International Journal of Scientific Research in Mathematical and Statistical Sciences
Research Paper
Vol.4, Issue.3, pp.12-18, June (2017)
E-ISSN: 2348-4519

Rough Set Based Rule Generation Techniques in Medical Diagnosis: With Reference to Identification of Heart Disease

Pragati Jain[1*], Kamini Agrawal[2], Deepensha Vaishnav[3]

[1]Department of Mathematics, SPIPS, Indore, India
[2]Department of Mathematics, IET, D.A.V.V., Indore, India
[3]Department of Mathematics, IET, D.A.V.V., Indore, India

[*]Corresponding Author: dr.pragatijain@gmail.com

Available online at: www.isroset.org

Received 24th Feb 2017, Revised 15th Mar 2017, Accepted 14th Apr 2017, Online 30th Jun 2017

Abstract — Rough Set Theory proposed by Pawlak in 1982 has now become very significant in the field of data mining and knowledge discovery. This led to his most widely recognized contribution to classifying objects with their attributes and his introduction of approximation spaces, which establish the foundation of granular computing and provide framework for perception and knowledge discovery in many areas. The theory of rough sets has been under continuous development and a fast growing group of researchers and practitioners are interested in this methodology. The theory has found many interesting applications in medicine, pharmacology, business, banking, market research, engineering design, meteorology, vibration analysis, switching function, conflict analysis, image processing, voice recognition, concurrent system analysis, decision analysis, character recognition, and other fields. The paper presents the concept of Rough Set Theory along with the Rule Generation Techniques in Medical Diagnosis in general and particularly for the identification of heart disease through the Core and Reduct of the concerned attributes of an information system.

Keywords — *Rough Set Theory, Data Mining, Heart Disease, QRA, LEM2 Algorithm.*

I. INTRODUCTION

Heart disease is one of the major issues arising in the world. The main risk factors are smoking, cholesterol, blood pressure, diabetes and age. But these factors lead to uncertainty and vagueness due to which experts have to face certain difficulties in diagnosing the patients of heart disease. To adjourn these difficulties regarding uncertainty, some computational techniques such as: neural network, fuzzy set theory, and rough set theory are being employed.

Since medical issues can be compiled on many factors or attributes; it is obvious to get large dataset with lots of vagueness. Rough Set Theory (RST) precisely works on uncertainty, imperfection among the dataset without any loss of information. The theory provides classification of objects with their attributes on the basis of approximation spaces. The lower and upper approximation space gives appropriate area/region of the decision attribute. The tendency of the theory to reduce attributes i.e. findings of reducts and core, helps in diagnosing the patients of heart disease based on various important conditions.

This paper presents some of the rule generation techniques based on RST. Firstly discernibility matrix is used to find reducts. Then some of the algorithms like LEM2, Johnson's Hueristic Algorithm, and Quick Reduct Algorithm (QRA) are employed on the same dataset. Finally, rules are depicted for classification of heart disease based on the reducts generated from the aforementioned techniques.

The paper is organized as follows:

Section I contains the introduction, Section II contains the review of literature, Section III contains importance of the study, Section IV contains material and methods, Section V contains result and discussion, Section VI contains conclusion.

II. REVIEW OF LITERATURE

The *Cantorian or Naïve set theory* introduced by Cantor leads to various contradictions or paradoxes. This concept is revised and classical sets, also known as *crisp sets* came into existence. Further two Non- Classical Set Theories become popular because of their characteristic to deal with real-valued data having vagueness and incompleteness. One of the two methods is Fuzzy Set; introduced by Lofti A. Zadeh in 1965 [14] and another one is introduced in 1982 by

