ISSN (Online) 2394-2320

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 7, Issue 10, October 2020

Heart Disease Prediction Using Machine Learning Techniques

[1] Krishna Kumar Yadav, [2]Dr.Anurag Sharma, [3]Dr.Abhishek Badholi
[1][2][3] Computer Science and Engineering, MATS University, Raipur,India

Abstract: Heart related diseases or Cardiovascular Diseases (CVDs) are the most reason for an enormous number of deaths within the world over the previous couple of decades and has emerged because the most life-threatening disease, not only in India but within the whole world. So, there's a requirement of reliable, accurate and feasible system to diagnose such diseases in time for correct treatment. Machine Learning algorithms and techniques are applied to varied medical datasets to automate the analysis of huge and sophisticated data. Many researchers, in recent times, are using several machine learning techniques to assist the health care industry and therefore the professionals within the diagnosis of heart related diseases. This paper presents a survey of varied models supported such algorithms and techniques and analyze their performance. Models supported supervised learning algorithms like Support Vector Machines (SVM), K-Nearest Neighbour (KNN), Naïve Bayes, Decision Trees (DT), Random Forest (RF) and ensemble models are found very fashionable among the researchers.

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

I. INTRODUCTION

The contents of this paper mainly specialize in various data processing practices that are valuable in heart condition forecast with the help of dissimilar data processing tools that are accessible. If the guts doesn't function properly, this may distress the opposite parts of the physical body like brain, kidney etc. heart condition may be a quite disease which effects the functioning of the guts . In today's era heart condition is that the primary reason for deaths. WHO-World Health Organization has anticipated that 12 million people die per annum due to heart diseases. Some heart diseases are cardiovascular, attack , coronary and knock. Knock may be a kind of heart condition that happens thanks to strengthening, blocking or lessening of blood vessels which drive through the brain or it also can be initiated by high vital sign [1]. the main challenge that the Healthcare industry faces now-a-days is superiority of facility. Diagnosing the disease correctly & providing effective treatment to patients will define the standard of service. Poor diagnosis causes disastrous consequences that aren't accepted. [2] Records or data of medical record is extremely large, but these are from many dissimilar foundations. The interpretations that are done by physicians are essential components of those data. the info in world could be noisy, incomplete and inconsistent, so data preprocessing is going to be required in directive to fill the omitted values within the database. albeit cardiovascular diseases are found because the important source of death in world in ancient years, these are announced because the most avoidable and manageable diseases. the entire and accurate management of a disease rest on the well-timed judgment of that disease. an accurate and methodical tool for recognizing high-risk patients and mining data for timely analysis of heart infection looks a significant want. Different person body can show different symptoms of heart condition which can vary accordingly. Though, they often include back pain, jaw pain, neck pain, stomach disorders and tininess of breath, pain, arms and shoulders pains. There is a spread of various heart diseases which incorporates coronary failure and stroke and arteria coronaria disease [3]. albeit heart condition is acknowledged because the supreme chronic kind of disease within the world, it are often most avoidable one also at an equivalent time. A healthy way of life (main prevention) and timely analysis (inferior prevention) are the 2 major origins of heart condition director. Conducting steady check-ups (inferior prevention) shows outstanding role within the judgment and early prevention of heart condition difficulties. Several tests comprising of angiography, chest X-rays, echocardiography and exercise tolerance test support to the present significant issue. Nevertheless, these tests are expensive and involve availability of accurate medical equipment. Heart experts create an honest and large record of patient's database and store them. It also delivers an excellent prospect for mining a valued knowledge from such kind of datasets [4]. there's huge

