International Journal of Computer Information Systems and Industrial Management Applications.
ISSN 2150-7988 Volume 5 (2013) pp. 462-471

Genetic Algorithm And Fuzzy-Rough Based Dimensionality Reduction Applied On Real Valued Dataset

Moumita Saha[1], Jaya Sil[1] and Nandita Sengupta[2]

[1]Computer Science and Engineering Department
Bengal Engineering and Science University
Shibpur, India
moumita.saha2012@gmail.com, js@cs.becs.ac.in

[2]Information Technology Program
University College of Bahrain
P.O. Box 55040, Manama, Kingdom of Bahrain
ngupta@ucb.edu.bh

Abstract: **Real-world datasets are often vague and redundant, creating problem to take decision accurately. Very recently, Rough-set theory has been used successfully for dimensionality reduction but is applicable only on discrete dataset. Discretisation of data leads to information loss and may add inconsistency in the datasets. The paper aims at applying fuzzy-rough concept to overcome the above limitations. However, handling of non discretized values increases computational complexity of the system. Therefore, to build an efficient classifier Genetic Algorithm (GA) has been applied to obtain optimal subset of attributes, sufficient to classify the objects. The proposed algorithm reduces dimensionality to a great extent without degrading the accuracy of classification and avoid of being trapped at local minima. Results are compared with the existing algorithms, demonstrate compatible outcome.**
Keywords: Fuzzy set, rough set, fuzzy-rough set, genetic algorithm, dimensionality reduction.

I. Introduction

$Dimensionality\ Reduction$ compromises of selection of most relevant features that are predictive of the class-outcome and rejection of irrelevant features with minimal information loss [27], [37], [38]. The computational efficiency of a classification problem depends on the selection of number of attributes, sufficient to build the classifier.
Dimensionality reduction using Rough Set Theory (RST) is widely applied in different domain, and produces satisfactory results. However, most often the values of attributes are continuous but RST is applicable only on discretized data. In addition, after discretisation it is not possible to judge the extent to which the attribute value belongs to the corresponding discrete levels. This is the source of information loss, and it affects the classification accuracy negatively. Therefore, it is essential to work with real-valued data for combating the information loss and this can be achieved by combining Fuzzy and Rough set theory.

