Effective System For Prediction Of Heart Disease By Applying Logistic Regression

Radha Mothukuri, Mallempudi Sai Satvik, Kolusu sri Balaji, Dodda Manikanta

Abstract: In the present current way of life of individuals are influencing by various medical problems, one among them is coronary illness which might be nascent from early age. Presently a day's machine learning is turning into a typical instrument in medical services field. AI technology helps in logical philosophy for recognizing significant data. Machine Learning provides various advantages in medical industry. Identification of the extortion in medical coverage, accessibility of restorative answers for the patients at less cost. Acknowledgement of reasons for infections and ID of medical treatment techniques. It additionally helps the social insurance analysts for making proficient medical services strategies, building drug suggestion frameworks, creating wellbeing profiles of people and so forth. The prevalent objective of this paper is to recognize the nearness or nonattendance of coronary illness for a person. In the medical industry, it is exceptionally hard to find whether an individual is influenced by coronary illness or not by a doctor. It requires a cautious comprehension of patient's information, and the distinguishing proof of those parameters which cause the ailment the entirety of this is considered as a troublesome assignment. Extra apparatuses are required for settling on the clinical choice of coronary illness. The data set for prediction of coronary illness, containing 303 cases, which have been utilized for the preparation and testing of the created framework. The consequences of this paper shows the regression technique like Logistic regression is being applied for the heart disease forecast so as to improve the framework productivity when contrasted with random forest and support vector machine(svm) algorithms.

Keywords: c Regression, Cost Function, Regularization, Gradient Descent, Artificial Intelligence (AI). c Regression, Cost Function, Regularization, Gradient Descent, Artificial Intelligence (AI). c Regression, Cost Function, Regularization, Gradient Descent, Artificial Intelligence (AI). Logistic regression, cost function, svm, random forest, Gradient descent

——————————————◆——————————————

INTRODUCTION
Cardiovascular disease is otherwise called as heart disease. Heart disease is a term used for covering any disorder of the heart that includes the conditions and problems with the blood vessels, circulatory framework, structural problems, blood clots and refers to issues and deformities in the heart. Heart attack occurs when coronary arteries get blocked. Arteriosclerosis which for the most part implies solidifying of supply routes, the supply routes become dense and never again are adaptable. Atherosclerosis, which is nothing but narrowing of supply routes, so not as much blood course through those fabricate ups called as plaque. Respiratory failures for the most part happen when the coagulation of blood or the obstruction of blood stream to the heart and from the heart. Coronary illness can influence the conduits, however the heart muscle, valves, beat, or other significant perspectives of a well-working heart. As showed by the centres for disease control (CDC), coronary illness is the predominant course of expiration in Australia, the UK, U.S.A and some more. Every One out of four passing's are happened because of coronary illness in the India, U.S.A. There are many driving reasons for death in India however cardiovascular illness turned out as topmost executioner that has influenced both the urban and provincial populace. As indicated by the Global Weight of infection study 28% of the passing's happened that is a fourth of passing's in our nationare because of cardiovascular infection. With 1,752 coronary illness related passing's out of 100,000 individuals Russia has the most noteworthy pace of coronary illness. Around 6,10,000 individuals kick the bucket of coronary illness in the US consistently. The normality of cardiovascular illness in 2016 was the most noteworthy in Punjab, Kerala and Tamilnadu–more than 5,000 for every

————————————————————

- ***Radha Mothukuri,*** *Assistant Professor, Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur-522502. radhahemanth12 @gmail.com*

populace of 100,000. Andhra Pradesh, Maharashtra, Himachal Pradesh, West Bengal and Goa are close second with commonness somewhere in the range of 4,500 and 4,999per 100,000. In India the greatest variables for heart illnesses are with the most noteworthy rating terrible nourishment propensities with over half, elevated cholesterol with almost 30% and tobacco utilization with 18%. Smoking is viewed as the significant hazard with 83% among tobacco clients.

