DOI: http://dx.doi.org/10.26483/ijarcs.v9i1.5257

Volume 9, No. 1, January-February 2018

ISSN No. 0976-5697

International Journal of Advanced Research in Computer Science

RESEARCH PAPER

Available Online at www.ijarcs.info

IMPROVING ACCURACY OF FUZZY RULE BASED MINING FOR HEART DISEASE DETECTION USING COST MATRIX

P. Sambasiva Rao
Ph.D. Scholar,
GITAM University
Visakhapatnam, India

Dr. T. Uma Devi
Associate Professor
GITAM University
Visakhapatnam, India

Abstract: Improvement in healthcare system and improvements in the pharmaceutical domain are capable of curing multiple life threatening diseases now a days. Nevertheless, the appropriate time for detection and accuracy in the detection method is the major bottleneck for preventing loss of life. One of the major life threatening diseases is the coronary disease, where the heart is tend to malfunction and cause human death. Correct detection in early stages of the disease can prevent stopping of the heart function, thus saving of the patient life. A number research attempts are made to detect the heart disease. The demonstrations of the predictive analysis of the patient's heart condition based on various body conditions are executed. Yet, the reports are biased with true negative results from those researchers impacting the wrong medication for the affected due to imperfect analysis. This ill medication is causing the side effects and effects of other severe diseases in the patient's body. Hence, the demand of the modern research is to improve the accuracy of the predictive analysis. This work analyses the popular predictive determination using fuzzy rule based methods of heart disease and apply novel cost based matrix to improve the accuracy of detection.

Keywords: Fuzzy Rules, Cost Matrix, Bayesian Rule, Clustering, Improved Correctness

1. INTRODUCTION

The core approaches of using fuzzy rules as demonstrated by various research approaches are to control the delinquents. The recent use of if –then – else fuzzy rules to classify the data is becoming highly popular [1]. The modern advancements have demonstrated the successful use of the automation in generating the fuzzy rules from the datasets and address the pattern based classification problems. [2] [3] [4] [5] [6].

The remarkable work by Zadeh et al. in demonstrating the linguistic statement [7] is a milestone in the research. Also yet another demonstration by Hu Y et al. in introducing trapezoidal membership function for evaluating the numeric values in fuzzy classifications [8] [9].

The principle approach in determining the structure of the classifier is to generate a set of master rules to be decided. The master rule sets are generally provided by the expert systems for each and every attributes in the datasets [10]. Furthermore, in order to improve the classification, the popular technique is to deploy weight based rules. The notable research by Mansoori E. G. et al. [11] has demonstrated the use of weight based rules alongside with the fuzzy rules. Conversely, the studies also made the point for further possibilities in improvement scopes. Thus, this work made an attempt to increase the classification accuracy by introducing cost matrix on the expert system.

The rest of the paper is furnished such that in section – II, the survey on the cost based matrix approaches are analysed, in section – III the fundamentals of rule based mining are demonstrated, in section – IV the novel cost based improvements are demonstrated, the data information is demonstrated in the section – V, results are been discussed in the section - VI and the work presents the conclusion in section – VII.

2. REVIEW OF THE LITERATURE

The total numbers of fuzzy rules accumulated by the classifier for each attributes are classified into m number of subsets and the pattern classification is defined as n^m [12]. This work proposes a weight matrix for each classification categories in order to reduce the wrong classifications. The method deploys an error correcting process and reduces the errors in the result.

The heart diseases are the highest rated mortality causing disease worldwide as recommended by Goldberg R. J. [13]. The notable research outcomes by Hatmi Z. N. et al. [14] have demonstrated the effect of sex, age, high blood pressure, diabetic conditions, obesity and smoking have significant effects on the heart diseases. It is natural to understand that the effect of sex and smoking are obvious whereas the effects of the other attributes are fuzzy. There are advanced methods demonstrated by various researchers to predict and detect the heart conditions as angiographic [15]. Ephzibah E. P. et al. [16] and Lahsasna A. et al [17] have demonstrated the global possibilities of applying the angiography methods for the same. Conversely, the associated cost for the method is not cheap and practically cannot cover the major populations of the world, especially the underdeveloped countries. Considering the same fact, Marateb H. R. et al. [17] has proposed multiple alternative techniques such as stress test or Single Photon Emission Computed Tomography or Echo.

