

ROLE OF CERTAINTY FACTOR IN GENERATING ROUGH-FUZZY RULE

Jyotirmoy Ghosh¹ and S. Mukhopadhyay²

¹Dept. of Computer Application, Heritage Institute of Technology, Kolkata-107
jyotirmoy.ghosh@yahoo.co.in

²Professor, Dept. of Computer Science, The University of Burdwan, Burdwan-104
dr.sripatim@gmail.com

ABSTRACT

The generation of effective feature-based rules is essential to the development of any intelligent system. This paper presents an approach that integrates a powerful fuzzy rule generation algorithm with a rough set-assisted feature reduction method to generate diagnostic rule with a certainty factor. Certainty factor of each rule is calculated by considering both the membership value of each linguistic term introduced at time of fuzzyfication of data as well as possibility values, due to inconsistent data, generated by rough set theory at time of rule generation. In time of knowledge inferencing in an intelligent system, certainty factor of each rule will play an important role to find out the appropriate rule to be selected. Experimental results demonstrate the superiority of our approach.

KEYWORDS

Rough set, Fuzzy set, Certainty Factor, Diagnostic rule.

1. INTRODUCTION

In developing expert systems, one of the most important problems is knowledge acquisition from experts [1]. Extraction of knowledge from data stored in the form of table in database may automate this problem. Knowledge extraction from data and combining it with available symbolic knowledge and refine it as rule-base for a rule-based expert systems is a great challenging work in computational intelligence. Recommendations given by black box systems [2] is good but reasoning with logical rules is more understandable and acceptable to users, because such reasoning is understandable, provides explanations and may be validate the system by inspecting the increasing confidence, shown the important relationships and features may be hidden in the data. Many methods have been developed to find logical depiction of the data in the past using statistical, pattern recognition [3] and machine learning [4] approaches. Rule-based systems should be preferred over other methods only in cases when the set of logical rules is not too complex and their prognostic accuracy is adequately high.

The rule generation techniques have been widely developed and used for data mining to developed expert system in many application areas [5], such as medical diagnosis, decision-making, classification and prediction. Many inductive learning methods, such as generation of decision trees [6], rule generation methods [7], soft computing tools in rule generation are: neural network [8], fuzzy systems [9], rough set theory [13], genetic algorithm etc, are introduced and applied to extract knowledge from databases. Every approach has some advantages and

disadvantages. In order to provide more flexible and robust information processing system, using only one approach is not enough. Hybridizations of soft computing methodologies for rule generation are introduced [22-34].

Typically in Rough-Fuzzy hybridization to generate rule utilized both the advantages of fuzzy set theory to deal with uncertainty and vagueness in data and power of rough set theory to generate dependency rule. Rough-Fuzzy hybridization is a method of hybrid intelligent system, where fuzzy set theory [9] is used for linguistic representation of patterns known as fuzzyfication of data and rough set theory [12] is used to obtain dependency rule generation. Data in data set may be continuous or discrete. Fuzzy set represent the continuous data in a linguistic [11] form with some membership value (MV). Knowledge may be incomplete or uncertain and fuzzy sets provide a natural framework for the process in dealing with uncertainty [10]. In this process a huge amount of data set is converted to a small one, which is helpful to generate the dependency rule using rough set theory. Knowledge may also be consistent or inconsistent, two records in a data set have same condition attribute values but have different decision attribute value, rough set theory can handle inconsistent knowledge also to generate dependency rule [14].

Rule generation using rough set theory may be done by different approach. One approach is first compute reducts [13], minimal set of attributes that preserve the indiscernibility relation and, consequently, set approximation of information system. An information system is a table, where each row represents an event and each column represents an attribute that can be measured for each event and acquired from the domain expert. In an information system, indiscernibility relation for a subset of attribute set presents the same attribute value for all attribute of that subset for all events. Then generate rule by calculating discernibility matrices [13]. If A is the attribute set of an information system with n events then the discernibility matrix of that information system is a $n \times n$ symmetric matrix with each entries is a set of attribute defined as

$$c_{ij} = \{a \in A \mid a(x_i) \neq a(x_j)\} \quad \forall i, j = 1, 2, \dots, n . \text{ Here } x\text{'s are events}$$

Other way is directly generate decision rule using decision matrix, which is a generalized from of rough set theory [14]. The decision matrix may also used to compute reducts of the information system.

But in all previously worked over rough-fuzzy hybridization to generate rule [14,15,34] the MV of the introduced linguistic terms at time of fuzzyfication of data were ignored at time of rule generation using rough set theory. In those work all fuzzyfied data are taken as certain. Thus finally when rules were generated all rules were defined as certain rule with respect to the fuzzyfied data and very few rules were generated with some possibility values, due to some inconsistency of data, defined as possible rules in rough set theory. When this rule set will be used as a knowledge base of a rule-based expert system then in time of inferencing, it will be difficult to select a rule among set of matching rules for execution. Only one way is there to consider the first matching rule for execution. This will not always found optimum result, because result will depend on physical arrangement of rule set.