Different rough set, fuzzy set and fuzzy-rough based approaches are already proposed to handle real life datasets. In paper [1], it has been mentioned that $Sammon's$ nonlinear projection methods is effective for smaller datasets but for large datasets, this method lacks predictability and becomes ineffective. This paper [1] proposes a method which combines $Sammon's$ nonlinear method with Fuzzy logic. The proposed method has been implemented and satisfactory result has obtained with compared to original $Sammon's$ method. In paper [2], dynamic dimension reduction is done for Fuzzy classifier. New dynamic data is considered as an input for which weight of each attribute varies from 0 to 1 is computed. If the value of this weight becomes near to 0, that attribute is considered as redundant attribute for Fuzzy classification. In paper [3], it is mentioned that gene selection is very difficult task because of its high dimensionality, redundant information and noise. This paper [3] introduces Fuzzy rules for dimension reduction of gene expression in two steps. In the first step, gene expression levels of a given dataset are transformed into fuzzy values using fuzzy inference rules. In the second step, applying similarity relation to these fuzzy values, fuzzy equivalence groups are formed, each group contains similar type of genes. Dimension reduction is achieved by considering each group of similar genes as a single representative value. In paper [4], the proposed method combines $Random\ Projections\ (RP)$ and $fuzzy\ k-means\ clustering\ (FKM)$ for dimension reduction. The proposed $RP-FKM$ is computationally less complex than $Single\ Value\ Decomposition\ (SVD)$ and $RP-SVD$. On the image data, the proposed $RP-FKM$ has produced less amount of distortion when compared with RP. The proposed $RP-FKM$ provides better text retrieval results when compared with conventional RP and performs similar to $RP-SVD$. By experimental results, proposed method proves that $RP - FKM$ produces better result for dimension reduction in image and text field. $Nick\ J.\ Pizzi\ et.\ al.$ mentioned in their paper [5] that due to high number features, pattern classification is becoming a challenging task and they presented a technique using an adaptive network of fuzzy logic connectives to combine class boundaries generated by sets of discriminant functions. They empirically evaluated the effectiveness of this classification technique by comparing it against two conventional benchmark approaches, both of which use feature averaging as a preprocessing phase. In paper [6], dimension reduction is done using $rough\ set\ theory$ (RST). Different algorithms are designed for finding the reducts from a set of attributes considering uncertainty, missing attribute values, inconsistencies. In paper [7], reducts are applied for two applications using neural network, one for diagnosing plant diseases and the other for intrusion detection. In both of these cases, after reduction of dimension using RST, performance of classifier increased considerably. In paper [8], the utility of $Rough\ Set\ Attribute\ Reduction$ $(RSAR)$ in supervised and unsupervised learning is investigated. A Fuzzy-Rough Estimator, combination of $RSAR$ and a $fuzzy\ Rule\ Induction\ Algorithm$ (RIA), is used in supervised learning system with dimensionality reduction capabilities. For dimension reduction of unsupervised learning system, $RSAR$ combined with $Multivariate\ Adaptive\ Regression\ Splines$ $(MARS)$ is applied. In both these cases $RSAR$ produces higher efficiency by providing efficient dimension reduction of the learning system. $A.\ Chouchoulas\ et.\ al.$ mentioned in their paper [9] that RST can be applied for dimension reduction in information filtering and information retrieval tasks by providing a measure of information content in the data set with respect to a given class. This paper compares the applicability of RST based dimension reduction technique with other text categorization techniques. $N.\ Zhong\ ei.\ al.$ in paper [10], proposed an algorithm for feature selection which uses RST with greedy heuristics. Selecting features is similar to the filter approach, but the evaluation criterion is related to the performance of induction. They have selected the features without damaging the performance of induction. In paper [11], it is mentioned that if dimension reduction is done only by using RST, there can be a major loss of information. To avoid such loss of information, fuzzy- rough is used for dimension reduction. This method is applied in categorization of email messages. Paper [12] proposes a system which combines a dimensionality reduction module (using principal component analysis), a feature extraction module (using independent component analysis), and a feature subset selection module (using rough set model). To reduce the effect of data inconsistency Rough set model is used and a fuzzy classifier is integrated into the system to label sub-images for classifying regions into normal and abnormal. The experimental results shows the accuracy of the system is 84.03%. Paper [13] implements a novel approach for dimension reduction by rough set theory followed by establishing a fuzzy discernibility matrix by using distance preserving strategy for attribute reduction. Experimental results show that classifiers produce better accuracy with reduced attribute compare to that of all attributes. This method can be applied in learning algorithms, like, PCA, SVM etc. This dimension reduction technique can be applied in many real life applications, like web categorization, image processing etc. $Qinghua\ Hu\ et.\ al.$ proposed in their paper [14], an information measure for computing discernibility power which is important for rough set or fuzzy rough set model.Significance of fuzzy attribute is considered based on this information measure.The independence of hybrid attribute subset, reduct, relative reduct is redefined. On the basis of this proposed method of independence measure, two greedy reduction algorithms have been framed for supervised and unsupervised data dimensionality reduction. It has been observed that after dimension reduction using proposed fuzzy rough approach produces better performance compared to classical rough set approach. Paper [15] reviews those techniques considering dimension reduction using fuzzy rough methodology and the same preserve the fundamental semantics of data. $Richard\ Jensen\ et.\ al.$ implemented fuzzy rough attribute reduction technique in the domain of website classification in paper [16]. They got promising classification accuracy with respect to the classification where RST based dimension reduction is applied. Paper [17] presents a novel approach, which integrates use of fuzzy and rough set theories, to greatly reduce data redundancy. It has been observed though experimental results that fuzzy-rough reduction is more powerful than the conventional rough set-based approach. In paper [18], it is mentioned that using RST dimension reduction can be done effectively. But as RST works only on discretized data, discretization is required for all continuous attributes. Due to discretization, loss of information occurs of the information system. To avoid that loss of information, $Richard\ Jensen\ et.\ al.$ implemented fuzzy rough set theory for dimension reduction without affecting the performance. In paper [19], a novel approach of fuzzy rough theory is applied for dimension reduction. This approach is based on the formulation of fuzzy-rough discernibility matrices, which is realised algorithmically by a modified version of a traditional satisfiability approach. This produces an efficient and optimal approach to data reduction and in terms of both time and classification accuracy this approach works well on a number of machine learning benchmarks. $Richard\ Jensen\ et.\ al.$ mentioned in their paper [20], RST is computationally efficient technique for addressing problems such as hidden pattern discovery from data, feature selection and decision rule generation. $Fuzzy - rough\ set\ theory$ improves upon this by enabling uncertainty and vagueness to be modelled more effectively. This paper proposes three novel methods for instance selection based on fuzzy-rough sets. The initial experimentation demonstrates that the methods can significantly reduce the number of instances which can maintain high classification accuracies. $Neil\ MacParthalain\ et.\ al.$ mentioned in their paper [21], that supervised learning needs those attributes which determine the class labels. Therefore, after dimension reduction, those features have to be there in the information system. In case of unsupervised learning, no such class label is required for determination of class. It becomes important to find out those attributes which can determine the class, so while reducing dimension, the attributes which create noise or which cannot contribute in classification can be reduced. In this paper, new approach of fuzzy rough has been proposed for unsupervised feature selection. These approaches require no thresholding or domain information, and produce significant reduction in dimensionality.

However, real value data handling is difficult and computational complexity of the system increases with the number of attributes in the dataset. Therefore, finding optimal set of attributes is the solution to classify the objects efficiently.
In order to handle vaugeness in data and obtain optimal feature set, the paper proposes an algorithm using fuzzy-rough set concept and genetic algorithm. $Fuzzy - Rough\ Quick\ Reduct(FRQR)$ [31] method, an efficient method of attribute reduction, overcoming the need of discretised attribute values but the search of most informative attributes may terminate at local optimum. Therefore, in the paper $Fuzzy - Rough\ Set$ concept is merged with $Genetic\ Algorithm$ to attain global optimum in the search space. Classification accuracy of the proposed algorithm demonstrate compatible result with the other existing methods.

Section II-A, II-B explains basic concepts of fuzzy set and rough sets, Section II-C explains fuzzy-rough concepts, Section II-D discusses the steps of genetic algorithm, Section III-A shows the preprocessing of data, Section III-B explains the proposed method and finally Section IV demonstrate the results.

II. THEORETICAL FOUNDATIONS

A. Fuzzy Sets

Fuzzy sets were introduced by Lotfi A. Zadeh in 1965 as an extension of the classical notion of set.
A fuzzy set A in the universe of discourse U can be defined as a set of ordered pairs,

$$A = \{(x, \mu_A(x)) | x \in U\} \tag{1}$$

Where $\mu_A(x)$ is the degree of membership [0,1] of x in A. The membership function $\mu_A(.)$ maps U to the membership space M, that is $\mu_A : U \to M$.