It is to be considered that, the medical methods are not precise and need an additional support to enhance the results in terms of enhancing the results. The uses of predictive approaches to enhance the reports are obvious and demonstrated the improvements [19] [20] [21]. The very popular researches by Cachin F. et al [22] are a considerable mile stone in this direction. Also the outcome of the research by Peters R. M. et al. [23] where it is clearly demonstrated that the combinative effect of clinical and computing methods are successful in delivering the higher accuracy.

3. RULE BASED MINING

In this section of the work, the understanding about the rule based classification is presented. Generally a rule based classification proposes a set of rules to be deployed in a specific order to classify the datasets. The operations of the classifier are to be incremental, where the generic processes are concatenation, union and difference. This work establishes a lemma to make the understanding strong.

Lemma: Considering a set of classifiers as “S”. Every subset of the classifier considered to be “s”. Every rule sets in all the subjects of the classifier are to be considered as a tuple which again can be denoted as {R, r}. “R” denotes the set of rules and “r” denotes the order of the rules to be deployed.

Assuming that, “s” is a classifier and “R” is a rule set, then “s” is equal to “r” and “r” is equal to “R”.

Proof:
Firstly, let it be assumed that there are two classifiers as s_1 and s_2, as following:

Where C_1 and C_2 are the Clusters

$$C_1 = \{R_1, r_1\} \quad \text{(Eq. 1)}$$

And

$$C_2 = \{R_2, r_2\} \quad \text{(Eq. 2)}$$

Secondly, the assumptions are also made as

$$C_1 \subseteq C_2 \quad \text{(Eq. 3)}$$

Hence, it is natural to understand as,

$$R_1 \subseteq R_2 \quad \text{(Eq. 4)}$$

And, subsequently

$$r_1 \subseteq r_2 \quad \text{(Eq. 5)}$$

Finally, considering “D” is the data set to be analysed by the same classifier, the three categories for the outcomes to be delivered by the classifier are as followings:

$$< s, D >= d \mid R = r \quad \text{(Eq. 6)}$$

$$< s, D >= d \mid R > r \quad \text{(Eq. 7)}$$

And,

$$< s, D >= d \mid R < r \quad \text{(Eq. 8)}$$

In the above three cases, the order of the rule set “r” is clearly controlling the total outcome for the classifier. Thus is natural to understand the key value of the order for correct classification.

4. COST BASED IMPROVEMENT

In this section, the work presents the composite knowledge on the cost based improvement of the classification results. The errors in classifications can be observed in many real world problems and the effects are significant. The generic solution for the problem is to deploy specific rule sets to classify the data. However, the effect of specific rule based mining is the method will lose the generalization property. Nevertheless, most of the real world classification situations are divided into two major categories as positive and negative. Often the positive class is understood as major category and the negative class is understood as minority class. The method of cost based matrix can significantly reduce the false negative classifications in the situation.

This work, briefly explains the use of cost matrix to reduce the class imbalance. First the example of classification types are understood [Table – 1].

TABLE I: CLASSIFICATION AND IMBALANCE

Class – A	**Class – B**
True Positive	False Positive
False Negative	True Negative

Henceforth, the applicability for the cost matrix is realized in here [Table – 2].

TABLE II: CLASSIFICATION AND IMBALANCE

Class – A	**Class – B**
P	Q
R	S

It is natural to understand that, the purpose of the cost matrix is to reduce the improper classification and classification errors. Hence, the costs are to be applied for wrong classification, rather the correct classifications.

Thus, the relationship between P, Q, R and S are to be realized as

$$P < Q \quad \text{(Eq. 9)}$$

And

$$R > S \quad \text{(Eq. 10)}$$

Where,

$$P = S \rightarrow 0 \quad \text{(Eq. 11)}$$

Also,

$$Q \rightarrow R \rightarrow \phi \quad \text{(Eq. 12)}$$

Here, ϕ is the optimal cost or weight for the classification.

5. DATASET INFORMATION

This work demonstrates the results and discussions on the widely popular UCI heart disease dataset [24] [25] [26] [27]. This database contains 76 attributes, but all published experiments refer to using a subset of 14 of them. In particular, the Cleveland database is the only one that has been used by ML researchers to this date. The "goal" field refers to the presence of heart disease in the patient. It is integer valued from 0 (no presence) to 4. Experiments with the Cleveland database have concentrated on simply attempting to distinguish presence (values 1,2,3,4) from absence (value 0). The names and social security numbers of the patients were recently removed from the database, replaced with dummy values. The dataset is contributed and periodically updated by Hungarian Institute of Cardiology. Budapest: AndrasJanosi, M.D., University Hospital, Zurich, Switzerland: William Steinbrunn, M.D. and University Hospital, Basel, Switzerland: Matthias Pfisterer, M.D.

