Rules Reduction Using Decision Matrix

Durgesh Srivastava
Computer Science & Engineering Dept.
BRCM CET, Bahal, Bhiwani, Haryana, INDIA
dkumar.bit@gmail.com

Shweta Bhalothia
Computer Science & Engineering Dept.
BRCM CET, Bahal, Bhiwani, Haryana, INDIA
shwetabhalothia252@gmail.com

Abstract-The rules are the prescribed standards on the basis of which decisions are made for specific purpose. The rule is a statement that establishes a principle or standard, and serves as a norm for guiding or mandating action or conduct. The rule can be a conditional statement that tells the system how to react to a particular situation.This paper presents an rule reduction approach to reduce rules using decision matrix.We have also used this approach on dengue data set.The result shows that this approach produces effective and minimal rules.

Keywords-Rule reduction:minimal:dengue data;decision matrix.

1.INTRODUCTION

Data classification is the categorization of data for its most effective and efficient use.In machine learning, classification problem is of identifying to which of a set of categories a new observation belongs, on the basis of a training set of data containing observations whose category membership is known.Here we are using the approach presented in the paper "Efficient Rule Set Generation Using K-Map & Rough Set Theory (RST)" to generate rule set.After generating the rule set we apply our rule reduction approach using decision matrix to reduce the rules.

Rough Set Theory, proposed in 1982 by Zdzislaw Pawlak, is concerned with the classification . It is also concerned with analysis of imprecise, uncertain or incomplete information and Knowledge , and of is considered one of the first non-statistical approaches in data analysis . Very large data sets, Mixed types of data (continuous valued, symbolic data), Uncertainty (noisy data), Incompleteness (missing, incomplete data), Data change, Use of background knowledge etc. Are real world issues. Rough set theory is very useful to deal with these issues. Rough sets can handle uncertainty and vagueness , discovering patterns in inconsistent data.

Let us say that we wish to find the minimal set of consistent rules (logical implications) that characterize our sample system. For a set of condition attributes $\mathcal{P} = \{ P_1, P_2, P_3, \ldots\ldots P_n\}$ and a decision attribute Q,Q is not element of $\mathcal{P}$, these rules should have the form $P_i^a P_j^b \ldots\ldots P_k^c \Rightarrow Q^d$, or, spelled out,

$$(P_i = a) \wedge (P_j = b) \wedge \ldots\ldots\ldots\ldots\ldots\ldots \wedge (P_k = c) \Rightarrow (Q = d)$$

Where { a,b,c,...} are legitimate values from the domains of their respective attributes. This is a form typical of association rules, and the number of items in U which match the condition/antecedent is called the support for the rule. The method for extracting such rules given in Ziarko & Shan (1995) is to form a decision matrix corresponding to each individual value d of decision attribute Q. Informally, the decision matrix for value d of decision attribute Q lists all attribute–value pairs that differ between objects having Q = d and Q ≠ d.

This paper is organized as follows. Section 2 describes some basic concepts of Rough Set Theory. Section 3, highlights the concept of decision matrix. Section 4, gives the introduction of proposed work. Experimental results are described in Section 5.Section 6 gives the conclusion.

2.ROUGH SET THEORY

Rough set theory proposed by Z. Pawlak in 1982, is a new mathematical approach to imperfect knowledge. Rough set theory has a close connections with many other theories. Despite of its connections with other theories, the rough set theory may be considered as own independent discipline. Rough sets have many advantages and because of this have been used in many applications such as Artificial Intelligence and cognitive sciences,especially in machine learning, knowledge discovery,data mining,expert systems,approximate reasoning and pattern recognition.

Basic concepts of Rough Set Theory

An **information system** is a pair $S = (U, P)$ where U is a non-empty finite set of *objects* called the *universe* and A is a non-empty finite set of *attributes* such that $a : U \rightarrow V_a$ for every $a \in P$.

The set V_a is called the *value set* of a.

A **decision system** is any information system of the form $S = (U, P \cup \{d\})$, where d A is the *decision attribute.* The elements of A are called *conditional attributes* or simply *conditions.*

Let Q □□ *P.Then Q-indiscernibility relation denoted by IND*(Q), is defined as following

$$IND(Q) = \{x, x' \in U^2 \mid \forall a \in Q\ a(x) = a(x')\}$$

If □*x, x*'□□ □□□□*IND* □*Q*□□then objects *x* and *x*' are indiscernible from each other by attributes from *Q*. The equivalence classes of the *Q-indiscernibility* relation are denoted $[x]_Q$.