B. Rough Sets

Another important parallel concept along with $Fuzzy\ Sets$ is $Rough\ Sets$. Rough Sets theory was introduced by Z. Pawlak (1982) as a mathematical approach to handle vagueness.Rough Set is a formal approximation of a crisp set (i.e., conventional set) in terms of a pair of sets which give the lower and the upper approximation of the original set [24].

1) Information and Desision System

An information system is a data table represented by $S = (U, R)$. S consisting of data universe U and set of attributes, R known as $condition\ attribute$. The attribute a (in R) characterizes each of the data elements x (in U) [26]. $Decision\ System$ is the information system represented by $S = (U, R \cup \{D\})$ where $D \notin R$ known as $decision\ attribute.$

2) Indiscernibility Relation

$Indiscernibility\ Relation\ IND\,(P)$ [26] is an equivalence relation defined below.

$$IND(P) = \{(e, f) \in U \times U, \quad \forall a \in R, \quad a(e) = a(f)\} \tag{2}$$

where e and f are indiscernible objects.

3) Lower Approximation

In U/R, the objects which are positively classified, called $lower\ approximation$ [29] of the set X and written as:

$$\underline{R}(X) = \{ x \quad \in U, \quad [x]_R \subseteq X \} \tag{3}$$

4) Upper Approximation

The $R - upper\ approximation$ [29] is the union of all equivalence classes in $[\ X\ ]_R$ which have non-empty intersection with the target set X. Mathematically, it is written as :

$$\overline{R}(X) = \{ x \in U, [x]_R \cap X \neq \emptyset \} \tag{4}$$

5) Positive Region

The $positive\ region$ [29] of a target set X is defined below .

$$POS_R\,(Q) = \cup_{X \in U/Q} \underline{R}(X) \tag{5}$$

where Q is the decision attribute.

6) Dependency

An important issue in data analysis is discovering dependencies between the attributes. Intuitively, a set of attributes D depends totally on a set of condition attributes R, denoted by $C \implies D$. D depends on R to a degree $k\,(0 \leq k \leq 1)$ as given below.
$k = \gamma(\ R,\ D\ ) = |\,(\ POS_R D\ )\,| / |U|$
The higher the dependency the more significant the attribute is.

C. Fuzzy-Rough Sets :

$Fuzzy\ set\ theory$ and $Rough\ set\ theory$ are useful computational intelligence tools in many real-world applications for dealing with vague information and to take important decision in uncertain domain [23]. Both of them works in different aspects in dealing with huge data and have their own merits and demerits. The selection of $appropriate\ membership\ function$ is the main bottleneck of fuzzy set. On the other hand, $Rough\ set\ theory$ is useful for decision making in situation where $indisernibility$ is present. As opposed to fuzzy sets, rough sets do not require experienced knowledge engineers to provide additional information about the membership functions for being processed. But the main bottleneck of applying rough set theory is that it only deals with discrete data values. The merits of $rough\ sets$ and $fuzzy\ sets$ are integrated to develop a much more powerful and efficient tool known as

$Fuzzy-Rough\ Set$, emerged as a new research area. $Fuzzy-rough\ set$ is a derivation of rough set theory in which the concept of crisp equivalence class is extended using fuzzy set theory to form fuzzy equivalence classes [28], [33]. Thus, every objects have degree of membership values to lower and upper approximation fuzzy sets. In fuzzy-rough sets the $equivalence\ class\ is\ fuzzy$. In addition, $fuzziness$ is introduced in the output classes too.
Let, the equivalence classes are in the form of fuzzy clusters $F_1, F_2...F_H$, which are generated by the fuzzy partitioning of the input set X into H number of clusters [22]. Each fuzzy cluster represents an equivalence class consisting of patterns of different output classes. The definite and possible number of output classes are identified using lower and upper approximations of the fuzzy equivalence classes.

1) Lower and Upper Approximation

The description of a fuzzy set X (output class) by means of the fuzzy partitions under the form of lower and upper approximations $\underline{R}X$ and $\overline{R}X$ is as follows [27],

$$\mu_{\underline{R}X}(F_j) = inf_x\{max(1-\mu_{F_j}(x), \mu_X(x))\}, \forall j \quad (6)$$

$$\mu_{\overline{R}X}(F_j) = sup_x\{min(\mu_{F_j}(x), \mu_X(x))\}, \forall j \quad (7)$$

R is an attribute subset, $\mu_{F_j}(x)$ and $\mu_X(x)$ are the fuzzy membership values of the object x in the fuzzy equivalence class F_j and output class X respectively.
Fuzzy-Rough lower and upper approximation can be defined more explicitly as:

$$\mu_{\underline{R}X}(x) = sup_{F\in U/R} min \\ (\mu_F(x), inf_{y\in U} max\{1-\mu_F(y), \mu_X(y)\}) \quad (8)$$

$$\mu_{\overline{R}X}(x) = sup_{F\in U/R} min \\ (\mu_F(x), sup_{y\in U} min\{\mu_F(y), \mu_X(y)\}) \quad (9)$$

The tuple $< \underline{R}X, \overline{R}X >$ is defined as $fuzzy-roughset$.
The lower and upper approximation of Fuzzy-Rough Set are fuzzy unlike the crisp value of Rough-Set, represented in figure 1.

