International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-5, March 2020

Cardiovascular Disease Prediction and Classification using Modified Teaching Learning Optimization Method

N. Rajinikanth, L. Pavithra

Abstract: ***Cardiovascular disease (CVD) is possibly the greatest reason for casualty and death rate among the number of inhabitants on the planet. Projection of cardiopathy is viewed as one of the most crucial subjects in the area of clinical records exploration. The measure of information in the social insurance industry is massive. The Data mining process transforms the huge range of unrefined medical service data into meaningful information that can lead to erudite decision and projection. Some recent investigations have applied data exploratory procedures too in CVD estimation. However, only very few studies have revealed the elements that play crucial role in envisioning CVDs. It is imperative to opt for the combination of correct and significant elements that can enhance the functioning of the forecasting prototypes. This study aims to ascertain meaningful elements and data mining procedures that can enrich the correctness of foretelling CVDs. Prognostic models were formulated employing distinctive blend of features selection modified teaching learning optimization techniques, SVM and boosting classification. Here the proposed strategy gives high precision outcomes with existing classification.***

Keywords**:** ***Feature Selection, Heart Disease Prediction, Modified Teaching Learning Optimization, Svm Classification, Adaboosting.***

I. INTRODUCTION

In the present current world, cardiovascular ailment is the most deadly one [1]. As indicated by World Health Organization about in excess of 12 million deaths happens around the world, consistently because of heart issues [2].With the turn of the century, cardiovascular maladies (CVDs) have become the main source of mortality in India [3]. The expression "cardiovascular disease" incorporates a wide scope of conditions that influence the heart and the veins, and the manner in which blood is siphoned and flowed through the body, additionally are viewed as types of heart disease [4]. This sickness assaults an individual so in a flash that it barely gets whenever to get treated with. Probably the most ideal approaches to analyze a heart disease are by utilizing *echo-cardio-graphy. Echo-cardio-graphy* or reverberation is an easy test that utilizes sound waves to make photos of the heart. The test gives data about the size and state of the heart and how well the heart chambers and valves are working [5].

Revised Manuscript Received on March 30, 2020.
* Correspondence Author
N. Rajinikanth, Department of Computer Applications, CMS College of Science & Commerce, Coimbatore, India. Email: rajni1402@yahoo.com
Dr. L. Pavithra, Department of Computer Science, Dr. N. G. P. Arts & Science College, Coimbatore, India. Email: paviloky@gmail.com

The test likewise can distinguish territories of heart muscles that are not contracting regularly because of poor blood stream or damage from a past heart attack [6, 7]. Along these lines, diagnosing patients accurately on auspicious premise is the most testing assignment for the therapeutic club. The Healthcare business today produces tremendous measures of complex information about patients, infection finding, emergency clinics assets and restorative gadgets, which is hard to process by manual strategies [8].

Information mining gives a lot of devices and methods to discover examples and concentrate information to give better patient consideration and it consolidates measurable investigation, AI and database innovation to extricate concealed examples and connections from enormous databases [9]. The recognition of heart disease from different elements or side effects is a multi-layered issue, which is not liberated from bogus assumptions frequently joined by eccentric impacts. Powerful and effective robotized heart disease forecast can profit human services part and this robotization will spare expense as well as time [10]. This exploration paper features the utility and utilization of three distinctive grouping models of information digging methods for forecast of cardiovascular malady to encourage specialists in the human services space [11-13].

Heart attack disease remains the fundamental explanation of death all through the world, and the bigger number of passing emerges in low and center salary nations like India. Medicinal professionals constantly produce huge measure of information in the field of biomedical. This information can be utilized for the early identification of the heart disease, which can support to lessen the quantity of cardiovascular failures. The medicinal services organizations and industry accumulate huge degree of information that is not extricated to discover shrouded information for effective selection. The fundamental focal point of information distinguishing proof is to find designs that were some time ago not known. When these examples are discovered, they can be utilized to settle on compelling selections.

The significant causes that lift the likelihood of cardiovascular failures are unsafe utilization of liquor, smoking, and hypertension, a lot of specific fats and cholesterol in the blood, absence of physical activities, undesirable eating regimen and high sugar level. The most widely recognized sort of heart disease is heart course disease, cardio-myopathy, cardio vascular infection, arrhythmias, preliminary Fibrillation, Irregular Heart Rhythm and innate heart absconds.