The notable works by P.Sambasiva Rao &Dr. T. Uma Devi [28] demonstrated the regularity/ regularizes the outcomes of predictive analyses by reducing the dataset size. As an outcome, the analysis time complexity has significantly reduced. The reduced attribute set is listed here [Table – 3].

TABLE III: LIST OF ATTRIBUTES AND DESCRIPTION

Attribute	Item Purchased Together/ Description
ID	Patient identification number
AGE	Age in years
SEX	Patient Sex, Male or Female
PAINLOC	Chest pain location
CP	Chest pain type • Value 1 typical angina • Value 2 atypical angina • Value 3 non-angina pain • Value 4 asymptomatic
SMOKE	Is or is not a smoker
CIGS	Cigarettes per day
YEARS	Number of years as a smoker
CA	Number of major vessels (0-3) colour by fluoroscopy
NUM	Denotes the severity of the heart disease

This work analyses the improvements based on the reduced attribute set and furnishes in the results and discussion section.

6. RESULTS AND DISCUSSION

In this section of the work, the cost best improvements are been demonstrated.

Firstly, the Cleveland Dataset is been analysed. The attribute influence are been recorded here [Table – 4]

TABLE IV: ATTRIBUTE INFLUENCE IN THE CLEVELAND DATASET

Attribute Name : AGE					
Mean	52.7312	55.1040	57.9484	55.8750	59.8000
std. dev.	9.4636	7.9702	6.9918	8.0383	9.1761
weight sum	157.0000	50.0000	31.0000	32.0000	12.0000
Precision	1.2000	1.2000	1.2000	1.2000	1.2000
Attribute Name : SEX					
mean	0.5478	0.8800	0.8065	0.8125	0.8333
std. dev.	0.4977	0.3250	0.3951	0.3903	0.3727
weight sum	157.0000	50.0000	31.0000	32.0000	12.0000
precision	1.0000	1.0000	1.0000	1.0000	1.0000
Attribute Name : PAINLOC					
mean	-9.0000	-9.0000	-9.0000	-9.0000	-9.0000
std. dev.	0.0017	0.0017	0.0017	0.0017	0.0017
weight sum	157.0000	50.0000	31.0000	32.0000	12.0000
precision	0.0100	0.0100	0.0100	0.0100	0.0100
Attribute Name : CP					
mean	2.8217	3.4000	3.7097	3.7500	3.6667
std. dev.	0.9204	0.9592	0.6812	0.5590	0.8498
weight sum	157.0000	50.0000	31.0000	32.0000	12.0000
precision	1.0000	1.0000	1.0000	1.0000	1.0000
Attribute Name : SMOKE					
mean	-9.0000	-9.0000	-9.0000	-9.0000	-9.0000
std. dev.	0.0017	0.0017	0.0017	0.0017	0.0017
weight sum	157.0000	50.0000	31.0000	32.0000	12.0000
precision	0.0100	0.0100	0.0100	0.0100	0.0100
Attribute Name : CIGS					
mean	14.6960	21.3055	15.8358	12.4261	24.1364
std. dev.	19.0934	19.8223	17.1743	18.6926	19.2707
weight sum	157.0000	50.0000	31.0000	32.0000	12.0000
precision	4.9091	4.9091	4.9091	4.9091	4.9091
Attribute Name : YEARS					
mean	13.6325	18.1849	17.3017	10.6951	24.8311
std. dev.	15.2531	14.4657	16.1422	15.8374	14.4221
weight sum	157.0000	50.0000	31.0000	32.0000	12.0000
precision	1.7027	1.7027	1.7027	1.7027	1.7027
Attribute Name : CA					
mean	0.1338	0.3600	0.9677	1.6875	1.5000
std. dev.	1.0350	1.7636	1.4024	1.4882	1.5000
weight sum	157.0000	50.0000	31.0000	32.0000	12.0000
Precision	3.0000	3.0000	3.0000	3.0000	3.0000

Furthermore, the summary of the Naïve Bayes classification is furnished [Table – 5].

