Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 44 No. 8
August 2023

Distance Based Outlier Detection (DOD) with Gradient Tree Boosting Classifier Algorithm for Heart Attack Prediction

A. Feza Naaz [1] , Dr. M. Anand [2] , Dr. T. Godhavari [3] , Dr. N. Gomathi [4]

[1] Research Scholar / ECE Dept., Dr.M.G.R. Educational & Research Institute, Chennai, Tamil Nadu.

[2,3] Professor / ECE Dept., Dr.M.G.R. Educational & Research Institute, Chennai, Tamil Nadu.

[4] Professor / CSE Dept., Vel Tech Rangarajan Dr.Sagunithala R & D Institute of Science and Technology, Chennai, Tamil Nadu.

Abstract — This research discusses the evaluation and development of OD and classification model depends on recognizing the outliers in clinical care and classifying it. Outliers are the activities which are uncommon and might denote errors in the patient- management. The benefits of this methods are (i) to model a detection system, it doesnot need expert input (ii) empirically, the relative medical outliers are extracted by employing a huge set of history of patient cases and updated continuously to denote common practice patterns and iii) coverage of alert might be deep and extensive. For positive and extensive impact over medical care, this novel method comprises major potential. In this research, a novel technique named DOD is introduced for obtaining the patient's care only for Heart Attack, which has all patient-management performance that is relied on the state of patient. It leads to analyze the patient-management operation for given data that is more abnormal and similar to predefined patients. Once the difference has been noted, it is employed in classifying the patient details that tends to find Heart Attack. Here, GBTC method is applied for classification process. In order to guarantee the accuracy of projected DOD-GBTC method, a set of 2 standard dataset such as Heart Attack Statlog as well as Cleveland dataset on the basis of various computing metrics.

Key Words: heart attack prediction, DOD-GBTC, outlier detection, deep learning, artificial intelligence, machine learning, ECG, arrhythmia

I. Introduction

The advanced methods in computing and communication technologies have enabled the healthcare sector to gather and save regular patient records which assists to make medical decisions. The saved medical information can be investigated to make the needed medical decisions that might be forecast, analysis, image examination, and line of treatment. Presently, various ML techniques have been commonly employed to classify and predict attacks. In this research, we made a review to study the existing ML models for Heart attack prediction. Besides, a review of CDSS takes place along with the survey of Outlier Detection (OD) based Heart Attack predictionmodels. A detailed comparative analysis is also made to identify the characteristics of the reviewed prediction models.

II. Proposed DOD-GBTC Method

The overall process involved in the presented DOD-GBTC method is shown in Figure 1. Initially, the medical dataset is processed by converting the format and replacing the missing values. Then, the DOD method using kNN gets executed to remove the outliers. Next, GBTC model is applied to classify the provided medical data. Finally, the performance validation takes place by the use of different classification measures.

a. DOD Technique Using kNN Graph

We describe kNN graph as weighted directed graph, in that each vertex denotes a single vector, and the edge correspond to adjacent vector pointers. Towards k-nearest vectors, each vertex comprises of k edges exactly in order to provided distance function. The distance among vectors v_i and v_j is the edge weight that is denoted through edge e_i. In computational geometry, the problem of producing kNN graph is represented as entire kNN problem. Through wide search, the graph might be built assuming entire pair wise distances at $O(N)^2$time cost. In O(kN+NlogN) time, the entire nearest neighbour problem might be resolved. For resolving the problem of clustering, the kNN graph might be employed.

The associated elements form the cluster within data and associated elements with one vertex is described as outlier. An outlier is near to inliers which might be misclassified with particular problem definition.

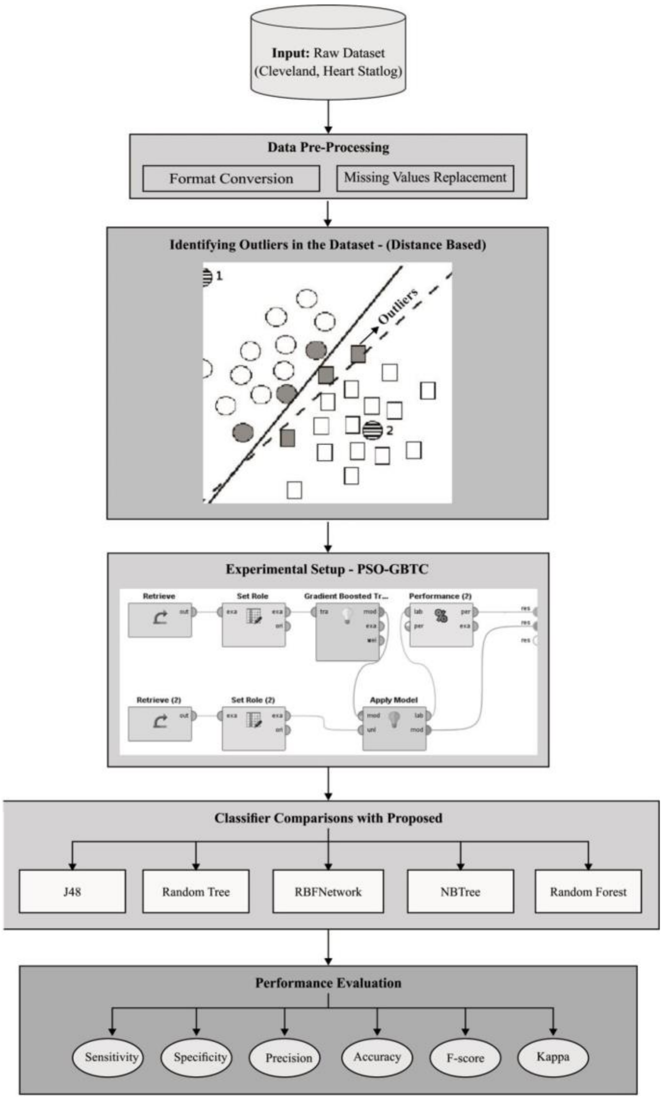


Figure 1 : Overall Process of the DOD-GBTC Method

b. GBTC Classifier

Boosting is presented through defining a technique for "transforming weak learning technique into one which attains high accuracy randomly". In “AdaBoost” technique, this scheme is used to solve the problems of classification. Using weak learners to re-weighted training data versions, boosting is used. The misclassified instance comprises of increased weights after every boosting iterations and precisely classified instances with decreased weights. Therefore, every subsequent classifier aims over instances which had been difficult to prior step classification. The weak classifiers series predictions are integrated through weighted majority vote into last prediction after the iteration count MGB. A variation of stochastic gradient boosting is presented. We sample without replacement of a training data subset over every iteration to suit the base learner. Various methods are used and chose at random mGBTC features from p features over each division.