These days numerous progressions are occur in ways of life of people groups in developing nations like India heart infections have become a significant explanation of passing. In the wellbeing, zone information mining methods assume significant job to find the sicknesses.

II. RELATED WORK

In past works, different creators have actualized various strategies on heart disease dataset from Cleveland Heart Disease databases, the effectiveness of the classifier is surveyed, and their yields are investigated. A portion of the creators have worked upon order procedures like Artificial Neural Network, Naive Bayes, K-closest neighbor, Decision tree and so forth on the heart informational collection and look at the productivity of various methods. The grouping techniques don't give better exactness and test results. In acquainted characterization, to coordinate arrangement rules and affiliation rules, most importantly, affiliation rules are created by the calculation and afterward little arrangement of these standards are chosen to shape the classifier and this methodology is not yet executed on heart maladies informational index [13].

Palaniappan and Awang [15] recommended a model of IHDPS (Intelligent Heart Disease Prediction System) executing information mining calculations, as Naive-Bayes, Decision Trees and Neural Network. The last yield of these calculations portrays that every strategy has its various abilities in the motivation behind the depicted mining objectives. Smart Heart Disease Prediction System is basic, simple to utilize, adaptable, expandable and solid electronic predication framework that can give yield of troublesome inquiries and the conventional selection emotionally supportive network neglect to do. In IHDPS creator's utilization the restorative profile traits like age run, sex, hypertension and high glucose to find the indications of patients.

Srinivas [16] planned use of information mining systems in disclosure of heart maladies. The creator utilized the Tanagra apparatus for usage of information mining, factual and AI calculations. Creator utilizes the preparation informational collection with 14 distinct traits and 3000 occasions. The examinations were performed on the preparation informational index to quantify calculation execution, as far as time taken and accuracy. The examples and characteristics in the informational index were depicting the results of different sorts of tests to ascertain the productivity of heart disease. Creator separated the informational index into two distinct parts, 30% of information was utilized as testing and 70% was utilized as preparing. The correlation was done on the bases of 10 overlap cross approvals. The proposed work best execution calculation results were 52.33% exactness utilizing Naive Bayes among of these order calculations on heart maladies informational collection.

The exploration paper [14] depicts the model utilizing guileless Bayes and weighted acquainted classifier (WAC) to foresee the likelihood of patients accepting heart attacks. The scientists [15] built up an online smart framework utilizing innocent Bayes calculation to answer complex inquiries for diagnosing heart disease and help therapeutic specialists with clinical selections.

The specialist [17] utilizes affiliation rules speaking to a system in information mining to improve sickness forecast with extraordinary possibilities. A calculation with search requirements was likewise acquainted with decrease the quantity of affiliation controls and approved utilizing train and test approach [17]. Three well-known information-mining calculations (bolster vector machine, counterfeit neural system and selection tree) were utilized by the specialists to build up a forecast model utilizing 502 cases. SVM turned into the best forecast model pursued by fake neural systems [18].

Two sorts of information mining calculations named developmental GA-KM (genetic algorithm - K-means method) and modified particle swarm optimization (MPSO)-KM group the heart malady informational collection and anticipate model exactness [20]. This half-breed strategy consolidates force type modified particle swarm optimization (MPSO) and K-means method. The examination was made in the exploration led utilizing C5, Naïve Bayes, K-means, GA-KM and MPSO-KM for assessing the accuracy of the strategies. The test results demonstrated that accuracy improved when utilizing GA-KM and MPSO-KM [20].

III. DATASET DETAILS

The heart disease information was gathered from UCI AI vault (Dua and Karra Taniskidou, 2017). There are four databases (for example Cleveland, Hungary, Switzerland, and the VA Long Beach). The Cleveland database was chosen for this examination since it is a usually utilized database by AI specialists with records that are generally finished. The dataset contains 303 records. Despite the fact that the Cleveland dataset has 76 attributes, the dataset gave in the vault just gives data to a subset of 14 traits. The information wellspring of the Cleveland dataset is Cleveland Clinic Foundation portrays the depiction and sort of traits. There are 13 attributes that component in heart disease forecast and one property fill in as the yield or the anticipated quality for the nearness of heart disease in a patient. The Cleveland dataset contains a trait named 'num' to show the finding of heart disease in patients on various scales, from 0 to 4. In this situation, 0 speaks to the nonappearance of heart disease and every one of the qualities from 1 to 4 speak to patients with heart disease , where the scaling alludes to the seriousness of the sickness (4 being the most elevated).