Zdzislaw Pawlak, known as Rough Set [7]. Former is based on membership functions whereas later one is based on the boundary region. Now-a-days Rough Set is being applied to various fields with implementation of its different forms. N.A. Setiawan et. al. [9] predicts the real missing attribute value in the heart disease database with the help of artificial neural network and RST. Aleksander and Todd gave a hypothesis generator for medical field explained on patients with spinal cord injury to predict their ambulation [6]. The aforesaid authors also presented a non-technical way to explicate the fundamentals of Rough Set Theory. In contingency management, Rough Set is employed to recognize the valve fault [3]. In 2010, the paper [8] provided an algorithm to solve constrained multi-objective optimization problems. The aforementioned algorithm is derived from the hybridization of a multi-objective evolutionary algorithm and local search method based on RST. Long-Jun Huang et al. [5] use Rough Set Theory (RST) to analyze the prognosis factors, remove the redundant attributes in the database and mine the useful information. The authors of [13] depict a technique for Medical Decision Making for the student suffering from study's anxiety based on data clustering and Variable Precision Rough Set (VPRS) Model. The authors of paper [1] proved that Rough Set based Supporting Vector Machine classifer (RS_SVM) is one of the high accuracy classifier. RS_SVM was experimented on the dataset of Wisconsin Breast Cancer.

III. IMPORTANCE OF THE STUDY

The theory allows to reduce the original data without any loss of information. It does not require any preliminary information about data and provides efficient methods for finding hidden patterns in data. The generated set of decision rules based on RS concept gives clear and precise knowledge of data on which study is being conducted. Medical datasets are large enough including uncertainties, imperfection and vagueness. When considering various risk factors related to heart disease while diagnosing the patient, experts often have to face hardship. These requirements of medical issues and characteristics of RST suggest that the combination of the two can be proved very beneficial. In our study we tried to generate a decision system for diagnosing the patient of heart disease.

IV. MATERIAL AND METHODS

Definition 4.1: Some basic notions of RST used in this paper are as follows [10, 12].

Definition 4.1.1: An information system (IS) is a pair *(U, A),* where U is a nonempty, finite set of objects called the universe and A is a nonempty, finite set of attributes, such that $a: U \to V_a$ for any $a \in A$, where V_a is called the domain of attribute a.

Definition 4.1.2: Indiscernibility Relation is a central concept of RST, and is considered as a relation between two objects or more, where all the values are identical in relation to a subset of considered attributes.
For every set of attribute R $\subset$ A, an indiscernibility relation IND(R) is defined in the following way:

$$IND(R) = \{(x_i, x_j) \in U, b(x_i) = b(x_j), \forall b \in R\}.$$

Indiscernibility relation is an equivalence relation, where all identical objects of set are considered as elementary sets of the universe U in the space R i.e., partition determined by R, will be denoted by $U/IND(R)$, or simple U/R. An equivalence class of IND(R), i.e., block of the partition U/R, containing x will be denoted by R(x).

Definition 4.1.3: The lower Approximation contains all objects which surely belong to the set.

$$R_*(X) = \{x_i \in U : R(x_i) \subseteq X\}$$

Definition 4.1.4: The Upper Approximation ion contains all objects which possible belong to the set.

$$R^*(X) = \{x_i \in U : R(x_i) \cap X \neq \phi\}$$

Definition 4.1.5: The difference between Upper and Lower Approximation is called Boundary Region.

$$BN_R(X) = R^*(X) - R_*(X)$$

If the boundary region of X is the empty set, i.e., $BN_R(X) = \emptyset$, then the set X is *crisp* (*exact*) with respect to R; in the opposite case, i.e., if $BN_R(X) \neq \emptyset$, the set X is *rough* (*inexact*) with respect to R.

Rough set can be also characterized numerically as follows:

$$\alpha_R(X) = \frac{card(R^*(X))}{card(R_*(X))}$$

and is known as *accuracy of approximation*. Obviously, $0 \leq \alpha_R(X) \leq 1$. If $\alpha_R(X) = 1$, X is *crisp* with respect to R (X is *precise* with respect to R), and otherwise, if $\alpha_R(X) < 1$, X is *rough* with respect to R (X is *vague* with respect to R).