ISSN (Online) 2394-6849

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 7, Issue 10, October 2020

research happening to work out heart condition risk factors in several patients, different researchers are using various statistical approaches and various programs of knowledge mining approaches. Statistical analysis has acknowledged the count of risk factors for heart diseases counting smoking, age, vital sign , diabetes, total cholesterol, and hypertension, heart condition training in family, obesity and lack of exercise. For prevention and healthcare of patients who are close to have addicted of heart condition it's vital to possess awareness of heart diseases [5]. Researchers make use of several data processing techniques that are accessible to assist the specialists or physicians identify the guts disease. Commonly used procedures used are decision tree, k-nearest and Naïve Bayes. Other different classification-based techniques used are bagging algorithm, kernel density, sequential minimal optimization and neural networks, straight Kernel self-organizing map and SVM (Support Vector Machine). subsequent section clearly provides details of techniques that were utilized in the study[6].

II. Related Material and Method

A. Support Vector Machine

Support vector machine (SVM) is employed in both classification and regression. In SVM model, the info points are represented on the space and are categorized into groups and therefore the points with similar properties falls in same group.

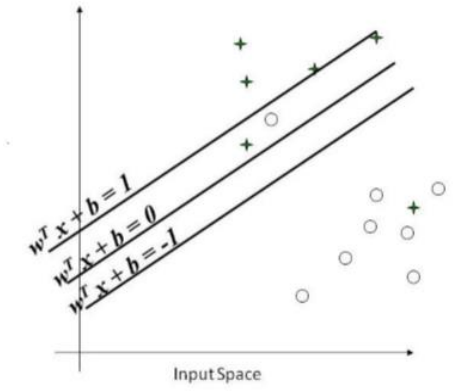


Figure 1: Representation of Support Vector Machine

In linear SVM the given data set is taken into account as p-dimensional vector which will be separated by maximum of p-1 planes called hyper-planes. These planes separate the info space or set the boundaries among the info groups for classification or regression problems as in Figure 2. the simplest hyper-plane are often selected among the amount of hyper-planes on the idea of distance between the 2 classes it separates. The plane that has the utmost margin between the 2 classes is named the maximum-margin hyper-plane[7].

For n data points is defined as:

$(X_1,Y_1)\ldots\ldots,(X_n,Y_n)\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 1$

Where X_1 is real vector and Y_1 can be 1 or -1, representing the class to which X_1 belongs.

A hyper-plane can be constructed so as to maximize the distance between the two classes y=1 and y=-1, is defined as:

$W.\ X - b = 0\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2$

Where W is normal vector and b is offset of hyper-plane along

.

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

Support vector machine has proven its efficiency on linear data and nonlinear data. Radial base function has been implemented with this algorithm to classify nonlinear data

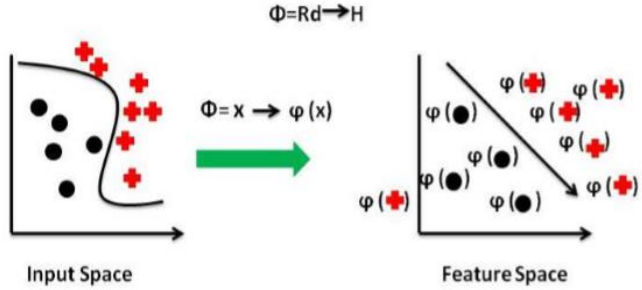


Figure 2: Representation of Radical Basis Function (RBF) Kernel support vector machine.

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

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

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:

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

By applying equation 3 in 4 we get new function, where

ISSN (Online) 2394-6849

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 7, Issue 10, October 2020

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)

k- Nearest neighbour may be a simple algorithm but yields excellent results. it's a lazy, nonparametric and instance-based learning algorithm. This algorithm are often utilized in both classification and regression problems. In classification, k-NN is applied to seek out out the category, to which new unlabeled object belongs [8]. For this, a 'k' is set (where k is number of neigh bours to be considered) which is usually odd and therefore the distance between the info points that are nearest to the objects is calculated by the ways like Euclidean's distance, Hamming distance, Manhattan distance or Minkowski distance. After calculating the space, 'k' nearest neighbours are selected the resultant class of the new object is calculated on the idea of the votes of the neighbours. The k-NN predicts the result with high accuracy [9].