LITERATURE SURVEY
There is a number of prediction systems proposed for different diseases and implemented using different techniques. Previous works on heart disease with different authors studied and implemented different methods and analyzed the results. For the implementation of the work, they have considered the data set from the UCI data repository which can also be collected from the kaggle. The authors performed the classi cation and prediction technique on the data set. SellappanPalaniappan, Ra ah Awang are the authors they have used three data mining classi cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a web-based prediction system that can be used easily and is reliable, expandable and scalable. The most effective model for heart disease prediction is Naïve Bayes followed by the Decision Tree. Naïve Bayes showed a better accuracy than the Decision Tree. [6]. Mai Showman, Tim Turner, Rob Stocker are the authors who applied the K-Means method that is combined with the Decision Tree for the Heart Disease prediction system. For the implementation of this work, they have applied Initial centroid selection techniques in order to boost the model

432

accuracy [7]. For diagnosing the heart disease they have used the Decision Tree classi☐er. In order to generate the initial centroids based on the number of actual samples in the data set the Random Row, There is a number of prediction systems proposed for different diseases and implemented using different techniques. Previous works on heart disease with different authors studied and implemented different methods and analyzed the results. For the implementation of the work, they have considered the data set from the UCI data repository which can also be collected from the kaggle. The authors performed the classi☐cation and prediction technique on the data set. SellappanPalaniappan, Ra☐ah Awang are the authors they have used three data mining classi☐cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi☐cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a web-based prediction system that can be used easily and is reliable, expandable and scalable. The most effective model for heart disease prediction is Naïve Bayes followed by the Decision Tree. Naïve Bayes showed a better accuracy than the Decision Tree. [6]. Mai Showman, Tim Turner, Rob Stocker are the authors who applied the K-Means method that is combined with the Decision Tree for the Heart Disease prediction system. For theimplementation of this work, they have applied Initial centroid selection techniques in order to boost the model accuracy [7]. For diagnosing the heart disease they have used the Decision Tree classi☐er. In order to generate the initial centroids based on the number of actual samples in the data set the Random Row, There is a number of prediction systems proposed for different diseases and implemented using different techniques. Previous works on heart disease with different authors studied and implemented different methods and analyzed the results. For the implementation of the work, they have considered the data set from the UCI data repository which can also be collected from the kaggle. The authors performed the classi☐cation and prediction technique on the data set. SellappanPalaniappan, Ra☐ah Awang are the authors they have used three data mining classi☐cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi☐cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a web-based prediction system that can be used easily and is reliable, expandable and scalable. The most effective model for heart disease prediction is Naïve Bayes followed by the Decision Tree. Naïve Bayes showed a better accuracy than the Decision Tree. [6]. Mai Showman, Tim Turner, Rob Stocker are the authors who applied the K-Means method that is combined with the Decision Tree for the Heart Disease prediction system. For theimplementation of this work, they have applied Initial

centroid selection techniques in order to boost the model accuracy [7]. For diagnosing the heart disease they have used the Decision Tree classi☐er. In order to generate the initial centroids based on the number of actual samples in the data set the Random Row, There are various expectation frameworks proposed for various infections and executed utilizing various strategies. Past takes a shot at coronary illness with various creators contemplated and actualized various strategies and dissected the outcomes. To execute the work, they considered the informational index from UCI information data stockpiling which is moreover gathered from the UCI repository. Therefore, the creators played out the arrangement and conjecture strategy on the informational index. SellappanPalaniappan, Ra☐ah Awang are the authors they have used three data mining classi☐cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi☐cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a webSellappanPalaniappan, Ra☐ah Awang are the authors they have used three data mining classi☐cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi☐cation matrix method for the evaluation of the effectiveness of the model. SellappanPalaniappan, Ra☐ah Awang are the authors they have used three data mining classi☐cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi☐cation matrix method for the evaluation of the effectiveness of the model. The creators Sellappan Palaniappan, Rafiah Awang have used three Machine learning grouping demonstrating procedures. These strategies concealed data from the coronary illness dataset [1]. To approach this model, they have used query language like DMX. This model is prepared on trained information and tried with test data to survey the results. These creators utilized Lift outline and order matrix technique for the appraisal of the sufficiency of the model. Thereby, the framework separates the concealed information from chronicle coronary illness database. The creators actualized the suggested framework subject to the.net framework which is an electronic expectation framework that can be utilized viably and is solid, versatile and expandable. Gullible Bayes or Naive Bayes seemed better than that of the Decision Tree. [2]. SellappanPalaniappan, Ra☐ah Awang are the authors they have used three data mining classi☐cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi☐cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden

knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a webThe K-Means technique that is joined with the classification Tree for the Heart Disease prediction framework is bided by Mai showman, Tim turner, Rob stocker. To actualize the work, they applied Initial centroid determination systems to expand the model precision [3]. To diagnose the coronary illness, they utilized the classification tree. For generating the underlying centroids dependent on the quantity of real examples in the data set the Random Row, Attributes, Inliers, anomaly and Range methods are utilized [4]. The creators collected the information from the UCI information archive. At long last, the creator contrasts the exhibition and recently applied choice tree with previously applied decision tree actualized earlier on similar information with Decision tree and k-means that is utilized to show the signs of improvement exactness. The proposed method demonstrated a precision of about 84% [5]. Kanika Pahwa and Ravinder Kumar applied the half and half procedure for choosing the highlights on the heart disease dataset for expectation. Right of the bat, the creator moved toward the feature determination technique utilizing the SVM-RFE alongside the addition proportion for getting the subset of highlight and expelling the unessential and repetitive element. Recognizing the highlights is significant for forecast. They have applied Naive Bayes and Random Forest on the subset of highlights for arranging the informational collection into present or nonappearance of coronary illness. The outcomes have indicated a superior precision for both when applied with chosen highlights. The precision accomplished by Naive Bayes is 84.15%, while the other Random Forest accomplished 84.16%, for which it got practically comparative precision when applied with chosen highlights for both the procedures [6]. Bommadevara, H. S. A., Sowmya.Y, & Pradeepini, G are the creators who have utilized machine learning algorithms for Heart disease prediction [7]. Praveena, M., Asha Deepika, R., & Sai Raghavendhar did their Analysis on forecast of heart disease utilizing data mining techniques [8]. Sai Krishnasree, K., &Narasinga Rao, M. R. made a comparative study Bayesian regularization with multiple regression models [9]. Siva Kumar, P., Anand, D., Uday Kumar, V., & Bhattacharyya clearly made a research on computational knowledge strategy for viable diagnosis of heart disease utilizing genetic algorithm [10].

Dataset:

SellappanPalaniappan, Ra□ah Awang are the authors they have used three data mining classi□cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi□cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a webSellappanPalaniappan, Ra□ah Awang are the authors they have used three data mining classi□cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results.

These authors used Lift chart and classi□cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a webSellappanPalaniappan, Ra□ah Awang are the authors they have used three data mining classi□cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi□cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a webSellappanPalaniappan, Ra□ah Awang are the authors they have used three data mining classi□cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi□cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a web-based prediction system that can be used easily and is reliable, expandable and scalable. The most effective model for heart disease prediction is Naïve Bayes followed by the Decision Tree. Naïve Bayes showed a better accuracy than the Decision Tree. [6] SellappanPalaniappan, Ra□ah Awang are the authors they have used three data mining classi□cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi□cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a web-based prediction system that can be used easily and is reliable, expandable and scalable. The most effective model for heart disease prediction is Naïve Bayes followed by the Decision Tree. Naïve Bayes showed a better accuracy than the Decision Tree. [6] SellappanPalaniappan, Ra□ah Awang are the authors they have used three data mining classi□cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi□cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a web-based prediction system that can be used easily and is reliable, expandable and scalable. The most effective model for heart disease prediction is Naïve Bayes followed by the Decision Tree. Naïve Bayes showed a better accuracy than the Decision Tree. [6] SellappanPalaniappan, Ra□ah Awang