In this paper we present a modified framework by introducing certainty factor (CF) of each rule by considering both the MV of each linguistic term introduced at time of fuzzyfication of data as well as possibility values, due to inconsistent data, generated by rough set theory at time of rule generation. In this modified framework of rough-fuzzy rule generation, CF will solve the problem of rule selection from a set of matching rule. One may select the rule with highest CF value. In this work we proceed through some steps. In step one we read data from database and perform the fuzzyfication of data by introducing some linguistic term with some MV in between 0 and 1. We also assign MV for all other data as 1 or certain membership. In steps two modified the data set by deleting the equal records and store these all MV for future use to calculate CF. In steps three we

use rough set theory to generate dependency rule. Finally in step four we calculate CF to each rule by considering the stored MV for certain rule and both stored MV and generated possibility value for possible rule and present CF as percentage form. This rule set may be used as a knowledge base of a rule-based expert system. In time of knowledge inferencing CF of each rule will play an important role to find out the appropriate rule to be fire. To generate rules and testing the rules we use the diabetes patient's data.

In this paper section 2 present the procedure of fuzzyfication of data i. e., the method of linguistic representation of pattern in fuzzy set theory. Section 3 describes the procedure of dependency rule generation using rough set theory. Section 4 presents the steps involve in the modified framework to compute dependency rule with CF using rough-fuzzy hybridization. Section 5 shows some results and comparison between exiting algorithms and the new modified rough-fuzzy framework and section 6 define conclusion.

2. FUZZYFICATION OF DATA USING FUZZY SET THEORY

Table of data are fuzzyfied using Fuzzy set theory [9, 11]. Fuzzy set theory has been introduced the concepts of degree of membership of elements to set. Previously elements could belong fully (membership 1) or not at all (membership 0) to set. The degree of membership allows an element to lie in a set with membership values anywhere in the range [0, 1]. A fuzzy set can be defined as a set of ordered pairs $\tilde{A} = \{(x, \mu_{\tilde{A}}(x)) / x \in \mathbb{U}\}$. The function $\mu_{\tilde{A}}(x)$ is called is called the membership function for \tilde{A} , mapping each element of the base set \mathbb{U} (universe) to a membership degree in the range [0.1]. The base set may be discrete or continuous.

In Rough-Fuzzy Hybridization fuzzyfication of data is performed to represent the linguistic patterns of the continuous data. Each linguistic pattern has membership degree in the range [0.1]. The type of the membership function is used depending on the base set patterns. If the base set contains many values, or if this set is continuous, then a parametric representation, which can be adapted by changing the parameters, is appropriate. Mostly this type of membership functions are triangular or trapezoidal functions that are defined by three and four parameters respectively. For some applications continuously differentiable curves requires for modeling and therefore smooth transitions, which is not possible using triangular or trapezoidal function. In those cases normalized Gaussian function, difference of two sigmoidal functions, generalized bell function, etc, and in some application π functions are used (8).

3. RULE GENERATION USING ROUGH SET THEORY

Rough set theory [12, 13] is used to generate dependency rule form table of data. Let us discuss some basic concepts of rough set which are used in this paper.

An Information system S is defined as $S = \langle U, A \rangle$, where U denote the domain of discourse formally Universe and A is the non-empty and finite set of attributes. Let $A=C \cup D$, where C is the non-empty and finite set of condition attributes and D is defined as the non-empty and finite set of decision attributes. An attribute $a \in A$, can be regarded as a function $a: U \rightarrow V_a$, where V_a is a value set.

An information system may be viewed as an attribute value-table known as decision table, where each row is labeled by object $\in U$ and each column by attribute $\in A$.

For all $B \subseteq C$, equivalent relation I_B on U is defined as

$$I_B = \{(x, y) \in U : \forall a \in B, a(x) = a(y) \}$$

$[x]_B$ is denoted as equivalence class of object $x \in U$ relative to I_B and defined as

$$[x]_B = \{y \mid y \in U, yBx\}$$

$\underline{B}X$ and $\overline{B}X$ are denoted as B-lower and B-upper approximations of $X \subseteq U$ in S , where $B \subseteq C$, and are defined as

$$\underline{B}X = \{x \in U : [x]_B \subseteq X\}$$

$$\overline{B}X = \{x \in U : [x]_B \cap X \neq \emptyset\}$$

$X \subseteq U$ will be B exact if $\overline{B}X = \underline{B}X$ and will be B rough if $\overline{B}X \neq \underline{B}X$.

Rule generation using rough set can be done by the following two methods:

3.1. Method 1

The main task in this method, used in [34], of rule generation is to find out the reducts relative to information system S . Let there are k number of decision attributes in D i.e., $D = \{d_1, d_2, \dots, d_k\}$. Then divide the decision table into k tables $S_i = \langle U_i, A_i \rangle, \forall i = 1, 2, \dots, k$, where $U = \bigcup U_i$ and $A_i = C \cup \{d_i\}$.