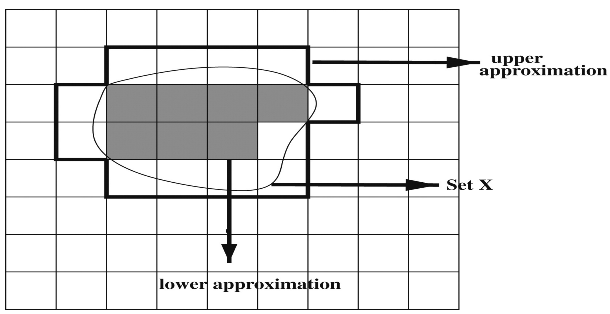

Figure. 1: Lower and Upper Approximation of Fuzzy-Rough Set

2) Positive Region

As an extension to crisp positive region in traditional rough set theory, the membership of an object $x \in U$, belonging to fuzzy positive region [27] is defined as:

$$\mu_{POS_R(Q)}(x) = sup_{X\in U/Q} \mu_{\underline{R}X}(x) . \quad (10)$$

3) Fuzzy-Rough Dependency

Fuzzy-Rough Dependency [27] can be defined with the aid of positive region as :
$\gamma'_R(Q) = |\mu_{POS_R(Q)}(x)| / |U|$
or,

$$\gamma'_R(Q) = \Sigma_{x\in U} \mu_{POS_R(Q)}(x) / |U| \quad (11)$$

D. Genetic Algorithm

Genetic algorithm (GA) is a search heuristic, used to generate solutions to optimization problems following the techniques inspired by natural evolution, such as inheritance, mutation, selection, and crossover. In the genetic algorithm [32],

- A population of strings (called chromosomes), which encode candidate solution to an optimization problem is taken.
- A proper fitness function is constructed, and fitness of the current population is evaluated.
- Two most fittest chromosomes are chosen as the parents and (a)crossing over between them or (b)mutation of a parent is performed to produce new children and a new population.
- Again the fitness function for the new population is estimated.
- The process recurs as long as the fitness function keeps on improving or termination condition is attained.

III. Dimensionality Reduction

Before presenting the proposed method, the decision table for three different datasets are prepared.

A. Data Preparation

Three different kinds of dataset are considered here: (1)Hypothyroidism dataset, (2)Pulmonary-Embolism dataset and (3)Wine dataset [35]. The fitness function of genetic algorithm used in the proposed method is $fuzzy-rough\ dependency\ factor$. To evaluate this factor we require two important parameters.

1. The membership values of each data objects in different clusters.
2. The membership values of each data objects in each of the classes.

The first parameter is calculated by partitioning the dataset using fuzzy-c-means clustering algorithm [36] which provides the degree of membership values of each data object belonging to c number of clusters. The number of cluster is kept same as the number of classes of the corresponding dataset.
The assignment of membership value to each object belonging to different class labels (i.e. the second parameter) are obtained by designing a $fuzzy\ inference\ system(FIS)$, as described below step by step.

- Input data is fuzzified based on the range of minimum and maximum value of each attribute, determines the spread of membership value of respective attribute.

- Objects are grouped based on the class labels of decision attribute.

- For each conditional attribute and for each respective class label $linguistic\ variables$ are assigned. The value of the linguistic variable is set based on the range of values to which an abtibute value is spread for a particular class label.

- For each of the $linguistic\ variables$ assigned for every different attributes, the membership function is identified by studying the data pattern in that region and corresponding membership curves are drawn.

- With the aid of the decision system, the membership curves for linguistic variable of each attribute and the rule-set has been designed. The rule-set is designed by randomly choosing data elements from the training set and evaluating their membership values using the membership curves built previously. Using the attribute values we judge in which linguistic label corresponding attribute value belongs and finally the rule is framed with the linguistic label along with the decision attribute as the class label of the rule.

- After generating the total rule-base, a $Fuzzy\ Inference\ System\ (FIS)$ has been build using $Mamdani\ model$.

- Finally the built FIS is utilized to produce the membership values of each objects belonging in different classes (second parameter) for evaluation of Fuzzy-rough dependency factor.

A dataset with two attributes and three output class-label is considered in table 1 to illustrate the above procedure.

Table 1: Decision Table

Objects	Attr1	Attr2	Class
O1	2	10	1
O2	7	5	2
O3	5	15	1
O4	6	8	2
O5	12	16	2
O6	8	20	1
O7	10	25	3
O8	15	22	3
O9	4	17	1
O10	13	21	3

step 1: The minimum and maximum ranges of the attributes for spread of corresponding membership curve are:
Attr1 = 2 - 15, Attr2 = 5 - 25.
step 2: Now reconstruct the decision table by grouping the class-label, shown in table 2.

Table 2: Output Of Step 2

Objects	Attr1	Attr2	Class
O1	2	10	1
O3	5	15	1
O6	8	20	1
O9	4	17	1
O2	7	5	2
O4	6	8	2
O5	12	16	2
O7	10	25	3
O8	15	22	3
O10	13	21	3

step 3: For each attributes, Fuzzy Values are assigned utilizing the range of attribute values corresponding to individual class-labels.
Attr1 : LOW (2 - 8), MED (6 - 12), HIGH (10 - 15).
Attr2 : VERY LITTLE (5 - 16), LITTLE (10 - 20), MORE(21 - 25).

step 4: The membership curves for the attributes are plotted in figure 2.