are the authors they have used three data mining classi␣cationmodeling techniques. These techniques dig out the hidden information from the heart disease database [5]. For accessing the model they have used DMX query language. The model is trained on train data and tested with test data to evaluate the results. These authors used Lift chart and classi␣cation matrix method for the evaluation of the effectiveness of the model. The system extracts the hidden knowledge from the historical heart disease database. The authors implemented the proposed system based on the.net framework which is a web-based prediction system that can be used easily and is reliable, expandable and scalable. The most effective model for heart disease prediction is Naïve Bayes followed by the Decision Tree. Naïve

Bayes showed a better accuracy than the Decision Tree. [6 Dataset:

ATTRIBUTE	DESCRIPTION
Age	Range [25-110]
Sex	1: Female 0: Male
CP	1: asymptomatic 2: non- angina 3: typical angina 4: non typical angina
Resting Blood Pressure RBS	blood pressure range [60 - 200] mm of Hg
Chol	cholesterol computation range [120 -600]mm/dl
fbps(Fasting Blood Sugar)	1: <=120 mg/dl 0: >120 mg/dl
ECG (Electro Cardio Graph)	0: normal 1: abnormal 2: showing probable
Exercise Induced Angina	0: NO 1: YES
Old peak	ST curve of rest range [0-6]
Slope(y)	Peak exercise ST 1: down sloping 2: upsloping 3: flat
Coronary Artery	Vessels shown in x-ray range [0-3]
Thalassemia	Bone marrow expand relative congestive heart failure 3: fixed defect 6: reversible defect 7: normal
Target	1: heart disease 0: No heart disease

Methodologies:
        Logistic Regression: Coronary illness forecast system utilizes one of the relapsing techniques like Logistic relapse approach for designing the preparation model. Determined It was initially created by David Cox. Strategic relapse is a factual examination procedure which has been utilized for foreseeing information esteem dependent on earlier perception of the informational collection. The calculated relapse model estimates the reliant information fluctuating by examining the connection in the middle of at least one existing free factor. Calculated Relapsing is best the significant instrument to expect the solution, which can likewise be utilized for grouping, anticipating the information dependent on authentic information. Thereby, the actualized model is a

twofold regression model which has subordinate factors with 2 potential results i.e., one is a favourable worth and other is the unfavourable worth which has 0 or 1 as class mark. The calculated relapse is executed on the coronary supply route ailment. The Python code is used for the execution of strategic relapse calculation done by Anthony rolled out not many improvements as indicated by need and prerequisites. The proposed model is executed as the accompanying: It for the most part comprises of two significant stages: cost work and inclination drop. Cost Function is utilized for ascertaining the greatest probability approximation. Angle plunge is a repetitive procedure for obtaining coefficients from preparing information. The procedure is rehashed till we obtain the ideal parameters of train information. At whatever point a tested information was passed to the model dependent on the limitations can perceive whether the individual is having coronary illness or not, it tests the information utilizing the s-shaped function. The minimization capacity is a technique utilized for lessening the blunders of the anticipated mark and genuine name. Slope plummet work is the strategy that is utilized for figuring the measurement until we acquire a base estimation of the group mark.

A. Log Loss Function

 It utilizes the Log Loss for example the logarithmic misfortune which quantifies the presentation of the model where the expectation input esteem is the likelihood between zero and one. The log misfortune is the vulnerability of the forecast which depends on the amount it differs from the real name. Cost work which causes the student to address or change the conduct to limit the expense. The minimization capacity can be assessed by iteratively running the model to think about the evaluated anticipated worth and the known or real worth. The regularized Log Loss function is a strategy that is utilized for conquering the danger of over fitting the information.   The minimization function is calculated by the following formula:

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}[y^{(i)}\log(h_\theta(x^{(i)})) + (1 - y^{(i)})\log(1 - h_\theta(x^{(i)}))]$$

  m = number of cases, n = number of features, y = target class, x = features of trained data, θ= value of coefficients,