Let us assume that there are n objects, i.e., $U = \{x_1, x_2, \dots, x_n\}$, and m condition attributes, i.e., $C = \{a_1, a_2, \dots, a_m\}$ in the information system S . Also assume that $U_i = \{x_{i1}, x_{i2}, \dots, x_{ip}\}$ that occur in $S_i, i=1, 2, \dots, k$. Now construct the discernibility matrix $M_{d_i}(B)$ for each d_i -reduct $B = \{b_1, b_2, \dots, b_l\}$ (say) from the d_i -discernibility matrix as defined follows

$$c_{ij} = \{a \in B : a(x) \neq a(y)\}, \forall i, j = 1, 2, \dots, n.$$

For each object $x_j \in \{x_{i1}, x_{i2}, \dots, x_{ip}\}$ the discernibility function denoted by $f_{d_i}(x_j)$ is defined as

$$f_{d_i}(x_j) = \bigwedge \{ \bigvee (c_{ij}) : 1 \leq i, j \leq n, j < i, c_{ij} \neq \emptyset \}$$

where \bigwedge and \bigvee are conjunction and disjunction operations respectively.

Now calculate R_i defined as

$$R_i = \bigvee (f_{d_i}(x_j)), \forall j = i_1, i_2, \dots, i_p$$

Thus the rule $r_i: R_i \rightarrow d_i$ is obtained and the dependency factor df_i is calculated as

$$df_i = \frac{\text{card}(\text{POS}_i(d_i))}{\text{card}(U_i)}$$

where $\text{card}(\cdot)$ define the cardinality and $\text{POS}_i(d_i)$ is defined as

$$\text{POS}_i(d_i) = \bigcup_{x \in U_i} k_i(X) \text{ where } k_i(X) \text{ is lower approximation of } X \text{ with respect to } I_i$$

3.2. Method 2

This method use decision matrix [14] for rule generation. Decision matrix is a generalization of rough set theory from where reduct and decision rule can be calculated. In this method first it is check that the information system is consistent or not. Information system is said to consistent if there is no two objects whose condition attributes are same but decision attributes are different. Similarly an information system will be inconsistent if there exists any two objects whose condition attributes are same and decision attributes are different. That means for any two object i and j

if $a_i(C) = a_j(C)$ and $a_i(D) \neq a_j(D)$ where $i \neq j$

then the information system is inconsistent.

Let the domain of discourse U of the information system S is divided into k classes $(c_1, c_2 \dots c_k)$ depending on equivalence relation defined on D . For any class $c_p \in (c_1, c_2 \dots c_k)$, the objects $\in U$ are belong in c_p are numbered by subscripts i ($i = 1, 2, \dots, m$) and those do not belong in c_p are subscripts j ($j = 1, 2, \dots, n$). The decision matrix M of the information system S for the class c_p is defined as $m \times n$ matrix with elements as a set {attribute-name, attribute-value}.

$$M_{ij}^p = \{(a, a(i)): a(i) \neq a(j)\} \quad \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$$

Where a is attribute name and $a(i)$ is attribute value.

3.2.1. For consistent information system:

The minimum-length decision rule for any object i ($i = 1, 2, \dots, m$) belong in class $c_p \in (c_1, c_2 \dots c_k)$ can be obtained as

$$|B_i^p| = \bigwedge_j \bigvee M_{ij}^p$$

Where \bigwedge and \bigvee are conjunction and disjunction operations respectively.

The decision rule for the class $c_p \in (c_1, c_2 \dots c_k)$ is calculated as

$$R^p = \bigvee |B_i^p| \quad \forall i = 1, 2, \dots, m$$

3.2.2. For inconsistent information system:

First find out the B-lower and B-upper approximation for each class $c_p \in (c_1, c_2 \dots c_k)$. Rule generated from the B-lower approximation are certain and rule generated from B-upper approximation is possible rule.

Thus the certain decision rule for any object i ($i = 1, 2, \dots, m$) belong in class $c_p \in (c_1, c_2 \dots c_k)$ can be obtained as

$$|B_i^p|_{\text{certain}} = \bigwedge_j \bigvee M_{ij}^p$$

where \bigwedge and \bigvee are conjunction and disjunction operations respectively.

The certain decision rule for the class $c_p \in (c_1, c_2 \dots c_k)$ is calculated as

$$(R^p)_{\text{certain}} = \bigvee |B_i^p|_{\text{certain}} \quad \forall i = 1, 2, \dots, m$$

And the possible decision rule for any object i ($i = 1, 2, \dots, m$) belong in class $c_p \in (c_1, c_2 \dots c_k)$ can be obtained as

$$|B_i^p|_{\text{possible}} = \bigwedge_j \bigvee M_{ij}^p$$

where \bigwedge and \bigvee are conjunction and disjunction operations respectively.

The possible decision rule for the class $c_p \in (c_1, c_2 \dots c_k)$ is calculated as

$$(R^p)_{\text{possible}} = \bigvee |B_i^p|_{\text{possible}} \quad \forall i = 1, 2, \dots, m$$

For possible rule the belief function can be defined as follows

$$\alpha^f = 1 - \frac{\text{card}(\bar{B}_c - B_c)}{\text{card}(U)}$$

Where $c \in (c_1, c_2, \dots, c_k)$ and $\text{card}(\cdot)$ define the cardinality of the set.

4. MODIFIED FRAMEWORK TO GENERATE ROUGH-FUZZY RULE WITH CF

In this proposed modified framework we associate a CF in each rough-fuzzy rule, which is more logical than considering each rough-fuzzy rule fully certain. The following steps describe the details procedure as well as modification performed in rough-fuzzy rule generation framework.

