HEART DISEASE PREDICTION USING MACHINE LEARNING TECHNIQUES

Krishna Kumar Yadav[1], Dr.Anurag Sharma[2], Dr.Abhishek Badholia[3]
[1,2] *Computer Science and Engineering, MATS University, Raipur,India*
[1] *kk.49681@gmail.com,*[2]*anusiraag@gmail.com*

Abstract: In few previous decades around the globe the reason for extensive number of deaths is cardiovascular disease or Heart related disease and not only in India but all around the world has emerged as a life-threatening disease. So for the correct treatment and in time diagnosis for this disease the need of feasible, accurate and reliable system is encountered. For automation of analysis of the sophisticated and huge data, to the various medical dataset of Machine Learning techniques and methods are applied. In recent times many researchers for the health care industry assistance with the help of various techniques of Machine Learning, this in turn helps the professionals in the procedure of the heart related disease diagnosis. A survey of various models that accepts such techniques and algorithms and their performance analysis is presented in this paper. Within the researchers few very fashionable Model supported supervised learning algorithms are Random forest (RF), Decision Tree (DT), Naïve Bayes, ensemble models, K-Nearest Neighbour (KNN) and Support Vector Machine (SVM).

Keyword: Ensemble Models, Random Forest; Decision Tree; Support Vector Machines; Machine learning, Naïve Bayes; K- Nearest Neighbour; Cardiovascular Diseases, Machine learning.

Introduction

With the use of the accessible data processing tools that are dissimilar for the different practices in data processing that are very important in the heart condition forecast are the main specialized content of this paper. The physical body opposite parts may get distress, if the guts does not work properly, opposite parts such as kidney, brain, etc. Functioning of guts might get effected by the disease of heart condition. One of the main reasons for death is heart condition in today's scenario. Nearly 12 million people per annum die because of the heart disease as anticipated by the WHO- World Health Organization [1]. Diseases like attack, cardiovascular, knock and coronary are few of the heart diseases. A heart condition that is caused because of the initiation of vital sign to be high or the blood vessels which passes across the brain, such blood vessels lessening, blocking or strengthening is known as Knock. In present situation facility superiority is one of the prime challenges faced by the Healthcare industry. The service standards are defined by the serving patient with the effective treatment and primarily correct diagnosis of the disease. Unacceptable consequences that are disastrous are caused by the poor diagnosis. Medical records of the data or records are extensively large, but these data or records are from various different foundations [2]. The important component of those data is physicians done interpretations. Inside the database directives to fill or complete the values that are omitted is needed by the data processing as the information in the world might be inconsistent, incomplete or noisy. As one of the important reasons if death in the ancient time was known as albeit cardiovascular diseases, due to these were the most manageable and avoidable diseases [3]. Disease well-time judgment on which the disease accurate and complete management is relayed on. There is an important need for identifying the patient with high-risk a methodological and accurate tool generation and heart infection timely analysis by mining data is required [4]. Heart condition different symptoms might be shown by the differential person bodies and can be accordingly varying. Stomach disorders and tininess of breath, pain, arms, shoulders pains, neck pain, jaw pain and back pain are included often. Stroke and arterial coronaries disease and coronary failure are included in the spread of different heart diseases [5]. Across the globe as a supreme chronic kind of disease is acknowledged as albeit heart condition and at the same time are also the most avoidable one. Two main origins of the heart condition director are timely analysis of the disease it is an inferior prevention or a life in a healthy way is the main prevention [6]. In the difficulties of heart condition early prevention and judgment a outstanding role is shown by executing steady check-ups that are inferior preventions. To support this significant issue various tests are included like exercise tolerance test, echocardiography, chest X-rays and angiography [7]. Nonetheless, accurate medical equipment availability is required and these tests are expensive. Storage and preparation of the patient's database large and honest record is created by the heart experts. From such a kind of the datasets, for valuable information mining excellent prospective is delivered by this [8]. In various patients to work out heart conditions risk factor a huge research is going no, various researchers are using different programs of knowledge mining

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

approaches and different statistical approaches [9]. For heart disease the count of risk factors is acknowledged by the statistical analysis by considering heart condition lack of exercise, obesity, training in family, total cholesterol and hypertension, diabetes, vital sign, age and smoking [10]. It is very important to have heart disease awareness to the patients that are near to have a permanent heart condition for there proper healthcare and prevention [11]. For identifying gut diseases various techniques of data processing that for assisting the physician or specialist are accessible is used by the researchers. Generally used procedures that are used are as following Naïve Bayes, k-nearest and decision tree. Other various classifications-based methods used are SVM (Support Vector Machine), straight Kernel self-organizing map, sequential minimal optimization and neural networks, kernel density and bagging algorithm [12]. The description of the methods that were used in the study are provided clearly jn the subsequent section.