Computational efficiency is enhanced using sub sampling process and commonly enhances performance and tree de-correlation. It employs H2O’s AdaBoost implementation, using shallow DTs as weak learners. We comprise four attributes to set: boosting iterations MGBTC, tree counts, learning rate λGBTC, tree depth JGBTC, feature subset to use at every split that is mGBTC. When MGBTC is high, boosting might over fit potentially, hence the iteration counts are fixed as 100 as highly conservative rates when comparing with offered examples in some surveys. Boosting depends on weak learners that are shallow trees which result commonly in huge performance. We fix JGBTC value as three as it only division enable for no parameter interaction effects for two-way interactions. The tree counts and learning rate are inversely proportion to the applied static error rates. It sets for upper spectrum end and we fix at 0.1 as λGBTC by considering low tree counts. It employs 15 which is the half part of feature space available for mGBTC.

The common idea of gradient boosting DT is merging a weak base classifiers series into strong one. Through the variation from conventional boosting technique which weight negative and positive samples, through following negative gradient direction, GBTC creates the global technique convergence.

III. Result Analysis

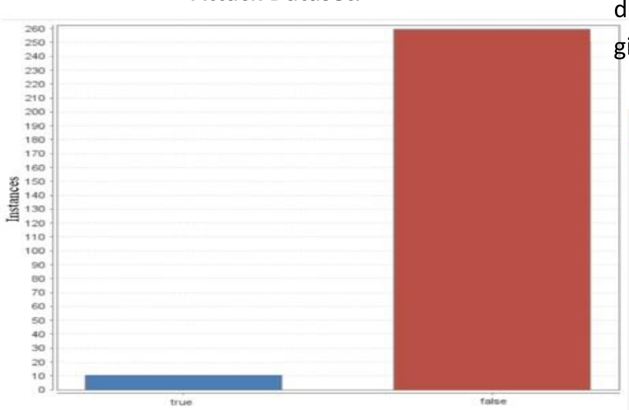
Various analysis were carried out in order to proposed DOD-GBTC for the Heart Attack prediction and the outcomes are discussed below. DOD is used to detect the outliers and the instances are subjected to classification further using GBTC. The confusion matrix, performance evaluation and comparison with recently proposed methods are provided in the following sections.

a. Outlier Detection Results For Heart Statlog Dataset

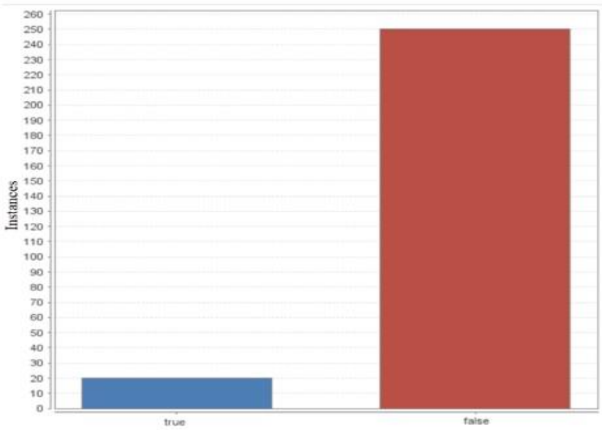
For various levels of Outliers Detection (δ), the confusion matrix for Heart Statlog Attack Dataset is shown in Table 1. By using the values gained using confusion matrix, classifier performances are computed. For δ=10, the number of 107 instances are under present category whereas it offers 140 instances under absent category. For δ=20, the number of 104 instances are under present category whereas it offers 137 instances under absent category. For δ=30, the number of 102 instances are under present category whereas it offers 129 instances under absent category.

Experts	δ=10		δ=20		δ=30	
	Present	Absent	Present	Absent	Present	Absent
Present	107	8	104	6	102	1
Absent	5	140	3	137	8	129

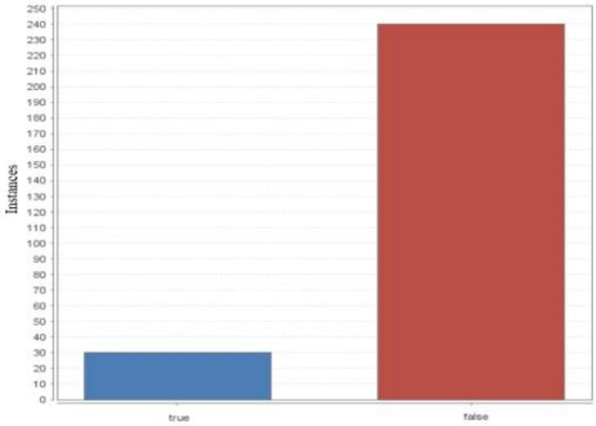
Table 1 : Confusion Matrix of Different Levels Of Outliers Detection δ=10, 20, 30 on Heart Statlog Attack Dataset.