Step-1: read the data set from database. Find out the attributes which have continuous values. Then perform the fuzzyfication operation over continuous attributes by introducing some linguistic variables like low, high, medium etc and calculate the MV of each linguistic variable. The MV is calculated by using triangular membership as described in section 2. According to the definition of MV describe in section 2, the MV must be in [0, 1]. We also assign MV 1 or certain membership to the other attribute values. We discard those parameters which have MV less than 0.25. Here we use seven parameter triangular function instead of regular three parameter triangular function defined as follows:

$$\mu(x, a1, a2, a3, a4, a5, a6, a7) = \begin{cases} 0 & x < a1, x > a7 \\ \frac{x - a1}{a2 - a1} & a1 \leq x \leq a2 \\ \frac{x - a2}{a3 - a2} & a2 < x \leq a3 \\ \frac{x - a3}{a4 - a3} & a3 < x \leq a4 \\ \frac{x - a4}{a5 - a4} & a4 < x \leq a5 \\ \frac{x - a5}{a6 - a5} & a5 < x \leq a6 \\ \frac{x - a6}{a7 - a6} & a6 < x \leq a7 \end{cases}$$

Step-2: Find out those records which have same attribute value but may have different MV of the attribute value in the data set. Then calculate the total membership value of those records by summing up the MV of each attribute value in each record. Only keep the record with maximum total membership value and delete other records. In this way we find out a modified data set with some linguistic variable and each attribute value has some membership degree. In this modified data set all attribute have discrete values. Now in this point we store all MV corresponding to the attribute values in the data set.

Step-3: Now we use rough set theory as described in section 3 with some modification over the modified data set constructed in step-2. Here we have used Method-2 as described in section 3.2. We take the fuzzyfied data set as information set S. Next we check that the information system S is consistent or not using the rule described in Method-2 in section 3.2. Construct the decision matrix as described in Method-2 in section 3.2 with a modification. We modified the decision matrix by adding MV of attributes as follows:

$$M_{ij}^p = \{(a, a(i), \mu(i)): a(i) \neq a(j)\} \forall i = 1, 2, \dots, m \text{ and } j = 1, 2, \dots, n$$

Where a is attribute name and a(i) is attribute value and $\mu(i)$ is the MV of attribute value.

For rule generation Method-2 describe in section 3.2 generate one rule for one class. The length of the rule is very large and the form of the rule is

If (X is a1 and Y is b1 and ...) or (X is b1 and Y is b2 and ...) or Then Z is c

These generated rules may be used as a rule-base of a rule-based Intelligent system and so for easier comparison we consider here the simple rule with the following form:

If X is a1 and Y is b1 and Then Z is c with CF m.

where **m** is the CF of the rule.

For consistent information system:

To construct the minimum-length decision rule as described in Method-2 of section 3.2.1 for any object *i* (*i* = 1, 2, ..., *m*) belong in class $c_p \in (c_1, c_2 \dots c_k)$ can be obtained as

$$|B_i^p| = \bigwedge_j \vee M_{ij}^p$$

Where \bigwedge and \vee are conjunction and disjunction operations respectively.

Next calculate the CF of each rule by

$$CF_i = \frac{\sum_{j=1}^k \mu_{ij}}{k}$$

Where *k* is the number of attributes present in the *i*th rule and μ_{ij} is the MV.

The decision rule is calculated as

$$R_j = |B_i^p| \quad \forall i = 1, 2, \dots, m; c_p \in (c_1, c_2 \dots c_k); j = \text{number of distinct rule}$$

If for any *i*, $|B_i^p|$ already exists in the rule set **R**, then compare the CF value of existing rule and new rule and store the rule with maximum CF value in rule set **R**.

For inconsistent information system:

First find out the B-lower and B-upper approximation for each class $c_p \in (c_1, c_2 \dots c_k)$. Rule generated from the B-lower approximation are certain and rule generated from B-upper approximation is possible rule.

To construct certain minimum-length decision rule for any object *i* (*i* = 1, 2, ..., *m*) belong in class $c_p \in (c_1, c_2 \dots c_k)$ can be obtained by performing the operation described in Method-2 of section 3.2.2 as

$$|B_i^p|_{\text{certain}} = \bigwedge_j \vee M_{ij}^p$$

where \bigwedge and \vee are conjunction and disjunction operations respectively.

Next calculate the CF of each rule as

$$CF_i = \frac{\sum_{j=1}^k \mu_{ij}}{k}$$

Where *k* is the number of attributes present in the *i*th rule and μ_{ij} is the MV.

The decision rule is calculated as

$$R_j = |B_i^p|_{\text{certain}} \quad \forall i = 1, 2, \dots, m; c_p \in (c_1, c_2 \dots c_k); j = \text{number of distinct rule}$$

If for any *i*, $|B_i^p|_{\text{certain}}$ already exists in the rule set **R** then compare the CF value of existing rule and new rule and store the rule with maximum CF value in rule set **R**.