A. Data preprocessing

The information was pre-processed after assortment. There were 6 records that have missing qualities in Cleveland data set. Every one of the records with missing qualities was expelled from the data set, in this way lessening the quantity of records from 303 to 297.

Next, the estimations of anticipated characteristic for the nearness of heart disease in the data-set was changed from multiclass values (0 for nonappearance and 1, 2, 3, 4 for nearness) to the parallel qualities (0 for nonattendance; 1 for nearness of heart disease). The information pre-processing task was performed by changing over all the analysis esteems from 2 to 4 into 1. The subsequent data set in this way contains just 0 and 1 as the analysis esteem, where 0 being the nonattendance and 1 being the nearness of heart disease. After the decrease and change, the circulation of 297 records for 'num' attributes brought about 160 records for '0' and 137 records for '1'.

IV. TEACHING-LEARNING-BASED OPTIMIZATION

Teaching Learning-Based Optimization (TLBO) calculation Teaching-learning is a significant procedure where each individual attempts to take in something from others to develop themselves. Proposed a calculation, known as Teaching-Learning-Based Optimization (TLBO), which reproduces the conventional showing-learning marvel of a study hall. The calculation reenacts two key methods of learning: (i) through the educator (known as the instructor stage) and (ii) associating with different students (known as the student stage). TLBO is a populace-based calculation, where a gathering of understudies (for example student) is viewed as the populace and the various subjects offered to the students are comparable to with the distinctive structure factors of the optimization issue. The after-effects of the student are practically equivalent to the wellness estimation of the improvement issue. The best arrangement in the whole populace is considered as the instructor. The activity of the TLBO calculation is clarified underneath with the educator stage and student stage.

A. Teacher phase

This period of the calculation recreates the learning of the understudies (for example students) through the instructor. During this stage, an instructor passes on information among the students and tries to build the mean after-effect of the class. Assume there are 'm' number of subjects (for example plan factors) offered to 'n' number of students (for example populace size, k = 1, 2. . . n). At any successive educating learning cycle i, Mj, i is the mean aftereffect of the students in a specific subject 'j' (j = 1, 2, . . . , m). Since an educator is the most experienced and proficient individual regarding a matter, the best student in the whole populace is viewed as an instructor in the calculation. Let X_{total}−k_{best}, i be the aftereffect of the best student considering every one of the subjects who is distinguished as an instructor for that cycle. The instructor will place most extreme exertion into expanding the information level of the entire class; however, students will pick up information as indicated by the nature of educating conveyed by an educator and the nature of students present in the class. Thinking about this reality, the distinction between the aftereffect of the instructor and the mean consequence of the students in each subject is communicated as:

$$Difference\_mean_{j,i}=r_i(X_{j,kbest,i}-T_f M_{j,i})$$

Where $X_{j,kbest,i}$ is the consequence of the educator (for example best student) in subject j. T_f is the showing factor, which chooses the estimation of intend to be changed, and r_i is the arbitrary number in the range [0, 1]. The estimation of T_f can be either 1 or 2. The estimation of T_f is chosen arbitrarily with equivalent likelihood as:

$$T_f = \text{round } [1+\text{rand } (0, 1) \{2-1\}]$$

Where rand is the irregular number in the range [0, 1]. T_f isn't a parameter of the TLBO calculation. The estimation of T_f is not given as a contribution to the calculation and its worth is randomly chosen by the calculation utilizing Eq. (2). In view of the $Difference\_mean_{j,i}$, the current arrangement is refreshed in the instructor stage as per the accompanying articulation:

$$X'_{j,k,i}=X_{j,k,i}+Difference\_mean_{j,i}$$

Where $X'_{j,k,i}$ is the refreshed worth of $X_{j,k,i}$. Accept $X'_{j,k,i}$ in the event that it gives superior capacity esteem. All the acknowledged capacity esteems toward the finish of the educator stage are kept up, and these qualities become the contribution to the student stage. It might be noticed that the estimations of r_i and T_f influence the presentation of the TLBO calculation. ri is the arbitrary number in the range [0, 1] and T_f is the instructing factor. Be that as it may, the estimations of r_i and T_f are produced randomly in the calculation and these parameters are not provided as contribution to the calculation (not at all like providing hybrid and transformation probabilities in GA, dormancy weight and psychological and social parameters in PSO, and province size and farthest point in ABC, and so forth.). In this manner, tuning of r_i and T_f isn't required in the TLBO calculation (dissimilar to the tuning of hybrid and change probabilities in GA, inactivity weight and intellectual and social parameters in PSO, and settlement size and cutoff in ABC, and so on.). TLBO requires tuning of just the regular control parameters, similar to populace size and number of ages, for its working, and these normal control parameters are required for the working of all populace based optimization calculations. In this way, TLBO can be called a calculation explicit parameter-less calculation.

B. Student stage

This period of the calculation recreates the learning of the understudies (for example students) through cooperation among themselves. The understudies can likewise pick up information by talking about and cooperating with different understudies. A student will adapt new data if different students have more information than the person in question. The learning marvel of this stage is communicated underneath.

Randomly select two students, P and Q, to such an extent that $X'_{total-p,i}=X'_{total-Q,i}$, where $X'_{total-p,i}$ and $X'_{total-Q,i}$ are the refreshed estimations of $X_{total-p,i}$ and $X_{total-Q,i}$,, individually, toward the finish of the educator stage.

$X''_{j,p,i}=X'_{j,p,i}+r_i(X'_{j,p,i}-X'_{j,Q,i})$, if $X'_{total-p,i}>X'_{total-Q,i}$,
$X''_{j,p,i}=X'_{j,p,i}+r_i(X'_{j,Q,i}-X'_{j,p,i})$, if $X'_{total-Q,i}>X'_{total-p,i}$

(The above conditions are for amplification issues, the turnaround is valid for minimization issues.) Accept $X''_{j,p,i}$

V. PROPOSED WORK FOR TLBO BASED FEATURE SELECTION

We structured feature selection TLBO to play out the procedure of feature selection for malady determination. The feature selection TLBO calculation impersonates the act of teaching the understudies in a study hall. To start with, educator enlightens the understudies with his insight, and a short time later, understudies are prepared with the partner understudies. Feature selection TLBO does not stipulate any express factors and just utilize general components like the quantity of examples in the populace alongside most extreme emphases.

The proposed calculation depends on the irregular hunt technique comprising of two phases i.e., Teaching Phase and Learning Phase. These phases utilized various classifiers in our wrapper based component selection strategy for finding the wellness of people in the populace.

In the principal stage, best individual goes about as an educator and trains the rest of the people by considering them as students to improve their insight dependent on his expertise. The person with the most noteworthy exactness or least blunder rate will be named as an instructor in this stage. Expectation models are utilized to assess the cost (wellness) of the populace in wrapper-based component determination techniques. Arrangement accuracy will be considered as wellness esteem for augmentation issues and order blunder rate for minimization issues. The instructor manages the students to carry their answers towards best with the assistance of contrast mean. We have utilized various classifiers like, Naïve Bayes, SVM, k-Nearest Neighbors (kNN), Decision Trees and Discriminant Analysis to discover the characterization blunders and exactnesses which goes about as wellness esteems during the assessment. The encouraging element will be chosen randomly between in the range of 1 and 2. Characterization accuracy (CA) can be characterized.

CA= Correctly Classified Instances / Total Instances.

In the subsequent stage, student's aptitude will be improved by interfacing among themselves in the populace. Every student in the arrangement space refreshes their answers (twofold strings) with the assistance of different students chosen randomly in the populace. The populace contains just the top people with the best wellness esteems and a lot of feature s after various cycles of instructor and student stages. The person with the least blunder rate or the most elevated accuracy will be treated as the last answer for the given issue of ideal component determination. The dataset with the new people can be utilized to prepare classifier models for better execution. We have tried our model with various populace estimates and got the best qualities with 30 people in the arrangement space for better assembly. We run the calculation with a extreme number of ages lastly restricted to 50 as a result of the quicker intermingling of characterization blunder. We additionally displayed the progression of the calculation .We created feature selection TLBO as an adjusted paired variation of the TLBO calculation to pick the best subspace of features from the huge dataset. The primary objectives of our proposed model are: a) Decrease the mark of features to fabricate an order model b) Boost the arrangement accuracy of a preparation model c) Reduce the preparation time of an AI model with a chosen subset of features.