TABLE V: SUMMARY OF NAÏVE BAYES

Correctly Classified Instances (%)	**Incorrectly Classified Instances (%)**
54.2553	45.7447

Here, the cost matrix is been listed [Table – 6].

TABLE VI: SUMMARY OF COST MATRIX

	Class 0	**Class 1**	**Class 2**	**Class 3**	**Class 4**
Class 0	0	5.0	5.0	5.0	5.0
Class 1	5.0	0	5.0	5.0	5.0
Class 2	5.0	5.0	0	5.0	5.0
Class 3	5.0	5.0	5.0	0	5.0
Class 4	5.0	5.0	5.0	5.0	0

Furthermore, the summary of the CostSensitiveClassifier classification is furnished [Table – 7].

TABLE VII: SUMMARY OF COST SENSITIVE CLASSIFIER

Correctly Classified Instances (%)	**Incorrectly Classified Instances (%)**
55.6738	44.3262

Secondly, the Hungarian Dataset is been analysed. The attribute influence are been recorded here [Table – 8]

TABLE VIII: ATTRIBUTE INFLUENCE IN THE HUNGARIAN DATASET

Attribute Name : AGE					
mean	47.0575	47.0212	49.5343	52.7085	50.6667
std. dev.	7.9235	8.5799	6.6821	5.8682	5.8660
weight sum	188.0000	37.0000	26.0000	28.0000	15.0000
Precision	1.0270	1.0270	1.0270	1.0270	1.0270
Attribute Name : SEX					
mean	0.6330	0.8649	0.9615	0.8929	0.8000
std. dev.	0.4820	0.3419	0.1923	0.3093	0.4000
weight sum	188.0000	37.0000	26.0000	28.0000	15.0000
Precision	1.0000	1.0000	1.0000	1.0000	1.0000
Attribute Name : PAINLOC					
mean	0.8989	0.9459	0.9615	0.9643	1.0000
std. dev.	0.3014	0.2261	0.1923	0.1856	0.1667
weight sum	188.0000	37.0000	26.0000	28.0000	15.0000
precision	1.0000	1.0000	1.0000	1.0000	1.0000
Attribute Name : CP					
mean	2.6170	3.6216	3.8077	3.3929	3.8000
std. dev.	0.8581	0.7480	0.6214	0.9761	0.5416
weight sum	188.0000	37.0000	26.0000	28.0000	15.0000
precision	1.0000	1.0000	1.0000	1.0000	1.0000
Attribute Name : SMOKE					
mean	-9.5213	-9.4595	-9.6154	-10.0000	-10.0000
std. dev.	2.1350	2.2612	1.9231	0.8333	0.8333
weight sum	188.0000	37.0000	26.0000	28.0000	15.0000
precision	5.0000	5.0000	5.0000	5.0000	5.0000
Attribute Name : CIGS					
mean	0.3670	0.0000	0.0000	0.0000	0.0000
std. dev.	11.5000	11.5000	11.5000	11.5000	11.5000
weight sum	188.0000	37.0000	26.0000	28.0000	15.0000
precision	69.0000	69.0000	69.0000	69.0000	69.0000
Attribute Name : YEARS					
mean	-9.0000	-9.0000	-9.0000	-9.0000	-9.0000
std. dev.	0.0017	0.0017	0.0017	0.0017	0.0017
weight sum	188.0000	37.0000	26.0000	28.0000	15.0000
precision	0.0100	0.0100	0.0100	0.0100	0.0100
Attribute Name : CA					
mean	-8.8085	-8.7568	-9.0000	-9.0000	-9.0000
std. dev.	1.5964	1.5000	1.5000	1.5000	1.5000
weight sum	188.0000	37.0000	26.0000	28.0000	15.0000
precision	9.0000	9.0000	9.0000	9.0000	9.0000

Furthermore, the summary of the Naïve Bayes classification is furnished [Table – 9].

TABLE IX: SUMMARY OF NAÏVE BAYES

Correctly Classified Instances (%)	**Incorrectly Classified Instances (%)**
47.619	52.381

The cost matrix is been listed same as [Table – 6].

Furthermore, the summery of the CostSensitiveClassifier classification is furnished [Table – 10].