Definition 4.1.6: If we distinguish in an information system two disjoint classes of attributes, called condition and decision attributes, respectively, then the system will be called a decision table and will be denoted by $S = (U, P, Q)$, where P and Q are disjoint sets of condition and decision attributes, respectively.

Definition 4.1.7: Reducts are the most precise way of discerning object classes, which are the minimal subsets provided that the object classification is the same as with the full set of attributes. The core is common to all reducts.
The reducts process for attributes reduces elementary set numbers, the goal of which is to improve the precision of

decisions. After the attribute dependence process, the reduct attribute sets are generated to remove superfluous attributes, so that the set of attributes is dependent. The complete set of attributes is called a reduct attribute set. There may be more than one reduct attribute set in an information system, but intersecting a number of reduct attribute sets yields a core attribute set.

$$RED(R) \subseteq A$$
$$CORE(R) = \bigcap RED(R)$$

To compute reducts and core, the discernibility matrix is used.

A discernibility matrix of a decision table $(U, P \cup Q)$ is symmetric $|U| \times |U|$ matrix with entire defined by $P_{ij} = \{a \in P | a(x_i) \neq a(x_j)\}$ for $i, j = 1, 2, 3, \ldots\ldots\ldots\ldots |U|$. The discernibility matrix assigns to each pair of objects x and y a subset of attributes $\delta(x, y) \subseteq R$ with the following properties.

1. $\delta(x, x) = \phi$
2. $\delta(x, y) = \delta(y, x)$
3. $\delta(x, z) \subseteq \delta(x, y) \cup \delta(y, z)$

Definition 4.1.8: Any decision table induces a set of decision rules. A decision rule is an implication in the form of "*if* P *then* Q "

($P \rightarrow Q$). It is also known as "association rules" or "production rules". Any set of mutually exclusive and exhaustive decision rules that cover all facts in information system and preserves the indiscernibility relation included by information system is referred as called a decision algorithm in *S*. The set of decision rules is called decision algorithm.

Definition 4.1.9: Each row of the decision table specifies a decision rule which determine decisions in terms of condition.

Let $S = (U, P, Q)$ be a decision table. Every $x \in U$ determines a sequence $P_1(x)$, $P_2(x)$,…….. $P_n(x)$, $Q_1(x)$, $Q_2(x)$……., $Q_n(x)$.

Where $\{P_1(x), P_2(x), \ldots\ldots\ldots\ldots\ldots\ldots, P_n(x)\} = P$ and $\{Q_1(x), Q_2(x), \ldots\ldots\ldots\ldots\ldots\ldots, Q_n(x)\} = Q$

The sequence will be called a decision rule induced by x (in S) and denoted by $P_1(x), P_2(x), \ldots\ldots, P_n(x) \rightarrow Q_1(x), Q_2(x), \ldots\ldots, Q_n(x)$ or $P \rightarrow_x Q$.

The number $supp_x(P, Q) = |P(x) \cap Q(x)|$ will be called a support of decision rule $P \rightarrow_x Q$. The number

$$\sigma_x(P, Q) = \frac{supp_x(P, Q)}{|U|}$$

Is referred to as the *strength* of the decision rule $P \rightarrow_x Q$, where $|X|$ denotes the cardinality of X.

With every decision rule $P \rightarrow_x Q$ we associate the certainty factor of the decision rule, denoted $cer_x(P, Q)$ and defined as follows:

$$cer_x(P, Q) = \frac{supp_x(P, Q)}{|P(x)|}$$

If $cer_x(P, Q) = 1$, then $P \rightarrow_x Q$ will be called a certain decision rule; if $0 < cer_x(P, Q) < 1$ the decision rule is referred to as an *uncertain decision rule*.

Besides, a *coverage factor* of the decision rule, denoted as $cov_x(P, Q)$ and defined as

$$cov_x(P, Q) = \frac{supp_x(P, Q)}{|Q(x)|}$$

If $P \rightarrow_x Q$ is a decision rule then $Q \rightarrow_x P$ is called an inverse decision rule. The inverse decision rules can be used to give *explanations (reasons)* for a decision.