A. SVM Classification

The Support-vector-machine (SVM) is perhaps one of the most significant and discussed AI calculations. The SVM is significant system for order of mutually direct and non-straight information. It uses a non-direct mapping to change the first preparing information into a higher measurement. With this new measurement, it scans for straight ideal isolating hyper-plane. With the appropriate non-straight mapping to a satisfactorily high measurement, information from two classes can generally be isolated by a hyper-plane. The SVM discover this hyper-plane utilizing bolster vectors and edges. SVM performs characterization assignments by expanding the edge particular, the two classes while limiting the grouping mistakes.

B. Adaptive Boosting

Adaptive Boosting (AdaBoost) is another troupe classifier strategy. The natural over fitting issue present in different Machine Learning methods can be diminished utilizing AdaBoost. Boosting is a strategy, which changes over feeble students into solid students. AdaBoost works by picking base classifiers and improve it by mulling over misclassified characteristics from the prepared dataset in an iterative technique. We allot equivalent loads to all the preparation features / qualities and pick a base classifier. After every emphasis, we apply the base classifier to the preparation set and increment the loads of the misclassified attributes. We emphasize ‘n’ times, each time applying base student on the preparation set with refreshed loads. In the last model, proposed approach joins the expectation of each powerless student by thinking about the normal or weighted normal of ‘n’ students.

In this proposed work, the blend of SVM and AdaBoost is utilized to characterize the heart disease expectation framework in prior stage. The hybridization technique is utilized initially to arrange ordinary and cardiovascular dataset utilizing straight SVM bit and afterward characterize the heart disease information into various stages through AdaBoost. In SVM there will be a hyper-plane between the arrangements of information focuses as the selection limit. For this situation, there are two arranged information of typical patients and heart disease patients are utilized by this SVM-AdaBoost cross breed calculation.

VI. RESULT AND DISCUSSION

The informational index utilized for heart maladies forecast has been taken from UCI Machine Learning Repository. The Cleveland dataset has been utilized to anticipate heart disease. This database contains 76 attributes, yet just a subset of 14 characteristics has been utilized for expectation.

A portion of the traits utilized are age, sex, cp (chest torment), trestbps (resting pulse), chol (cholestoral), fbs (fasting glucose), restecg (resting electrocardiographic), thalach (maximum pulse accomplished), exang (exercise instigated angina), num and so forth. Among the features utilized in heart disease prediction, only 13 features utilized to project CVDs. The other components 'age' and 'sex' implies to the individual data of every patient. The remaining 11 features are on the whole clinical attributes gathered from different therapeutic assessments. In this examination, a mix of features was chosen based on 'altered instructing learning enhancement' feature to be utilized with characterization methods, Support Vector Machine and Boosting calculation, to make the arrangement model.

The changed training learning advancement based component selection all out of 14 elements in dataset. Utilizing MTLBO strategy features diminished component (characteristics), only 11 chosen attributes are (1,2,3,4,5,7,8,10,11,12,13) adjusted and TLBO based element is chosen.

Table 1: Feature selection and its time period for existing and proposed method

Algorithm	Feature (attributes) selection	Time period
PSO	14	3.51
TLBO	13	3.14
MTLBO	11	2.86

Table 2: parameters analysis between existing algorithm and proposed algorithm

Algorithm	Accuracy	precision	recall	F measure	Time period	Error rate
SVM	85	81	80	82	4.6	0.452
Adaboosting	87	83	85	83	4.3	0.421
Adaboosting_svm	90	85	89	87	3.8	0.354

