Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 85 (2016) 962 – 969

Procedia Computer Science

Efficient Heart Disease Prediction System

Purushottam[a,c]* , Prof. (Dr.) Kanak Saxena[b], Richa Sharma[c]

[a]Research Scholar , R.G.P.V. Bhopal(M.P),India.
[b]Prof & Head, Department of Computer Application S.A.T.I. Vidisha (M.P), India
.[a] [c]Assistant Prof, Amity University Uttar Pradesh, Noida, India.

Abstract

Cardiovascular sickness is a major reason of dreariness and mortality in the present living style. Distinguishing proof of cardiovascular ailment is an imperative yet an intricate errand that should be performed minutely and proficiently and the right robotization would be exceptionally attractive. Each individual can't be equivalently skilled thus as specialists. All specialists can't be similarly talented in each sub claim to fame and at numerous spots we don't have gifted and authority specialists accessible effortlessly. A mechanized framework in therapeutic analysis would upgrade medicinal consideration and it can likewise lessen costs. In this exploration, we have planned a framework that can proficiently find the tenets to foresee the risk level of patients in view of the given parameter about their health. The main contribution of this study is to help a non-specialized doctors to make correct decision about the heart disease risk level. The rules generated by the proposed system are prioritized as Original Rules, Pruned Rules, Rules without duplicates, Classified Rules, Sorted Rules and Polish . The execution of the framework is assessed as far as arrangement precision and the outcomes demonstrates that the framework has extraordinary potential in anticipating the coronary illness risk level all the more precisely.

Keywords: *Heart disease prediction System , Polish, CVD, CAD ,C4.5*.

1. Introduction

In today's opportunity at numerous spots clinical test outcomes are regularly made in light of specialists' instinct and experience as opposed to on the rich data accessible in numerous expansive databases. Numerous a times this procedure prompts inadvertent predispositions, lapses and a tremendous medicinal expense which influences the nature of administration gave to patients.

Today numerous doctor's facilities introduced some kind of quiet's data frameworks to man-age their social insurance or patient information. These data frameworks commonly produce a lot of information which can be in distinctive organization like numbers, content, diagrams and pictures yet sadly, this database that contains rich data is once in a while utilized for clinical choice making.

Like business knowledge and examination, the term information mining can mean diverse things to distinctive individuals. In exceptionally straightforward way we can characterize information mining as this is the investigation of substantial information sets to discover examples and utilize those examples to foresee or fore-cast the probability

Peer-review under responsibility of the Organizing Committee of CMS 2016
doi:10.1016/j.procs.2016.05.288

of future occasions. The motivation to do this problem comes from World Health Organization estimation. According to the World Health Organization estimation till 2030, very nearly 23.6 million individuals will pass on because of Heart malady. So to minimize the danger, expectation of coronary illness ought to be finished. Analysis of coronary illness is typically in view of signs, manifestations and physical examination of a patient. The most troublesome and complex assignment in medicinal services area is finding of right ailment. This colossal entirety huge of rough data is the rule resource that can be capably pre-taken care of and inspected for key information extraction that direct or by suggestion influences the remedial society for cost sufficiency and reinforce decision making. Authentic determination of coronary sickness can't be possible by using simply human understanding. There are heaps of parameters that can impacts the accu-rate conclusion like less exact results, less experience, time subordinate execu-tion, data up degree and whatnot. Packs of headway and examination happened in this field using multi-parametric qualities with nonlinear and direct parts of Heart Rate Variability (HRV).A novel framework was proposed by Heon Gyu Lee et al. [4]. To fulfill this, various experts have used various classifiers e.g. CMAR (Classification Multiple Association Rules), SVM (Support Vector Machine), Bayesian Classifiers and C4.5). A latest's rate techniques in this field depicted in [8].Some plausible strategies and technique we recommended incorporates the clinical information institutionalization, examination and the information sharing over the related industries to improve the precision & viability of information mining applications in social insurance. [5] It is likewise prudent to investigate the utilization of content digging and picture digging for extension the nature and extent of information mining applications in medicinal services part. Information mining application can likewise be investigated on computerized indicative pictures for application viability. Some advancement has been made in these areas. [6][7].

There is a lot of data put away in stores that can be utilized viably to guide a medical practitioners in decision making in human services. This brings up an essential issue:

"By what means would we be able to transform information into helpful data that can empower medicinal services practitioners to settle on viable clinical decision?" This is the primary goal of this research.

