IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 5, May 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.379

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

Heart Disease Prediction and Doctor Recommendation System Using Machine Learning

Juned Hakim Shaikh, Abhishek Ganesh Takale, Mrs. K.T. Mohite

Student, Department of Computer Engineering, Sinhgad Academy Of Engineering, Pune, Maharashtra, India.

Professor, Department of Computer Engineering, Sinhgad Academy Of Engineering, Pune, Maharashtra, India.

ABSTRACT: In recent years, heart disease has become the most dangerous disease worldwide and accounts for a large proportionof deaths worldwide. Not in India but everywhere. Reliable, timely and accurate disease diagnosis and treatment is essential for proper care. Machine learning techniques to automate diagnostic and treatment processes using clinical data as a test. As technology developed, machine learning techniques were used by researchers to diagnose heart problems to help the medical industry.Saying "true" or "false" can be expected when the person's disease is associated with heart disease. Future treatments would greatly benefit from such predictions.

KEYWORDS: Disease Prediction, Machine Learning Classification, Doctor Recommendation, Python, SVM Algorithm.

I. INTRODUCTION

Heart disease is currently one of the leading causes of death, abnormal changes in cholesterol, blood pressure,Heart rate, etc. Heart disease is one of the most common conditions in major obstacles; there are many parameters and correctly predicting this state requires complexity. There is a huge need for study in the field of predicting heart diseases in humans, as according to the latest WHO survey, doctors were able to correctly predict only 67% of heart diseases. Researchers are using data mining techniques to help health professionals identify heart disease. There is a vast amount of clinical data available, but where key information is hidden, it is rarely accessed and remains unused. A support vector machine algorithm is used to predict the risk level of heart disease. Age, gender, type of chest discomfort, resting blood pressure, cholesterol in mg/dL, fasting blood sugar and 14 other variables are entered into the algorithm. After predicting the level of risk, the user would be advised on preventive measures that would allow him to temporarily manage the level of risk. The topic of heart related diseases is covered in detail in this article. The user's understanding of the many parameters related to prediction and their options for controlling these factors are described in more detail later in the scope.The disease prediction system contains data sets collected from various health-related websites. With the help of Disease Predictor, the user will be able to know the probability of the disease with the given symptoms. As the use of internet is increasing day by day, people are always curious about different new things. People always try to refer to the internet if there is any problem. People have internet access than hospitals and doctors. People do not have an immediate option when they suffer from a particular disease. So this system can be useful for people because they can access the internet 24 hours a day. The prediction of such a system may have very great potential in the medical treatment of the future. Once the system predicts the disease, it will recommend which type of doctor to visit.

II. MOTIVATION

We found a solution to the problems with the current framework where we designed accuracy, reliability. Productivity by developing the main characteristics of various diseases, where we find most common diseases in individual well-being. In one application, we introduced disease expectations by examining symptoms collected from patient records and taking the Reliable data set from different hospitals. Thus reducing costs and saving time. If certain anomalies are found in the diagnosis, then the recommendation of a nearby specialist and hospital according to the user's preferences will facilitate quick and appropriate treatment

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

III. PROBLEM STATEMENT

It is difficult to have regular personal access to the hospital and doctors. Going to hospitals for routine advice is time-consuming and expensive. There is a need for localized people to easily connect with GPs, which is possible with a machine learning approach. Several times, it is difficult for common people to get advice from health experts in time and at an affordable cost. This can be reduced using existing technologies.

IV. LITERATURE SURVEY

[1]The available studies that are surveyed concentrate on the forecast of cardiovascular illness. The transformed data set is the system is trained using data obtained from a variety of sources and are subsequently put to use for forecasting. T. Princy and J. Thomas

[2] use data mining techniques that are currently being used in research for the prediction of heart disease to provide specifics regarding various knowledge abstraction strategies. In this study, algorithms are used to assess three types of data mining methods support vector machine Algorithm on sets of medical data. While Priyanka and Naveen

[3] concentrate primarily on developing a decision support system for heart disease prediction using the Naive Bayes algorithm. The Cleveland dataset is used to construct a web application that collects user input and can access concealed information about heart disease from a historical database. M. Gandhi and S. Singh's

[4] focus on data mining classification techniques for data discovery. For data classification and knowledge extraction, several data mining classification algorithms offer advantages and disadvantages. Richa Sharma, Purushottam, and K. Saxena

[5] illustrated how to obtain the risk level from the heart disease database. The screening of clinical data for cardiac patients is contained in the input database. To increase the effectiveness of the mining process, the database is first pre-processed. The prediction is made using a decision tree. As reliable predictions are required in the medical field, the studies mostly concentrate on determining the effectiveness of various algorithms.