The **discernibility matrix** of S is a symmetric $n \times n$ matrix with entries c_{ij} as given below.

$$c_{ij} = \{a \in P \mid a(x_i) \neq a(x_j)\} \text{ for } i,j = 1,.....,n$$

Each entry thus consists of the set of attributes upon which objects x_i and x_j differ. Since discernibility matrix is symmetric and $c_{ii} = \emptyset$□(the empty set) for$i$=1,.,n. Thus, this matrix can be represented using only elements in its lower triangular part, i.e. for$1 \leq j < i \leq n$.

A **discernibility function** f_S for an information system S is a Boolean function of m Boolean variables a_1^*,.....,a_m^*(corresponding to the attribute a_1,...,a_m) defined as follows.

$$f_s(a_1, ..., a_m) = \forall\{\exists c_{ij}^* \mid 1 \leq j \leq i \leq n, c_{ij} \neq \varphi\}$$

where $c_{ij}^* = \{a^* \mid a \in c_{ij}\}$. *The set of all prime implicants of* f_s *determines the set of all* reducts of *P*.The discernibility function is an boolean function in POS form.

Approximations of set

Let X be a subset of U, i.e. $X \subseteq U$.

Lower Approximation: For a given concept, its lower approximation refers to the set of observations that can all be classified into this concept.

$$Q_*(X) = \{x : [x]_Q \subseteq X\}$$

Upper Approximation: For a given concept, its upper approximation refers to the set of observations that can be possibly classified into this concept.

$$Q^*(X) = \{x : [x]_Q \cap X \neq \emptyset\}$$

Once the reducts have been computed, the rules are easily constructed by overlaying the reducts over the originating decision table and reading off the values.

3. DECISION MATRIX

How decision matrix works is best explained by example .Consider the following table:

TABLE 1 SAMPLE TABLE

Objects	P	Q	R	S
O_1	1	2	0	1
O_2	1	2	0	1
O_3	2	0	0	1
O_4	0	0	1	0
O_5	2	1	0	0
O_6	0	0	1	0
O_7	2	0	0	1
O_8	0	1	2	0
O_9	2	1	0	0
O_{10}	2	0	0	1

S is the decision variable (i.e., the variable on the right side of the implications) and P, Q, R are the condition variables (on the left side of the implication). We note that the decision variable S takes on two different values, namely {1,0}. We treat each case separately.

First, we look at the case S= 1 , and we divide up U into objects that have S= 1 and those that have have S = 0.In this case, the objects having S = 1 are $\{O_1, O_2, O_3, O_7, O_{10}\}$ while the objects which have S = 0 are $\{O_4, O_5, O_6, O_8, O_9\}$ .The decision matrix for S= 1 lists all the differences between the objects having S = 1 and those having S = 0 ; that is, the decision matrix lists all the differences between $\{O_1, O_2, O_3, O_7, O_{10}\}$ and $\{O_4, O_5, O_6, O_8, O_9\}$. We put the "positive" objects (S = 1) as the rows, and the "negative" objects S = 0 as the columns.

TABLE 2 DECISION MATRIX FOR S=1

Decision matrix for S = 1					
Object	**O_4**	**O_5**	**O_6**	**O_8**	**O_9**
O_1	P^1,Q^2,R^0	P^1,Q^2	P^1,Q^2,R^0	P^1,Q^2,R^0	P^1,Q^2
O_2	P^1,Q^2,R^0	P^1,Q^2	P^1,Q^2,R^0	P^1,Q^2,R^0	P^1,Q^2
O_3	P^2,R^0	Q^0	P^2,R^0	P^2,Q^0,R^0	Q^0
O_7	P^2,R^0	Q^0	P^2,R^0	P^2,Q^0,R^0	Q^0
O_{10}	P^2,R^0	Q^0	P^2,R^0	P^2,Q^0,R^0	Q^0

To read this decision matrix, look, for example, at the intersection of row O_3 and column O_6 , showing P^2 ,R^0 in the cell. This means that with regard to decision value S = 1, object O_3 differs from object O_6 on attributes P and R , and the particular values on these attributes for the positive object O_3 are P=2 and R=0. This tells us that the correct classification of O_3 as belonging to decision class S=1 rests on attributes P and R; although one or the other might be dispensable, we know that at least one of these attributes is indispensable.