Algorithm 4.2: Following are few algorithms used in this paper.

Algorithm 4.2.1: Quick Reduct Algorithm (QRA)

The QUICKREDUCT algorithm attempts to calculate a reduct without exhaustively generating all possible subsets. It starts off with an empty set and adds, in turn, one at a time, those attributes that result in the greatest increase in the rough set dependency metric, until this produces its maximum possible value for the dataset [11].

This, however, is not guaranteed to find a minimal subset. Using the dependency function to discriminate between candidates may lead the search down a non minimal path. It is impossible to predict which combination of attributes will lead to an optimal reduct based on changes in dependency with the addition or deletion of single attributes. It does result in a close-to-minimal subset, though, which is still useful in greatly reducing dataset dimensionality.

QUICK_REDUCT (P, Q)

P, the set of all conditional features
Q, the set of decision features
$A \leftarrow \{\ \}$
Do
$T \leftarrow A$
$\forall x \in (P - A)$
If $\gamma_{A\cup\{x\}}(Q) > \gamma_T(Q)$
Where
$\gamma_A(Q) = card(POS_A(Q))/card(U)$
$T \leftarrow A \cup \{x\}$
$A \leftarrow T$
Until $\gamma_A(Q) = \gamma_P(Q)$
return A

Algorithm 4.2.2: Johnson's Heuristic Algorithm
Johnsons _Reduct (F_P, f_Q, C)

Input F_P **:** conditional features, f_Q : decision feature, C:cases
Output R: Reduct $R \subseteq F_p$

1. $R \leftarrow \phi, F' \leftarrow F_p$.
2. $M \leftarrow$ Compute Discernibility Matrix (C, F', f_Q)
3. Do
4. $f_h \leftarrow$ Select Highest Scoring Feature (M)
5. $R \leftarrow R \cup \{f_h\}$
6. For $(i = 0\, to\, |C|,\ j = i\, to\, |C|)$
7. $m_{ij} \leftarrow \phi$ if $f_h \in m_{ij}$
8. $F' \leftarrow F' - \{f_h\}$
9. Until $m_{ij} = \phi \ \ \forall i, j$
10. Return R

The Johnson's Heuristic Algorithm is used to calculate reduct for a decision problem. It sequentially selects features by finding those that are most discernible for given decision feature [2]. It computes a discernibility matrix M, where

$$m_{ij} = \left\{ \begin{array}{l} \{ f \in F_P : f(x_i) \neq f(x_j) \} \\ for\ f_Q(x_i) \neq f_Q(x_j), and\ \phi\ otherwise \end{array} \right\}$$

For the standard Johnson's Algorithm, this is typically a count of the number of appearances an attribute makes within clauses; attributes that appear more frequently are considered to be more significant. The attribute with the highest heuristic value is added to the reduct candidate, and all clauses in the discernibility function containing this attribute are removed. As soon as all clauses are removed, the algorithm terminates and returns to reduct R.

Algorithm 4.2.3: LEM 2 Algorithm

In general, LEM2 compute local covering and then converts it into a rule set. It is based on an idea of an attribute-value pair block [4]. Let U be the set of all cases of the data set. For an attribute-value pair $(a, v) = t$, a block of t, denoted by $[t]$, is a set of all cases from U , such that for attribute a has value v . Let B be a nonempty lower or upper approximation of a concept represented by a decision-value pair (d, w). Set B depends on a set T of attribute-value pair $t = (a, v)$ if and only if:

$$\phi \neq [T] = \bigcap_{t\in T} [t] \subseteq B$$

Set T is a minimal complex of B if and only if B depends on T and no proper subset T' of T exists such that B depends on T' . Let Φ be a nonempty collection of nonempty sets of attribute-value pairs. Then Φ is a local covering of B if and only if the following conditions are satisfied:

1. Each member T of Φ is a minimal complex of B .
2. $\bigcup_{T\in\Phi} [T] = B$
3. Φ is minimal, that is, Φ has the smallest possible number of members.