D. Artificial neural network (ANN)

Artificial neural network mimics the functionality of human brain. It is often seen as a set of nodes called artificial neurons. All of those nodes can transmit information to at least one another. The neurons are often represented by some state (0 or 1) and every node can also have some weight assigned to them that defines its strength or importance within the system. The structure of ANN is split into layers of multiple nodes; the info travels from first layer (input layer) and after passing through middle layers (hidden layers) it reaches the output layer, every layer transforms the info into some relevant information and eventually gives the specified output [10]. Transfer and activation functions play important role in functioning of neurons. The transfer function sums up all the weighted inputs as: [10].

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

……6

Where b is bias value, which is usually 1.

The activation function basically flattens the output of the transfer function to a specific range. It could be either linear or non linear. The simple activation function is:

$$f(z) = z$$

……7

Since this function does not provide any limits to the data, sigmoid function is used which can be expressed as:

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

……8

E. Multifactor Dimensionality Reduction (MDR)

Multifactor dimensionality reduction is an approach for locating and representing the consolidation of independent variables which will somehow influence the dependent variables. it's basically designed to seek out the interactions between the variables which will affect the output of the system [11]. It doesn't depend upon parameters or the sort of model getting used, which makes it better than the opposite traditional systems. It takes two or more attributes and converts it into one. This conversion changes the space representation of knowledge This leads to improvement of the performance of system in predicting the category variable. Several extensions of MDR are utilized in machine learning [12]. a number of them are fuzzy methods, odds ratio, risk scores, covariates and far more. Based on the above review, it are often concluded that there's an enormous scope for machine learning algorithms in predicting cardiovascular diseases or heart related diseases. Each of the above-mentioned algorithms have performed extremely well in some cases but poorly in another cases [13]. Alternating decision trees when used with PCA, have performed extremely well but decision trees have performed very poorly in other cases which might be thanks to overfitting. Random Forest and Ensemble models have performed alright because they solve the matter of overfitting by employing multiple algorithms (multiple Decision Trees just in case of Random Forest). Models supported Naïve Bayes classifier were computationally in no time and have also SVM performed extremely well for many of the cases. Systems supported machine learning algorithms and techniques are very accurate in predicting the guts related diseases but still there's tons scope of research to be done on the way to handle high dimensional data and overfitting. tons of research also can be done on the right ensemble of algorithms to use for a specific sort of data[14].

III. PROPOSED METHODOLOGY

during this paper, we deploy a model "Optimized DNN using Talos" and compare the tactic to others it's more

ISSN (Online) 2394-6849

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 7, Issue 10, October 2020

efficient to others. This model provided a high accuracy compared to others. during this model, we are following some steps. Dimensionality Reduction involves selecting a mathematical representation such one can relate the bulk of, but not all, the variance within the given data, thereby including only most vital information [15]. the info considered for a task or a drag, may consists of tons of attributes or dimensions, but not all of those attributes may equally influence the output. an outsized number of attributes, or features, may affect the computational complexity and should even cause overfitting which results in poor results. Thus, Dimensionality Reduction may be a vital step considered while building any model. Dimensionality Reduction is usually achieved by two methods -Feature Extraction and have Selection [16].
A. Feature Extraction **during**
this, a replacement set of features springs from the first feature set. Feature extraction involves a change of the features. This transformation is usually not reversible as few, or even many, useful information is lost within the process [17]. In and Principal Component Analysis (PCA)is used for feature extraction. Principal Component Analysis may be a popularly used linear transformation algorithm. within the feature space, it finds the directions that maximize variance and finds directions that are mutually orthogonal. it's a worldwide algorithm that provides the simplest reconstruction [18].
B. Feature Selection during this, a subset of original feature set is chosen. In, key features are selected by CFS (Correlation based Feature Selection) Subset Evaluation combined with Best First Search method to scale back dimensionality. In [19] chi-square statistics test is employed to pick the foremost significant features.
IV. Algorithms and Techniques Used.
K – Nearest Neighbour in 1951, Hodges et al. introduced a nonparametric technique for pattern classification which is popularly known the K-Nearest Neighbour rule [20]. K-Nearest Neighbour technique is one among the foremost elementary but very effective classification techniques. It makes no assumptions about the info and is usually be used for classification tasks when there's very less or no prior knowledge about the info distribution. This algorithm involves finding the k nearest data points within the training set to the info point that a target value is unavailable and assigning the typical value of the found data points thereto. In [21] KNN gives an accuracy of 83.16% when the worth of k is adequate to 9 while using 10-cross validation technique. In [22] KNN with Ant Colony Optimization performs better than other techniques with an accuracy of 70.26% and therefore the error rates are 0. 526.. have obtained an efficiency of 87.5% [23], which is extremely good.
(i) K-NN:-KNN may be a non- parametric machine learning algorithm. it's a supervised learning algorithm. It means to predict the output from the input file.

