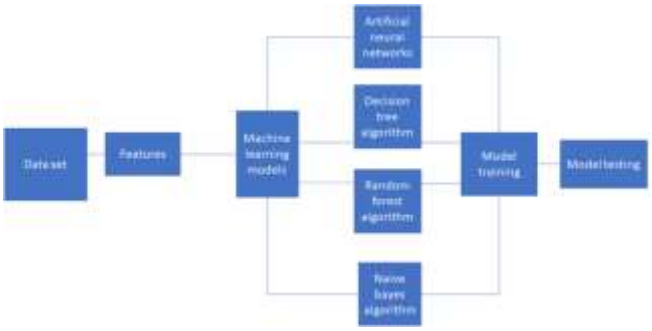

Figure 1: Proposed methodology block diagram

Dataset is collected from Kaggle. This dataset consists of 303 records. Features are extracted from the dataset. With the help of the extracted features, we have implemented the above-mentioned four algorithms such as Artificial neural network, Decision tree, Random Forest, and Naive Bayes. And finally concluded that Random forest provides great accuracy.

B. Heart Disease Types

Table 1: Varying heart diseases

Type of heart disease	Description
Coronary artery	Coronary artery disease develops when arteries in the heart become blocked.
Vascular disease	When the blood supply to the heart is restricted, the vascular disease develops.
Heart rhythm disorder	It's a different sort of cardiac illness that involves a disturbance in heart rhythm, such as heartbeats that are too fast, too slow, or in an aberrant pattern.
Structural heart disease	It denotes the presence of disordered muscles, blood vessels, valves, and walls near the heart. Heart failure will result from this disorder of the heart anatomy.
Heart failure	When the heart is destroyed, cardiac failure occurs. Heart failure is caused by two factors: heart attacks and excessive blood pressure.

C. Dataset

A dataset containing attributes The patient's gender is indicated by sex, and his or her age is indicated by age. trestbps stands for resting blood pressure, cp stands for chest pain, fbs stands for fast blood sugar, chol stands for cholesterol, thalach indicates the maximum heart rate achieved, restecg stands for resting electroc. result (1 anomaly), old peak stands for ST depression inducing ex, exang stands for exercise-induced angina. The letter ca denote the number of major vessels. slope denotes the peak exercise slope Thalassemia is indicated by the letters ST, pred attribute, and thal.

D. Proposed Models

Artificial Neural Network(ANN):
Artificial neural networks are in general meant for computational purposes; the principle subject matter of this version is to finish a assignment quicker than the conventional version.The biological structure of neurons in the human brain is similar to this model. The same way neurons connect in the brain, neurons (nodes) will connect here. This model is made up of a large number of interconnected elements (neurons) that work together to complete a task. A perceptron is a single-layer neural network that produces a single output.

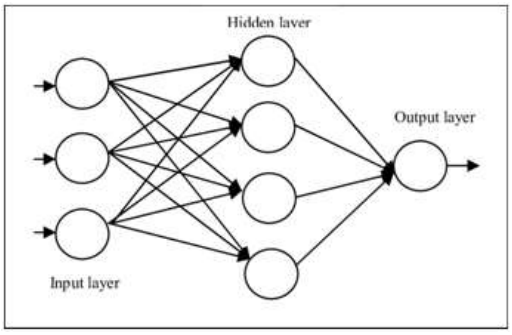

Figure 2: Architecture of ANN

Decision Tree(DT):
In machine learning algorithms, a decision tree is one of the supervised learning techniques. It is used for classification as well as regression. The parameters will be used to divide the data in this algorithm. A decision tree is a type of tree that has nodes and leaves. We will get outcomes or decisions at the leaves, and data will be split at the nodes.
There are two kinds of decision trees:
1. Tree of classification
2. Tree of Regression
Tree of classification: As in the previous example, we will get a decision (outcome) variable that is categorical.
Tree of Regression: A decision (outcome) variable as continuous will be obtained here.