Authors Ahmed M. Alaa and Senthil Kumar Mohan experimented with a variety of variables and found that a random hybrid forest produced results with an accuracy of 88.7%. In order to make the decision support system simpler, the goal of this work is to enhance the savvy treatment. The diagnosis of heart illness by keeping an eye on a person's heartbeat is the subject of this extensive paper. You are allowed by the framework to specify your pulse's criteria. After establishing these limitations, a person can start monitoring their heart rate, and whenever it exceeds a certain threshold, they are warned that they are at risk of having a heart attack or coronary failure and are having a high pulse.
This study examines traditional supervised binary classification, where the dataset contains a number of attributes. Plasma glucose concentration is included in the dataset. Body mass index and blood pressure in mm Hg Years of age, etc. In order to identify those who have the condition, a number of factors are used, each with distinct characteristics. To solve the issue, we must analyze data, make any necessary adjustments, apply ML, train a model, examine the output of the trained model, and continue with several techniques until we get the most accurate result. When developing software or websites, it's important to understand the framework requirements and obtain the necessary data to communicate with customers and providers.

V. PROJECT REQUIREMENT

TABLE 1. HARDWARE REQUIREMENTS

SR.NO	Hardware	Description
1.	Processor	Pentium-IV
2.	Speed	1.1GHz
3.	Hard Disk Space	40GB
4.	Device	HP Pavilion
5.	Other	Other required standard computer peripherals, such as keyboard and mouse.

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

Table 2. Software Requirements

Sr. No.	Software	Description
1.	Operating System	Microsoft Windows 10
2.	IDE	Spyder
3.	Frontend	Tkinter
4.	Backend	Python
5.	Database	DBSQ Lite

VI. SYSTEM ARCHITECTURE

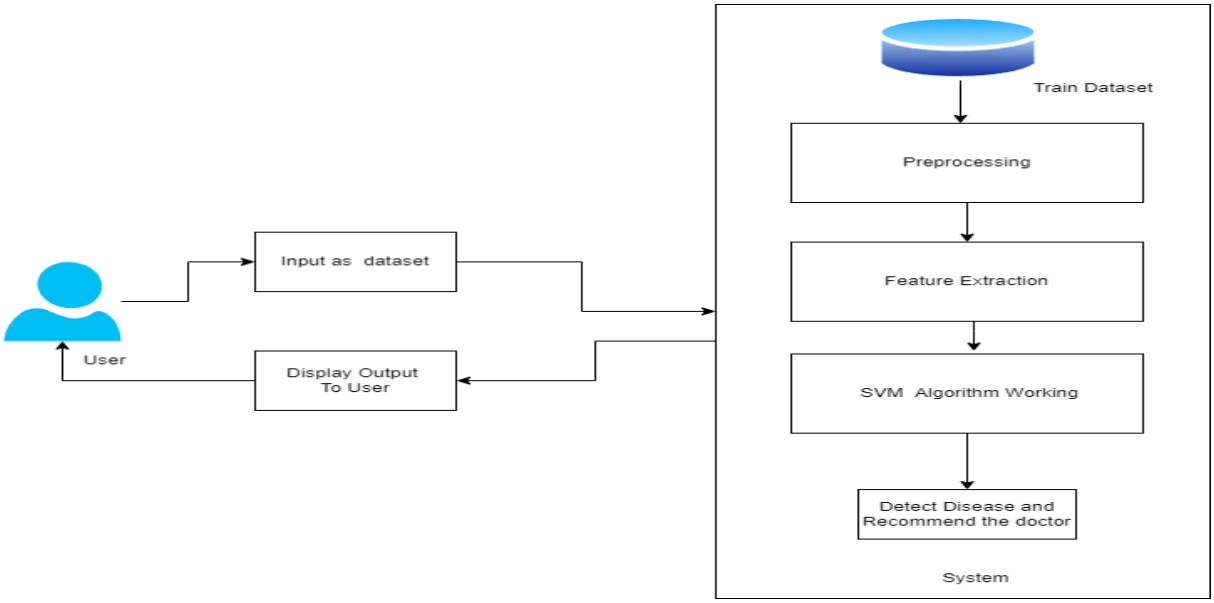

Fig 5.1 System Architecture

VII. METHODOLOGY

Medical Dataset:

The Cleveland Heart Disease dataset is obtained from the UCI Machine Learning Repository [5] for use in this paper.Although there are 303 records with 76 attributes, we only used the 14 most relevant attributes.In the data set used, HD is present in 45.54% of cases.Table 1 shows the 14 attributes that were used along with their descriptions.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

Heart Disease Attribute Information:

Attributes	Description
Age	Age in years(29-77)
Sex	Male=1, Female=0
Chol	Serum Cholestrol in mg/dl
Fbs	Fasting blood sugar >120 mg/dl, Value 1=yes, Value 0=no
Cp	Chest pain types Value 1= typical angina, Value 2= atypical angina Value 3= non-angina, Value 4= asymptomatic
Trestbps	Resting blood pressure in mm Hg [94-200]
Chol	Serum cholesterol in mg/dl [126--564]
Restecg	Resting electrocardiographic , Value 0=normal, Value1= having ST-T wave abnormality

Series	Age	Sex	ChestPain	RestBP	Chol	Fbs	RestECG	MaxHR	ExAng	Oldpeak	Slope	Ca	Thal	AHD
1	63	1	typical	145	233	1	2	150	0	2.3	3	0	fixed	No
2	67	1	asymptom	160	286	0	2	108	1	1.5	2	3	normal	Yes
3	67	1	asymptom	120	229	0	2	129	1	2.6	2	2	reversable	Yes
4	37	1	nonangina	130	250	0	0	187	0	3.5	3	0	normal	No
5	41	0	nontypical	130	204	0	2	172	0	1.4	1	0	normal	No
6	56	1	nontypical	120	236	0	0	178	0	0.8	1	0	normal	No
7	62	0	asymptom	140	268	0	2	160	0	3.6	3	2	normal	Yes
8	57	0	asymptom	120	354	0	0	163	1	0.6	1	0	normal	No
9	63	1	asymptom	130	254	0	2	147	0	1.4	2	1	reversable	Yes

Fig 6.1 Heart Disease Attribute Information

Data Pre-processing:

The most important first step in any project is to prepare documents. prediction model Helps transform data into inform ation to understand the treatment modelmost of the data is incomplete, lacks important features, and is uninformative or irrelevant. The estimated behavior (number) in the original data has five values: 0 for no HD, 1 to 4 for variable HD, a nd 5 for no HD. In this study, we are only interested in the presence or absence of HD, not the actual disease distributio n. As a result, the attribute class is assigned a binary value of 0 or 1, indicating whether the patient has HD.

Data Classification:

A Support Vector Machine (SVM) is a learning model often used for classification. It is defined as a set of dimension s in which each dimension represents a feature of a particular model. The head system is supported to find the best ch osen energy plane with the best separation boundary between the two classes. Due to their computational power over l arge datasets, DVMs show promise as reliable methods for solving large-scale problems.