II. Related Material and Method

A. Support Vector Machine

In both of the regression and classification employment of the Support Vector Machines (SVM) is done. Representation of the info point is done of the space and into group's categorization and hence, in the same group the points with similar properties fall in the SVM model [13].

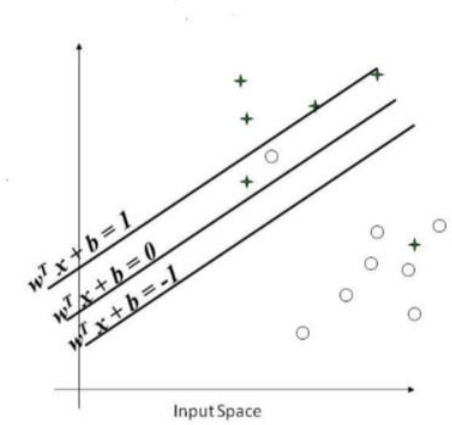


Figure 1: Representation of Support Vector Machine

In linear SVM, the provided data set as p-dimension vector is taken into the account, that by the maximum of p-1 planes will be separated also known as hyper-plane [14]. As in the figure 2 among the info groups for regression or classification problems, these planes set the boundaries or separate the info space. The amount of hyper plane on the thought of distance between the classes it separates, the selection of the simple hyper-plane is done. Namely maximum-margin hyper-plane is between 2 classes the plan that has the utmost margined14].

Definition of n data points:

(X_1,Y_1)……,(X_n,Y_n)……1

Here real vector is represented by X1 and 1 or -1 is Y1, the class to which X1 belongs is represented.

Construction of hyper-plane can be done so to the distance between to classes minimization y=1 and y=-1, is defined as follows -

W. X- b = 0 ……2

Here Normal vector is represented by W and offset of hyper-plane is b.

.

B.Radial Basis Function (RBF) Kernel Support Vector Machine

On the linear and non-linear data prove of the efficiency of the Support vector machine is shown. To classify nonlinear data implementation of the Radial base function is done[15].

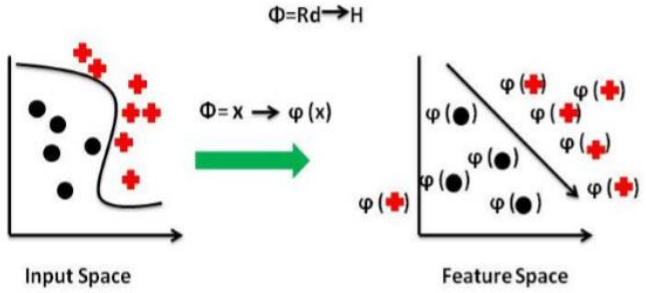


Figure 2: Representation of Radical Basis Function (RBF) Kernel support vector machine.

Kernel function plays very important role to put data into feature space. Mathematically, kernel trick (K) is defined as:

$$K(x_1, x_2) = exp\left(-\frac{|x_1-x_2|^2}{2\sigma^2}\right)$$ ……3

A Gaussian function is also known as Radial basis function (RBF) kernel. In Figure 3, the input space separated by feature map (Φ). By applying equation 1& 2 we get:

$$f(X) = \sum_i^N \alpha_i y_i k(X_i, X) + b$$ ……4

By applying equation 3 in 4 we get new function, where N represents the trained data.

$$f(X) = \sum_i^N \alpha_i y_i exp\left(-\frac{|x_1-x_2|^2}{2\sigma^2}\right) + b$$ ……5

C. k-Nearest Neigh bour (k-NN)

Yielding excellent results although it is a single algorithm by the k- Nearest neighbour. It's an instance-based, nonparametric and lazy learning algorithm [16]. In both the regression as well as the classification problem utilization of this problem is often done. To the seek out out the class k-NN which the belongs to the unlabeled object, in the classification. Setting a 'k' for this where k is number of neighbours to be considered that in general is odd and hence nearest to the object info points distance is calculated by the methods such as

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Minkowski distance, Manhattan distance or Hamming distance, Euclidean's distance [17]. The 'k' nearest neighbour after the space calculation are selected as the resultant class of new object that is calculated or analyzed on the thought of the votes of the neighbors. The result with the highest accuracy is predicted by the k-NN[18].