TABLE X: SUMMERY OF COST SENSITIVE CLASSIFIER

Correctly Classified Instances (%)	**Incorrectly Classified Instances (%)**
63.9456	36.0544

Thirdly, the Switzerland Dataset is been analysed. The attribute influence are been recorded here [Table – 11]

TABLE XI: ATTRIBUTE INFLUENCE IN THE SWITZERLAND DATASET

Attribute Name : AGE					
mean	54.6875	55.5868	52.6823	58.3722	52.0333
std. dev.	9.8898	10.1162	7.3156	7.7123	5.0370
weight sum	8.0000	48.0000	32.0000	30.0000	5.0000
Precision	1.1667	1.1667	1.1667	1.1667	1.1667
Attribute Name : SEX					
mean	1.0000	0.8750	0.9063	0.9667	1.0000
std. dev.	0.1667	0.3307	0.2915	0.1795	0.1667
weight sum	8.0000	48.0000	32.0000	30.0000	5.0000
Precision	1.0000	1.0000	1.0000	1.0000	1.0000
Attribute Name : PAINLOC					
mean	0.5000	0.9375	0.8750	1.0000	1.0000
std. dev.	0.5000	0.2421	0.3307	0.1667	0.1667
weight sum	8.0000	48.0000	32.0000	30.0000	5.0000
Precision	1.0000	1.0000	1.0000	1.0000	1.0000
Attribute Name : CP					
mean	3.0000	3.6667	3.7500	3.8333	4.0000
std. dev.	0.7071	0.7169	0.7500	0.4534	0.1667
weight sum	8.0000	48.0000	32.0000	30.0000	5.0000
Precision	1.0000	1.0000	1.0000	1.0000	1.0000
Attribute Name : SMOKE					
mean	-10.0000	-8.1250	-8.4375	-7.6667	-6.0000
std. dev.	0.8333	3.9031	3.6309	4.2295	4.8990
weight sum	8.0000	48.0000	32.0000	30.0000	5.0000
precision	5.0000	5.0000	5.0000	5.0000	5.0000
Attribute Name : CIGS					
mean	0.0000	2.0417	3.0625	4.0833	4.9000
std. dev.	4.0833	8.4180	10.1572	11.1079	9.8000
weight sum	8.0000	48.0000	32.0000	30.0000	5.0000
precision	24.5000	24.5000	24.5000	24.5000	24.5000
Attribute Name : YEARS					
mean	0.0000	0.6146	0.0000	1.9667	0.0000
std. dev.	4.9167	4.9167	4.9167	10.5908	4.9167
weight sum	8.0000	48.0000	32.0000	30.0000	5.0000
precision	29.5000	29.5000	29.5000	29.5000	29.5000
Attribute Name : CA					
mean	-9.6250	-10.5417	-11.0000	-10.2667	-11.0000
std. dev.	3.6379	2.1981	0.9167	2.7439	0.9167
weight sum	8.0000	48.0000	32.0000	30.0000	5.0000
Precision	5.5000	5.5000	5.5000	5.5000	5.5000

Furthermore, the summary of the Naïve Bayes classification is furnished [Table – 12].

TABLE XII: SUMMERY OF NAÏVE BAYES

Correctly Classified Instances (%)	**Incorrectly Classified Instances (%)**
39.0244	60.9756

The cost matrix is been listed same as [Table – 6].

Furthermore, the summary of the CostSensitiveClassifier classification is furnished [Table – 13].

TABLE XIII: SUMMARY OF COST SENSITIVE CLASSIFIER

Correctly Classified Instances (%)	**Incorrectly Classified Instances (%)**
39.0244	60.9756

Henceforth, the final analysis of the classification task is furnished here [Table – 14].

TABLE XIV: SUMMARY OF IMPROVEMENTS

Naïve BayesClassification			
	Cleveland	**Hungarian**	**Switzerland**
Correctly Classified Instances (%)	54.2553	47.619	39.0244
Incorrectly Classified Instances (%)	45.7447	52.381	60.9756

Cost Sensitive Classification			
Correctly Classified Instances (%)	55.6738	63.9456	39.0244
Incorrectly Classified Instances (%)	44.3262	36.0544	60.9756

Finally, the results are been analysed graphically [Fig – 1] and [Fig – 2].