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

VIII. RESULTS

Fig 7.1 Main Page

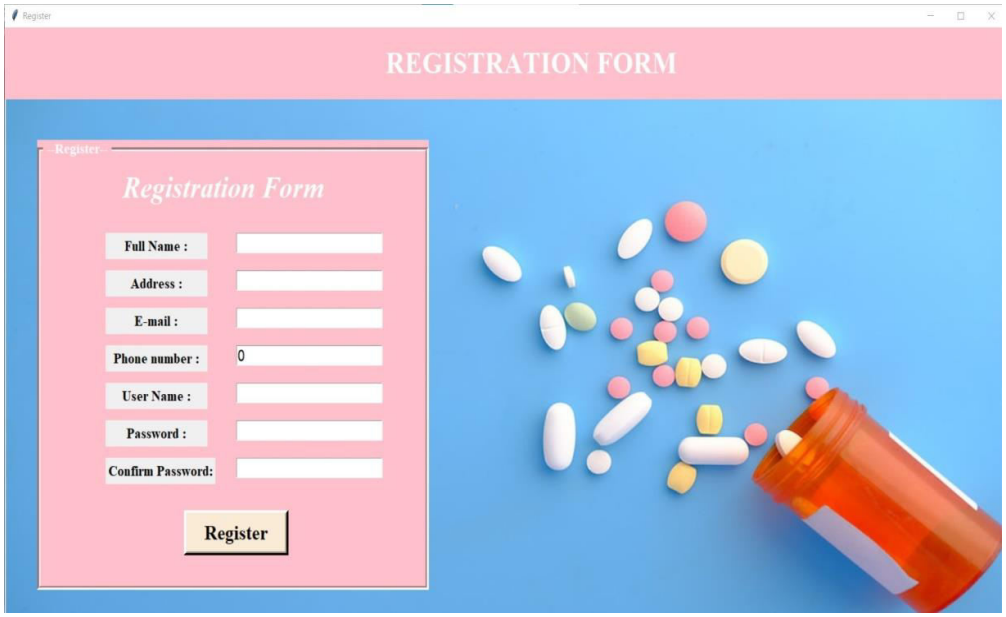

Fig 7..2Registration Page

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

Fig 7.3 Login Page

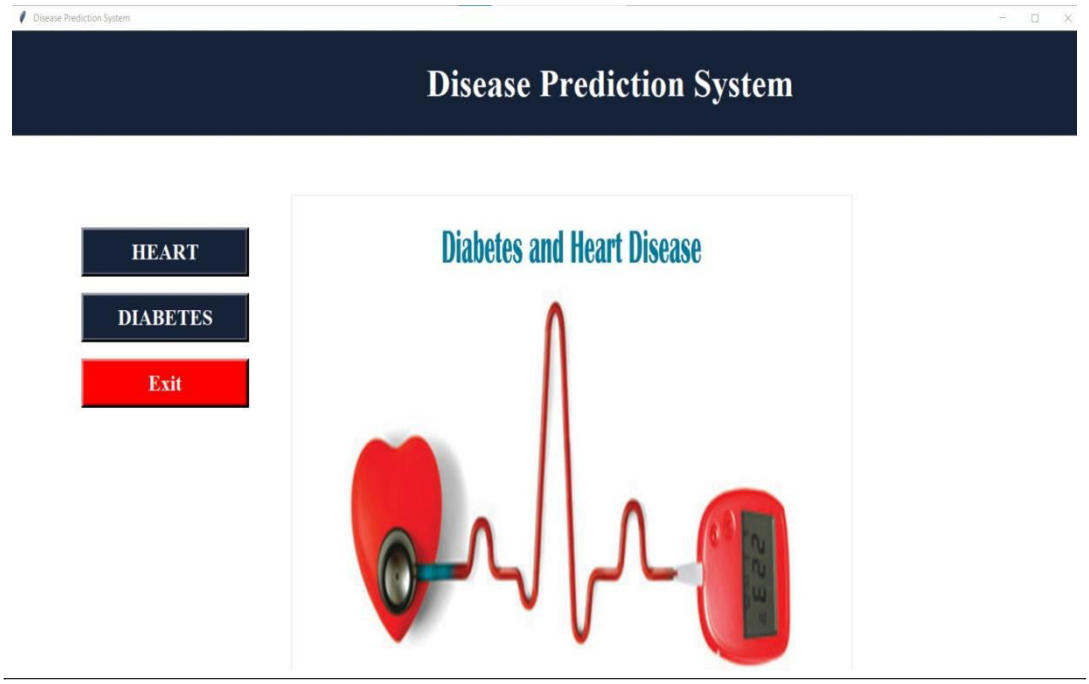

Fig 7.4 Home Page

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

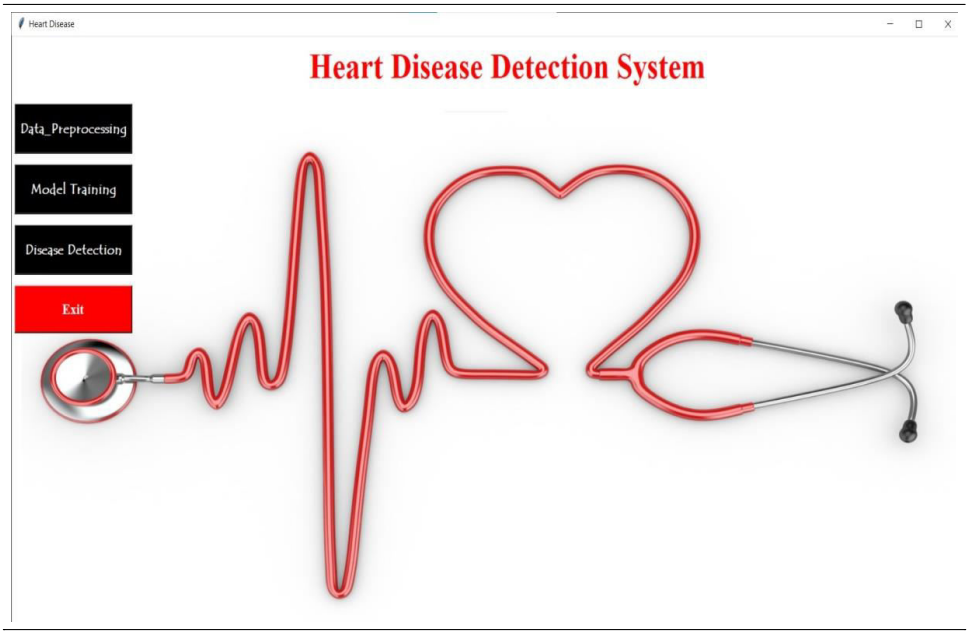

Fig 7.5 Heart disease Detection Page

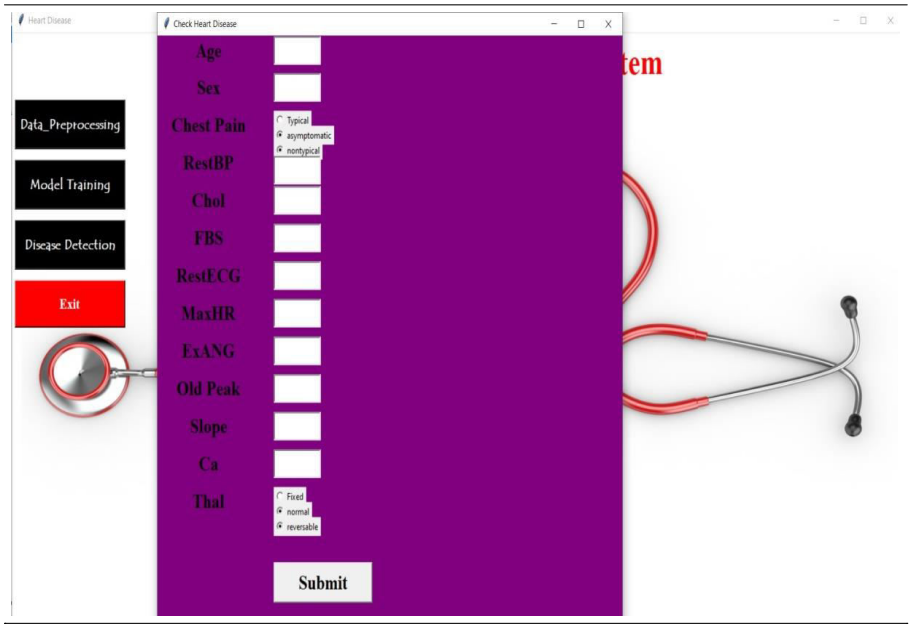

Fig 7.6 Heart disease Detection data entry Page

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

Fig 7.7 Diabetes disease Detection Page

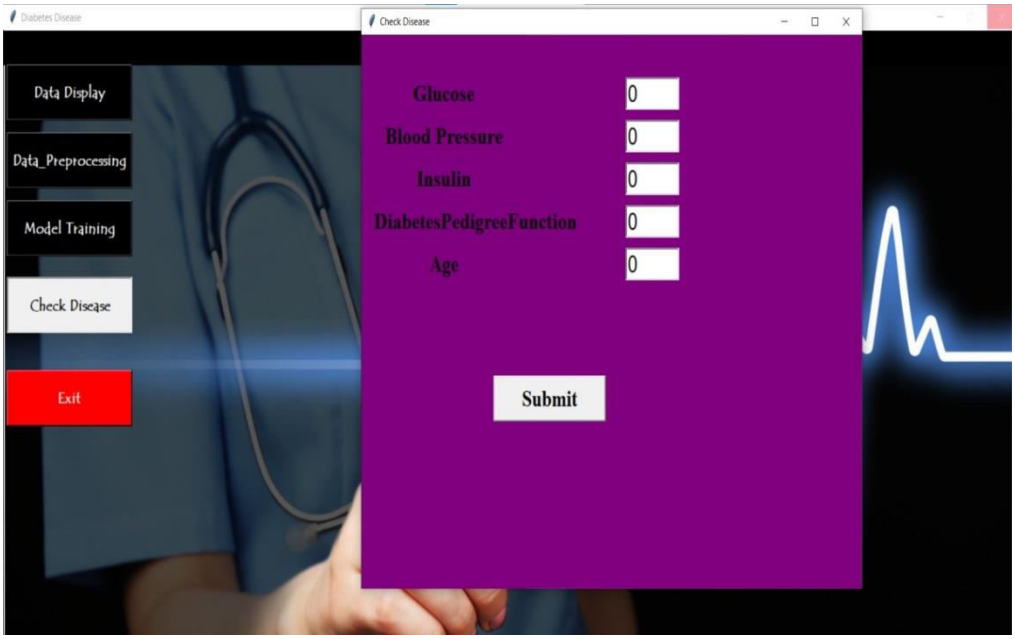

Fig 7.8 Diabetes disease Detection data entry Page

IX. CONSOLATION

We set out to develop a system that can predict disease based on the symptoms that are provided to it. According to the prior Papers we analyzed, this was never implemented before, making the Project more definitive. Such a framework can store the data input by the user in the database which can be utilized in the future to help in constructing better versions of such systems. The disease episodes have been assessed using data analysis and machine learning techniques, which have been combined to determine each algorithm's output expectation yield and apply the