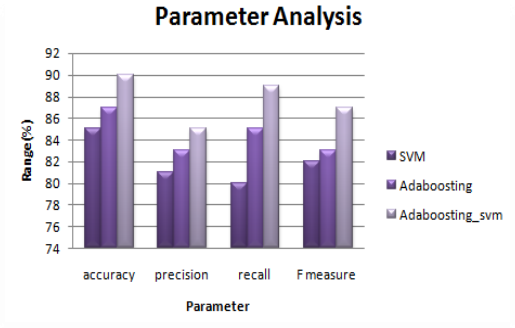

Figure 1: Parameter Analysis with SVM and Adaboosting

Figure1 presents the performance of the proposed scheme evaluated using SVM, Adaboosting and Adabossting_SVM with the feature performance based on accuracy, precision, recall and F-measure. The ranges of accuracy, precision, recall and F-measure facilitated by the proposed scheme with Adabossting_SVM is determined to be comparatively higher than the benchmarked.

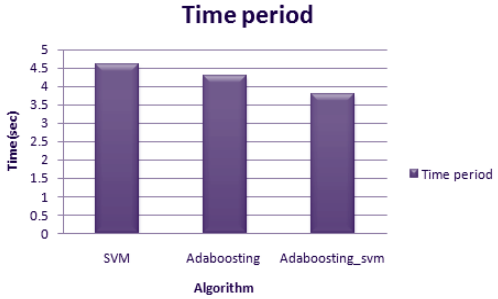

Figure 2: Time Period Analysis with SVM

Figure2 shows time period comparison for existing, proposed classification methods without feature selection

Table 3: parameters analysis between existing algorithm and proposed algorithm

Algorithm	accuracy	precision	recall	F measure	Time period	Error rate
MTLBO_SVM	93	91	92	93	3.6	0.34
MTLBO Adaboosting	95	92	94	94	3.1	0.28
MTLBO _Adaboosting_SVM	97	93	95	96	2.5	0.24

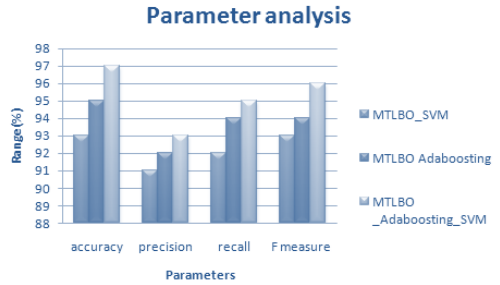

Figure 3: Parameter Analysis with MTLBO

Figure 3 presents the performance of the proposed scheme evaluated using SVM, Adaboosting and Adabossting_SVM with the with feature performance based on accuracy, precision, recall and F-measure.

The ranges of accuracy, precision, recall and F-measure facilitated by the proposed scheme with Adabossting_SVM is determined to be comparatively higher than the benchmarked

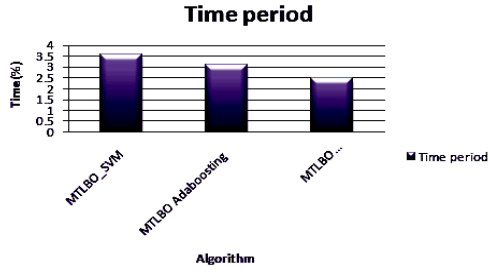

Figure 4: Time Period Analysis with MTLBO

Figure 4 shows time period comparison for existing, proposed classification methods with feature selection

VII. CONCLUSION

Medical industry is having enormous measure of valuable information. From this, the information is utilized with various objectives; here the heart diseases prediction information is utilized for identifying the performance of classifiers. The conclusive outcome shows the performance of classifier algorithms utilizing prediction accuracy. The outcome also shows the assessment of preciseness of classifiers. The correlation result shows that the adjusted instructing learning optimization feature selection based characterization is having the most efficient prediction accuracy compared to other existing techniques.