2. Background

In late time, numerous associations in human services division utilizes data mining applications seriously and broadly on substantial scale. In information mining we can utilize diverse master cess and innovation to change this colossal measures of information into helpful data for solid and exact choice making. Another reason is that the social insurance exchanges created by this part are excessively voluminous and perplexing, making it impossible to be broke down and prepared by customary systems. Choice using so as to make can be enhanced majorly by using mining applications in finding patterns and examples in substantial volumes of ordinary data.[1] In late patterns investigation on these extensive dataset has gotten to be fundamental because of monetary weights on medicinal services commercial enterprises. This separated data can be utilized for choices making taking into account the relapse examination of restorative and money related information. Learning extraction can impact industry working proficiency, income and expense maintaining so as to utilize learning disclosure from database with at most care[2].Research demonstrates that on the off chance that we utilizes information mining applications as a part of social insurance organizations then these associations would be in better position to meet their fleeting objectives and long haul needs, Benko and Wilson argue.[3] We can get extremely valuable results from human services crude information by changing crude information into helpful data. An extraordinary reason that empowers analysts in this field is that this is exceptionally helpful for all partner included in the human services division. Like, in the event that we consider Insurance supplier, they can identify misuse and extortion, expert in human services can pick up help with choices making, similar to in client relationship administration. Social insurance suppliers (doctor's facilities, doctor, test research centres and patient and so forth.) can likewise utilize information mining applications in their separate master zone for master choice finding so as to make for instance, best practices and right & compelling medicines.

3. Heart Disease risk level prediction

The Heart disease database contains the screening clinical information of heart patients. At first, the database pre-processed to make the mining handle more able.

Database Details

(a) Database Creators: V.A. Therapeutic Center, Long Beach and Cleveland Clinic Foundation: Robert Detrano, M.D., Ph.D.
(b) Database Donor: David W. Aha (aha@ics.uci.edu) (714) 856-8779
The registry contains a database related with coronary disease. Data can be collected from uci.
Cleveland Clinic Foundation (Cleveland. Data) [12].

Inputs attributes

Age, Sex, Chest Pain, Resting blood pressure, Serum cholesterol, Fasting blood sugar, Resting electrocardiographic results, Maximum heart rate achieved, Exercise induced angina, ST depression, Slope of the peak exercise ST segment, Number of major vessels colored by fluoroscopy and thal.

Outputs class attribute

num (the predicted attribute)
The proposed study used covering rules model for classification (taking into account decision trees) as C4.5Rules [10], [11], [13] on the pre-processed database and discover the created rule sets with various need. Additionally pruned and ordered standards are also calculated.
We have utilized WEKA device [15] for dataset examination and KEEL [13],[14] to discover the order choice principles.

4. KEEL Experiment Implementation:

KEEL (Knowledge Extraction based on Evolutionary Learning) is being utilized for implementation. KEEL is an open source (GPLv3) Java programming apparatus to implement developmental process for Data Mining issues.
In the proposed study an implementation is being done using the dataset from Cleveland [12].In the pre-processing stage an AllPossible-MV [13][14] algorithm used for calculation to fill the missing values in the data set.
Missing Values Handling-for each instance in the data set, the presence of missing data is tested, and if exists any, all seen values of the attribute are imputed, resulting in 1 or more instances

5. Classification Decision Rules generated in our experiments:

The decision tree is constructed top-down. In each step a test for the actual node is chosen (starting with the root node), which best separates the given examples by classes.

A hill climbing algorithm is then performed in order to find the best subset of rules (according to the MDL heuristic).

PARAMETERS

- Confidence: is the confidence level. It is a float value that determines what is the minimal confidence that must has a leaf in order to be considered in the tree. Confidence value for our study is 0.25

- MinItemsets: is the minimum number of item-sets per leaf. It is an integer value that determines how much data instances must contain a leaf in order to be created. This value is 2 for our study

- Threshold: determines which algorithm to use in order to find the best subset of rules. For rule sets with sizes under the threshold, an exhaustive algorithm is performed; for sets above the threshold, a hill climbing algorithm is used. We have set the Threshold value to 10.