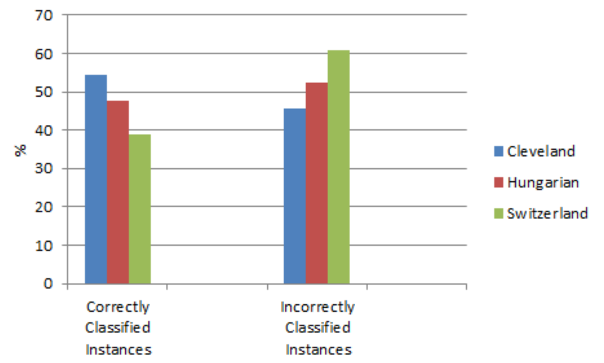

Fig.1Comparative Study of Naïve Bayes Classification

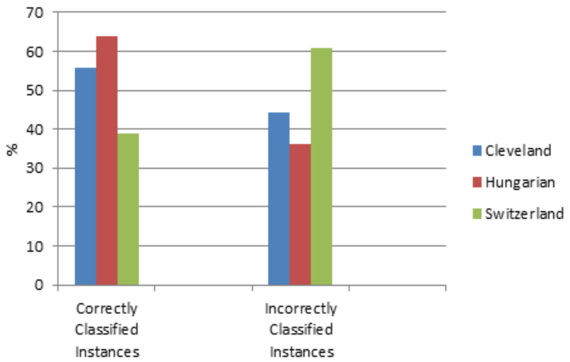

Fig.2 Comparative Study of Cost Sensitive Classification

I. Conclusion

Motivated by the current improvements in the field of pharmaceutical applicability and the supremacy of curing the life threating diseases, the need for the medical research is to improve the disease detection with the help of computing capabilities. Thus, a number of predictive approaches are been carried out by various researchers to detect and prevent the disease, especially the heart disease being highest mortality rated in growing countries. This work analyses the popular predictive techniques to detect the heart diseases in the early stages. Nevertheless, the accuracy is always been a challenge for the researches. Hence, this work introduces the cost matrix oriented fuzzy rule based predictive analysis to improve the results. The results are significant in terms of accuracy improvement and been demonstrated in three popular datasets from UCI source. The outcome of the research is a novel framework to improve the detection accuracy, which can be generalized targeting to save precious human life and made it available in low cost for the underdeveloped countries.

REFERENCES

[1] Aabdeh M. S., Habibi J., Soroush E. Induction of fuzzy classification systems via evolutionary ACO-BASED algorithms. International Journal of Simulation, Systems, Science and Technology. 2008;9(3):1–8.

[2] Nakashima T., Schaefer G., Yokota Y., Ishibuchi H. A weighted fuzzy classifier and its application to image processing tasks. Fuzzy Sets and Systems. 2007;158(3):284–294. doi: 10.1016/j.fss.2006.10.011.

[3] Mansoori E. G., Zolghadri M. J., Katebi S. D. Using distribution of data to enhance performance of fuzzy classification systems. Iranian Journal of Fuzzy Systems. 2007;4(1):21–36.

[4] Nakashima T., Yokota Y., Ishibuchi H., Bargiela A. Constructing fuzzy classification systems from weighted training patterns. Proceedings of the IEEE International Conference on Systems, Man and Cybernetics; October 2004; IEEE; pp. 2386–2391.

[5] Lekova A., Mikhailov L., Boyadjiev D., Nabout A. Redundant fuzzy rules exclusion by genetic algorithms. Fuzzy Sets and Systems. 1998;100(1–3):235–243. doi: 10.1016/s0165-0114(97)00056-0.

[6] Ishibuchi H., Murata T., Gen M. Performance evaluation of fuzzy rule-based classification systems obtained by multi-objective genetic algorithms. Computers and Industrial Engineering. 1998;35(3-4):575–578. doi: 10.1016/s0360-8352(98)00162-4.

[7] Zadeh L. A. Fuzzy sets. Information and Computation. 1965;8:338–353.

[8] Hu Y.-C., Chen R.-S., Tzeng G.-H. Finding fuzzy classification rules using data mining techniques. Pattern Recognition Letters. 2003;24(1–3):509–519. doi: 10.1016/s0167-8655(02)00273-8.

[9] Hu Y.-C. Sugeno fuzzy integral for finding fuzzy if-then classification rules. Applied Mathematics and Computation. 2007;185(1):72–83. doi: 10.1016/j.amc.2006.07.010.

[10] Jahromi M. Z., Taheri M. A proposed method for learning rule weights in fuzzy rule-based classification systems. Fuzzy Sets and Systems. 2008;159(4):449–459. doi: 10.1016/j.fss.2007.08.007.