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 8.379 |

|| Volume 11, Issue 5, May 2023 ||

| DOI: 10.15680/IJIRCCE.2023.1105058 |

right approach suggested for the area needed. Additionally, it presents information acquired and accomplishments achieved in a different visual manner. Additionally, it features numerous graphic depictions.

REFERENCES

1. T. Princy, J. Thomas, "Human Heart Disease Prediction System using Data Mining Techniques," International Conference on Circuit, Power and Computing Technologies [ICCPCT], 5090-1277, 2016.
2. Priyanga, Naveen, "Web Analytics Support System for Prediction of Heart Disease Using Naive Bayes Weighted Approach (NBwa)," Asia Modelling Symposium (AMS), 2376-1172, 2017.
3. M. Gandhi, S. Singh, "Predictions in Heart Disease using Techniques of Data Mining," 1st International Conference on Futuristic trend in Computational Analysis and Knowledge Management, 4799-8433, pp. 520-525, June 2015.
4. Purushottam, K. Saxena, R. Sharma, "Efficient Heart Disease Prediction System using Decision Tree", 4799-8890, International Conference on Computing, Communication and Automation (ICCCA), pp. 72-77, July 2015.
5. Mrs. Shital. P. Chattar, PrathameshWadile, BhushanSaraf, OmkarKardile and AdityaPawar." IOT Based Heart Disease Prediction." International Journal for Research in Applied Science & Engineering Technology (IJRASET) (2021):2321-9653.
6. Megha Vyas, Priya Kunwar, Vinod Joshi, Mihir Maheshwari and Mr. GirishAmeta. "Software for Disease Prediction." International Journal for Research in Applied Science & Engineering Technology(IJRASET) :2321-9653,2020
7. Fayyad U.M, Piatetsky-Shapiro G, Smyth P., "Knowledge discovey of data mining : Towards a unifying framework in KDD", 1996 august , Vol 96, pp.82-88.
8. Kurgan L.A, MusilekP ,"A survey of knowledge discovery and data mining process model," The knowledge engineering review 2006, 21(1), 1-24.
9. Das R Turkoglu I, Sengur A 2009, "Effective diagnosis of heart disease through neural network ensembles", Expert system with applications 36(4), 7675-7680.
10. Dewan A, Sharma M, " Prediction of heart disease using hybrid technique in data mining classification," In computing for sustainable global development (INDIA Com), 2015 2nd international conference on (pp. 704-706). IEEE March 2015
11. .Chandna D, "Diagnosis of heart disease using data mining algorithm", International journal of computer science and information technologies, 5(2), 1678-1680.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.379

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details