D. Artificial neural network (ANN)

The functionality of the human brain is mimicked by the artificial neural network. In generally as a set of nodes it is seen called as artificial neurons. At least to one another all of that node can transmit information. By some state 0 or 1 neurons are represented [19]. To every node assignment of some weight might be done to define within the system the importance or strength. Into the layers of multiple nodes; it reaches the output layer by the info travels from first layer input layer and after passing through middle layers hidden layers, into some important information the info is transformed by every layer and the specified output is given eventually [20]. In neurons functioning important role is played by activation and transfer function. All the weighted input is summed up by transfer function:

$$z = \sum_{x=1}^{n} w_i x_i + w_b b \quad \ldots\ldots\ldots\ldots 6$$

To specific range the output of the transfer function is flattened by the activation function. In can be linear as well as nonlinear. Simple activation function is:

$$f(z) = z \quad \ldots\ldots\ldots\ldots 7$$

As no limitations are applied on the data by the function, use of sigmoid function is done that can be expressed as follows:

$$a = \sigma(z) = \frac{1}{1+e^{-z}} \quad \ldots\ldots\ldots\ldots 8$$

E. Multifactor Dimensionality Reduction (MDR)

The approach for representation and location of the independent variables consolidation that will in some way influence the variable that are dependent is known as Multifactor Dimensionality Reduction (MDR). In general, it's designed to get the interaction in-between the variables that will affect the system output. It is not dependent on the sort of used model or parameters that as compared to the tradition system will make it better [21]. The two or more attributes are taking and combined into one attribute. The knowledge space representation is changed by this conversion. The category variable prediction by system performance improvement is lead by this. In Machine Learning utilization of the various MDR extensions is done. Few of them are as follows- covariates, risk scores, odds ratio, fuzzy methods and many more. In prediction of the heart related diseases or cardiovascular diseases Machine learning algorithm has a huge scope as frequently concluded by this [22]. In some cases, the above-mentioned methods performance extremely well where as in few cases the methods show very poor performance. It shows an extremely well performance when used together with the PCA and the Alternating decision tree whereas due to overfitting in other cases decision trees perform very poorly. As by solving the problem of overfitting performance of Ensemble model and Random Forest is alright as they employ multiple algorithms, just in case of Random Forest multiple Decision Trees are used [23]. The Naïve Bayes classifier supported models where show no time computation and along with this for many cases extremely well performance was shown by the SVM. In the guts related disease prediction Support Vector Machines shows very accurate results but there is still a huge scope for the research work to be done for the high dimensional data handling and overfitting. On the right ensemble of algorithms huge scope of research is possible to be used on a specific data sort [24].

III. PROPOSED METHODOLOGY

In this research paper a model "Optimized DNN using Talos" was deployed and to others the tactics were compared it was found that this was more efficient as compared to others. As compared to others it also provided high accuracy. Some steps were followed by us in this model execution. Selecting a mathematical representation was involved in Dimensional Reduction in a way such that bulk of can be related by one but not with all, hence only inclusion of vital information was done within the given data with variance [25]. It might consist tons of dimensions and attributes by the task the info that is considered, but equal influence on the output is not shown by all the attributes. Computational complexity might be affected by the number of features or attributes being outsized and may also lead to overfitting that gives poor results [26]. Hence as a vital step Dimensional Reduction can be seen, in general by two methods this can be achieved that are- have selection and Feature Extraction [27].

A. Feature Extraction during, springing from the first feature set is the replacement set of features. Involving a change of features in feature extraction. In general, non-reversible is the transformation done, as within the process few or many useful information are lost [28]. For feature extension, in and principal component analysis or PCA is used. With the linear transformation algorithm may be popularly used in Principal Component Analysis. It searches the direction to maximize the variance within the feature space and mutually orthogonal directions are also searched. The simplest reconstruction is provided by the worldwide algorithm[29].