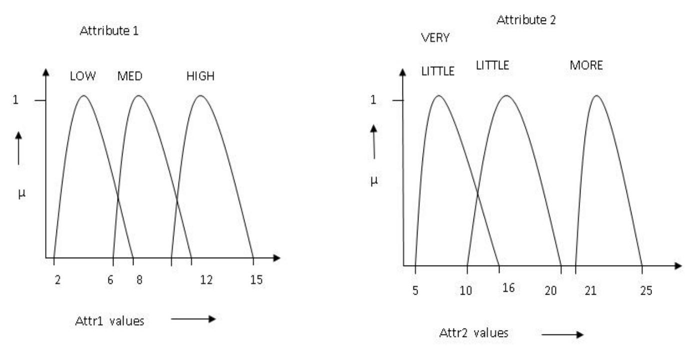

Figure. 2: Membership Curve

step 5: The rule-set corresponding to the decision system is :

1. IF Atr1 is LOW and Atr2 is VERY LITTLE THEN class is 1.
2. IF Atr1 is LOW and Atr2 is LITTLE THEN class is 1.
3. IF Atr1 is LOW and Atr2 is MORE THEN class is 1.
4. IF Atr1 is MED and Atr2 is LITTLE THEN class is 1.
5. IF Atr1 is MED and Atr2 is MORE THEN class is 1.
6. IF Atr1 is LOW and Atr2 is VERY LITTLE THEN class is 2.
7. IF Atr1 is MED and Atr2 is VERY LITTLE THEN class is 2.
8. IF Atr1 is MED and Atr2 is LITTLE THEN class is 2.
9. IF Atr1 is HIGH and Atr2 is VERY LITTLE THEN class is 2.
10. IF Atr1 is HIGH and Atr2 is LITTLE THEN class is 2.
11. IF Atr1 is MED and Atr2 is MORE THEN class is 3.

12. IF Atr1 is HIGH and Atr2 is MORE THEN class is 3.

Finally, Mamdani model has been applied for evaluating degree of membership value of each object in different classes.

B. DIM-RED-GA Algorithm

In the proposed method, the dependency of the decision attribute on different set of conditional attribute is calculated and attributes with highest dependency value is selected as optimum reduct by applying genetic algorithm.
In order to apply the DIM-RED-GA, population size is considered as the total dataset of the decision table. The $fuzzy-rough\ dependency\ factor$ is considered as the fitness function. The chromosomes are built up by taking the attribute values of the objects of the decision table. Two parents are chosen and $crossing\ over$ between them is performed by randomly choosing a crossing over point with a probability of 0.10. $Mutation$ of a data object is performed with a probability of 0.02. In each generation, combination of different attributes are formed and from them few are selected based on the fitness value. Termination condition is kept as combination of two conditions: (1) Number of generation is greater than $MAX-NUMBER-OF-GENERATION$ or, (2) Number of times same Dependency factor appears greater than $MAX-NUMBER-OF-ITERATION$.
Following functions are applied on data objects represented by x.

function crossing-over (x, no-of-objects, no-of-attributes)
(1) Choose a cross-over point randomly within the attribute set.

cp ⟵ rand() % no-of-attributes.

(2) Choose two parents within the data objects of the decision system.

P1 ⟵ rand() % no-of-objects.
P2 ⟵ rand() % no-of-objects.

(3) Develop new generation children by cross-over.

for (i=0 ; i < cp ; i++)
temp [i] ⟵ x [P1] [i] .
x [P1] [i] ⟵ x [P2] [i] .
x [P2] [i] ⟵ temp [i] .

function mutation-flip (x, no-of-objects, no-of-attributes)
(1) Choose the attribute column to be mutated.

mut-col ⟵ rand() % no-of-attributes.

(2) Calculate the value to be added to the mutated attribute column.

(a) for (i=0 ; i < no-of-objects ; i++)
mut-val += x [i] [mut-col].

(b) Find the maximum of the mutated attribute values in max.
(c) Final mutation value is calculated as:

mut-val /= (no-of-objects * max).

(3) Update the mutated attribute column .

for (i=0 ; i < no-of-objects ; i++)
x [i] [mut-col] += mut-val.

function variation (column, no-of-attributes)
(1) Compute the attributes absent in column set and store them in left set.
(2) Generate a random attribute to be replaced from column set and a random attribute from left set that will be utilized for replacement.

replace ⟵ rand() % no-of-attributes in column set.
select ⟵ rand() % no-of-attributes in left set.

(3) Replace the corresponding attribute from column set by selected attribute of the left set.

for (j=0 ; j < no-of-attributes-of-column-set ; j++)
if (column [j] == column [replace])
column [j] = column [select].

DIM-RED-GA (x)
BEGIN
(1) Take input the information system along with membership values of each data objects in every classes.
(2) Initialize γ_{prev} = 0.0, γ_{best} = 0.0, flag = 0, count-of-generation = 0.
(3) *do repeat until (flag == 1) :*

(a) Count-of-generation += 1.
(b) Select randomly number of attributes to be taken.

number-attr ⟵ rand() % NO-OF-ATTRIBUTES
number-attr += 1.

(c) Generate the combination set containing all combinations of $number-attr$ number of attributes.
(d) Select a combination from the combination set.

comb-num ⟵ rand() % total-number-of-elements in combination set.