δ=10



δ=20



δ=30

Figure 2: 5 OD With Varying Delta (δ)

Figure 2 and Table 2 shows that the OD with varying delta (δ) which was detected through by the DOD method. For δ=10, 10 are detected as outlier whereas 262 is detected as false. For δ=20, 20 are detected as outlier whereas 250 is detected as false. For δ=30, 30 are detected as outlier whereas 240 is detected as false. The representations of outliers are given in Figure 3.

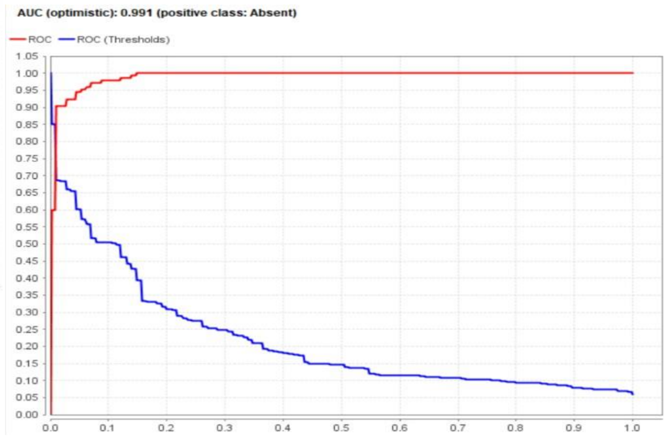


Figure 3: AUC of Outliers δ=10

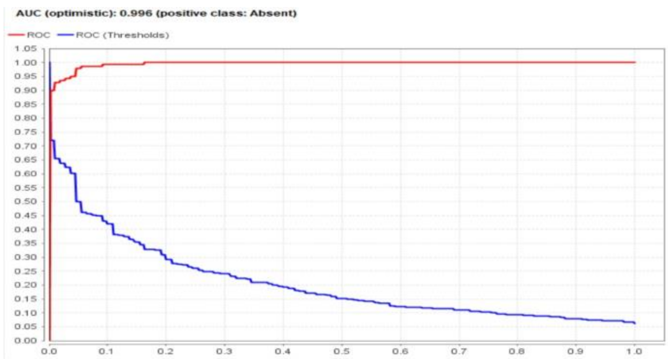


Figure 4 : AUC of Outliers δ=20

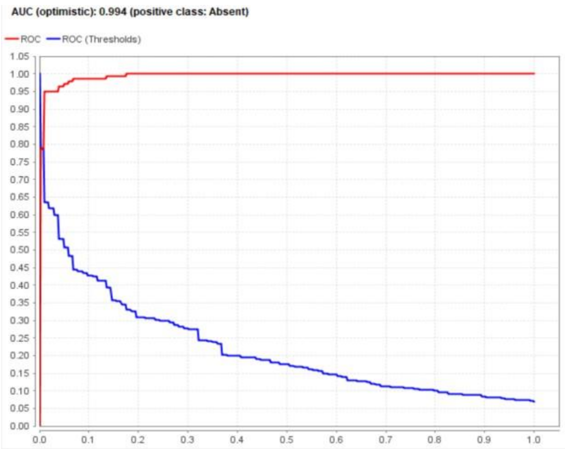


Figure 5: AUC of Outliers δ=30

Levels of Outliers	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)	F-Score (%)	Kappa (%)
δ=10	95.53	94.59	93.04	95.00	94.27	89.84
δ=20	919	95.80	94.54	96.40	95.85	92.67
δ=30	92.72	99.23	99.02	96.25	95.77	92.41

Table 2: Performance Evaluation of Different Levels of Outliers Detectionδ=10, 20, 30 on GBTC model on Heart Statlog Attack Dataset.

Figure. 3-5 show the AUC curve of outliers with varying delta (δ). For δ=10, the attained AUC curve for the projected DOD-GBTC method is 0.991. For δ=20, the attained AUC curve for the projected DOD-GBTC method is 0.996.For δ=30, the attained AUC curve for the projected DOD-GBTC method is 0.994. Among all the compared δ values, superior AUC outcomes is demonstrated through δ=20.

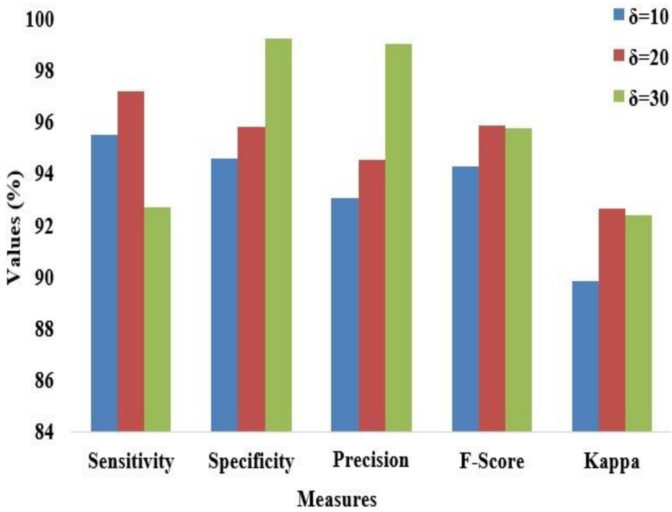
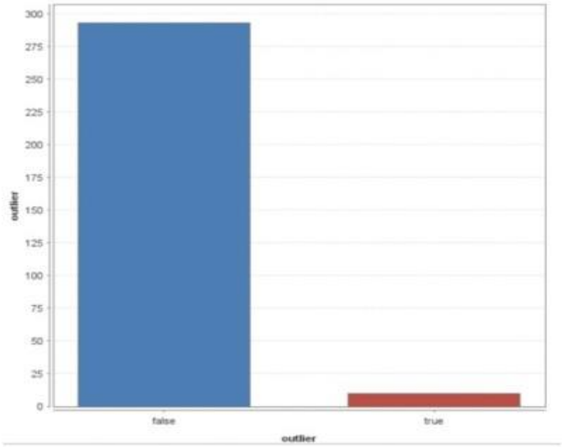