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

B. Feature Selection during this, original features subset is selected. By Correlation based Feature Selection or CFS key features are chosen with Best First Search method Subset Evaluation is combined for back dimensionality scaling. To select the foremost significant feature employment of chi-square statistics test is done.

IV. Algorithms and Techniques Used.

In 1951 K – Nearest Neighbour, introduced by Hodges et al. In which for pattern classification a nonparametric technique that popularly was called as the K-Nearest Neighbour rule. One amongst the foremost elements is K-Nearest Neighbour technique but is also a very useful and effective technique for classification. It is generally used for classification purposes with no or very less priority knowledge and about the info makes no assumptions about the info distribution. The algorithm includes searching within the training set to info point, the k nearest data point, the typical value of the found data points thereto is assigning and in which the targets value is unavailable. Accuracy of 83.16% in 21 KNN when the adequately to 9 is the worth of k during using the technique of 10-cross validation. With the accuracy of 70.26% better than any other technique it performs in 22 KNN along with the Ant Colony Optimization and hence 0.526 is the error rate. An extremely good accuracy of 87.5% is obtained[30].

(i)K-NN:- Machine learning algorithm that is non- parametric might be a K-NN. It's a supervised learning algorithm. Form the input file it predicts the output.

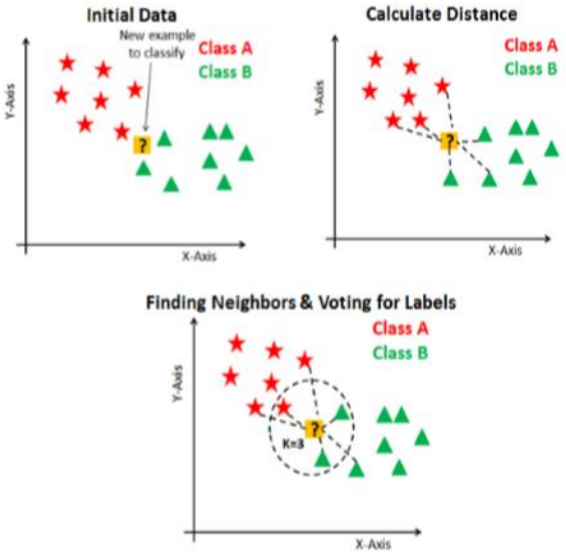


Fig.3- KNN predict the output from the input data

K-NN Algorithms: K-NN algo following few steps.

K-NN algo following few steps are:

1. Pick K a worth.
2. From all cases calculate the unknown case space.
3. In the training data choose the k-observations to the unknown datum that is nearest.
4. From the KNN using the foremost popular responds predicting the responds of the unknown datum.
5. Stop.

Into test and training dataset the splitting is done of this algorithm data. For model training and building the employment of the training dataset is done. In general the root of the amount of the observation is set as the value of k. On the model built based the test data [31].

(ii) Decision Tree might be of supervised learning algorithm is Decision tree. In the classification problem this system is generally utilized. With categorical and continuous attributes it effortlessly performs. Into two or more similar sets the population is divided population by this algorithm that supported the foremost significant predictors. Each and every attributes entropy is first calculated by the Decision tree algorithm. With maximum entropy or maximum information gained the dataset is spilt with the predictors or variables assistance. With the left out attributes these two steps are performed recursively[32].

With the 77.55% of accuracy is the worst performance of the decision tree but with the 82.17% accuracy it shows better performance when the decision tree is employed with the technique of boosting. With the 42.8954% a correctly specified percentage the poor performance of decision tree can be seen whereas it also uses dataset that is equivalent but for implementing decision tree uses the J48 algorithm and hence receives the accuracy of 67.7% that is a small amount but show an improvement on the previously accuracy obtained of 71.43%, it has used with principal component analysis alternating decision tree with the 92.2% accuracy. The simplest result on using the combination of with the forward selection and decision tree-based classifier that secures the weighted accuracy of the 78.4604% and this achieve by the Kamran Farooq et al.