Random Forest(RF):
It is a supervised device mastering set of rules that may be used for each type and regression. However, it's far usually used for type. The call means that it's far a woodland; a woodland is a set of bushes; similarly, in a random woodland set of rules, we are able to have bushes; those bushes are the choice bushes. The prediction consequences might be greater correct if we've got a bigger quantity of choice bushes. The random woodland set of rules works as follows: first, it collects random samples from the dataset, after which it creates choice bushes for every pattern from which we pick the tree that produces the quality prediction consequences.

Naive Bayes(NB):
One of the supervised system mastering category algorithms is Naive Bayes. Previously, it turned into used for textual content category. It is worried with the datasets with the best dimensionality. Some examples encompass sentiment analysis, unsolicited mail filtration, and so on. This naive Bayes set of rules is primarily based totally at the Bayes theorem and assumes that attributes are unbiased of 1 another. It is not anything greater than attributes in a unmarried elegance which might be unbiased of some other attributes within side the identical elegance.

4. RESULT ANALYSIS

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

%matplotlib inline

import os
print(os.listdir())

import warnings
warnings.filterwarnings('ignore')
```

```
['.config', 'drive', 'sample_data']
```

Figure 3: Implementing Essential Libraries

Figure 4 shows the essential libraries for implementing our machine learning algorithms.

```
heart_data=pd.read_csv('/content/drive/MyDrive/heart.csv')
print(heart_data)
```

	age	sex	cp	trestbps	chol	fbs	...	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	...	0	2.3	0	0	1	1
1	37	1	2	130	250	0	...	0	3.5	0	0	2	1
2	41	0	1	130	204	0	...	0	1.4	2	0	2	1
3	56	1	1	120	236	0	...	0	0.8	2	0	2	1
4	57	0	0	120	354	0	...	1	0.6	2	0	2	1
..	...	...	..	...	...	...	...	...	...	...	..	...	...
298	57	0	0	140	241	0	...	1	0.2	1	0	3	0
299	45	1	3	110	264	0	...	0	1.2	1	0	3	0
300	68	1	0	144	193	1	...	0	3.4	1	2	3	0
301	57	1	0	130	131	0	...	1	1.2	1	1	3	0
302	57	0	1	130	236	0	...	0	0.0	1	1	2	0

[303 rows x 14 columns]

Figure 4: dataset

Figure 5 shows the dataset for our project it contains 14 features and 303 records.

```
from sklearn.model_selection import train_test_split

predictors = heart_data.drop("target",axis=1)
target = heart_data["target"]

X_train,X_test,Y_train,Y_test = train_test_split(predictors,target,test_size=0.20,random_state=0)
```

Figure 5: Train dataset

Figure 6 shows the train test and split dataset. It contains 242 trained data and 61 test data.

```
scores = [score_nb,score_dt,score_rf,score_nn]
algorithms = ["Naive Bayes","Decision Tree","Random Forest","Neural Network"]

for i in range(len(algorithms)):
    print("The accuracy score achieved using "+algorithms[i]+" is: "+str(scores[i])+" %")
```