Then construct the possible minimum-length decision rule for any object i ($i = 1, 2, \dots, m$) belong in class $c_p \in (c_1, c_2 \dots c_k)$ by performing the operation described in Method-2 of section 3.2.2 as

$$|B_i^p|_{\text{possible}} = \bigwedge_j \bigvee M_{ij}^p$$

where \bigwedge and \bigvee are conjunction and disjunction operations respectively.

For possible rule the belief function of i^{th} rule can be defined as follows

$$df_i = 1 - \frac{\text{card}(\overline{B}c_p - Bc_p)}{\text{card}(U)}$$

Where $c_p \in (c_1, c_2 \dots c_k)$ and $\text{card}(\cdot)$ define the cardinality of the set. and then

$$CF_i = \frac{\sum_{j=1}^k \mu_{ij}}{k} * df_i$$

Where k is the number of attributes present in the i^{th} rule and μ_{ij} is the MV.

The decision rule is calculated as

$$R_j = |B_i^p|_{\text{possible}} \quad \forall i = 1, 2, \dots, m; c_p \in (c_1, c_2 \dots c_k); j = \text{number of distinct rule}$$

If for any i , $|B_i^p|_{\text{possible}}$ already exists in the rule set \mathbf{R} then compare the CF value of existing rule and new rule and store the rule with maximum CF value in rule set \mathbf{R} .

5. APPLICATION ON MEDICAL DATA SET

We have applied this framework over the following medical data-sets of diabetes patients for rheumatological manifestations of Diabetes Mellitus like

- Diabetic cheiroarthropathy or Limited joint mobility (LJM)
- Adhesive capsulitis of shoulder (ADH)
- Clinical Carpal tunnel syndrome (CTS CL)
- NCV finding of Carpal tunnel syndrome (CNCV)
- Dupuytren's contracture (DUPY)
- Flexor tenosynovitis (FTS)
- Diffuse interstitial skeletal hyperostosis (DISH)
- Gout and Hyperuricaemia (GOUT)
- Hand Osteoarthritis (OAH)
- Knee Osteoarthritis (OAK)

The datasets are contained nine attributes with values as follows

- Age (Integer Numbers)
- Sex (0 = MALE, 1 = FEMALE)
- Type of diabetes (1 = TYPE1, 2 = TYPE2)
- Duration of Diabetes (Integer Number in year)
- Use of Insulin (1 = YES, 0 = NO)

- Fasting Blood Sugar(Integer Numbers)
- Post Prandial Blood Sugar(Integer Numbers)
- Albuminuria (1 = MICROALBUMINURIA , 2= PROTEINURIA)
- Uric acid(Floating Point Number)

We find out the result considering the best rule i.e. maximum CF value as well as consider the maximum occurrence (vote) of a class for rules of top three CF value and top five CF value. We also compare the result with the previously established rough-fuzzy framework shown in *table-1* which is self explaining. The datasets contains 100 instances. The datasets are presents in form of two files .nam and .dat file. In .data file present the data and .nam file represent the data structure about data. Here we used 5-fold data. Only 20% data are used to generate rule (c.f. Appendix) and other 80 % data is used for testing.

Data Set	Existing Algorithm	Modified Algorithm		
		Best Rule	Vote of 3 Rule	Vote of 5 Rule
LJM	72.50%	75.00%	77.5%	73.75%
ADH	78.75%	83.75%	83.75%	81.25%
CTS CL	92.50%	93.75%	93.75%	93.75%
CNCV	95.00%	96.25%	96.25%	95.00%
DUPY	96.25%	98.75%	98.75%	97.50%
FTS	85.00%	93.75%	91.25%	90.00%
DISH	95.00%	98.75%	96.25%	98.75%
GOUT	95.00%	96.25%	96.25%	96.25%
OAH	82.50%	85.00%	87.50%	83.75%
OAK	72.50%	73.75%	73.75%	72.50%

Table-1

6. CONCLUSION

We have presented a rule generation methodology based on rough-fuzzy hybridization. Fuzzy set is used to represent a pattern in terms of its membership to linguistic variables. Rough sets are used to generate diagnostic rules. CF of each rule is calculated by combining the fuzzy MV and rough possibility values. We obtained good results but several issues still remain unexplored, like aggregation of large number of input features, construction of hierarchical systems when a large number of features contain missing data, automatization of the whole process of logical data description and creation of complete intelligent systems [35]. Hybridization with neural networks may increase performance of the system considerable.

ACKNOWLEDGEMENT

We acknowledge the help received from Dr. Subhobrata Roy, MD, North Bengal Medical College, for providing the dataset.

REFERENCES

- [1] B.G. Buchanan, E.H. Shortliffe, "Rule-Based Expert Systems", Addison-Wesley, New York, 1984.