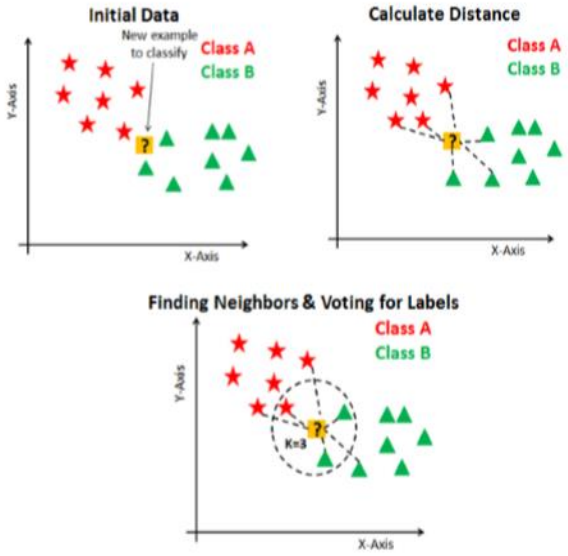


Figure 3.- KNN predict the output from the input data

K-NN Algorithms: - K-NN algo following some steps.
K-NN Algorithms: - K-NN algo following some steps.
1. Pick a worth for K.
2. Calculate the space of unknown case from all cases.
3. Select the K- observations within the training data that nearest to the unknown datum.
4. Predict the response of unknown datum using the foremost popular response value from the KNN
5. Stop.
In this algorithms data is split into training and test data sets. The training dataset is employed for model building and training. K- value is set which is usually the root of the amount of observations. Now the test data is based on the model built [23].
(ii) Decision Tree Decision tree may be a of supervised learning algorithm. this system is usually utilized in classification problems. It performs effortlessly with continuous and categorical attributes. This algorithm divides the population into two or more similar sets supported the foremost significant predictors [24].

ISSN (Online) 2394-6849

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 7, Issue 10, October 2020

Decision Tree algorithm, first calculates the entropy of every and each attribute. Then the dataset is split with the assistance of the variables or predictors with maximum information gain or minimum entropy [25]. These two steps are performed recursively with the remaining attributes.