The rules generated by our experiments:

5.1 Original Rules:

i. If MHR<=3 and ST depression <=2.1 and Chest Pain <=3 ➔ 0
ii. If (MHR<=3 and ST depression <=2.1 and Chest Pain >3 and ca <=0 ➔ 0
iii. If (MHR<=3 and ST depression <=2.1 and Chest Pain >3 and ca >0 and serum cholesterol <=282) ➔ 1

iv. If (MHR<=3 and ST depression <=2.1 and Chest Pain >3 and ca >0 and serum cholesterol >282) ➔ 0
v. If (MHR<=3 and ST depression >2.1 and exercise induced angina <=0 and slope of the peak exercise ST segment <=2) ➔ 1
vi. If (MHR<=3 and ST depression >2.1 and exercise induced angina <=0 and slope of the peak exercise ST segment >2) ➔ 0
vii. If (MHR<=3 and ST depression >2.1 and exercise induced angina >0 and ca <=0) ➔ 0
viii. If (MHR<=3 and ST depression >2.1 and exercise induced angina >0 and ca >0) ➔ 3
ix. If (MHR>3 and Sex <=0 and serum cholesterol <=295) ➔ 3
x. If (MHR>3 and Sex <=0 and serum cholesterol >295) ➔ 1
xi. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR<=6 and resting blood pressure <=135) ➔ 0
xii. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR<=6 and resting blood pressure >135) ➔ 2
xiii. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol <=205 and Age<=54) ➔ 3
xiv. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol <=205 and Age>54) ➔ 2
xv. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting blood pressure <=114 and Age<=53) ➔ 1
xvi. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting blood pressure <=114 and Age>53) ➔ 2
xvii. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting blood pressure >114 and serum cholesterol <=266 and ca <=0) ➔ 0
xviii. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting blood pressure >114 and serum cholesterol <=266 and ca >0 and Age<=61) ➔ 1
xix. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting blood pressure >114 and serum cholesterol <=266 and ca >0 and Age>61) ➔ 0
xx. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting blood pressure >114 and serum cholesterol >266 and resting electrocardiographic<=1) ➔ 1
xxi. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting blood pressure >114 and serum cholesterol >266 and resting electrocardiographic>1) ➔ 3

5.2 Pruned Rules:

i. If (MHR<=3 and ST depression <=2.1 and Chest Pain <=3) ➔ 0
ii. If (MHR<=3 and ST depression <=2.1 and ca <=0) ➔ 0
iii. If (MHR<=3 and ST depression <=2.1 and Chest Pain >3 and ca >0 and serum cholesterol <=282) ➔1
iv. If (MHR<=3 and ST depression <=2.1) ➔ 0
v. If (MHR<=3 and ST depression >2.1 and exercise induced angina <=0 and slope of the peak exercise ST segment <=2) ➔ 1
vi. If (MHR<=3 and exercise induced angina <=0) ➔ 0
vii. If (MHR<=3 and ca <=0) ➔ 0
viii. If (MHR<=3 and ST depression >2.1 and exercise induced angina >0 and ca >0) ➔ 3
ix. If (MHR>3 and Sex <=0 and serum cholesterol <=295) ➔ 3
x. If (MHR>3 and serum cholesterol >295) ➔1
xi. If (ST depression <=2.4 and fasting blood sugar <=0 and MHR<=6) ➔ 0
xii. If (MHR>3 and fasting blood sugar <=0 and MHR<=6 and resting blood pressure >135) ➔ 2
xiii. If (fasting blood sugar <=0 and MHR>6 and serum cholesterol <=205 and Age<=54) ➔ 3

xiv. If (Sex >0 and fasting blood sugar <=0 and MHR>6 and serum cholesterol <=205 and Age>54) ➔ 2
xv. If (ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and Age<=53) ➔ 1
xvi. If (ST depression <=2.4 and MHR>6 and resting blood pressure <=114 and Age>53) ➔2
xvii. If (ST depression <=2.4 and resting blood pressure >114 and serum cholesterol <=266 and ca <=0) ➔0