REFERENCES

1. Reddy RV, Raju KP, Kumar MJ, Sujatha CH, Prakash PR (2016) Prediction of heart disease using decision tree approach. Int J Adv Res Comput Sci Softw Eng 6: 530-532.
2. Soni J, Ansari U, Sharma D, Soni S (2011) Predictive data mining for medical diagnosis: An overview of heart disease prediction. Int J Comput Appl 17: 43-48.
3. Prabhakaran D, Jeemon P, Roy A (2016) cardiovascular diseases in India: Current epidemiology and future directions Circulation 133: 1605-1620.
4. Taneja A (2013) Heart disease prediction system using data mining techniques. Orient J Comp Sci & Technol 6: 457-466.
5. Dey M, Rautaray SS (2014) Study and analysis of data mining algorithms for healthcare decision support system. Int J Comput Sci Inf Technol 5:470-477.
6. Kaddoura S (2009) Echo made easy. (2nd edn). Churchill Livingstone, Elsevier pp:1-219.
7. Kirmani MM, Ansarullah SI (2016) Classification models on cardiovascular disease detection using neural networks, naïve Bayes and J48 data mining techniques. Int J Adv Res Comput Sci 7: 52-61.
8. Rohilla J, Gulia P (2013) Analysis of data mining techniques for diagnosing heart disease. Int J Adv Res Comput Sci Softw Eng 3: 717-724.
9. Abdar M, Kalhori SR, Sutikno T, Subroto IM, Arji G (2015) Comparing performance of data mining algorithms in prediction heart diseases. International Journal of Electrical and Computer Engineering 5:1569-1576.
10. Bhatla N, Jyoti K (2012) an analysis of heart disease prediction using different data mining techniques. Int J Adv Res Techno 1: 1-4.
11. Sudhakar K, Manimekalai DM (2014) Study of heart disease prediction using data mining. Int J Adv Res Comput Sci Softw Eng 4: 1157-1160.
12. Srinivas K, Rani BK, Govrdhan A (2010) Applications of data mining techniques in healthcare and prediction of heart attacks. International Journal on Computer Science and Engineering 2: 250-255.
13. Kajal ES, Nishika P (2016) Prediction of heart disease using data mining techniques. International Journal of Advance Research, Ideas and Innovations in Technology 2: 1-7
14. S. Palaniappan and R. Awang, "Intelligent heart disease prediction system using data mining techniques," in Proc.of IEEE/ACS Int. Conf., Doha, 2008, pp. 108-115.
15. R. Das, I. Turkoglu, and A. Sengur, "Effective diagnosis of heart disease through neural networks ensembles." Expert Systems with Applications, Elsevier, vol. 36, no. 4, pp.7675-7680, 2009.
16. K. Srinivas, B.K. Rani and D.A. Govrdhan, "Application of data mining techniques in healthcare and prediction of heart attacks." International Journal on Computer Science and Engineering, vol. 2, no. 2, pp. 250-255, 2011.
17. N. A. Sundar, P. P. Latha, and M. R. Chandra, "performance analysis of classification data mining techniques over heart disease data base," International Journal of Engineering Science & Advanced Technology, vol. 2, no. 3, pp. 470–478, 2012.
18. S. A. Pattekari and A. Parveen, "Prediction System for Heart Disease Using Naive Bayes," International journal of Advanced Computer and Mathematical Sciences, vol. 3, no. 3, pp. 290–294, 2012.
19. C. Ordonez, "Association rule discovery with the train and test approach for heart disease prediction.," IEEE transactions on information technology in biomedicine : a publication of the IEEE Engineering in Medicine and Biology Society, vol. 10, no. 2, pp. 334–43, Apr. 2006.
20. Y. Xing, J. Wang, Z. Zhao, and A. Gao, "Combination Data Mining Methods with New Medical Data to Predicting Outcome of Heart Disease," in 2007 International Conference on Convergence Information Technology (ICCIT 2007), 2007, pp. 868–872.
21. Nahar, J., Imam, T., Tickle, K.S., Chen, Y.P.P., 2013. Computational intelligence for heart disease diagnosis: a medical knowledge driven approach. Expert Syst. Appl.40 (1), 96–104.

AUTHORS PROFILE

N. Rajinikanth, is currently pursuing Ph.D in Computer Science. He has more than 10 years of teaching and research experience and he is currently working with CMS College of Science & Commerce, Coimbatore. He has handful of contributions to national and international publications. His main interest includes Data Mining, Machine Learning

Dr. L. Pavithra, received her Ph.D. Degree in 2010 from Bharathiar Univeristy. She has more than 12 years of experience and presently she is working with Dr. N. G. P Arts & Science College, Coimbatore. She has contributed her research work to many international publications. She has also guided many research candidates. Her research interest includes Data Mining, Machine Learning, Artificial Intelligence.