In [10] decision tree has the worst performance with an accuracy of 77.55% but when decision tree is employed with boosting technique it performs better with an accuracy of 82.17%.In [9] decision tree performs

$$Entropy(S) = \sum_{i=1} -p_i \log_2 p_i$$

$$Gain(S, A) = Entropy(S) - \sum_{v \in Values(A)} \frac{|S_v|}{|S|} Entropy(S_v)$$

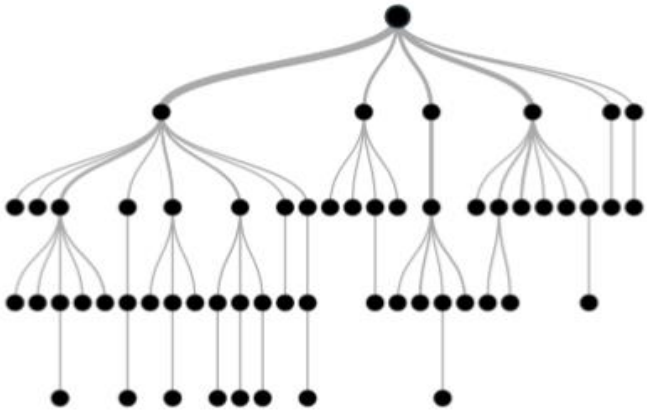


Figure. 4: Decision Tree

very poorly with a correctly classified instance percentage of 42.8954%whereas in [16] also uses an equivalent dataset but used the J48 algorithm for implementing Decision Trees and therefore the accuracy thus obtained is 67.7% which is a smaller amount but still an improvement on the previous obtained an accuracy of 71.43% [17] have used alternating decision trees with principle component analysis to get an accuracy 92.2%[18].Kamran Farooq et al. have achieved the simplest results on using decision tree-based classifier combined with forward selection which achieves a weighted accuracy of 78.4604% [19].

(iii) Random Forest Random Forest is additionally a popularly supervised machine learning algorithm. this system are often used for both regression and classification tasks but generally performs better in classification tasks. because the name suggests, Random Forest technique considers multiple decision trees before giving an output. So, it's basically an ensemble of decision trees. this system is predicated on the assumption that a greater number of trees would converge to the proper decision. For classification, it uses an electoral system then decides the category whereas in regression it takes the mean of all the outputs of every of the choice trees. It works well with large datasets with high dimensionality. In [5], random forest performs exceptionally well. In Cleveland dataset, random forest features a significantly higher accuracy of 91.6% than all the opposite methods.

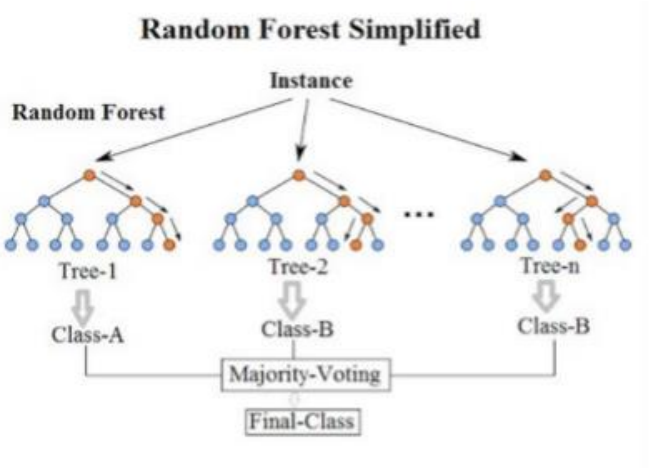


Figure 5: Random Forest

In People's Hospital dataset, it achieves an accuracy of 97%. In [20] random forest has achieved an f-measure of 0.86. In [21], random forest is employed to predict coronary heart condition and it obtains an accuracy of 97.7%.