(iii) Random Forest an additional supervised machine learning algorithm is known as Random Forest. For both the task of classification as well as the regression this system are used frequently but in the classification task it generally performs better. As the name suggests before giving an output multiple decision trees are considered in the Random forest techniques. In general it is decision tree's ensemble. On the assumption that there will be convergence to the proper decision by the greater numbers of trees, this is the prediction on which this system is based. Electoral system is used for categorization then it decides the class whereas in the regression for every choice trees it takes mean of all the output. With the high dimensionality with large datasets it works well.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

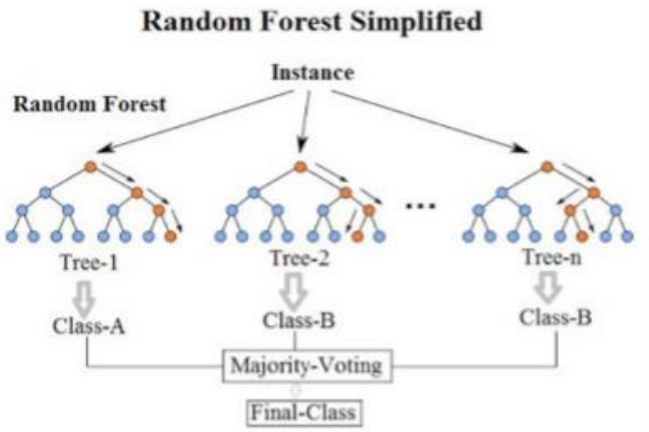


Fig.4- Random Forest

In figure 5, exceptionally well performance is shown by the random forest. A significant high accuracy of 91.6% is seen in the Random forest features of Cleveland dataset than any of the methods that are opposite. An accuracy of 97% is achieved in the People's Hospital dataset. A 0.86 f-measure is achieved in random forest method in 20. To predict coronary heart condition Random forest in 21 and it receives an accuracy of 97.7%.

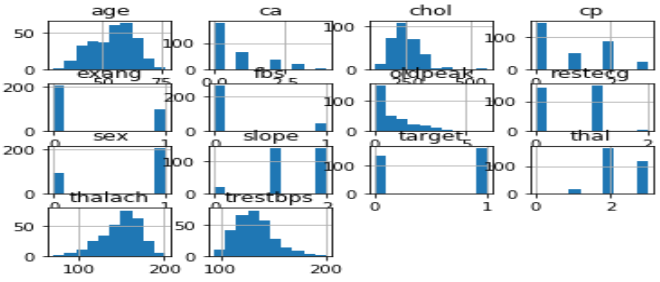
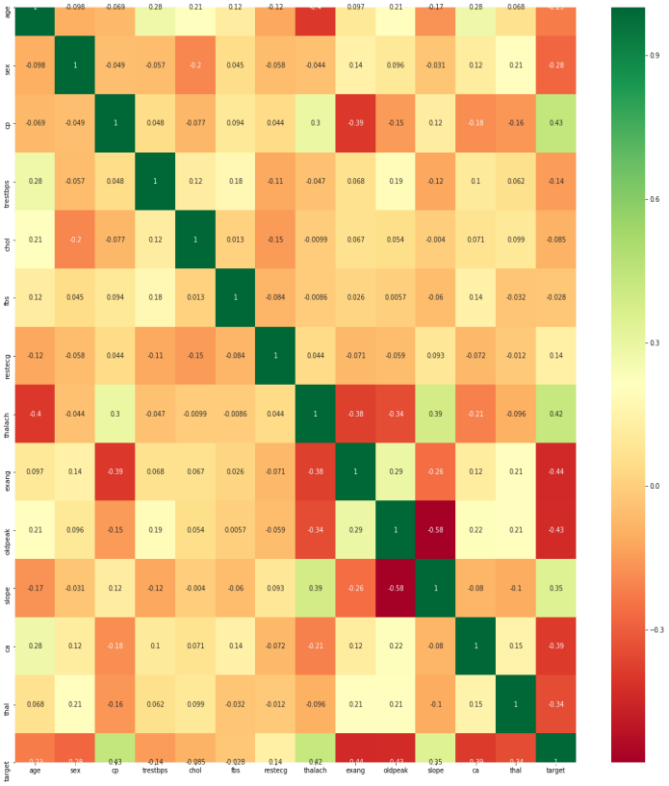


Fig.5: attribute of measure

(iv) Ensemble Model Two or more related yet different models of analysis are used in the ensemble model and the results produced are combined in one score in 22 used a ANN, KNN and SVM ensemble for the accuracy of 94.12%. Demonstration of the bulk vote based model that includes Support Vector Machines, Decision tree and Naïve Bayes classifiers giving 74% sensitive, 82% accuracy and for UCI heart condition dataset 93% of specifications. In 24, a Naïve Bayes classifier, SVM and Gini Index are consisted in this ensemble that in predicting syncope disease provides 98% accuracy.