(e) Take the reduced information system $number-attr$ number of attributes for $comb-num^{th}$ combination of the combination set.
(f) Find out the crossing over probability.

if (cross-over probability is 10 %)
Call function *cross-over (x , no-of-objects , no-of-attributes)*

(g) Modify the information system (x) as required after crossing over.
(h) Find out the mutation probability.

if (mutation probability is 2 %)
Call function *mutation (column , no-of-attributes)*

(i) Modify the information system (x) as required after mutation.
(j) Call function *variation (x , no-of-objects , no-of-attributes)* for generating different combination of attributes.
(k) Evaluate the membership values of each objects in every clusters by $fuzzy-c-means\ clustering\ algorithm.$
(l) Calculate the $fuzzy-rough\ lower-approximation$ of each data objects in each class.

$$\mu_{\underline{R}X}(x) = sup_{F\in U/R} min$$
$$(\mu_F(x), inf_{y\in U} max\{1-\mu_F(y), \mu_X(y)\})$$

(m) Then evaluate $fuzzy-rough\ positive\ region$ of each data objects.

$$\mu_{POS_R(Q)}(x) = sup_{X\in U/Q}\mu_{\underline{R}X}\ (x)\ .$$

(n) Finally, evaluate the $fuzzy-rough\ dependency$ factor for the information system with specified number of attributes.

$$\gamma^{'}_{R}(Q) = \Sigma_{x\in U}\mu_{POS_R(Q)}(x)/|U|$$

(o) check if ($\gamma^{'} > \gamma_{prev}$)

Update : reduct ⟵ present set of attributes.
$\gamma_{best} = \gamma^{'}$.

(p) check if ($\gamma_{best} == \gamma_{prev}$)
iteration += 1 .
(q) Termination condition .
check if ((iteration == MAX-ITERATION-TERMINATION) || (count-of-generation == MAX-GENERATION))
flag = 1.
end do while

(4) Display the final reduced set of attributes in $reduct$ and the dependency degree of the reduced set as it is the best dependency achieved .

END

IV. Results And Analysis

It is very likely that all the attributes of a decision system are not required to determine the class-label. Different attributes have different weight and evaluating the most informative attributes among them is the main aim of $Dimensionality\ reduction$ method. In order to evaluate the efficiency of the proposed algorithm, two key factors must be observed:

- The extent of dimensionality reduction, i.e observing the number of attributes present in the $reduct$.
- The accuracy of classification using the reduct set .

The proposed algorithm is applied on three datasets and comparisons between DIM-RED-GA() and other rough-set and fuzzy-rough set based methods [27],[30] are summarized in table 3.

Table 3: Dimensionality Reduction

Data sets	Actual no-of-atr	FRDR-BE	DIM-RED-GA
Hypothyroidism	3	3	3
Pulmonary Embolism	4	4	4
Wine	13	8	3

The accuracy of classification is judged using different classifiers [39], shown in table 5.

The accuracy and coverage of reduct formation by DIM-RED-GA is compared with standard Rough-Set-Exploration-System(RSES), given in table 4.

Table 4: Comparison of Accuracy and Coverage

Data Set	Accuracy		Coverage	
	RSES	DIM-RED-GA	RSES	DIM-RED-GA
Hypothyroidism	75.80	92.00	48.00	100
Pulmonary Embolism	71.40	75.00	15.20	100
Wine	88.10	99.43	69.50	100

It has been observed that the accuracy of classification and the coverage of **DIM-RED-GA** is even better than standard **RSES**. The algorithm uses fuzzy-rough concepts so discretization of attribute values of the information system is not required, hence overcoming the problem of information loss as in case of Rough Set approach. Genetic algorithm in optimizing the number of attributes of the reduct set is very efficient in two aspects – (1) We need not to search exhaustively all combination of attributes for evaluating the reduct set, thus improving the run-time efficiency (2) In fuzzy-rough quick-reduct, as the fuzzy-rough dependency factor is non-monotonic, it is possible that the search terminates by reaching a local optimum whereas the global optimum may lie elsewhere in the search space, this is overwhelmed by random search and updation process of optimization of genetic algorithm.

V. Conclusions

In this paper, the shortcomings of traditional rough set attribute reduction has been highlighted and a new algorithm **DIM-RED-GA** based on fuzzy-rough sets has been proposed. Before applying the proposed algorithm, three types of datasets are processed to build the decision table. The new approach incorporates the information usually lost in crisp discretization by utilizing fuzzy-rough sets to provide a more informed technique. $DIM-RED-GA$ also utilizes the concept of genetic algorithm to obtain the optimal reduct set by incorporating randomness in search process. This helps in surmounting the problem of getting struct into the local optimum as in case of fuzzy-rough Quick Reduct(FRQR). It is ascertained that in $DIM-RED-GA$ the average length of reduct are less or equal than those found in other traditional methods and the accuracy of classification is also compatible.

References

[1] Narendra S. Chaudhari, Avishek Ghosh. Feature Extraction Using Fuzzy Rule Based System, *International Journal of Computer Science and Applications*, Volume 5, No. 3, pages 1 - 8, 2008.

[2] Edwin Lughofer. On dynamic soft dimension reduction in evolving fuzzy classifiers, In *Proceedings of the Computational intelligence for knowledge-based systems design, and 13th international conference on Information processing and management of uncertainty*,$IPMU'$10,pages 79-88,Springer-Verlag, 2010.

[3] Edmundo Bonilla Huerta, Beatrice Duval, Jin-Kao Hao. Fuzzy Logic for Elimination of Redundant Information of Microarray Data, *Journal of Genomics, Proteomics & Bioinformatics, Elsevier*, Volume 6, Issue 2, pages 61 - 73, 2008.

[4] Ch. Aswani Kumar. Reducing data dimensionality using random projections and fuzzy k-means clustering, *International Journal of Intelligent Computing and Cybernetics*, Volume 4 Issue 3, pages 353 - 365, 2011.

[5] Nick J. Pizzi, WitoldPedrycz. Classifying High-Dimensional Patterns Using a Fuzzy Logic Discriminant Network, *Advances in Fuzzy Systems*, vol. 2012, Article ID 920920, 7 pages, 2012.

[6] K. Thangavel, A. Pethalakshmi. Dimensionality reduction based on rough set theory : A review, *Journal of*

Applied Soft Computing, Volume 9, Issue 1, pages 1 - 12, 2009.