- [2] R. Andrews, J. Diederich, A.B. Tickle, "A Survey and Critique of Techniques for Extracting Rules from Trained Artificial Neural Networks," *Knowledge-Based Systems* vol. 8, pp. 373–389, 1995.
- [3] R. Schalkoff, "Pattern Recognition. Statistical, Structural and Neural Approaches". Wiley 1992.
- [4] T. Mitchell, "Machine learning". McGraw Hill 1997.
- [5] S. Mitra, Sankar K. Pal, and P. Mitra, "Data Mining in Soft Computing Framework: a Survey", *IEEE Trans.on neural networks*, vol. 13, no. 1, pp. 3–14, 2002.
- [6] J.R. Quinlan, "C4.5—Programs for Machine Learning", Morgan Kaufmann, Palo Alto, 1993.
- [7] P.S. Michalski, I. Mozetic, J. Hong, N. Lavrac, "The multi-purpose incremental learning system AQ15 and its testing application to three medical domains", in: *Proceedings of the Fifth National Conference on Artificial Intelligence*, AAAI Press, Menlo Park, 1986, pp. 1041–1045.
- [8] S.T. Wang, "Fuzzy system and Fuzzy Neural Networks", Shanghai Science and Technology Press, 1998, Edition 1.
- [9] L.A. Zadeh, "Fuzzy sets", *Inform. Contr.* 8(1965), 338-353
- [10] L. A. Zadeh, "Fuzzy logic as a basis for the management of uncertainty in expert systems, *Fuzzy Set and Systems* 11 (1983c), 199-227.
- [11] L.A. Zadeh. "A computational approach to fuzzy quantifiers in natural languages", *Computers and Mathematics with Applications*, 9:149–184, 1983.
- [12] Z. Pawlak, "Rough sets", *International Journal of Computer and Information Sciences* **11** (1982), 341–356.
- [13] Z. Pawlak, "Rough Sets: Theoretical Aspects of Reasoning about Data, *System Theory*", *Knowledge Engineering and Problem Solving* vol. 9, Kluwer Academic Publishers, Dordrecht, The Netherlands (1991).
- [14] Shi-tong Wang, Dong-jun Yu and Jing-yu Yang, "Integrating rough set theory and fuzzy neural network to discover fuzzy rules", *Intelligent Data Analysis* 7 (2003) 59–73
- [15] Sankar K. Pal, Pabitra Mitra, "Case Generation: A Rough-fuzzy Approach",
- [16] Gang Xie , Fang Wang , Keming Xie, "RST-Based System Design of Hybrid Intelligent Control", *IEEE International Conference on Systems, Man and Cybernetics*, 2004
- [17] S. K. Pal and A. Skowron, Eds., "Rough Fuzzy Hybridization: A New Trend in Decision Making" Singapore: Springer-Verlag, 1999
- [18] Shusaku Tsumoto, "Mining diagnostic rules from clinical databases using rough sets and medical diagnostic model", *Information Sciences* 162 (2004) 65–80
- [19] Brenda Mak, Toshinori Munakata, "Rule extraction from heuristic: A comparative study of rough sets with neural network and ID3" *European Journal of Operational Research* 136(2002) 212-229
- [20] S. Tsumoto, H. Tanaka, "PRIMEROSE: Probabilistic rule induction method based on rough sets and resampling methods", *Computational Intelligence* 11 (1995) 389–405.
- [21] V. Dhar, A.Tuzhilin, "Abstract-driven pattern discovery in databases", *IEEE Transactions on Knowledge and Data Engineering* 5(6)(1993) 926-937
- [22] S. Mitra and Y. Hayashi, "Neuro-fuzzy Rule Generation: Survey in Soft Computing Framework," *IEEE Trans. On Neural Network*, vol. 11, no. 3, pp. 748–768, 2000.
- [23] L. X. Wang and J. M. Mendel, "Generating Fuzzy Rules by Learning from Examples", *IEEE Trans. Systems, Man, and Cybernetics*, vol. 22, pp. 1414–1427, 1992
- [24] R. Rovatti and R. Guerrieri, "Fuzzy Sets of Rules for System Identification," *IEEE Trans. Fuzzy Syst.*, vol. 4, pp. 89–102, 1996.
- [25] H. Ishibuchi, K. Nozaki, N. Yamamoto, and H. Tanaka, "Selecting Fuzzy If–Then Rules for Classification Problems Using Genetic Algorithms," *IEEE Trans. Fuzzy Syst.*, vol. 3, pp. 260–270, 1995.
- [26] L. Wang and J. Yen, "Extracting Fuzzy Rules for System Modeling Using a Hybrid of Genetic Algorithms and Kalman Filter," *Fuzzy Sets Syst.*, vol. 101, pp. 353–362, 1999.
- [27] P. Mitra and S. K. Pal, "Staging of Cervical Cancer with Soft Computing", *IEEE Trans. On Biomedical Engineering*, vol. 47, no. 7, pp. 934–940, 2000.
- [28] Y. Jung-heurn, Y. Seung-moo, and J. Hong-Tae, "Structure Optimization of Fuzzy-Neural Network Using Rough Set Theory", the *Proceedings of 1999 IEEE International Conference on Fuzzy Systems*, Aug 1999, pp. 1666-1670.
- [29] B. S. Ahn, S. S. Cho, and C. Y. Kim, "The Integrated Methodology of Rough Set Theory and Artificial Neural Network for Business Failure Prediction," *Expert Systems with Applications*, vol. 18, pp. 65–74, 2000.
- [30] R. Yasdi, "Combining Rough Sets Learning and Neural Learning Method to Deal with Uncertain and Imprecise Information", *Neurocomputing*, vol. 7, pp. 61–84, 1995.