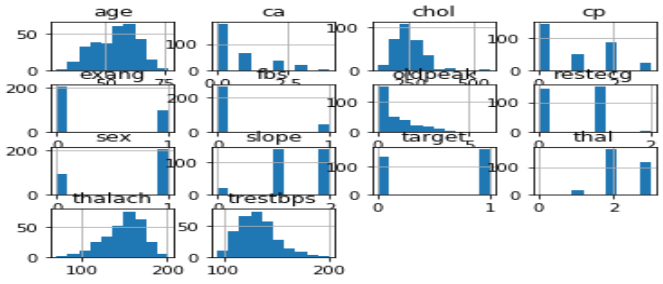


Fig. 6: attribute of measure

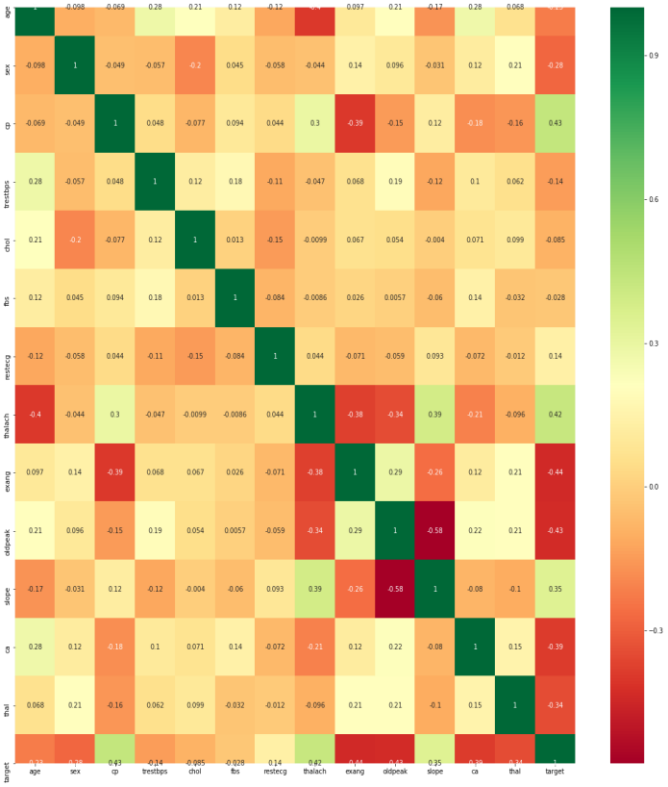
ISSN (Online) 2394-6849

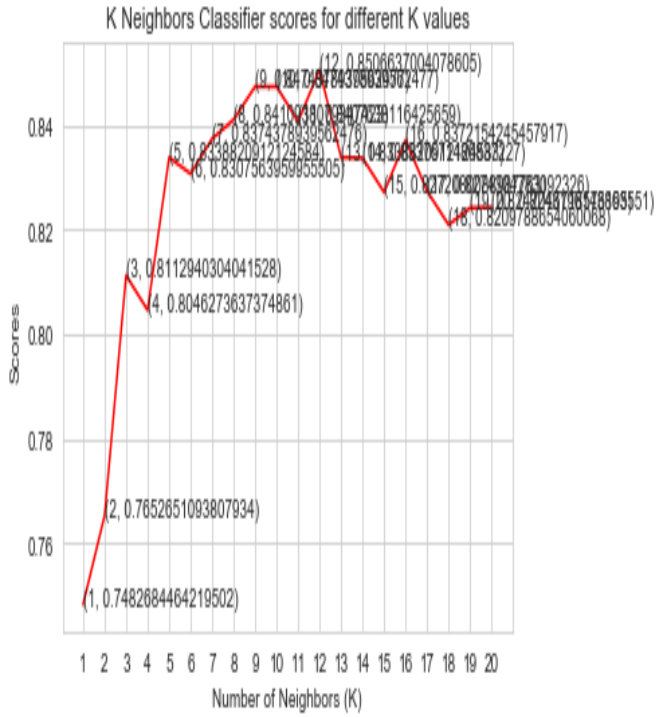
International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 7, Issue 10, October 2020

(iv) Ensemble Model In ensemble modeling two or more related but different analytical models are used and produce their results are combined into one score [22] have used an ensemble of SVM, KNN and ANN to realize an accuracy of 94.12%. the bulk vote-based model as demonstrated [23] which comprises of Naïve Bayes, Decision Tree and Support Vector Machine classifiers, gave an accuracy of 82%, sensitivity of 74% and specificity of 93% for UCI heart condition dataset. In [24] an ensemble model, consisting of Gini Index, SVM and Naïve Bayes classifiers, has been proposed which gave an accuracy of 98% in predicting Syncope disease.