Next, from each decision matrix we form a set of Boolean expressions, one expression for each row of the matrix. The items within each cell are aggregated disjunctively, and the individuals cells are then aggregated conjunctively. Thus, for the above table we have the following five Boolean expressions:

$(P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2) \wedge (P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2)$

$(P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2) \wedge (P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2)$

$(P^2 \vee R^0) \wedge (Q^0) \wedge (P^2 \vee R^0) \wedge (P^2 \vee Q^0 \vee R^0) \wedge (Q^0)$

$(P^2 \vee R^0) \wedge (Q^0) \wedge (P^2 \vee R^0) \wedge (P^2 \vee Q^0 \vee R^0) \wedge (Q^0)$

$(P^2 \vee R^0) \wedge (Q^0) \wedge (P^2 \vee R^0) \wedge (P^2 \vee Q^0 \vee R^0) \wedge (Q^0)$

Each statement here is essentially a highly specific (probably too specific) rule governing the membership in class S=1 of the corresponding object. For example, the last statement, corresponding to object O_{10} , states that all the following must be satisfied:

- Either P must have value 2, or R must have value 0, or both.
- Q must have value 0.
- Either P must have value 2, or R must have value 0, or both.
- Either P must have value 2, or Q must have value 0, or R must have value 0, or any combination thereof.
- Q must have value 0.

It is clear that there is a large amount of redundancy here, and the next step is to simplify using traditional Boolean algebra.The statement $(P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2) \wedge (P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2 \vee R^0) \wedge (P^1 \vee Q^2)$ corresponding to objects { O_1 ,O_2 }simplifies to $P^1 \vee Q^2$, which yields the implication

$$(P = 1) \vee (Q = 2) \Rightarrow (S = 1)$$

Likewise, the statement $(P^2 \vee R^0) \wedge (Q^0) \wedge (P^2 \vee R^0) \wedge (P^2 \vee Q^0 \vee R^0) \wedge (Q^0)$ corresponding to objects { O_3,O_7,O_{10}} simplifies to $P^2Q^0 \vee R^0Q^0$. This gives us the implication

$$(P = 2 \wedge Q = 0) \vee (R = 0 \wedge Q = 0) \Rightarrow (S = 1)$$

The above implications can also be written as the following rule set:

(P = 1) => (S =1)

(Q = 2) => (S =1)

(P = 2) ∧ (Q = 0) => (S = 1)

(R = 0) ∧ (Q = 0) => (S = 1)

It can be noted that each of the first two rules has a support of 1 (i.e., the antecedent matches two objects), while each of the last two rules has a support of 2. To finish writing the rule set for this knowledge system, the same procedure as above (starting with writing a new decision matrix) should be followed for the case of S = 0 , thus yielding a new set of implications for that decision value (i.e., a set of implications with S = 0 as the consequent). In general, the procedure will be repeated for each possible value of the decision variable.

4. PROPOSED WORK

The presented work is to deal with the reduction of rules. The basic methodology of the proposed work is defined under the following steps.

1. Representation of the information system
2. Discernibility matrices
3. Discernibility functions
4. Reduction of attributes using K-map
5. New reduct table
6. Set approximation
7. Rule generation
8. Rule reduction using decision matrix
9. Validation phase of generated rules

5. EXPERIMENTAL RESULT

Dengue Data Set is used in applying the proposed work.The data is discretized by using Rough Set Theory and K-map.The data is splitted into training set and test set.In Dengue Data Set we have taken twentyfive patients data.The data about twenty patients is used as training data and the data about five patients is used as test data.The attributes for this data are blotched-red skin,muscular-pain-articulation,temperature and dengue where blotched-red skin,muscular-pain-articulation,temperature are conditional attributes and dengue is decision attribute.Samples of decision rules for Dengue data set generated from reduct are given below:

Here several patients data set is taken with possible dengue symptoms.By using k-map & Rough set approach ,we analyze the data, eliminate the redundant data , reduce the attributes and develop a set of rules .Below the table is shown with the patients data set and respective symptoms.