5.3 Rules without duplicates:

i. If (MHR<=3 and ST depression <=2.1 and Chest Pain <=3) ➔ 0
ii. If (MHR<=3 and ST depression <=2.1 and ca <=0) ➔0
iii. If (MHR<=3 and ST depression <=2.1 and Chest Pain >3 and ca >0 and serum cholesterol <=282) ➔ 1
iv. If (MHR<=3 and ST depression <=2.1) ➔ 0
v. If (MHR<=3 and ST depression >2.1 and exercise induced angina <=0 and slope of the peak exercise ST segment <=2) ➔ 1
vi. If (MHR<=3 and exercise induced angina <=0) ➔ 0
vii. If (MHR<=3 and ca <=0) ➔ 0
viii. If (MHR<=3 and ST depression >2.1 and exercise induced angina >0 and ca >0) ➔ 3
ix. If (MHR>3 and Sex <=0 and serum cholesterol <=295) ➔ 3
x. If (MHR>3 and serum cholesterol >295) ➔ 1
xi. If (ST depression <=2.4 and fasting blood sugar <=0 and MHR<=6) ➔ 0
xii. If (MHR>3 and fasting blood sugar <=0 and MHR<=6 and resting blood pressure >135) ➔ 2
xiii. If (fasting blood sugar <=0 and MHR>6 and serum cholesterol <=205 and Age<=54) ➔ 3
xiv. If (Sex >0 and fasting blood sugar <=0 and MHR>6 and serum cholesterol <=205 and Age>54) ➔ 2
xv. If (ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and Age<=53) ➔ 1
xvi. If (ST depression <=2.4 and MHR>6 and resting blood pressure <=114 and Age>53) ➔ 2
xvii. If (ST depression <=2.4 and resting blood pressure >114 and serum cholesterol <=266 and ca <=0) ➔ 0
xviii. If (MHR>3 and ST depression <=2.4 and fasting blood sugar <=0 and serum cholesterol >205 and resting blood pressure >114 and serum cholesterol <=266 and ca >0 and Age<=61) ➔ 1
xix. If (ST depression <=2.4 and serum cholesterol <=266 and Age>61) ➔ 0
xx. If (ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting electrocardiographic<=1) ➔ 1
xxi. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and serum cholesterol >266 and resting electrocardiographic>1) ➔3
xxii. If (MHR>3 and Sex >0 and ST depression >2.4 and MHR>124 and Chest Pain >3 and exercise induced angina >0) ➔ 2
xxiii. If (MHR>3 and Sex >0 and MHR>124 and Chest Pain >3 and resting electrocardiographic>1 and resting blood pressure <=142) ➔1

5.4 *Classified Rules:*

Rule Set 0:
i. If (MHR<=3 and ST depression <=2.1 and Chest Pain <=3) ➔ 0 ➔t:10.1312574862847
ii. If (MHR<=3 and ST depression <=2.1 and ca <=0) ➔0 ➔t:10.1312574862847
iii. If (MHR<=3 and ST depression <=2.1) ➔ 0 ➔t:6.762569063545247
iv. If (MHR<=3 and exercise induced angina <=0) ➔0 ➔t:3.5897011441351196
v. If (MHR<=3 and ca <=0) ➔ 0 ➔t:4.1217701197746
vi. If (ST depression <=2.4 and fasting blood sugar <=0 and MHR<=6) ➔0 ➔t:10 31158272161852

vii. If (ST depression <=2.4 and resting blood pressure >114 and serum cholesterol <=266 and ca <=0) ➔ 0 ➔t:16.840367272392715
viii. If (ST depression <=2.4 and serum cholesterol <=266 and Age>61) ➔0 ➔t:13 9746836919829
ix. If (fasting blood sugar >0 and exercise induced angina <=0 and resting blood pressure <=156) ➔ 0 ➔t:9.75855933260464

Rule Set 1:

i. If (MHR<=3 and ST depression <=2.1 and Chest Pain >3 and ca >0 and serum cholesterol <=282) ➔ 1 ➔t:19.648509913498945
ii. If (MHR<=3 and ST depression >2.1 and exercise induced angina <=0 and slope of the peak exercise ST segment <=2) ➔ 1 ➔t:12.888640199119738
iii. If (MHR>3 and serum cholesterol >295) ➔ 1 ➔t:8.130109337967596
iv. If (ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and Age<=53) ➔ 1
v. ➔t:20.28699049907931
vi. If (MHR>3 and ST depression <=2.4 and fasting blood sugar <=0 and serum cholesterol >205 and resting blood pressure >114 and serum cholesterol <=266 and ca >0 If (ST depression <=2.4 and fasting blood sugar <=0 and MHR>6 and serum cholesterol >205 and resting electrocardiographic<=1) ➔1 ➔t:19.575519434255018