- [31] R. W. Swiniarski and L. Hargis, "Rough Sets as a Front End of Neural Networks Texture Classifiers", *Neurocomputing*, vol. 36, pp. 85–102, 2001.
- [32] R. R. Hashemi et. al., "A Hybrid Intelligent System for Predicting Bank Holding Structures", *European Journal of Operational Research*, vol. 109, pp. 390–402, 1998.
- [33] P. J. Lingras, "Rough Set Clustering for Web Mining." *Proceedings of 2002 IEEE International Conference on Fuzzy Systems*, Hawaii, May 2002, pp. 12-17.
- [34] S. K. Pal, S. Mitra, and P. Mitra, "Rough-Fuzzy MLP: Modular Evolution, Rule Generation, and Evaluation", *IEEE Trans. On Knowledge and data engineering*, vol. 15, no. 1, pp. 14-25, Jan/Feb 2003.
- [35] Wlodzislaw Duch, Rafal Adamczak and Krzysztof Grabczewski, "A new methodology of extraction, optimization and application of crisp and fuzzy logical rules", *IEEE Transactions on Neural Networks*, vol. 11, no. 2, pp. 1-31, march 2000

APPENDIX

Following are the portion of the rule-base generated by modified method for the different diabetic mellitus. Here decision is 1 means that diabetic mellitus is present and 0 means it is absent.

LJM

- RULE-24: If TYPE is 1 and PPB is Extrm High Then Decision is 1 With CF 100
- RULE-25: If TYPE is 1 and ALB is 1 Then Decision is 1 With CF 100
- RULE-26: If INS is 1 and PPB is Extrm High Then Decision is 1 With CF 100
- RULE-27: If PPB is Extrm High and ALB is 1 Then Decision is 1 With CF 100
- RULE-28: If AGE is Very High and SEX is 0 Then Decision is 0 With CF 98
- RULE-29: If AGE is Very High and INS is 1 Then Decision is 0 With CF 98
- RULE-30: If SEX is 0 and FB is Medium Then Decision is 0 With CF 97
- RULE-31: If DUR is Extrm Low and FB is Medium Then Decision is 0 With CF 97
- RULE-32: If FB is Medium and ALB is 0 Then Decision is 0 With CF 97
- RULE-33: If AGE is Very High and DUR is Extrm High Then Decision is 0 With CF 96
- RULE-34: If AGE is Very High and ALB is 1 Then Decision is 0 With CF 96

ADH

- RULE-07: If PPB is Extrm High Then Decision is 0 With CF 100
- RULE-08: If AGE is Extrm Low Then Decision is 0 With CF 100
- RULE-09: If DUR is Extrm High Then Decision is 0 With CF 100
- RULE-10: If SEX is 1 and PPB is Extrm Low and ALB is 1 Then Decision is 1 With CF 100
- RULE-11: If PPB is Low Then Decision is 0 With CF 98
- RULE-12: If AGE is Very High and SEX is 1 and ALB is 1 Then Decision is 1 With CF 95
- RULE-13: If FB is Low Then Decision is 0 With CF 93
- RULE-14: If FB is Extrm Low Then Decision is 0 With CF 90
- RULE-15: If AGE is High Then Decision is 0 With CF 87
- RULE-16: If PPB is Very Low Then Decision is 0 With CF 86
- RULE-17: If FB is Medium Then Decision is 0 With CF 84

CTS CL

- RULE-05: If FB is Extrm Low Then Decision is 0 With CF 100
- RULE-06: If TYPE is 1 Then Decision is 0 With CF 100
- RULE-07: If DUR is Extrm Low Then Decision is 0 With CF 100
- RULE-08: If INS is 1 Then Decision is 0 With CF 100
- RULE-09: If FB is Extrm High and ALB is 2 Then Decision is 1 With CF 100
- RULE-10: If PPB is Very Low Then Decision is 0 With CF 98
- RULE-11: If FB is Extrm High and ALB is 0 Then Decision is 0 With CF 96

RULE-12: If AGE is Very High and SEX is 1 Then Decision is 0 With CF 96

RULE-13: If AGE is Very High and ALB is 0 Then Decision is 0 With CF 96

RULE-14: If SEX is 1 and DUR is Medium and INS is 0 Then Decision is 1 With CF 94

RULE-15: If AGE is High and SEX is 1 and DUR is Medium Then Decision is 1 With CF 92

CNCV

RULE-11: If AGE is Extrm High Then Decision is 0 With CF 100

RULE-12: If AGE is Extrm Low Then Decision is 0 With CF 100

RULE-13: If TYPE is 1 Then Decision is 0 With CF 100

RULE-14: If INS is 0 and ALB is 2 Then Decision is 1 With CF 100

RULE-15: If PPB is High and ALB is 2 Then Decision is 1 With CF 100

RULE-16: If PPB is Very High Then Decision is 0 With CF 98

RULE-17: If FB is Very Low Then Decision is 0 With CF 94

RULE-18: If AGE is Very High Then Decision is 0 With CF 90

RULE-19: If PPB is Medium Then Decision is 0 With CF 87

RULE-20: If PPB is Low Then Decision is 0 With CF 85

RULE-21: If SEX is 0 and FB is High and PPB is High Then Decision is 1 With CF 84