Figure 6 : Comparison of various measures on different levels of outliers on HeartStatlog Attack Dataset

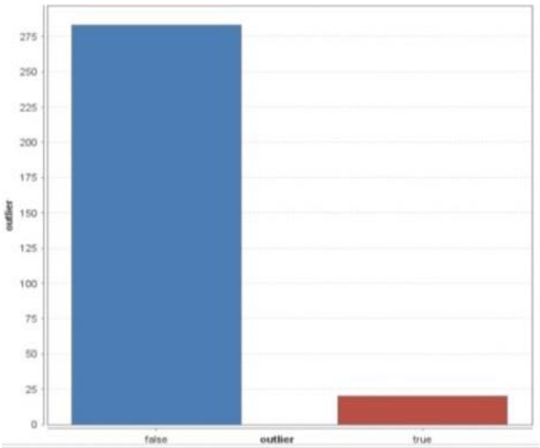
Figure 6 shows the performance evaluation of different levels of Outliers Detection δ=10, 20, 30 on GBTC model on Heart Statlog Attack dataset. The performance is measured by specificity, F-Score, sensitivity, precision, accuracy and kappa. For sensitivity, higher rate of 919% is achieved when δ=20. The high specificity rate is attained while δ=30. From the different outlier levels, the highest precision rate is attained when δ=30. The highest accuracy rate of 96.40% is achieved while δ=20. The highest F-Score rate of 95.85% is achieved while outlier level is 20. The highest kappa rate of 92.67% is achieved while outlier level is 20.

b. Outlier Detection Results for Cleveland Heart Attack Dataset

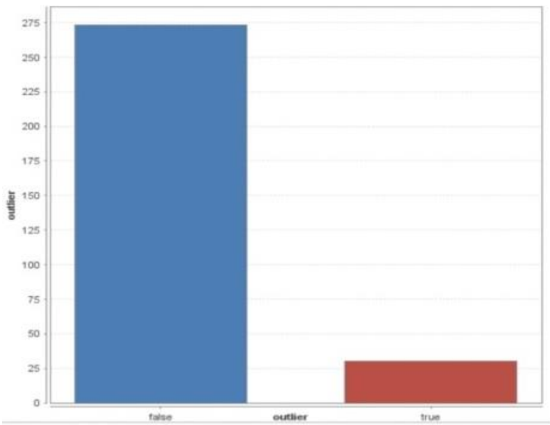
For various levels of Outliers Detection (δ), the confusion matrix for Cleveland Heart Attack Dataset is shown in Table 3. By using the values gained using confusion matrix, classifier performances are computed.



δ=10



δ=20



δ=30

Figure 7 : OD with varying delta (δ)

Through the DOD scheme, Figure 7 demonstrates the OD with variation in δ. For δ=10, 10 are detected as outlier whereas 290 is detected as false. For δ=20, 20 are detected as outlier whereas 280 is detected as false. For δ=30, 30 are detected as outlier whereas 273 is detected as false.

Experts	δ=10		δ=20		δ=30	
	Present	Absent	Present	Absent	Present	Absent
Present	153	6	148	7	144	7
Absent	6	128	3	125	2	120

Table 3 : Confusion Matrix of Different Levels of Outliers Detectionδ=10, 20, 30 on Cleveland Heart Attack Dataset

Figure 8-10 show the AUC curve of outliers with varying delta (δ). For δ=10, the attained AUC curve for the projected DOD-GBTC method is 0.992. For δ=20, the attained AUC curve for the projected DOD-GBTC method is 0.994. For δ=30, the attained AUC curve for the projected DOD-GBTC method is 0.993. Among all the compared δ values, superior AUC outcomes are demonstrated through δ=20.