Rule Set 2:

i. If (MHR>3 and fasting blood sugar <=0 and MHR<=6 and resting blood pressure >135) ➔2➔t:12.619989401751669
ii. If (Sex >0 and fasting blood sugar <=0 and MHR>6 and serum cholesterol <=205 and Age>54) ➔2➔t:19.4758425613074
iii. If (ST depression <=2.4 and MHR>6 and resting blood pressure <=114 and Age>53) ➔ 2 ➔t:15.441499694596182
iv. If (ST depression <=2.4 and fasting blood sugar >0 and exercise induced angina >0) ➔ 2 ➔t:9.979316860749648
v. If (MHR>3 and Sex >0 and ST depression >2.4 and MHR>124 and Chest Pain >3 and exercise induced angina >0) ➔ 2 ➔t:21.681708712080138

Rule Set 3:

i. If (MHR<=3 and ST depression >2.1 and exercise induced angina >0 and ca >0) ➔3 ➔t:12.952068134551817
ii. If (MHR>3 and Sex <=0 and serum cholesterol <=295) ➔3 ➔t:12 26913357183703
iii. If (fasting blood sugar <=0 and MHR>6 and serum cholesterol <=205 and Age<=54) ➔ 3 ➔t:16.18425003790777
iv. If (MHR>3 and Sex >0 and ST depression <=2.4 and fasting blood sugar <=0 and serum cholesterol >266 and resting electrocardiographic>1) ➔ 3 ➔t:22.695159196669785
v. If (MHR>3 and resting blood pressure >156) ➔3 ➔t:6.612535146046203
vi. If (Sex >0 and ST depression >2.4 and MHR<=124) ➔ 3 ➔t:12 87169988355576

Rule Set 4:

i. If (MHR>3 and ST depression >2.4 and MHR>124 and Chest Pain <=3) ➔ 4 ➔t:15.570914546225168
ii. If (MHR>3 and Sex >0 and ST depression >2.4 and exercise induced angina <=0) ➔ 4 ➔t:12.823786323183375

5.5 Polish:

Class wise Rule distribution

0- Rule Set: 3

i. If (MHR>3 and Sex <=0 and serum cholesterol <=295) ➔ 3
ii. If (Sex >0 and ST depression >2.4 and MHR<=124) ➔ 3

1- Rule Set: 2

i. If (ST depression <=2.4 and fasting blood sugar >0 and exercise induced angina >0) ➔ 2

2- Rule Set: 4

i. If (MHR>3 and Sex >0 and ST depression >2.4 and exercise induced angina <=0) ➔ 4

3- Rule Set: 0

i. If (MHR<=3 and ST depression <=2.1 and Chest Pain <=3) ➔ 0

ii. If (MHR<=3 and ca <=0) ➔ 0

iii. If (fasting blood sugar >0 and exercise induced angina <=0 and resting blood pressure <=156) ➔ 0

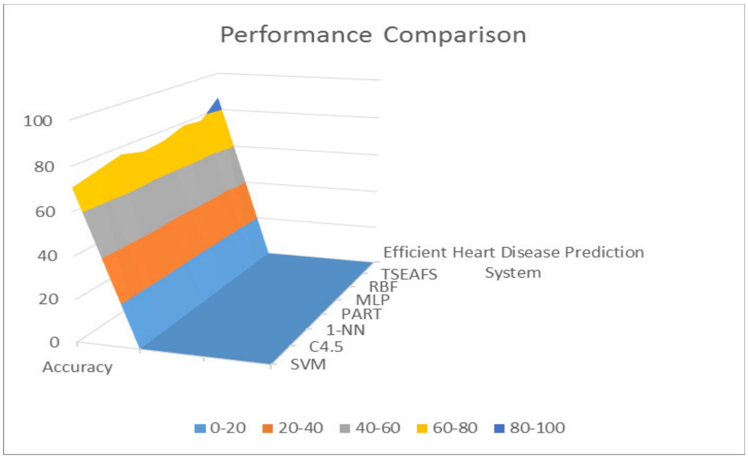
4- Rule Set: 1

i. If (MHR>3 and serum cholesterol >295) ➔ 1

6. Performance Evaluation

The performance of various well known algorithms on Heart Disease data set [12] is listed in Table 1 and it shows that Efficient Heart Disease Prediction System have the better accuracy than other given classifiers.