Step1:Taken Dengue Data Set:

TABLE 3 DENGUE DATA SET

Patient	Blotched-red skin	Muscular-pain articulations	Temperature	Dengue
P1	No	No	Normal	No
P2	No	No	High	No
P3	No	No	Very high	Yes
P4	No	Yes	High	Yes
P5	No	Yes	Very high	Yes
P6	Yes	Yes	High	Yes
P7	Yes	Yes	Very high	Yes
P8	No	No	High	No
P9	Yes	No	Very high	Yes
P10	Yes	No	High	No
P11	Yes	No	Very high	No
P12	No	Yes	Normal	No
P13	No	Yes	High	Yes
P14	No	Yes	Normal	No
P15	Yes	Yes	Normal	No
P16	Yes	No	Normal	No
P17	Yes	No	High	No
P18	Yes	Yes	Very high	Yes
P19	Yes	No	Normal	No
P20	No	Yes	Normal	No
P21	No	Yes	High	Yes
P22	Yes	Yes	Very high	Yes
P23	No	Yes	Normal	No
P24	No	Yes	Very high	Yes
P25	Yes	No	High	No

Here Q is the set of objects(patients),given set

Q={P1,P2,P3,P4,P5,P6,P7,P8,P9,P10,P11,P12,P13,P14,P15,P16,P17,P18,P19,P20}

A is the set of conditional attributes,given set A={Blotched-red skin,Muscular-pain-articulations pain,Temperature} and the set D represents the decision attribute ,where D={Dengue}.The table shown below represents the nominal values of the attributes:

	Attributes	Nominal Values
Conditional Attributes	Blotched-red skin	Yes,No
	Muscular-pain-articulations	Yes,No
	Temperature	Normal,High,Very High
Decision Attributes	Dengue	Yes,No

Step2 :Discernibility Matrix:

Discernibility matrix of an information system is a symmetric n×n matrix with entries cij. Each entry consists of the set of attributes upon which objects xi and xj differ.Since discernibility matrix is symmetric and cii=Φ for i=1,.....,n. Thus this matrix can be represented using only elements in its lower triangular part i.e. for 1≤j<i≤n.Since in the above given information system we have 20 objects,so the discernibility matrix for this system is a 20×20 matrix.The complete discernibility matrix for above information system is given below:

TABLE 4 DISCERNIBILITY MATRIX FOR DDS

	P1	**P2**	**P3**	**P4**	**P5**	**P6**	**P7**	**P8**	**P9**	**P10**	**P11**	**P12**	**P13**	**P14**	**P15**	**P16**	**P17**	**P18**	**P19**	**P20**
P1																				
P2	T																			
P3	T	T																		
P4	M,T	M	M,T																	
P5	M,T	M,T	M	T																
P6	B,M,T	B,M	B,M,T	B	B,T															
P7	B,M,T	B,M,T	B,M	B,T	B	T														
P8	T	-----	T	M	M,T	B,M	B,M,T													
P9	B,T	B,T	B	B,M,T	B,M	M,T	M	B,T												
P10	B,T	B	B,T	B,M	B,M	M	M,T	B	T											
P11	B,T	B,T	B	B,M,T	B,M	M,T	M	B,T	-----	T										
P12	M	M,T	M,T	T	T	B,T	B,T	M,T	B,M,T	B,M,T	B,M,T									
P13	M,T	M	M.T	-----	T	B	B,T	M	B,M,T	B,M	B,M,T	T								
P14	M	M,T	M,T	T	T	B,T	B,T	M,T	B,M,T	B,M,T	B,M,T	------	T							
P15	B,M	B,M,T	B,M,T	B,T	B,T	T	T	B,M,T	B,M,T	M,T	B,M,T	B	B,T	B						
P16	B	B,T	B,T	B,M,T	B,M,T	M,T	M,T	B,T	T	T	T	B,M	B,M,T	B,M	M					

P17	B,T	B	B	B,M	B,M,T	M	M,T	B	T	-----	T	B,M,T	B,M	B,M,T	M,T	T			
P18	B,M,T	B,M,T	B,M,T	B,T	B	T	------	B,M,T	M	M,T	M	B,T	B,T	B,T	T	M,T	M,T		
P19	B	B,T	B,T	B,M,T	B,M,T	M,T	M,T	B,T	T	T	T	B,M	B,M,T	B,M	M	M,T	T	MT	
P20	M	M,T	M,T	T	T	B,T	B,T	M,T	B,MT	B,MT	B,M,T	-----	T	-----	B	B,M	B,M,T	B,T	B,M

In this table,B,M,T denote Blotched-red skin,Muscular-pain-articulations and Temperature respectively.