[7] Mohammed Sammany, T. Medhat. Dimensionality Reduction Using Rough Set Approach for Two Neural Networks-Based Applications, In *Proceedings of the international conference on Rough Sets and Intelligent Systems Paradigms*, pages 639-647, Springer-Verlag, 2007.

[8] Qiang Shen, Alexios Chouchoulas. Rough Set-Based Dimensionality Reduction For Supervised And Unsupervised Learning, *International Journal Appl. Math. Comput. Sci.*, Volume 11, No. 3, pages 583-601, 2001.

[9] A. Chouchoulas, Q. Shen. Rough Set-Aided Keyword Reduction for Text Categorisation, *Applied Artificial Intelligence*, volume 15, no. 9, pages 843-873, 2001.

[10] N. Zhong, J. Dong, S. Ohsuga. Using Rough Sets with Heuristics for Feature Selection, *Journal of Intelligent Information Systems*, volume 16, pages 199-214, 2001.

[11] Richard Jensen and Qiang Shen. Fuzzy-Rough Sets for Descriptive Dimensionality Reduction, In *Proceedings of the 11th International Conference on Fuzzy Systems*, pages 29-34, 2002.

[12] Fadi Abu-Amara and Ikhlas Abdel-Qader. Hybrid Mammogram Classification Using Rough Set and Fuzzy Classifier, *International Journal of Bio-medical Imaging*, 2009.

[13] Ming Yang, Songcan Chen and Xubing Yang. A novel approach of rough set-based attribute reduction using fuzzy discernibility matrix, In *Proceedings of the Fourth International Conference on Fuzzy Systems and Knowledge Discovery*, Volume 03, pages 96-101, IEEE Computer Society, 2007.

[14] Qinghua Hu, Daren Yu, Zongxia Xie. Information-preserving hybrid data reduction based on fuzzy-rough techniques, *Pattern Recognition Letters*, Volume 27, pages 414-423, 2006.

[15] Richard Jensen, Qiang Shen. Semantics-Preserving Dimensionality Reduction: Rough and Fuzzy-Rough-Based Approaches, *IEEE Transactions On Knowledge And Data Engineering*, Volume 17, No. 1, 2005.

[16] R. Jensen, Q. Shen. Fuzzy-Rough Attribute Reduction with Application to Web Categorization, *Fuzzy Sets and Systems*, volume 141, no. 3, pages 469-485, 2004.

[17] Hu Guohua, Shi Yuemei. An Attribute Reduction Method Based on Fuzzy-Rough Sets Theories, In *First International Workshop on Education Technology and Computer Science*, $ETCS - 2009$, volume 3, pages 828-831, 2009.

[18] Richard Jensen, Qiang Shen. Rough and Fuzzy Sets for Dimensionality Reduction, In *Proceedings of 2001 UK Workshop on Computational Intelligence*, pages 69-74, 2001.

[19] Richard Jensen, Andrew Tuson, Qiang Shen. Extending propositional satisfiability to determine minimal fuzzy-rough reducts, In *2010 IEEE International Conference on Fuzzy Systems*, FUZZ-IEEE pages 1-8, 2010.

[20] Richard Jensen, Chris Cornelis. Fuzzy-rough instance selection, In *2010 IEEE International Conference on Fuzzy Systems*, FUZZ-IEEE pages 1-7, 2010.

[21] Neil Mac Parthalain, Richard Jensen. Measures for unsupervised fuzzy-rough feature selection, In *Proceedings of International Conference on Intelligent Systems Design and Applications*, $ISDA - 2009$, pages 560-565, 2009.

[22] Rajen B. Bhatt, M.Gopal. On Fuzzy Rough Set Approach To Feature selection, *IEEE Transaction on Fuzzy Systems*, volume 15, pages 73-89, 2007.

[23] CHEN Haishan, WANG Meihong, QIAN Feng, JIANG Qingshan. Research on Combined Rough Sets with Fuzzy Sets, In *Proceeding of the International Symposiums on Information Processing*, ISIP-2008, 2008.

[24] (Pawlak,1982) Zdzislaw Pawlak. Rough sets, *International Journal of Computer and Information Sciences*, volume 11, pages 341-356, 1982.

[25] Pramod Kumar P., PrahladVadakkepat, Ai PohLoh. Fuzzy-rough discriminative feature selection and classification algorithm, with application to microarray and image datasets. *Appl. Soft Comput.* 11(4): 3429-3440, 2011.

[26] Sankar K. Pal, Pabitra Mitra. Case Generation Using Rough Sets with Fuzzy Representation, *IEEE Trans. on Knowl. and Data Eng.*,volume 16,number 3,pp292300, 2004.

[27] Richard Jensen, Qiang Shen. Fuzzy-Rough Sets Assisted Attribute Selection, *IEEE Transaction on Fuzzy Systems*, volume 15, pages 73-89, 2007.

[28] Richard Jensen, Qiang Shen. New Approaches to Fuzzy-Rough Feature Selection, *IEEE Transactions on Fuzzy Systems*, volume 17, No. 4, pages 824-838, 2009.

[29] C.P.Chandran. Feature Selection from Protein Primary SequenceDatabase using Enhanced QuickReduct Fuzzy-Rough Set, *IEEE International Conference on Granular Computing*, GrC 2008, pp 111-114, 2008.

[30] Hu Guohua, Shi Yuemei. An Attribute Reduction Method based on Fuzzy-Rough Sets Theories, In *First International Workshop on Education Technology and Computer Science*, $ETCS - 2009$, volume 3, pages 828-831, 2009.