The Algorithm Used	SVM	C4.5	1-NN	PART	MLP	RBF	TSEAFS	**Efficient Heart Disease Prediction System**
Accuracy (%)	70.59	73.53	76.47	73.53	74.85	78.53	77.45	**86.7**



7. Conclusion

In this research paper, we have presented an Efficient Heart Disease Prediction System using data mining. This system can help medical practitioner in efficient decision making based on the given parameter. We have train and test the system using 10 fold method and find the accuracy of 86.3 % in testing phase and 87.3 % in training phase and because this model demonstrates the better results and helps the area specialists and even individual related with the field to get ready for a superior determine and give the patient to have early determination results as it performs sensibly well even without retraining.

References

1. Biafore, S. (1999). Predictive solutions bring more power to decision makers. Health Management Technology, 20(10), 12-14. 2

2. Silver, M. Sakata, T. Su, H.C. Herman, C. Dolins, S.B. & O'Shea, M.J. (2001). Case study: how to apply data mining techniques in a healthcare data warehouse. Journal of Healthcare Information Management, 15(2), 155-164.
3. Benko, A. & Wilson, B. (2003). Online decision support gives plans an edge. Managed Healthcare Executive, 13(5), 20
4. Heon Gyu Lee, Ki Yong Noh, Keun Ho Ryu, "Mining Biosignal Data: Coronary Artery Disease Diagnosis using Linear and Nonlinear Features of HRV," LNAI 4819:
5. Cody, W.F. Kreulen, J.T. Krishna, V. & Spangler, W.S. (2002). The integration of business intelligence and knowledge management. IBM Systems Journal, 41(4), 697-713
6. Ceusters, W. (2001). Medical natural language understanding as a supporting technology for data mining in healthcare. In Medical Data Mining and Knowledge Discovery, Cios, K. J. (Ed.), PhysicaVerlag Heidelberg, New York, 41-69.
7. Megalooikonomou, V. & Herskovits, E.H. (2001). Mining structure function associations in a brain image database. In Medical Data Mining and Knowledge Discovery, Cios, K. J. (Ed.), Physica-Verlag Heidelberg, New York, 153-180.
8. Chhikara, S & Sharma,P Data Mining Techniques on Medical Data for Finding Locally Frequent Diseases, I JRASET 2014,PP 396-402.
9. 12Tallón-s, Antonio J., César Hervás- Martínez, JoséC. Riquelme, and Roberto Ruiz. (2013)"Feature selection to enhance a two-stage evolutionary algorithm in product unit neural networks for complex classification problems",Neurocomputing J.R. Quinlan. 1993, C4.5: Programs for Machine Learning. Morgan Kauffman Publishers, San Mateo-California.
10. J.R. Quinlan. 1995, MDL and Categorical Theories (Continued). In Machine Learning: Proceedings of the Twelfth International Conference. Lake Tahoe, California. Morgan Kaufmann, , 464-470.
11. Eibe Frank and Ian H. Witten. 1998 Generating accurate rule sets without global optimization. In Proc 15th International Conference on Machine Learning, Madison, Wisconsin, pages 144-151. Morgan Kaufmann.
12. Heart attack dataset from http://archive.ics.uci.edu/ml/datasets/Heart Disease
13. J. Alcalá-Fdez, L. Sánchez, S. García, M.J. del Jesus, S. Ventura, J.M. Garrell, J. Otero, C. Romero, J. Bacardit, V.M. Rivas, J.C. Fernández, F. Herrera, 2009 KEEL: A Software Tool to Assess Evolutionary Algorithms to Data Mining Problems. Soft Computing 307-318
14. J. Alcalá-Fdez, A. Fernandez, J. Luengo, J. Derrac, S. García, L. Sánchez, F. Herrera. KEEL Data-Mining Software Tool: Data Set Repository, Integration of Algorithms and Experimental Analysis Framework. Journal of Multiple-Valued Logic and Soft Computing 17:2-3 (2011) 255-287
15. Mark Hall, Eibe Frank, Geoffrey Holmes, Bernhard Pfahringer, Peter Reutemann, Ian H. Witten 2009; The WEKA Data Mining Software: An Update; SIGKDD Explorations, Volume 11, Issue 1.
16. Sharma Purushottam, Dr Kanak Saxena, Richa Sharma" Efficient Heart Disease Prediction System using Decision Tree" in IEEE International Conference on Computing Communication and Automation (ICCCA-2015),May 2015
17. Sharma Purushottam, Dr Kanak Saxena, Richa Sharma" Heart Disease Prediction System Evaluation Using C4.5 Rules and Partial Tree" in Springer, Computational Intelligence in Data Mining, 2015, pp-285-294, DOI 10.1007/978-81-322-2731-1_26.