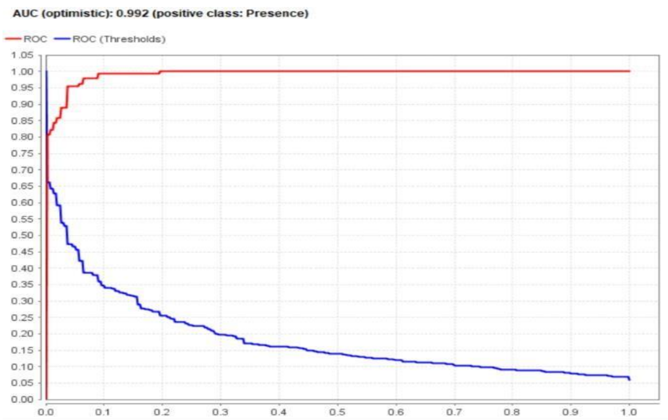


Figure 8: AUC of Outliers δ=10

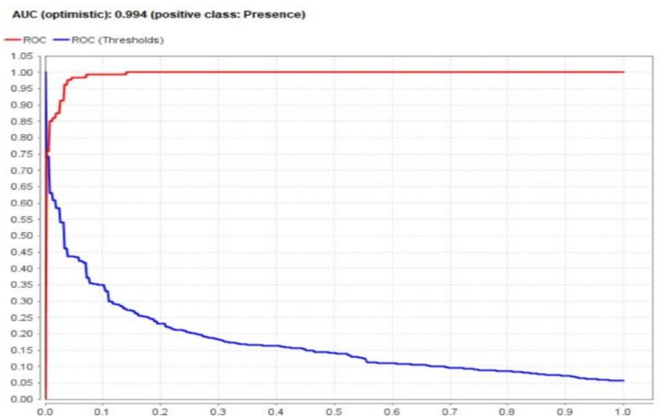


Figure 9: AUC of Outliers δ=20

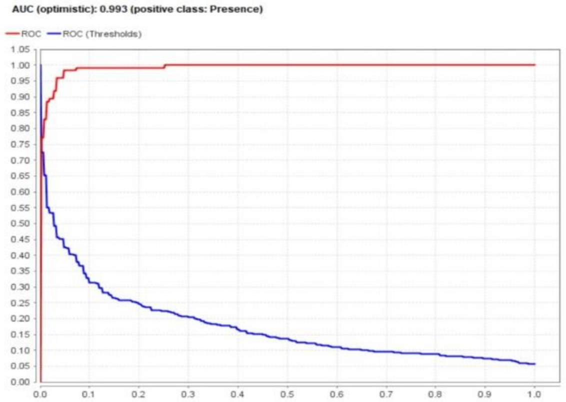


Figure 10: AUC of Outliers δ=30

Levels of Outliers	Sensitivity %	Specificity %	Precision %	Accuracy %	F-Score %	Kappa %
δ=10	96.22	95.52	96.22	95.90	96.22	91.74
δ=20	98.01	94.69	95.48	96.47	96.73	92.89
δ=30	98.63	94.49	95.36	96.70	96.97	93.36

Table 4: Performance Evaluation of Different Levels of Outliers Detection δ=10, 20, 30 on GBTC Model on Cleveland Heart Attack Dataset

Figure 11 and Table 4 show the performance evaluation of different levels of Outliers Detection δ=10, 20, 30 on GBTC model on Heart Statlog Attack dataset. The performance is measured by specificity, F-Score, sensitivity, precision, accuracy and kappa. For sensitivity, higher rate of 98.63% is achieved when δ=30. The high specificity rate is attained while δ=10. From the different outlier levels, the highest precision rate of 96.22% is attained when δ=30. The highest accuracy rate of 96.70% is achieved while δ=30. The highest F-Score rate of 96.97% is achieved while outlier level is 30. The highest kappa rate of 93.36% is achieved while outlier level is 30.

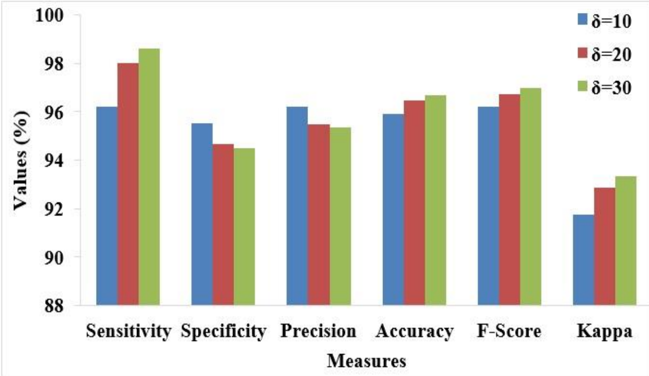


Figure 11: Comparison Of Various Measures On Different Levels Of Outliers On Cleveland Heart Attack Dataset

c. Comparative Analysis of Heart Statlog Dataset

In examining the proposed DOD-GBTC model over the applied datasets, few measures like sensitivity, precision, accuracy, kappa, specificity, and F-score are used. To represent the projected classifier efficiency, it is compared with five other classifiers namely, GBTC, J48, RT (RT), NB Tree, RF (RF) and RBF Network. The confusion matrix of the different classifiers is represented in Table 5 over the given heart-statlog dataset. By employing the values gained using confusion matrix, classifier performance is computed. It is absolute from the table, J48 classifies 88 instances are under the present category, and 119 are under the absent category. RT classifies 89 samples as a present and 117 samples as absent types of Heart Attack. NB Tree classifies 90 samples as a present and 127 samples as absent types of Heart Attack. RF classifies 94 samples as present and 127 samples as absent types of Heart Attack. RBF classifies the 97 of the instances as present and 130 as absent cases in Heart Attack prediction out of the 130 instances. The Gradient boost classifier provides 115 samples under the present type of Heart Attack and 142 under the absent type of Heart Attack. The proposed DOD-GBTC provides 104 samples under the present type of Heart Attack and 137 under the absent type of Heart Attack.

Confusion matrix for different classifiers is represented in Table 6 over the given Cleveland dataset. By using the values gained using confusion matrix, classifier performances are computed. It is absolute from the table, RT classifies 125 instances are under present category, and 96 are under the absent case. NB Tree classifies 134 samples as a present and 106 samples as absent types of Heart Attack. J48 classifies 135 samples as a present and 103 samples as absent types of Heart Attack. RBF Network classifies the 138 of the instances as present and 115 as absent cases in Heart Attack prediction out of the 303 instances. RF (RF) classifies 141 samples asa present and 108 samples as absent types of Heart Attack. The Gradient boost classifier provides 151 samples under the present type of Heart Attack and 137 under the absent type of Heart Attack.

Experts	Proposed (DOD-GBTC)		GBTC		J48		RT		RBF Network		NB Tree		RF	
	Present	Absent	Present	Absent	Present	Absent	Present	Absent	Present	Absent	Present	Absent	Present	Absent
Present	104	6	115	5	88	32	89	31	97	23	90	30	94	26
Absent	3	137	8	142	31	119	33	117	20	130	23	127	23	127

Table 5 : Confusion Matrix of Different Classifiers on Heart Statlog Dataset

Experts	Proposed (DOD-GBTC)		GBTC		J48		RT		RBF Network		NB Tree		RF	
	Present	Absent	Present	Absent	Present	Absent	Present	Absent	Present	Absent	Present	Absent	Present	Absent
Present	144	7	151	13	135	29	125	39	138	26	134	30	141	23
Absent	2	120	2	137	36	103	43	96	24	115	33	106	31	108

Table 6 : Confusion Matrix of Different Classifiers on Cleveland Heart Attack Dataset

The proposed DOD-GBTC provides 144 samples under the present type of Heart Attack and 120 under the absent type of Heart Attack. It also provides minimized counts of false positive and true negative values while comparing with conventional classifiers.