```
The accuracy score achieved using Naive Bayes is: 85.25 %
The accuracy score achieved using Decision Tree is: 81.97 %
The accuracy score achieved using Random Forest is: 90.16 %
The accuracy score achieved using Neural Network is: 81.97 %
```

Figure 6: Accuracy scores

Figure 7 shows the accuracy score for four algorithms.

Table 2 Accuracy Percentage

Algorithms	Accuracy
Random Forest	90.16%
Naive Bayes	85.25%
Decision Tree	81.97%
Artificial Neural Network	81.91

The table shows the accuracy of Naive Bayes, ANN, RF, and Decision Tree

Figure 7: Graphical Representation

Figure 8 shows the graphical comparison of four algorithms.

5. CONCLUSION AND DISCUSSION

Heart disorder is a chief problem in today`s unexpectedly converting world. As a result, an automatic machine to are expecting coronary heart disorder at an in advance degree is required. As a result, the health practitioner could be capable of diagnose sufferers greater efficiently, and the general public will gain as properly due to the fact they'll be capable of music their fitness problems the usage of this automatic machine. This paper summarised a number of the professional automatic systems. Feature choice and prediction are crucial additives of any automatic machine. We can enhance our prediction of coronary heart disorder via way of means of deciding on capabilities wisely. We have summarised a few algorithms which might be beneficial while deciding on capabilities, inclusive of the hybrid grid seek set of rules and the random seek set of rules. So, withinside the future, it's far leading to apply seek algorithms to choose capabilities, accompanied via way of means of device studying strategies for prediction, so as to yield higher outcomes withinside the prediction of coronary heart disorder.

6. FUTURE WORK

In the scientific sector, predicting coronary heart ailment is tough however crucial. However, if the situation is diagnosed early and preventative measures are taken as quickly as feasible, the fatality charge may be dramatically reduced.Deep learning algorithms will become increasingly important in healthcare applications in the future. As a result, using deep learning algorithms to forecast heart disease may yield superior results. To determine the disease's level, we'd like to classify it as a multi-class problem.

7. REFERENCES

[1] Bo Jin, Chao Che, Zhen Liu, Shulong Zhang, Xiaomeng Yin, And Xiaopeng Wei, "Predicting the Risk of Heart Failure With EHR Sequential Data Modeling", IEEE Access 2018.

[2] Aakash Chauhan, Aditya Jain, Purushottam Sharma, Vikas Deep, "Heart Disease Prediction using Evolutionary Rule Learning", "International Conference on "Computational Intelligence and Communication Technology" (CICT 2018).

[3] Senthilkumar Mohan, Chandrasegar Thirumalai, and Gautam Srivastava, "Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques", IEEE Access 2019.

[4] K.Prasanna Lakshmi, Dr. C.R.K.Reddy, "Fast Rule-Based Heart Disease Prediction using Associative Classification Mining", IEEE International Conference on Computer, Communication, and Control (IC4-2015).

[5] M.Satish, D Sridhar, "Prediction of Heart Disease in Data Mining Technique", International Journal of Computer Trends & Technology (IJCTT), 2015.

[6] Lokanath Sarangi, Mihir Narayan Mohanty, Srikanta Pattnaik, "An Intelligent Decision Support System for Cardiac Disease Detection", IJCTA, International Press 2015.

[7] Boshra Bahrami, Mirsaeid Hosseini Shirvani, "Prediction and Diagnosis of Heart Disease by Data Mining Techniques", Journal of Multidisciplinary Engineering Science and Technology (JMEST) ISSN: 3159-0040 Vol. 2 Issue 2, February–2015.

[8] Mamatha Alex P and Shaicy P Shaji, "Prediction and Diagnosis of Heart Disease Patients using Data Mining Technique", International Conference on Communication and Signal Processing 2019.

[9] Megha Shahi, R. Kaur Gurm, "Heart Disease Prediction System using Data Mining Techniques", Orient J. Computer Science Technology, vol.6 2017, pp.457-466.

[10] Mr. Chala Beyene, Prof. Pooja Kamat, "Survey on Prediction and Analysis the Occurrence of Heart Disease Using Data Mining Techniques", International Journal of Pure and Applied Mathematics, 2018.