Input: a set B
Output: a single local covering Φ of set B
Begin
$G = B$
$\Phi = \phi$
While $G \neq \phi$
Begin
$T = \phi$
$T(G) = \{t \mid \{t\} \cap G \neq \phi\}$
While $T = \phi$ or not $([T] \subseteq B)$
Begin
Select a pair $t \in T(G)$ with the highest attribute priority, if tie occurs,
Select a pair $t \in T(G)$ such that $|[t] \cap G|$ is maximum; if another tie occurs,

Select a pair $t \in T(G)$ with the smallest cardinality of $[t]$; if further tie occurs, select first pair,

$T = T \cup \{t\}$
$G = [t] \cap G$
$T(G) = \{t \mid \{t\} \cap G \neq \phi\}$
$T(G) = T(G) - T$
End {while}

For each $t \in T$ do
If $[T - \{t\}] \subseteq B$ then
$T = T - \{t\}$
$\Phi = \Phi \cup [T]$
$G = B - \bigcup_{T \in \Phi} [T]$

End {while}

For each $T \in \Phi$ do
If $\bigcup_{s \in \Phi - \{T\}} [s] = B$ then $\Phi = \Phi - \{T\}$
End {Procedure}

V. RESULT AND DISCUSSION

To illustrate the above mentioned concepts of RST, dataset of eight objects i.e., patients $U = \{X_1, X_2, X_3, X_4, X_5, X_6, X_7, X_8\}$ with four conditional attributes A_1 (Chest Pain type), A_2 (Serum Cholesterol), A_3 (Heart rate), A_4 (Fasting blood sugar) and one decision attribute Q which represents whether heart disease is present or absent with domain values 1 & 2 respectively. Columns of the table are labeled by symptoms (attributes) and row by patients (objects), where each cell of the table are attribute values.

Table 1: Heart Disease Dataset

U	A_1	A_2	A_3	A_4	Q
X_1	1	0	2	0	1
X_2	1	0	2	1	2
X_3	1	2	0	0	1
X_4	1	2	2	1	1
X_5	2	1	0	0	2
X_6	2	1	1	0	2
X_7	2	1	0	0	1
X_8	1	2	1	0	2

Equivalence classes of conditional attribute P i.e., the partition determined by set of all attributes P, denoted by U/P or $I(P)$. Hence,

$$U/P = \{\{X_1\}, \{X_2\}, \{X_3\}, \{X_4\}, \{X_5, X_7\}, \{X_6\}, \{X_8\}\}$$

Similarly Equivalence classes of decision attribute Q are $U/Q = \{\{X_1, X_3, X_4, X_7\}, \{X_2, X_5, X_6, X_8\}\}$.

Lower approximation for decision $(Q,1)$ is $R_*(Q) = \{X_1, X_3, X_4\}$

Upper approximation for decision $(Q,1)$ is $R^*(Q) = \{X_1, X_3, X_4, X_5, X_7\}$

Boundary region $BNR(Q) = \{X_5, X_7\}$

Accuracy of approximation is 0.6.

Table 2: Discernibility Matrix

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8
X_1	-	-	-	-	-	-	-	-
X_2	A_4	-	-	-	-	-	-	-
X_3		A_2, A_3, A_4	-	-	-	-	-	-
X_4		A_2		-	-	-	-	-
X_5	A_1, A_2, A_3		A_1, A_2	A_1, A_2, A_3, A_4	-	-	-	-
X_6	A_1, A_2, A_3		A_1, A_2, A_3	A_1, A_2, A_3, A_4		-	-	-
X_7		A_1, A_2, A_3, A_4				A_3	-	-
X_8	A_2, A_3		A_3	A_3, A_4			A_1, A_2, A_3	-

The discernibility matrix is used to find reduct which leads to the same partition of the data as the whole set of attributes P. To do this, one has to construct the discernibility function $f(P)$, a Boolean function. The minimal subset of attributes (reduct) of dataset is obtained by converting the Boolean expression from conjunctive normal form to disjunctive normal form. Hence the reduct generated from the above discernibility matrix is $\{A_2, A_3, A_4\}$.