Table.1 In ensemble modeling two or more related but different analytical models.

Elements	Age	Sex	Cp	Test Bps	Chol	Fbs	Restecg	Thalach	Exang	Oldpeak	Slope	CA	THAL	Target
Count	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000	303.000000
mean	54.366337	0.683168	0.966997	131.623762	246.264026	0.148515	0.528053	149.646865	0.326733	1.039604	1.399340	0.729373	2.313531	0.544554
std	9.082101	0.466011	0.466011	17.538143	51.830751	0.356198	0.525860	22.905161	0.469794	1.161075	0.616226	1.022606	0.612277	0.498835
min	29.000000	0.000000	0.000000	94.000000	126.000000	0.000000	0.000000	71.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000
25%	47.500000	0.000000	0.000000	120.000000	211.000000	0.000000	0.000000	133.500000	0.000000	0.000000	1.000000	0.000000	2.000000	0.000000
50%	55.000000	1.000000	1.000000	130.000000	240.000000	0.000000	1.000000	153.000000	0.000000	0.800000	1.000000	0.000000	2.000000	1.000000
75%	61.000000	1.000000	2.000000	140.000000	274.500000	0.000000	0.000000	166.000000	1.000000	1.600000	2.000000	1.000000	3.000000	1.000000
max	77.000000	1.000000	3.000000	200.000000	274.500000	1.000000	2.000000	202.000000	1.000000	6.200000	2.000000	4.000000	3.000000	1.000000



ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

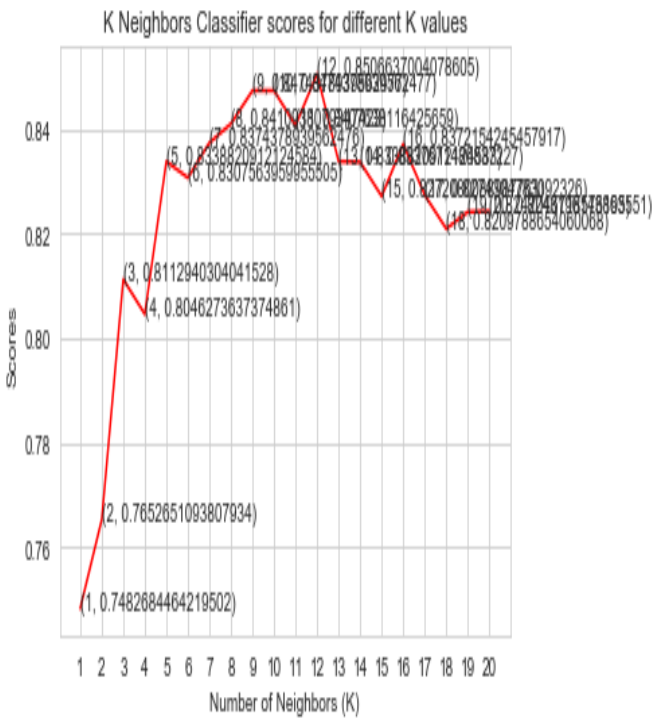


Fig. 6: K neighbors Classifier scores for different values

S.NO.	Classification Algorithms	Accuracy
1	Logistic Regression	83.25004050021001
2	SVM	81.78000529001220
3	Naïve Bayes	
4	Hyper-parameter optimization	84.16000511002407
5	K-NN	85.06637004078605
6	Random forest	83.41045606229145

V.RESULT:

Few classification algorithms are applied in this paper like Random forest, KNN, etc. On the dataset of heart disease and all accuracy of classification is measured in the table mentioned below. If the target classes are of equal size approximately it's a good practice to work with such dataset.

Table.2 Classification Algorithms and Accuracy.