Step3:Discernibility Function:

A discernibility function for an information system is a boolean function of boolean variables. With every discernibility matrix one can associate a discernibility function.For above discernibility matrix, discernibility function is given below:

f(B,M,T)=

T(T)(M∨T)(M∨T)(B∨M∨T)(B∨M∨T)T(B∨T)(B∨T)(B∨T)M(M∨T)M(B∨M)B(B∨T)(B∨M∨T)(B)M T(M)(M∨T)(B∨M)(B∨M∨T)(B∨T)B(B∨T)(M∨T)M(M∨T)(B∨M∨T)(B∨T)B(B∨M∨T)(B∨T)(M∨T) (B∨M)

Each row in the above discernibility function corresponds to one column in the discernibility matrix.Each paranthesized tuple is a sum in the Boolean expression and one-letter Boolean variables correspond to attribute names.

Step4:K-map for this function:

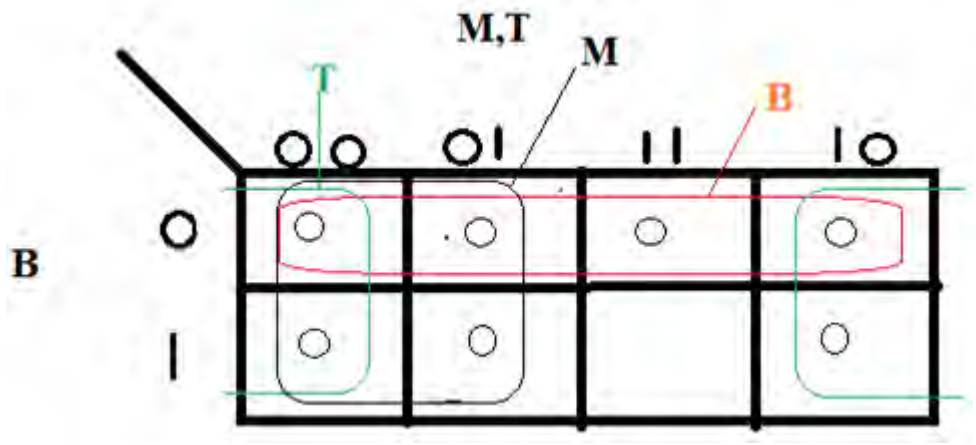

Figure 1 K-Map For Discernibility Function

Reduct=BMT

Step5 : Reduct Table:

After the simplification of K-map we obtain the following expression

BMT

Which says that there is only one reduct{B,M,T} in the data table and it is the core.That means no attribute can be eliminated from the table .The new reduct table is same as the original table.

Step6: Set Approximation:

The lower and the upper approximations of a set are interior and closure operations in a topology generated by a indiscernibility relation. Indiscernibility Relation is the relation between two objects or more, where all the values are identical in relation to a subset of considered attributes.The indiscernibility relation for given information system is following:

IND(Q)={{P1},{P2,P8},{P3},{P4,P13},{P5},{P6},{P7,P18},{P9,P11},{P10,P17},{P12,P14,P20},{P15},{P16,P19}}

$$X=\{x:Dengue(x)=Yes\}=\{P3,P4,P5,P6,P7,P9,P13,P18\}$$

A={Blotched-red skin,Muscular-pain-articulations,Temperature}

Lower Approximation set $A_*(X)$ of the patients that are definitely have Flu are identified as

$$A_*(X)=\{P3,P4,P5,P6,P7,P13,P18\}$$

Upper Approximation set A*(X) of the patients that possibly have Flu are identified as

$$A^*(X)=\{P3,P4,P5,P6,P7,P9,P11,P13,P18\}$$

Boundary region is defined as follows

$$AN^A(X)=\{P9,P11\}$$

Boundary Region (BR), the set constituted by elements P9 and P11, which cannot be classified, since they possess the same characteristics, but with differing conclusions differ in the decision attribute.