The graphical representation of precision, sensitivity, accuracy, specificity, kappa,and F-score is given below in Figure 12 -14 and the values are given in Table 7. The performance values are offered using percentage. Figure 12 demonstrates the comparison among different classifiers over classifying result of the dataset Statlog using Sensitivity, Specificity. For sensitivity, RT classifier attains the poorperformance of 72.95%.

Classifiers	Sensitivity %	Specificity %	Precision %	Accuracy %	F-Score %	Kappa %
Proposed (DOD-GBTC)	919	95.80	94.54	96.40	95.85	92.67
GBTC	93.49	96.59	95.83	95.19	94.65	90.27
J48	73.94	78.81	73.33	76.66	73.64	52.71
RT	72.95	79.05	74.16	76.29	73.55	52.08
RBF Network	82.91	84.97	80.33	84.07	81.86	667
NB Tree	79.64	80.89	75.00	80.37	77.25	60.02
RF	80.34	83.00	78.33	81.85	79.32	63.16

Table 7 : Performance Evaluation of Different Classifiers on Heart StatlogDataset

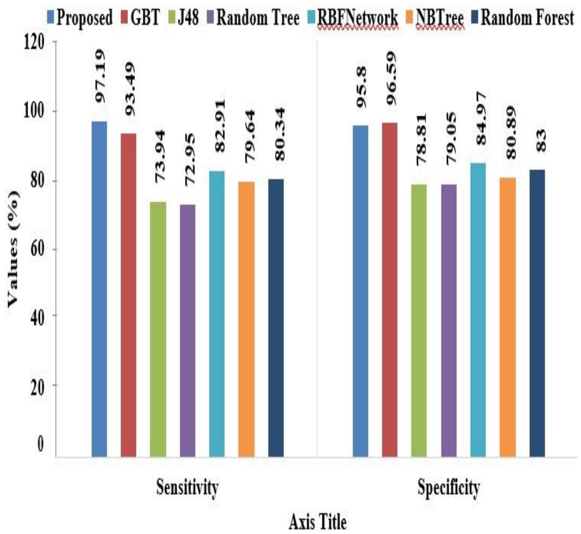


Figure 12 : Comparison of Different Classifiers on Heart Statlog Dataset in termsof Sensitivity, Specificity

J48 offers the sensitivity rate of 73.94% which is more or less demonstrates a same rate as RT. NB Tree classifier exceeds the above mentioned method through attainingthe sensitivity rate of 79.64% but it fails to outperform the GBTC model which attains the higher sensitivity rate of 93.45%. However, the proposed DOD-GBTC method attains the highest sensitivity rate of 91.9% which shows that it is the better method using sensitivity rate. For specificity, as similar to sensitivity, RT and J48 demonstrate more or less the same specificity rate of 79.05% and 78.81%. However, RF classifier outperforms by obtaining the specificity rate of 83.00% which is higher than the other two methods except for the proposed method. The GBTC achieves 96.59% of specificity rate when classifying the Heart Attack dataset. However, the proposed DOD-GBTC method attains the highest sensitivity rate of 91.9% which shows that it is the better method using sensitivity rate.

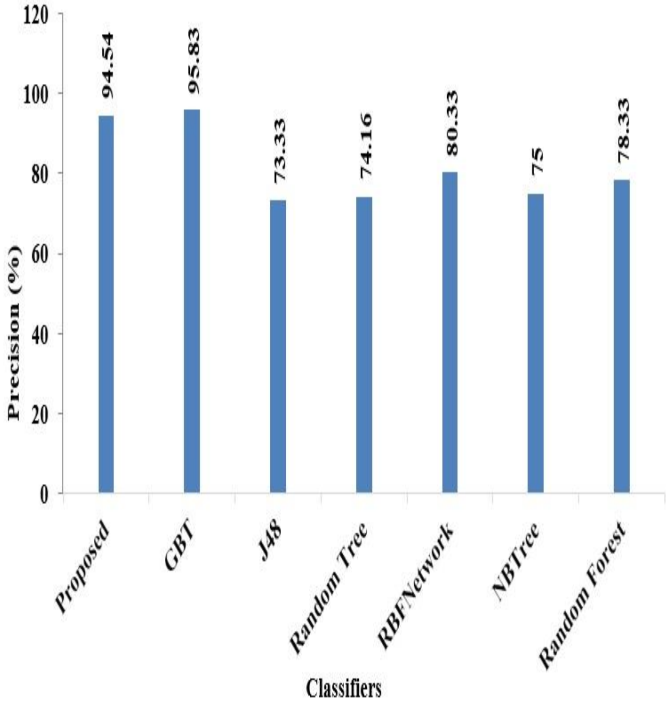


Figure 13 : Comparison of Different Classifiers on Heart Statlog Dataset in termsof Precision

As shown in Figure 13, on the applied Statlog dataset for precision, J48 classifier gives the poor performance of 73.33%. RT gives the precision rate of 74.16% which is more or less demonstrates a similar rate as J48. RBF classifier outperforms the above mentioned method by achieving the precision rate of 80.33% but it fails to outperform the Gradient boost classifier which attains the maximum precision rate of 95.83%.

Figure 14 shows the comparison of various classifiers in terms of accuracy, F-score and kappa value. For F-score, RT and J48 demonstrate more or less the similar rate of F-score of 73.5% and 73.6% respectively. The classifier RBF attains 81.86% of F-score rate. Above all, the projected model attains the F-score rate of 95.8% which is best among all.