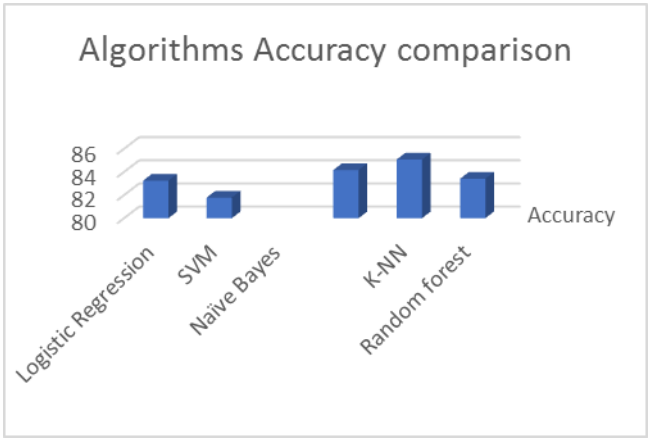


Fig. 7: comparative result Classifier scores for different values.

VI. CONCLUSION AND FUTURE WORK

In general, that follows your methodology section in a logical sequence is the summary of this paper. In comparison to the other optimization and algorithm, it for prediction provides good results. Using a Talos optimization we deploy machine learning. In DNN, a new optimization technique is Talos optimization. To other optimization Talos optimization provides better accuracy. To datasets of heart disease Talos optimization is applied and gives predictions that are good. With the help of Talos optimization.

References

[1]. Tarun Dhar Diwan, Dr.H.S.Hota, Dr. Siddhartha Choubey "Multifactor Authentication Methods: A Framework for Their Selection and Comparison" accepted for publication in International Journal of Future Generation Communication and Networking Vol. 13, No. 3, (2020), pp. 2522–2538, ISSN: 2233-7857 (Web of Science).

[2]. Tarun Dhar Diwan, Dr.H.S.Hota, Dr. Siddhartha Choubey "A Novel Hybrid Approach for Cyber Security in IoT network Using Deep Learning Techniques" accepted for publication in International Journal of Advanced Science and Technology ISSN:2394-5125, ISSN: 2005-4238 (Scopus indexed Journal).

[3]. Tarun Dhar Diwan, Dr.H.S.Hota, Dr. Siddhartha Choubey entitled "Development of Real Time Automated Security System for Internet of Things (IoT)" accepted for publication in International Journal of Advanced Science and Technology Vol. 29, No. 6s, (2020), pp. 4180 – 4195, ISSN: 2005-4238 (Scopus indexed Journal).

[4]. Tarun Dhar Diwan, Dr.H.S.Hota, Dr. Siddhartha Choubey "A Proposed Security Framework for Internet of Things: An Overview" presented in international Conference held on 20-22

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

December,2019, MTMI, Inc. USA in Collaboration with at amity Institute of Higher Education, Mauritius.

[5]. Tarun Dhar Diwan, Dr.H.S.Hota, Dr. Siddhartha Choubey "Control of Public Services for Public Safety through Cloud Computing Environment" presented in international Conference held on 04-05 January,2020, Organized by Atal Bihari Vajpayee University, Bilaspur in association with MTMI, USA and sponsored by CGCOST, Raipur (C.G), India.

[6]. Tarun Dhar Diwan, Dr.H.S.Hota, Dr. Siddhartha Choubey "A Study on Security and Data Privacy issues of IoT based Application in Modern Society" presented in international Conference held on 04-05 January,2020, Organized by Atal Bihari Vajpayee University, Bilaspur in association with MTMI, USA and sponsored by CGCOST, Raipur (C.G), India.

[7]. C.BeulahChristalinLatha,S.CarolinJeeva,Improvingtheaccuracyofpredictionofheartdiseaseriskbasedonensembleclassificationtechniques,https://doi.org/10.1016/j.imu.2019.100203, July 2019.

[8]. Avinash Golande, Pavan Kumar T, Heart Disease Prediction Using Effective Machine Learning Techniques, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019.

[9]. S.Nandhini, Monojit Debnath, Anurag Sharma, Pushkar. Heart Disease Prediction using Machine Learning, International Journal of Recent Engineering Research and Development (IJRERD) ISSN: 2455-8761, www.ijrerd.com Volume 03 – Issue 10,,PP. 39-46, October 2018.

[10]. A.Sahaya Arthy, G. Murugeshwari "A survey on heart disease prediction using data mining techniques" (April 2018).