Step7: Rule Generation:

After computing the reduct and set approximations we divide the data set into classes on the bases of the values of the attributes i.e. find out the indiscernibility relation.

IND(Q)={{P1},{P2,P8},{P3},{P4,P13},{P5},{P6},{P7,P18},{P9,P11},{P10,P17},{P12,P14,P20},{P15},{P16,P19}}

Since {P9,P11} is the boundary region,it represents uncertainty.So we discard this class,

IND(Q)={{P1},{P2,P8},{P3},{P4,P13},{P5},{P6},{P7,P18},{P10,P17},{P12,P14,P20}, {P15}, {P16,P19}}

Now we have 11 classes,generate one rule for each class.The generated rules are:

1. (B=No)&(M=No)&(T=Normal)=>(D=No)
2. (B=No)&(M=No)&(T=High)=>(D=No)
3. (B=No)&(M=No)&(T=Very High)=>(D=Yes)
4. (B=No)&(M=Yes)&(T=High)=>(D=Yes)
5. (B=No)&(M=Yes)&(T=Very High)=>(D=Yes)
6. (B=Yes)&(M=Yes)&(T=High)=>(D=Yes)
7. (B=Yes)&(M=Yes)&(T= Very High)=>(D=Yes)
8. (B=Yes)&(M=No)&(T= High)=>(D=No)
9. (B=No)&(M=Yes)&(T= Normal)=>(D=No)
10. (B=Yes)&(M=Yes)&(T= Normal)=>(D=No)
11. (B=Yes)&(M=No)&(T= Normal)=>(D=No)

Step8: Rules reduction using decision matrix

First of all we convert the rules in the form of table.

TABLE 5 TABLE GENERATED FROM RULES

Rules	B	M	T	D
1	0	0	1	0
2	0	0	2	0
3	0	0	3	1
4	0	1	2	1
5	0	1	3	1
6	1	1	2	1
7	1	1	3	1
8	1	0	2	0
9	0	1	1	0
10	1	1	1	0
11	1	0	1	0

Here for attributes B,M and D : 0 represents No and 1represents Yes

For attribute T: 1represents Normal, 2 represents High and 3 represents Very high

For D=1 : (3,4,5,6,7)

For D=0 : (1,2,8,9,10,11)

TABLE 6 DECISION MATRIX FOR D=1

	1	**2**	**8**	**9**	**10** **11**
3	T^3	T^3	B^0T^3	M^0T^3	$B^0M^0T^3$ B^0T^3
4	M^1T^2	M^1	B^0M^1	T^2	B^0T^2 $B^0M^1T^2$
5	M^1T^3	M^1T^3	$B^0M^1T^3$	T^3	B^0T^3 $B^0M^1T^3$
6	$B^1M^1T^2$	B^1M^1	M^1	B^1T^2	T^2
7	$B^1M^1T^3$	$B^1M^1T^3$	M^1T^3	B^1T^3	T^3 M^1T^3

Solve the matrix row wise:

For first row :

$(T^3) \wedge (T^3) \wedge (B^0 \vee T^3) \wedge (M^0 \vee T^3) \wedge (B^0 \vee M^0 \vee T^3) \wedge (B^0 \vee T^3)$

After simplify this by applying the laws of boolean algebra we get the result: T^3

Like this after simplify for all row we have the results($M^1 \wedge T^2$), T^3,($M^1 \wedge T^2$), T^3

So from the decision matrix for D=1 we get two rules:

(T=3)=>(D=1)

(M=1)&(T=2)=>(D=1)

i.e.