[11] R. Sharmila, S. Chellammal, "A conceptual method to enhance the prediction of heart diseases using the data techniques", International Journal of Computer Science and Engineering, May 2018.

[12] S.Prabhavathi, D.M.Chitra, "Analysis and Prediction of Various Heart Diseases using DNFS Techniques", International Journal of Innovations in Scientific and Engineering Research, vol.2, 1, January 2016, pp.1-7.

[13] Sarabi H.Mujawar, P.R.Devale, "Prediction of Heart Disease using Modified K-means and by using Naïve Bayes", International Journal of Innovative Research in Computer and Communication Engineering, vol.3, October 2015, pp.10265-10273.

[14] Sharan Monica. L, Sathees Kumar. B, "Analysis of CardioVascular Disease Prediction using Data Mining Techniques", International Journal of Modern Computer Science, vol.4, 1 February 2016, pp.55-58.

[15] Mr.P.Sai Chandrasekhar Reddy, Mr.Puneet Palagi, S.Jaya, "Heart Disease Prediction using ANN Algorithm in Data Mining", International Journal of Computer Science and Mobile Computing, April 2017, pp.168- 172.

[16] S. Shylaja, & Muralidharan S. (2019). 2R. International Journal of Engineering Development and Research (www.ijedr.org), Volume 7, (Issue 3 | ISSN: 2321-9939), p.365.

[17] Ty - Jourti - A Hybrid Intelligent System Framework for the Prediction of Heart Disease Using Machine Learning Algorithms VL - 2018PY - 2018DA - 2018/12/02DO -

10.1155/2018/3860146 UR - https://doi.org/10.1155/2018/3860146AB

[18] Data Made Simple. (2019). Top 6 use cases of data science in healthcare. [online] Available at: https://bigdata-madesimple.com/top 6-use-cases-of-data-science-in-healthcare/ [Accessed 29 Jan. 2020].

[19] Mai Shouman, Tim Turner, Rob Stocker, " Using data mining techniques in heart disease diagnosis and treatment", IEEE Japan-Egypt Conference on Electronics, Communications and Computers, 2012.

[20] Syed Umar Amin, Kavita Agarwal, Dr. Rizwan Beg "Genetic Neural Network based Data mining in the prediction of Heart disease using risk factors" IEEE Conference on Information and Communication Technologies (ICT 2013).

[21] K. Polaraju, D. Durga Prasad, "Prediction of Heart Disease using Multiple Linear Regression Model", International Journal of Engineering Development and Research Development, ISSN:2321-9939, 2017.

[22] Marija Sultana, Afrin Haider, "Heart Disease Prediction using WEKA tool and 10-Fold cross-validation", The Institute of Electrical and Electronics Engineers, March 2017.

[23] Dr.S.SeemaShedole, Kumari Deepika, "Predictive analytics to prevent and control chronic disease", https://www.researchgate.net/punlication/316530782, January 2016.

[24] Ashok Kumar Dwivedi, "Evaluate the performance of different machine learning techniques for prediction of heart disease using ten-fold cross-validation", Springer, 17 September 2016.

[25] Datasciencebowl. 2020. Transforming how I measure heart disease. [ONLINE] Available at: https://datasciencebowl.com/transforminghow-we-diagnose-heart-disease/. [Accessed 2 February 2020]

[26] Rashmee U. Shah, John S. Rumsfeld, Big Data in Cardiology, European Heart Journal, Volume 38, Issue 24, 21 June 2017, Pages 1865–1867, https://doi.org/10.1093/eurheartj/ehx284

[27] Alexander, Cheryl & Wang, Lidong. (2017). Big Data Analytics in Heart Attack Prediction. Journal of Nursing & Care. 06. 10.4172/2167-1168.1000393

[28] N. K. S. Banu and S. Swamy,(2016) "Prediction of heart disease at an early stage using data mining and big data analytics: A survey," 2016 International Conference on Electrical, Electronics, Communication, Computer and Optimization Techniques (ICEECCOT), Mysuru, 2016, pp. 256-261.DOI: 10.1109/ICEECCOT.2016.7955226.

[29] Dua, D. and Graff, C. (2019). UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: the University of California, School of Information and Computer Science.

[30] Flint, Alan J et al. "Body mass index, waist circumference, and risk of coronary heart disease: a prospective study among men and women." Obesity research & clinical practice vol. 4,3 (2010): e171-e181. doi:10.1016/j.orcp.2010.01.001