[11] Mansoori E. G., Zolghadri M. J., Katebi S. D. A weighting function for improving fuzzy classification systems performance. Fuzzy Sets and Systems. 2007;158(5):583–591. doi: 10.1016/j.fss.2006.10.004.

[12] Nakashima T., Nakai G., Ishibuchi H. Improving the performance of fuzzy classification systems by membership function learning and feature selection. Proceedings of the IEEE International Conference on Fuzzy Systems (FUZZ-IEEE '02); May 2002; Honolulu, Hawaii, USA. pp. 488–493.

[13] Goldberg R. J., Goff D., Cooper L. S., Luepker R., Zapka J., Bittner V. Age and sex differences in presentation of symptoms among patients with acute coronary disease: the REACT trial. Coronary Artery Disease. 2000;11(5):399–407. doi: 10.1097/00019501-200007000-00004.

[14] Hatmi Z. N., Tahvildari S., GafarzadehMotlag A., SabouriKashani A. Prevalence of coronary artery disease risk factors in Iran: a population based survey. BMC Cardiovascular Disorders. 2007;7, article 32 doi: 10.1186/1471-2261-7-32.

[15] Pal D., Mandana K. M., Pal S., Sarkar D., Chakraborty C. Fuzzy expert system approach for coronary artery disease screening using clinical parameters. Knowledge-Based Systems. 2012;36:162–174. doi: 10.1016/j.knosys.2012.06.013.

[16] Ephzibah E. P., Sundarapandian V. A neuro-fuzzy expert system for heart disease diagnosis. Computer Science & Engineering. 2012;2(1):17–23.

[17] Lahsasna A., Ainon R. N., Zainuddin R., Bulgiba A. Design of a fuzzy-based decision support system for coronary heart

disease diagnosis. Journal of Medical Systems. 2012;36(5):3293–3306. doi: 10.1007/s10916-012-9821-7.

[18] Marateb H. R., Goudarzi S. A. A noninvasive method for coronary artery diseases diagnosis using a clinically-interpretable fuzzy rule-based system. Journal of Research in Medical Sciences. 2015;20:214–223.

[19] Peters R. M., Shanies S. A., Peters J. C. Predicting extent of coronary disease: fuzzy cluster analysis vs. Duke treadmill score. Journal of Clinical and Basic Cardiology. 2000;3(1):39–41.

[20] Mohammadpour R. A., Babaee G., Mahmodi M., Memariani A., Mohseni A. Evaluation of exercise test variables in diagnosis and prediction of the severity of coronary artery disease by fuzzy discriminant analysis. Journal of Mazandaran University of Medical Sciences. 2002;11(33)

[21] Pop H. F., Pop T. L., Sarbu C. Assessment of heart disease using fuzzy classification techniques. The Scientific World Journal. 2001;1:369–390.

[22] Cachin F., Lipiecki J., Mestas D., et al. Preliminary evaluation of a fuzzy logic-based automatic quantitative analysis in myocardial SPECT. Journal of Nuclear Medicine. 2003;44(10):1625–1632.

[23] Peters R. M., Shanies S. A., Peters J. C. Fuzzy cluster analysis off positive stress tests, a new method of combining exercise test variables to predict extent of coronary artery disease. The American Journal of Cardiology. 1995;76(10):648–651. doi: 10.1016/s0002-9149(99)80190-8

[24] Zhi-Hua Zhou and Yuan Jiang. NeC4.5: Neural Ensemble Based C4.5. IEEE Trans. Knowl. Data Eng, 16. 2004.

[25] Detrano, R., Janosi, A., Steinbrunn, W., Pfisterer, M., Schmid, J., Sandhu, S., Guppy, K., Lee, S., &Froelicher, V. International application of a new probability algorithm for the diagnosis of coronary artery disease. American Journal of Cardiology, 64,304--310. 1989.

[26] David W. Aha & Dennis Kibler. "Instance-based prediction of heart-disease presence with the Cleveland database." 1989.

[27] Gennari, J.H., Langley, P, & Fisher, D. Models of incremental concept formation. Artificial Intelligence, 40, 11--61. 1989.

[28] P.SambasivaRao and Dr.T.Uma Devi, "A Parameter Based Heart Disease Detection Technique using Mining Technique". International Journal of Latest Trends in Engineering and Technology Vol.(7)Issue(3), pp. 077-087. 2016.