[11]. Hamid Reza Marateb and Sobhan Goudarzi, "A noninvasive method for coronary artery diseases diagnosis using a clinically interpretable fuzzy rule-based system," Journal of Research in Medical Sciences, Vol. 20, Issue 3, pp.214-223, March 2015.

[12]. K.R. Lakshmi, M.Veera Krishna, and S.Prem Kumar, "Performance Comparison of Data Mining Techniques for Predicting of Heart Disease Survivability," International Journal of Scientific and Research Publications, Vol.3, Issue 6, pp.1-10, June 2013.

[13]. 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.

[14]. Purushottam, Prof. (Dr.) Kanak Saxena, Richa Sharma, "Efficient Heart Disease Prediction System", 2016, pp.962-969.

[15]. Ashwini Shetty A, Chandra Naik, "Different Data Mining Approaches for Predicting Heart Disease", International Journal of Innovative in Science Engineering and Technology, Vol.5, May 2016, pp.277-281.

[16]. 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.

[17]. V. Krishnaiah, G. Narasimha, N. Subhash Chandra, "Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review" IJCA 2016.

[18]. K.Sudhakar, Dr. M. Manimekalai "Study of Heart Disease Prediction using Data Mining", IJARCSSE 2016.

[19]. NagannaChetty, Kunwar Singh Vaisla, NagammaPatil, "An Improved Method for Disease Prediction using Fuzzy Approach", ACCE 2015.

[20]. VikasChaurasia, Saurabh Pal, "Early Prediction of Heart disease using Data mining Techniques", Caribbean journal of Science and Technology,2013

[21]. ShusakuTsumoto," Problems with Mining Medical Data", 0-7695- 0792-1 I00@ 2000 IEEE.

[22]. Y. Alp Aslandoganet. al.," Evidence Combination in Medical Data Mining", Proceedings of the international conference on Information Technology: Coding and Computing (ITCC'04) 0-7695-2108-8/04©2004 IEEE.

[23]. Carlos Ordonez, "Improving Heart Disease Prediction Using Constrained Association Rules," Seminar Presentation at University of Tokyo, 2004.

[24]. Franck Le Duff, CristianMunteanu, Marc Cuggiaa, Philippe Mabob, "Predicting Survival Causes After Out of Hospital Cardiac Arrest using Data Mining Method", Studies in health technology and informatics, Vol. 107, No. Pt 2, page no. 1256-1259, 2004.

[25]. Boleslaw Szymanski, Long Han, Mark Embrechts, Alexander Ross, KarstenSternickel,Lijuan Zhu, "Using Efficient Supanova Kernel For Heart Disease Diagnosis", Proc. ANNIE 06, intelligent engineering systems through artificial neural networks, vol. 16,page no. 305-310, 2006. 20. [26]. Kiyong Noh, HeonGyu Lee, Ho-Sun Shon, Bum Ju Lee, and Keun Ho Ryu, "Associative Classification

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731

Approach for Diagnosing Cardiovascular Disease", Springer 2006,Vol:345, page no. 721- 727.

[26]. Hongyu Lee, Ki Yong Noh, Keun Ho Ryu, “Mining Bio signal Data: Coronary Artery Disease Diagnosis using Linear and Nonlinear Features of HRV,” LNAI 4819: Emerging Technologies in Knowledge Discovery and Data Mining, May 2007, page no. 56-66.

[27]. Niti Guru, Anil Dahiya, NavinRajpal, "Decision Support System for Heart Disease Diagnosis Using Neural Network", Delhi Business Review, Vol. 8, No. 1, January - June 2007.

[28]. Hai Wang et. al.,”Medical Knowledge Acquisition through Data Mining”, Proceedings of 2008 IEEEInternational Symposium on IT in Medicine and Education 978-1-4244- 2511-2/08©2008 Crown.

[29]. SellappanPalaniappan, RafiahAwang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques", (IJCSNS), Vol.8 No.8, August 2008.

[30]. LathaParthiban and R.Subramanian, "Intelligent Heart Disease Prediction System using CANFIS and Genetic Algorithm", International Journal of Biological, Biomedical and Medical Sciences, Vol. 3,Page No. 3, 2008.

[31]. Chaitrali S. Dangare, Sulabha S. Apte, Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques”, International Journal of Computer Applications (0975888) Volume 47No.10, June 2012.

ISSN (Print): 2204-0595
ISSN (Online): 2203-1731