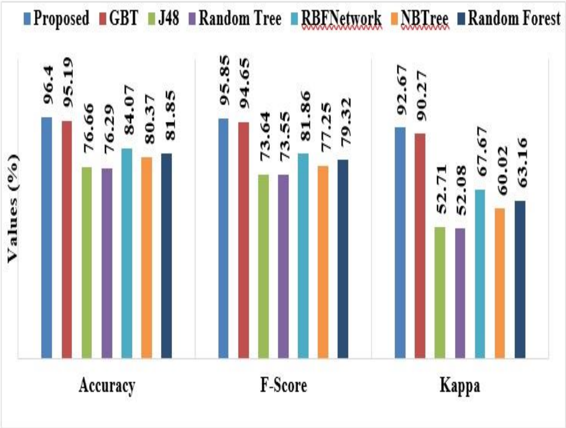


Figure 14 : Comparison of Different Classifiers on Heart Statlog Dataset in termsof Accuracy, F-score, Kappa

Figure 14 shows the classifier performance using accuracy for the given Heart Attack dataset. RT and J48 demonstrate more or less the similar rate of accuracy of 76.29% and 76.66% correspondingly. The classifier RBF attains 84.07% of accuracy rate. Above all, the proposed model attains the accuracy rate of 96.40% which is best among the compared methods. For kappa-value, RT gives the poor performance of 52.08%, and J48 gives the Kappa value rate of 52.71%. RBF classifier outperforms the above mentioned method by attaining the Kappa value of 667%, however it fails to outperform the projected DOD-GBTC which attains the maximum Kappa value of 90.27% whereas GBTC attains 90.27% as kappa rate. Therefore, for the given Heart Attack dataset, the proposed method attains the enhanced performance for all metrics like precision, sensitivity, accuracy, kappa, specificity, and F-score.

d. Comparative analysis of Cleveland Heart Attack Dataset

The graphical representation of performance evaluation in terms of precision, sensitivity, accuracy, specificity, kappa, and F-score over the given Cleveland Heart Attack Dataset is given below in Figure 15-17 and the values are given in Table 8.

Classifiers	Sensitivity (%)	Specificity (%)	Precision (%)	Accuracy (%)	F-Score (%)	Kappa (%)
Proposed (DOD-GBTC)	98.63	94.49	95.36	96.70	96.97	93.36
GBTC	98.69	91.33	92.07	95.05	94.81	90.09
J48	78.94	78.03	82.32	78.55	80.59	56.64
RT	74.41	71.11	76.21	72.94	75.30	45.38
RBF Network	85.19	81.56	84.15	83.49	84.66	66.81
NB Tree	80.24	794	81.71	79.21	80.97	58.06
RF	81.98	82.44	85.98	82.18	83.93	63.95

Table 8 : Performance Evaluation of Different Classifiers on Cleveland HeartAttack Dataset

Figure 15 demonstrates the comparison among different classifiers over classifying result of the dataset Cleveland using Sensitivity, Specificity. For sensitivity, RT classifier gives the poor performance of 74.41%. J48 gives a sensitivity rate of78.94%.

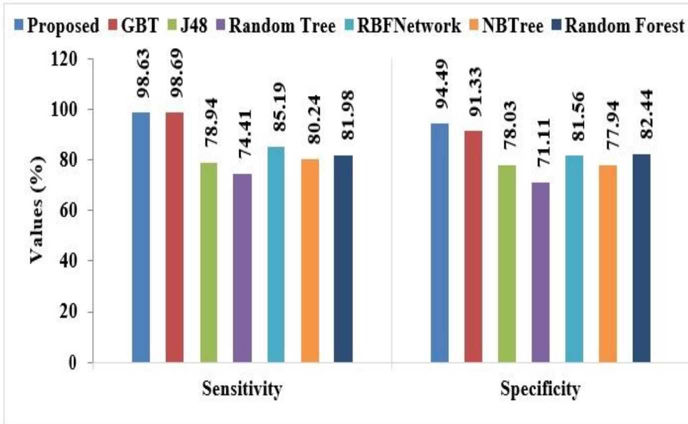


Figure 15: Comparison of Different Classifiers on Cleveland Heart AttackDataset in terms of Sensitivity, Specificity

RBF classifier outperforms the abovementioned method through achieving the sensitivity rate of 85.19% but it fails to outperform the GBTC which attains the sensitivity rate of 98.69%. The DOD-GBTC model attains sensitivity rate of 98.63% which shows that it is the better method using sensitivity rate. For specificity, as similar to sensitivity, NB Tree and J48 demonstrate more or less the similar specificity rate of 794% and 78.03%. However, RF classifier outperforms by obtaining the specificity rate of 82.44% which is higher than the other two methods except for the GBTC and DOD-GBTC models. The GBT model attains 91.33% of specificity rate while classifying the Heart Attack dataset. Whereas the DOD-GBTC model attains the specificity rate of 94.49% which is the highest among all while classifying the given dataset.

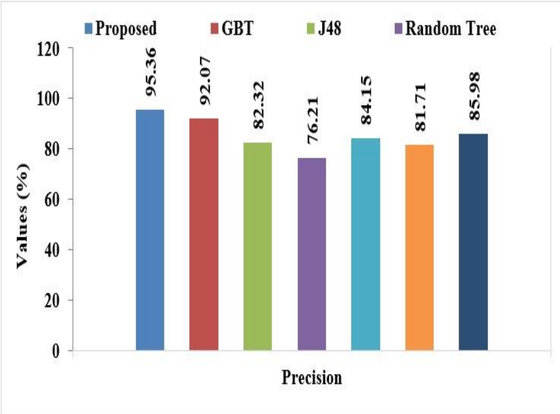


Figure 16 : Comparison of Different Classifiers On Cleveland Heart AttackDataset In Terms Of Precision

For precision, RT classifier gives the poor performance of 76.21% as shown inFigure 16. NB Tree gives the precision rate of 81.71%. RF classifier outperforms the above mentioned method by attaining the precision rate of 85.98% but it fails to outperform the GBC which attains the maximum precision rate of 92.07%. This method is overcome by the DOD-GBTC model which shows higher performance of 98.63% which it is the better method using precision rate. For F- score, NB Tree and J48 demonstrate more or less the similar rate of F-score of 80.97% and 80.59% respectively as depicted in Figure 16. The classifier RBF attains 84.66% of F-score rate. Above all, the proposed model achieves the F-score rate of 93.36% which is superior among all.