(T=Very high)=>(Dengue=Yes)

(M=Yes)&(T=High)=>(Dengue=Yes)

TABLE 7 Decision matrix for D=0

	3	**4**	**5**	**6** **7**
1	T^1	M^0T^1	M^0T^1	$B^0M^0T^1$ $B^0M^0T^1$
2	T^2	M^0	M^0T^2	B^0M^0 $B^0M^0T^2$
8	B^1T^2	B^1M^0	$B^1M^0T^2$	M^0 M^0T^2
9	M^1T^1	T^1	T^1	B^0T^1 B^0T^1
10	$B^1M^1T^1$	B^1T^1	B^1T^1	T^1 T^1
11	B^1T^1	$B^1M^0T^1$	$B^1M^0T^1$	M^0T^1 M^0T^1

Solve the matrix row wise:

For first row :

$(T^1) \wedge (M^0 \vee T^1) \wedge (M^0 \vee T^1) \wedge (B^0 \vee M^0 \vee T^1) \wedge (B^0 \vee M^0 \vee T^1)$

After simplify this by applying the laws of boolean algebra we get the result: T^1

Like this after simplify for all row we have the results($M^0 \wedge T^2$),$(B^1 \wedge M^0) \vee (M^0 \wedge T^2)$, T^1, T^1, $T^1 \vee (B^1 \wedge M^0)$

So from the decision matrix for D=0 we get two rules:

(T=1)=>(D=0)

(M=0)&(T=2)=>(D=0)

(B=1)&(M=0)=>(D=0)

(B=1)&(T=2)=>(D=0)

i.e.

(T=Normal)=>(Dengue=No)

(M=No)&(T=High)=>(Dengue=No)

(B=Yes)&(M=No)=>(Dengue=No)

(B=Yes)&(T=High)=>(Dengue=No)

So after performing the rule reduction approach we have the folliwing rules:

1. (T=Very high)=>(Dengue=Yes)
2. (M=Yes)&(T=High)=>(Dengue=Yes)
3. (T=Normal)=>(Dengue=No)
4. (M=No)&(T=High)=>(Dengue=No)
5. (B=Yes)&(M=No)=>(Dengue=No)
6. (B=Yes)&(T=High)=>(Dengue=No)

6.CONCLUSION

An efficient rule reduction approach has been presented to reduce the rules by using Decision Matrix.We applied this approach to Dengue Data Set which has 25 objects.Out of 25 objects we took 20 objects as training data and 5 objects as testing data.Before applying this approach we have 11 rules but after applying this approach we have 6 rules.We reduced 5 rules by applying this approach.So this approach is an efficient approach.

REFERENCES

[1] Durgesh Srivastava, Shweta Bhalothia: “Efficient Rule Set Generation Using K-Map & Rough Set Theory (RST)”

[2] Junbo Zhang , Tianrui Li , Da Ruan b, Zizhe Gao, Chengbing Zhao :”A parallel method for computing rough set approximations”. Information Sciences 194 (2012) 209–223 Elsevier 2012

[3] Prasanta Gogoi, Ranjan Das, B Borah & D K Bhattacharyya:”Efficient Rule Set Generation using Rough Set Theory for Classification of High Dimensional Data”. International Journal of Smart Sensors and Ad Hoc Networks (IJSSAN) ISSN No. 2248-9738 (Print) Volume-1, Issue-2, 2011

[4] Puniethaa Prabhu,Duraiswamy:“Feature selection for HIV database using Rough system”. Second International conference on Computing, Communication and Networking Technologies,978-1-4244-6589-7/10/ IEEE(2010)

[5] Prof. Ayesha Butalia.Prof. M.L.Dhore, Ms. Geetika Tewani:” Applications of Rough Sets in the field of Data Mining”.First International Conference on Emerging Trends in Engineering and Technology.978-0-7695-3267-7/08 IEEE(2008)

[6] Xiuping Xu,James F. Peters:”Rough Set Methods in Power System Classification”

a. Canadian Conference on Electrical and Computer Engineering 0-7803-7514-9 IEEE(2002).

[7] Zdzisław Pawlak:” Rough set theory and its applications”.Journal of Telecommunications and Information Technology 3/2002.

[8] J. J. Alpigini, J. F. Peters, J. Skowron and N.Zhong. Rough Sets and Current Trends in Computing, Third International Conference, USA, 2002.

[9] Daijin Kim:”Data classification based on tolerant rough set”. 2001 Pattern Recognition Society. Published by Elsevier Science Ltd

[10] J.F. Peters, A. Skowron, “A rough set approach to reasoning about data”, International Journal of Intelligent Systems, vol. 15,2001, 1-2.

[11] Z.Pawlak, “Rough Sets”, International Journal of Computer and Information Sciences, Vol.11, 341-356(1982)