[31] Richard Jensen, Qiang Shen. Fuzzy-Rough Sets for Descriptive Dimensionality Reduction, In *Proceeding 11th International Conference on Fuzzy Systems*, pages 29-34, 2002.

[32] Pengfei Guo, Xuezhi Wang, Yingshi Han. The Enhanced Genetic Algorithms for the Optimization Design, In *3rd International Conference of Biomedical Engineering and Informatics*, BMEI-2010, volume 7, pages 2990-2994, 2010.

[33] Richard Jensen, Qiang Shen. Semantics-Preserving Dimensionality Reduction: Rough and Fuzzy-Rough-Based Approaches, *IEEE Transactions on Knowledge and Data Engineering*, volume 16, No. 12, pages 1457-1471.

[34] Richard Jensen, Qiang Shen. Are More Features Better? A Response to Attributes Reduction Using Fuzzy Rough Sets.

[35] UCI : Machine Learning Repository, *http://archive.ics.uci.edu/ml/*.

[36] Weina Wang, Yunjie Zhang, Yi Li and Xiaona Zhang. The Global Fuzzy C-Means Clustering Algorithm, In *Sixth world Congress on Intelligent Control and Automation*, $WCICA - 2006$, volume 1, pages 3604-3607.

[37] Anna M. Bartkowiak. Anomaly, Novelty, One-Class Classification:A Comprehensive Introduction, *International Journal of Computer Information Systems and Industrial Management Applications*, ISSN 2150-7988, Volume 3, pages 061-071, 2011.

[38] Tatiane M. Nogueira, Heloisa A. Camargo,Solange O. Rezende. Fuzzy Rules for Document Classification to Improve Information Retrieval, *International Journal of Computer Information Systems and Industrial Management Applications*, ISSN 2150-7988, Volume 3, pages 210-217, 2011.

[39] WEKA software classifier rules *http://weka.sourceforge.net/doc/weka/classifiers/rules/package-summary.html* .

Author Biographies

Moumita Saha was born in Kolkata on 4^{th} novenber, 1988. She has completed her Bachelor of Technology from Meghnad Saha Institute Of Technology, WBUT. She is pursuing Master Of Engineering from Bengal Engineering And Science University(BESUS), Shibpur. She is performing her research work in the field of Soft Computing.

Jaya Sil, an alumnus of BESUS(Bengal Engineering and Science University, Shibpur) and JU(Jadavpur University), completed her Ph.D. in Engineering from JU, Kolkata, India. She has been working in Academics for last 25 years and presently working as Professor in the Department of Computer Science and Technology Department in BESUS. She has more than 80 publications in International Conferences and Journals. Dr. Sil worked as Post-Doc Fellow in Nanyang Technological University, Sngapore on 2002-03 and visited Heidelburg University, Germany on 2007. Dr. Sil contributed several Book Chapters on different areas like data mining, image processing. Her areas of research include Image Processing, Soft Computing Techniques, Multi-agent Systems and Bio-Informatics.

Nandita Sengupta has done her Bachelor and Master of Engineering from Bengal Engineering and Science University, Shibpur, India. She has 21 years of working experience and last 10 years with academics in University College of Bahrain, Bahrain. Her area of interest is Analysis of Algorithm, Theory of Computation, and Network Computing.

Table 5: Comprehensive Comparison Of Classification Accuracy With The Proposed $DIM-RED-GA$

Classifier	**Accuracy**								
	Hypothyroidism			**Pulmonary-Embolism**			**Wine**		
	Or-Atr	FRQR-BE	DIM-RED-GA	Or-Atr	FRQR-BE	DIM-RED-GA	Or-Atr	FRQR-BE	DIM-RED-GA
Bayes Net	91.2	91.2	91.2	71.5	71.5	71.5	94.38	96.06	99.43
Navie Bayes	88	88	88	75.5	75.5	75.5	92.69	93.82	93.82
Navie Bayes Updatable	88	88	88	75.5	75.5	75.5	92.69	93.82	93.82
Logistic	92	92	92	75	75	75	99.43	99.43	99.43
Multilayer Perceptron	92.4	92.4	92.4	75	75	75	96.62	96.62	98.31
RBF Network	92.4	92.4	92.4	79	79	79	93.82	93.25	97.19
SMO	88	88	88	73	73	73	94.38	94.38	95.50
IBK	87.2	87.2	87.2	80.5	80.5	80.5	94.94	94.44	97.19
K-Star	88.4	88.4	88.4	81.5	81.5	81.5	90.44	89.88	97.19
Bagging	96.4	96.4	96.4	85	85	85	99.43	99.43	99.43
Decision Table	92.4	92.4	92.4	74.5	74.5	74.5	99.43	99.43	99.43
J-Rip	95.6	95.6	95.6	80	80	80	98.87	98.31	98.87
NNge	96.4	96.4	96.4	80.5	80.5	80.5	98.87	99.43	99.43
PART	96.8	96.8	96.8	79.5	79.5	79.5	98.87	98.87	98.87
Ridor	96.8	96.8	96.8	82	82	82	98.87	98.87	98.87
J48	96.4	96.4	96.4	86.5	86.5	86.5	98.87	98.87	98.87
LMT	93.6	93.6	93.6	85	85	85	99.43	99.43	99.43
NB-Tree	97.2	97.2	97.2	80	80	80	99.43	99.43	99.43
Random Forest	96.4	96.4	96.4	87	87	87	99.43	99.43	99.43
Random Tree	87.2	87.2	87.2	68.5	68.5	68.5	93.25	95.50	98.31