B. Inclination drop

Inclination drop is an optimized strategy which is utilized to discover the parameters or the coefficient of the cost function. Inclination drop is otherwise called as gradient descent. Inclination drop is a repeated process in order to get the coefficients to minimize the cost function. The inclination drop is determined for both the classes to get the pair of coefficients for both class names. The objective here is to proceed with the technique to attempt the diverse incentive for the coefficient, assessing their expense and choosing the new coefficient that is marginally having the lesser cost.   Inclination drop is calculated by:

Gradient Descent

Remember that the general form of gradient descent is:

$$Repeat \ \{$$
$$\theta_j := \theta_j - \alpha \frac{\partial}{\partial \theta_j} J(\theta)$$
$$\}$$

We can work out the derivative part using calculus to get:

$$Repeat \ \{$$
$$\theta_j := \theta_j - \frac{\alpha}{m} \sum_{i=1}^{m} (h_\theta(x^{(i)}) - y^{(i)})x_j^{(i)}$$
$$\}$$

m = number of cases,
x=features of trained data, y = target class
θ = value of coefficients,

C. Activation function (s-shape curve)
Activation function is explicitly characterized as a squashing function. Squashing function limit the yield somewhere in the range of zero and one. It is a scientific capacity which have a trait of s-shaped curve. The quality decision for activation function is logistic function. Generally, activation function is utilized to delineate to likelihood it is characterized as:

$$h(x) = \frac{1}{1 + e^{-x}}$$

x = features of test data

At whatever point a test information is passed it computes the worth dependent on parameters put away in the model. It forecast he likelihood of each class name.

Experimental Results:
The precision acquired is about 69.93% ensuing to apply the Logistic Regression on the dataset. The dataset is separated into an extent of 70:30 for example 211 occurrences for preparing and 92 examples for testing the information. The precision got for the proposed technique is about 69.93% when contrasted and the Support vector machine which got 57.38% and 55.74% for random forest. The applied technique is proficient than the Random Forest and support vector machine.
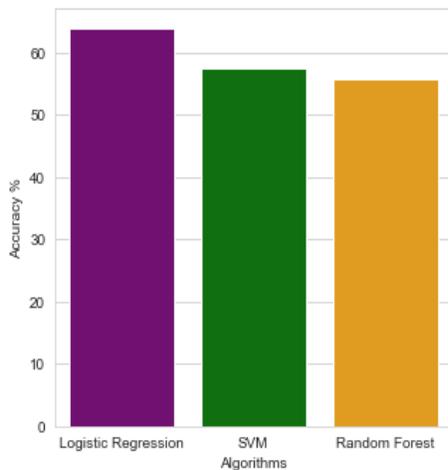
Fig: *Comparison of three machine learning algorithms*

Confusion matrix:
Confusion matrix generally used for measuring the performance

of the algorithm. It shows the way in which the model is confused
when it makes predictions. Each row indicates the instances of
the actual class label and each column indicates the instances of
the predicted class label.
We illustrate the results with confusion matrix which shows the
complete details
Confusion matrix generally used for measuring the performance
of the algorithm. It shows the way in which the model is confused
when it makes predictions. Each row indicates the instances of
the actual class label and each column indicates the instances of
the predicted class label.
We illustrate the results with confusion matrix which shows the
complete details
Confusion matrix generally used for measuring the performance
of the algorithm. It shows the way in which the model is confused
when it makes predictions. Each row indicates the instances of
the actual class label and each column indicates the instances of
the predicted class label.
We illustrate the results with confusion matrix which shows the
complete details
It is a table that is frequently utilized to depict the exhibition on an arrangement model on a lot of test information for which the genuine values are known values are known. Each row indicates the occurrences of the authentic class label and each column indicates the instances of the forecast class label. We constructed the confusion matrix for 3 algorithms as shown below.
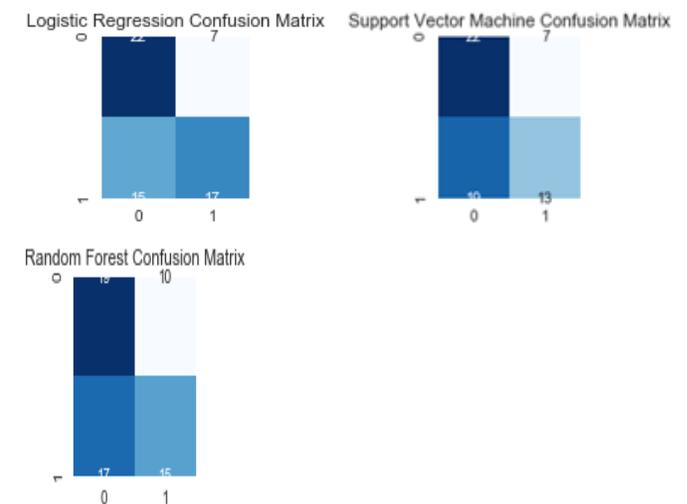
CONCLUSION:
In this paper, we introduce about the effective system for heart disease forecast by using logistic regression. Forecast of heart