According to QRA, the dependency of each attribute is calculated, and the best candidate is chosen. In the give dataset attribute A_3 generates the highest dependency degree and hence the subsets $\{A_1, A_3\}$, $\{A_2, A_3\}$ and $\{A_3, A_4\}$ are evaluated. Further with these sets, the process continues until the dependency of any of the subsets become equal to the consistency of the conditional attribute set P. Finally, the algorithm terminates after evaluation of the reduct $\{A_2, A_3, A_4\}$.

According to Johnson's algorithm, the feature with the highest frequency is added to the reduct. In our discernibility matrix, the attribute A_3 appears with the highest frequency i.e., 12 times therefore it is the first attribute added to the reduct. Now, remove all of the cells containing A_3 from discernibility matrix and search for the next highest frequency feature and add it to the previously obtained reduct. This process continues till the discernibility matrix becomes empty. After the complete execution of the Johnson's Algorithm reduct obtained is $\{A_2, A_3, A_4\}$.

For application of LEM2 algorithm compute all attribute-value pair blocks. Let the input set be $B=\{X_1, X_3, X_4, X_7\}$. The best attribute-value pair t i.e. $(A_1,1)$ is considered according to the condition of algorithm. Now, look for the next t, There are three attribute-value pairs with same cardinality of the intersection of $[t]$ & G. Again select the first pair among them i.e. $(A_2,2)$ and it can be observed that $[\mathrm{T}]\subseteq \mathrm{B}$, so $\{(A_1,1),(A_2,2)\}$ is the first element in T of Φ. The new set will be $G=B-[T]$ that is $\{X_1, X_7\}$ same process continue until G becomes empty. Hence, the LEM2 algorithm induces the rule set for $(Q,1)$ as $(A_1,1)\wedge(A_2,2)\rightarrow(Q,1)$ and so on. For finding the rule set for $(Q,2)$ input set will be $\{X_2, X_5, X_6, X_8\}$ and follow the same procedure as above.

The decision rules generate from the given decision table are as follows:

$$(A_3,1)\rightarrow(Q,2)$$
$$(A_2,2)\wedge(A_3,0)\rightarrow(Q,1)$$
$$(A_2,2)\wedge(A_3,2)\rightarrow(Q,1)$$
$$(A_3,2)\wedge(A_4,0)\rightarrow(Q,1)$$
$$(A_2,0)\wedge(A_4,1)\rightarrow(Q,2)$$
$$(A_2,1)\wedge(A_3,0)\wedge(A_4,0)\rightarrow(Q,1)$$
$$(A_2,1)\wedge(A_3,0)\wedge(A_4,0)\rightarrow(Q,2)$$

The strength, certainty and coverage factor for decision rule are show in Table 3.

Table 3: Strength, certainty and coverage factor

Rules	Support	Strength	Certainty	Coverage
R1	2	0.25	1	0.5
R2	1	0.125	1	0.25
R3	1	0.125	1	0.25
R4	1	0.125	1	0.25
R5	1	0.125	1	0.25
R6	1	0.125	0.5	0.25
R7	1	0.125	0.5	0.25

VI. CONCLUSION

In this paper rule generation techniques of RST are applied in the field of medical diagnosis for identification of heart disease. Many factors (attributes) affect human body in different ways so it becomes difficult to diagnose the diesease easily on the basis of one or two factors. The diagnosis requires different set of features with high risk. The paper serves a decision system for heart disease identification. Reducts, the minimal subset of attributes, are generated by various RS techniques like Discernibilty matrix, Quick Reduct algorithm and Johnson's Heuristic algorithm. On the basis of the reducts, decision rules are generated and the optimality of rules obtained are shown with the help of strength, certainty factor and coverage factor. Lastly, LEM2 algorithm, an attribute-value pair block method, is used on heart disease dataset to demonstrate how rules can be generated without any calculation of reduct. Out of the alorithms considered in this study, QRA is found to be less time consuming.