Figure 17 shows the classifier performance using accuracy for the given Heart Attack dataset. NB Tree and J48 demonstrate more or less the similar rate of accuracy of 79.21% and 78.55% respectively. The classifier RBF attains 83.49% of accuracy rate. Above all, the DOD-GBTC model achieves the accuracy rate of 96.70% which is best among the compared methods because of the implication of OD method prior to classification.

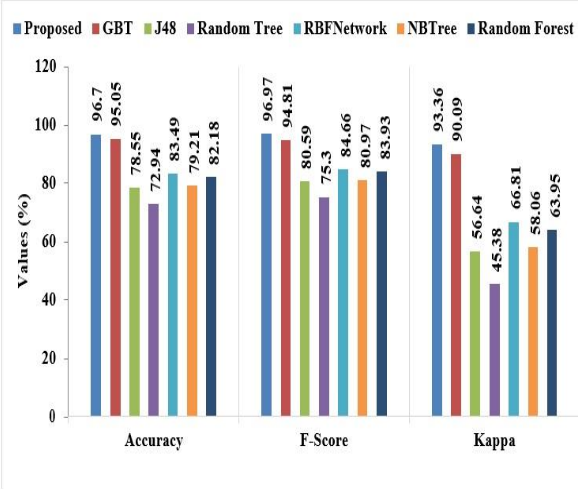


Figure 17 : Comparison of Different Classifiers on Cleveland Heart Attack Dataset in terms of Accuracy, F-score, Kappa

Figure 17 also shows the classifier performance using Kappa Value for the given Heart Attack dataset. For kappa-value, RT gives the poor performance of 45.38%, and J48 gives the Kappa Value rate of 56.64%. RBF classifier outperforms the above mentioned method by attaining the Kappa Value of 66.81%, but it fails to outperform the GBTC which attains the maximum Kappa Value of 90.09%.

Above all, the projected DOD-GBTC attains the higher kappa value of 93.36%. Therefore, for the given Heart Attack dataset, the DOD-GBTC model attains the enhanced performance for all metrics like precision, sensitivity, accuracy, kappa, specificity, and F-score due to the inclusion of OD technique prior to classification.

IV. Conclusions

Finally new DOD model was implemented from the given patient's care particularly Heart Attack, where every patient-management action are mainly based on the condition of the patient. For classification purposes, GBTC model is employed. Various analysis were carried out in order to proposed DOD-GBTC for the Heart Attack prediction and the outcomes are discussed above. Two standard datasets such as Heart-Statlog and Cleveland were used for the performance evaluation of DOD- GBTC. The DOD-GBTC model offers superior classifier results with a maximum of 96.40% of accuracy and 96.70% of accuracy on the applied Statlog and Cleveland dataset. The experimental results verified that the presented DOD-GBTC model is effective over the compared methods in a significant way. The presented DOD-GBTC model offers superior classifier results with a maximum of 96.40% of accuracy and 96.70% of accuracy on the applied Statlog and Cleveland dataset.

V. Future Developments

In future, the performance of the presented works can be further enhanced using following ways -

- The clustering techniques can be incorporated before the data classification process to improve the classifier results.
- At the same time, Deep learning based classifier models can be employed in place of ML based classifier models to enhance the classification performance.
- Also, the proposed models can be applied to detect various attacks apart in real time apart from Heart Attack.
- Finally, the performance of the proposed models can be further improved by properly adjusting the parameter values using appropriate tuning models.

References

[1] “Predictive analysis of cardiovascular disease using gradient boosting based learning and recursive feature elimination technique”, Prasannavenkatesan Theerthagiri, Intelligent Systems with Applications, Volume 16, November 2022, www.elsevier.com

[2] “An Ensemble Basedheart Disease Predictionusing Gradient Boosting Decision Tree”, S.Irin Sherly, G. Mathivanan, Turkish Journal of Computer and Mathematics Education Vol.12 No.10 (2021)

[3] “A Gradient Boosted Decision Tree with Binary Spotted Hyena Optimizer for cardiovascular disease detection and classification”,Siripuri Kiran , GantaRaghotham Redd y, Girija S.P., Venkatramulu S., Kumar Dorthi, Chandr a Shekhar Rao V., Healthcare Analytics, Volume 3, 2023, www.elsevier.com

[4] “Heart Disease Prediction Model based Ongradient Boosting Tree (GBT) Classification Algorithm”, R. Bhuvaneeswari, P. Sudhakar, G. Prabakaran, International Journal of Recent Technology and Engineering (IJRTE), Volume-8, Issue-2S11, September 2019.

[5] “Prediction of Heart Disease Using a Combination of Machine Learning and Deep Learning”, Rohit Bharti, Aditya Khamparia, Mohammad Shabaz, Gaurav Dhiman, Sagar Pande and Parneet Singh, Computational Intelligence and Neuroscience, 2021

[6] “An Artificial Intelligence Model for Heart Disease Detection Using Machine Learning Algorithms”, Victor Chang , Vallabhanent Rupa Bhavani, Ariel Qianwen Xu and MA Hossain, Healthcare Analytics 2 (2022) 100016, www.elsevier.com

[7] “Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques”, Jyoti Maurya, Shiva Prakash, C. Beulah Christalin Latha, S. Carolin Jeeva, Informatics in Medicine Unlocked, Volume 16, 2019, www.elsevier.com

[8] “An artificial intelligence model for heart disease detection using machine learning algorithms”, Victor Chang, Vallabhanent Rupa Bhavani, Ariel Qianwen Xu, MA Hossain, Healthcare Analytics, Volume 2, November 2022, www.elsevier.com

[9] “Artificial Intelligence in Predicting Cardiac Arrest: Scoping Review”, Asma Alamgir, Osama Mousa, Zubair Shah, JMIR Med Inform. Publications, 2021 Dec, PubMed Central, www.ncbi.nlm.nih.gov