disease spare the life of people and can have significant sway on its treatment. The techniques are random forest, svm. we have analysed that the logistic regression has better exactness when contrasted with random forest and svm. An investigative correlation has been accomplished for discovering best available calculation for medical dataset. In future our point is to convey forward crafted by fleeting restorative dataset, where dataset changes with the time and retraining of dataset is required.

REFERENCES

[1] UCI Repository https://archive.ics.uci.edu/ml/datasets/Heart+Disease

[2] R. Awang, S.Palaniappan- "Intelligent heart disease prediction system using data mining techniques."

[3] Seyedamin pouriyeh, Juan Gutierrez, Hamid Arabnia-A Comprehensive Investigation and Comparison of Machine Learning Techniques in the Domain of Heart Disease.

[4] Prediction of heart disease using machine learning algorithms-Sonam Nikhar1, A.M. Karandikar2:.

[5] Evaluation of different machine learning techniques for prediction of heart diseases-dwivedi ashok kumar

[6] A review on heart diagnosis based on decision support systems: Saima Safdar1 · Saad Zafar1 · Naurin Frooq Khan · Nadeem Zafar2

[7] Prediction of heart diseaseusing hybrid technique for selecting features: Pahwa, Kanika, Ravinder Kumar

[8] Heart disease prediction using machine learning algorithms: Bommadevara, H. S. A., Sowmya, Y., & Pradeepini, G. (2019).

[9] Srinivas,V., Aditya, K., Prasanth, G., Babukarthik, R. G., Satheeshkumar, S., & Sambasivam, G. (2018). A novel approach for Heart disease prediction: Machine learning techniques.

[10] Analysis of single and data mining techniques for heart disease prediction using real time dataset: Yasin, S. A., & Prasad Rao, P. V. R. D.

[11] A survey on development of pattern evolving model for discovery of patterns in text mining using data mining techniques: Changala.R, D.Rajeswara rao.

[12] A computational intelligence method for effective diagnosis of heart disease using genetic algorithm: Siva kumar .p, D.Anand, V. Uday kumar, D. Bhattacharya

[13] K. SaiKrishnaSree, Narasing Rao.M.R: Diagnosis of heart disease using neural networks-comparative study of Bayesian regularization with multiple regression model.

[14] Chapman, P., Clinton, J., Kerber, R. Khabeza, T., Reinartz, T., Shearer, C.,Wirth, R.SPSS, 1-78, 2000:CRISP-DM 1.0: Step by step data mining guide.

[15] Heart disease using machine learning algorithms: Himanshu Sharma, M A Rizvi.

[16] K. Chandana, Y. Prashanth, J. Prabhudas: decision support system for predicting diabetic retinopathy using neural networks.

[17] Performance Prediction of Chronic Kidney Disease using various Data Mining Techniques,by Radha Mothukuri.International Journal of Advanced in Management,Technology and Engineering Sciences,2018,7(12),2249-7455.