DUPY

RULE-10: If SEX is 0 and INS is 1 Then Decision is 1 With CF 100

RULE-11: If INS is 1 and ALB is 0 Then Decision is 1 With CF 100

RULE-12: If FB is Extrm High Then Decision is 1 With CF 100

RULE-13: If PPB is Low Then Decision is 0 With CF 95

RULE-14: If FB is Low Then Decision is 0 With CF 95

RULE-15: If SEX is 0 and DUR is High Then Decision is 1 With CF 94

RULE-16: If DUR is High and ALB is 0 Then Decision is 1 With CF 94

RULE-17: If DUR is Very High Then Decision is 0 With CF 94

RULE-18: If FB is Medium Then Decision is 0 With CF 93

RULE-19: If AGE is Very Low Then Decision is 0 With CF 92

FTS

RULE-08: If AGE is Extrm High Then Decision is 0 With CF 100

RULE-09: If ALB is 2 Then Decision is 0 With CF 100

RULE-10: If DUR is Extrm High Then Decision is 0 With CF 100

RULE-11: If FB is Extrm High Then Decision is 0 With CF 100

RULE-12: If PPB is Extrm High Then Decision is 0 With CF 100

RULE-13: If SEX is 1 and PPB is Extrm Low Then Decision is 1 With CF 100

RULE-14: If PPB is Medium Then Decision is 0 With CF 95

RULE-15: If AGE is High and SEX is 1 Then Decision is 1 With CF 93

RULE-16: If PPB is High Then Decision is 0 With CF 92

RULE-17: If AGE is Very Low Then Decision is 0 With CF 91

DISH

RULE-10: If PPB is Extrm High Then Decision is 0 With CF 100

RULE-11: If ALB is 2 Then Decision is 0 With CF 100

RULE-12: If AGE is Extrm High Then Decision is 1 With CF 100

RULE-13: If SEX is 1 and PPB is High Then Decision is 0 With CF 98

RULE-14: If SEX is 0 and PPB is High and ALB is 1 Then Decision is 1 With CF 97

RULE-15: If DUR is Medium and PPB is High Then Decision is 1 With CF 96

RULE-16: If AGE is High and PPB is High and ALB is 1 Then Decision is 1 With CF 94

RULE-17: If SEX is 0 and PPB is Low Then Decision is 0 With CF 92

RULE-18: If DUR is Medium and PPB is Low Then Decision is 0 With CF 92

RULE-19: If PPB is Low and ALB is 1 Then Decision is 0 With CF 92

GOUT

RULE-06: If FB is Extrm High Then Decision is 0 With CF 100

RULE-07: If DUR is Extrm High Then Decision is 0 With CF 100

RULE-08: If ALB is 2 Then Decision is 0 With CF 100

RULE-09: If AGE is Extrm High and FB is Medium Then Decision is 1 With CF 96

RULE-10: If AGE is High Then Decision is 0 With CF 95

RULE-11: If AGE is Very Low Then Decision is 0 With CF 94

RULE-12: If PPB is Low Then Decision is 0 With CF 94

RULE-13: If INS is 0 and FB is Medium and PPB is Very Low Then Decision is 1 With CF 93

RULE-14: If FB is Medium and PPB is Very Low and ALB is 0 Then Decision is 1 With CF 93

RULE-15: If PPB is Medium Then Decision is 0 With CF 93

OAH

RULE-09: If DUR is Extrm High Then Decision is 1 With CF 100

RULE-10: If AGE is Extrm High Then Decision is 1 With CF 100

RULE-11: If SEX is 1 and TYPE is 2 and INS is 1 Then Decision is 1 With CF 100

RULE-12: If TYPE is 2 and INS is 1 and ALB is 1 Then Decision is 1 With CF 100

RULE-13: If AGE is High and INS is 0 and ALB is 1 Then Decision is 0 With CF 97

RULE-14: If AGE is High and SEX is 1 and INS is 0 Then Decision is 0 With CF 97

RULE-15: If AGE is High and SEX is 1 and INS is 1 Then Decision is 1 With CF 94

RULE-16: If AGE is High and INS is 1 and ALB is 1 Then Decision is 1 With CF 94

RULE-17: If SEX is 0 and DUR is High Then Decision is 0 With CF 94

RULE-18: If DUR is High and ALB is 0 Then Decision is 0 With CF 94

OAK

RULE-06: If INS is 1 Then Decision is 0 With CF 100

RULE-07: If AGE is Extrm High and SEX is 1 Then Decision is 1 With CF 100

RULE-08: If DUR is Extrm Low Then Decision is 1 With CF 100

RULE-09: If AGE is Extrm High and PPB is Very Low Then Decision is 1 With CF 99

RULE-10: If SEX is 0 and FB is High Then Decision is 0 With CF 98

RULE-11: If PPB is High Then Decision is 1 With CF 97

RULE-12: If SEX is 0 and DUR is Very High Then Decision is 0 With CF 96

RULE-13: If FB is Extrm High Then Decision is 0 With CF 95

RULE-14: If AGE is Low Then Decision is 0 With CF 95

RULE-15: If AGE is High and SEX is 1 Then Decision is 1 With CF 95