References

[1]. H. L. Chen, B. Yang, J. Liu, D. Y. Liu, "*A Support Vector Machine Classifier with Rough Set based Feature Selection for Breast Cancer Diagnosis*", Expert Systems with Applications, Vol.38, Issue.7, pp.9014-9022, 2011.

[2]. A. H. El-Baz, "*Hybrid Intelligent System-based Rough Set and Ensemble Classifier for Breast Cancer Diagnosis*", Neural Computing Applications, Vol.26, Issue.2, pp.437-446, 2015.

[3]. E. H. Francis, S. Lixiang, "*Fault Diagnosis based Rough Set Theory*", Engineering Applications of Artificial Intelligence, Vol.16, Issue.1, pp.39-43, 2003.
[4]. J. W. Grzymala-Busse, "*Rule Induction*", The Data Mining and Knowledge Handbook, Springer –Heidelberg, US, pp. 249-265, 2010.
[5]. L. Huang, L. Dai, C. Zhou, *"Prognosis System for Lung Cancer Based on Rough Set Theory*", Third International Conference on Information and Computing, China, pp.7-10, 2010.
[6]. A. Ohrn, T. Rowland, *"Rough Sets: A Knowledge Discovery Technique for Multifactorial Medical Outcomes*", American Journal of Physical Medicine and Rehabilitation, Vol.79, Issue.1, pp.100-108, 2000.
[7]. Z. Pawlak, *"Rough Sets*", International Journal of Computer and Information Sciences, Vol.11, Issue.5, pp.341-356, 1982.
[8]. L. V. Santana-Quintero, A. G. Hernandez-Diaz, , J. Molina, C. A. Coello, R. Caballero, "*DEMORS: A Hybrid Multi- Objective Optimization Algorithm using Differential Evolution and Rough Set Theory for Constrained Problems*", Computers and Operations Research, Vol.37, Issue.3, pp.470-480, 2010.
[9]. N. A. Setiawan, P. A. Venkatachalam, A. F. Hani, "*Missing Data Estimation on Heart Disease Using Artificial Neural Network and Rough Set Theory*", International Conference on Intelligent and Advanced Systems, India, pp.129-133, 2007.
[10]. Q. Shen, R. Jensen, *"Rough Sets, their Extensions and Applications*", International Journal of Automation and Computing, Vol.4, Issue.1, pp.100-106, 2007.
[11]. K. Thangeval, M. Karnan, A. Pethalakshmi, "*Performance Analysis of Rough Reduct Algorithms in Mammogram*", International Journal on Graphics Vision and Image Processing, UK, Vol.8, Issue.4, pp.13-21, 2005.
[12]. B. Walczak, D. L. Massart, "*Rough Set Theory*", Chemometrics and Intelligent Laboratory Systems, Vol.47, Issue.1, pp.1-16, 1999.
[13]. I. T. Yanto, P. Vitasari, T. Herawan, M. M. Deris, "*Applying Variable Precision Rough Set Model for Clustering Student Suffering Study's Anxiety*", Expert Systems with Applications,Vol.39, Issue.1, pp.452-459, 2012.
[14]. L. A. Zadeh, *Fuzzy Sets*, Information and Control , Vol.8, Issue.3, pp.338-353, 1965.

Authors Profile

Dr. Pragati Jain has done her Ph.D. in Mathematics from Choudhary Charan Singh University Meerut in 2007. She is currently working as Associate Professor and Head- Department of Science in St. Paul Institute of Professional Studies, Indore. She has 16 years of teaching experience and 9 years of research experience. Her research area is computational mathematics, big data analysis, rough set and soft set theory. Ms. Deepensha Vaishnav and Ms. Kamini Agrawal are research scholars at IET, DAVV, Indore and both have teaching experience of 5 years.