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 (Online) 2394-6849

International Journal of Engineering Research in Computer Science and Engineering (IJERCSE)
Vol 7, Issue 10, October 2020

Fig. 6: K neighbors Classifier scores for different values

V.RESULT:

In this paper we applied some classification algorithms (like – K-NN, Random forest) on Heart diseases data set and measure the all classification accuracy is available on below
mention table.

Table.2 Classification Algorithms and Accuracy.

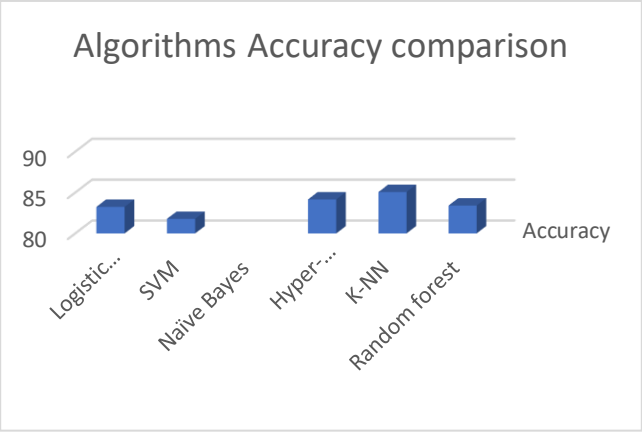


Fig. 7: comparative result Classifier scores for different values.

It's always a good practice to work with a dataset where

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

the target classes are of approximately equal size

VI. CONCLUSION AND FUTURE WORK

A summary of this paper arranged in a logical sequence that generally follows your methodology section. Compare to other algorithms and optimization, it is proved good results for prediction. In this paper, we deploy a Machine learning using Talos optimization. Talos optimization is newly optimization techniques in DNN. Talos provide better accuracy to other optimizations. It is applied on the Heart disease datasets and find out the good prediction. Using the Talos optimization.

REFERENCES

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

2. 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.

3. 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.

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

5. 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.

6. 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.

7. 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.

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

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

IJCA 2016.
12.]K.Sudhakar, Dr. M. Manimekalai "Study of Heart Disease Prediction using Data Mining", IJARCSSE 2016.
13. NagannaChetty, Kunwar Singh Vaisla, NagammaPatil, "An Improved Method for Disease Prediction using Fuzzy Approach", ACCE 2015.
14. VikasChaurasia, Saurabh Pal, "Early Prediction of Heart disease using Data mining Techniques", Caribbean journal of Science and Technology,2013
15. ShusakuTsumoto," Problems with Mining Medical Data", 0-7695- 0792-1 I00@ 2000 IEEE.
16. 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.
17. Carlos Ordonez, "Improving Heart Disease Prediction Using Constrained Association Rules," Seminar Presentation at University of Tokyo, 2004.
18. 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.
19. 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. Kiyong Noh, HeonGyu Lee, Ho-Sun Shon, Bum Ju Lee, and Keun Ho Ryu, "Associative Classification Approach for Diagnosing Cardiovascular Disease", Springer 2006,Vol:345, page no. 721- 727.
21. Hongyu Lee, Ki Yong Noh, Keun Ho Ryu, "MiningBiosignal 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.
22. 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.
23. 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.
24. SellappanPalaniappan, RafiahAwang, "Intelligent Heart Disease Prediction System Using Data Mining Techniques", (IJCSNS), Vol.8 No.8, August 2008.
25. 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. 16. Chaitrali S. Dangare, Sulabha S. Apte, Improved Study of Heart Disease Prediction System using Data Mining Classification Techniques", International Journal of Computer Applications (0975 888)Volume 47No.10, June 2012.