Multiple Reducts Computation in Rough Sets with applications to Ensemble Classification

Abhimanyu Bar and P. S. V. S. Sai Prasad

February 27, 2021

Multiple Reducts Computation in Rough Sets with applications to Ensemble Classification

Abhimanyu Bar[1] and P. S. V. S. Sai Prasad[2]

School of Computer and Information Sciences, University of Hyderabad, Hyderabad, India-500046
[1]abhi16@uohyd.ac.in,[2]saics@uohyd.ernet.in

Abstract. Rough set theory has emerged as an influential soft-computing approach for feature subset selection (reduct computation) in the decision system amidst incompleteness and inconsistency. Multiple reducts computation using rough sets provide an elegant way for construction of ensemble classifier for better and stable classification. The existing approaches for multiple reducts computation are primarily based on the genetic algorithm and select diverse multiple reducts after generation of abundant candidate reducts. This work proposes an $MRGA\_MRC$ algorithm for multiple reducts computation by utilizing the systematically evolving search space of all reducts computation in the MRGA algorithm without generation of many candidate reducts. A novel heuristic is introduced for selection of diverse multiple reducts. Experiments conducted on the benchmark decision systems have established the relevance of the proposed approach in comparison to the genetic algorithm based multiple reducts computation approach REUCS.

Keywords: Rough sets · Reduct · Multiple reducts computation · Discernibility matrix · Ensemble classification

1 Introduction

In 1982, Polish mathematician Prof. Z. Pawlak introduced a mathematical formulism to analyze, represent and manipulate knowledge using the information within the data, named as rough set theory(RST) [22]. The RST deals with vague and uncertain concepts for incomplete and inconsistent decision systems. Since then in the various phases of the knowledge identification process, RST has witnessed great achievement in the domains of data mining, machine learning, pattern recognition [14], for feature selection, data reduction, pattern extraction, rule generation, etc. Many researchers from different fields like mathematics, soft computing, artificial intelligence, engineering sciences, etc., have focused on feature selection, and more precisely on the attribute reduction process. The RST based dimensionality reduction is taking place with the minimum set of attributes, called reduct, which retains the information as that of the whole set of attributes. The process selects the relevant attributes and drops the redundant/irrelevant ones for specific applications, and rank the relative importance of features which improves the required data quality and enhanced comprehensibility, as well as help in the construction of better interpretable classification models with reduced time complexity.

One of the unique aspects of rough set theory is, it has produced the methods with which one can compute all possible reducts of a given decision system. This is in particular made possible because of the work done by A. Skowron et.al [27], in 1992, using the concept of discernibility matrix. The prime implicants of induced discernibility function will result in all possible reducts for a given system.

There are many advantages, for computing all possible reducts [1,18,23,27] in a given decision system. They can be used in the construction of efficient ensemble classifier, construction of all possible rule induction, to identify the single best reduct with any kind of optimality criteria such as being minimum length or maximum strength rule inducer, etc. But it is simultaneously proved that the computation of all possible reducts is an NP-hard problem [17,32]. Hence most of the researchers in the area of the rough sets focused on developing single reduct algorithms [2,3,11,16,24,25,31], using which a single classifier can be induced. But it has been observed that for the same combination of the training and the test data, for different reducts, the generalizability of the acquired classification performance varies from one reduct to other. Thereby it can be seen that any particular single reduct algorithm can't achieve the best possible accuracy which is feasible from the model constructed. Hence researchers in rough set theory, investigated the possibility for multiple reducts computation [6,7,10,12,21,26,30,33,34,35]. Multiple reducts refers to selection of a subset of reducts in such a way that, near-optimal advantages of all possible reducts can be achieved with reduced time and space complexity and achieves better accuracies that obtained from single reduct approaches. Hence this work focuses on the area of multiple reducts computation using classical rough set theory.

In machine learning, multiple classifier ensembles become the propitious method for its better generalizability, less over-fitting, and comparable performance. As the ensemble of classifiers constructs a distinct predictor by integrating the predictions of multiple base classifiers, it needs two different phases. In the first phase, train a set of weak learners (i.e., the classifier whose performance is comparable with the random sample but not considerably good) as base classifiers, then in the second phase assimilate the predictions of base classifiers using some mechanism like majority voting for classification or weighted averaging for regression. The improvement in accuracy of the ensemble model depends upon the performance of base classifiers and diversity(i.e., the degree of disagreements among the base classifiers) [4,15].

The traditional ensemble integrates the prediction of all possible base classifiers. Sometimes, it is infeasible as decisions of the all-base classifier will need much time and lots of CPU and memory for massive datasets or even on moderate size datasets, which affects comprehensibility and understandability of rule sets. Whereas running an ensemble with fewer classifier will not improve the classification accuracy either [9]. On the conception of "Many Could be Better than

All" [36], the selective ensemble outperforms over traditional ensemble by contributing the predictions of subsets of the base classifiers, as well as much more effective than single trained classifier. The classifier trained by rough set reducts has a comparable generalization power over the traditional classifiers. The rough set reducts characterize the problem in distinct sub-spaces and capture different information about the classification of objects. As rough set reducts facilitate the trade-off between the conditions of ensemble classifiers and training base classifiers with multiple reducts, having better generalizability, can improve the performance.

The literature [7,6,13,30] shows that the rough set reducts based ensemble perform comparable or better than the single trained classifiers. This has motivated us to investigate methods for efficient computation of multiple reducts and to reach their suitability for ensemble classifier construction.

A review of exiting approaches for multiple reducts computation is described below. In 1995, J. Wroblewski [33] proposed three approaches using a genetic algorithm(GA), for finding short reducts. The first method use bit string representations to find the global minimal reducts among the possible reducts, where as it deteriorates the global optimal solution some times. In the rest, two methods the author used greedy heuristic with permutational coding for achieving the short reducts. Method-3 achieves better performance than others, as it results in reducts where as others in subreducts.

In 1997, Xi. Hu et al. [12] proposed a discernibility matrix based multiple reducts approach. Where the process begins with the most common element among the entries of the matrix, then use step wise forward selection and backward elimination on the most significant and the least significant conditional attributes respectively for generating reducts. The process terminates when all indispensable attributes are part of the reducts.

In 2005, QX. Wu et al. [34] proposed an approach for calculating multiple reducts where the attributes are categorized as indispensable and dispensable with some degree of importance. Then on every go, they dropped the least important dispensable attribute as a superfluous feature and reassigned the degree ranking among the rest. The process continues until no non-significant attributes are left.

In 2010, X. Pan et al. [21] proposed a genetic algorithm based forward selection approach, where GA computes the minimal length reducts set without accessing all possible reducts with the help of fitness values of attributes. Then selects the highest fitness attribute as the candidate reduct in each iteration.

In 2013, E. Debie et al. [7] proposed a GA based diverse reduct heuristic for computing multiple reducts where the multiple reducts are computed as post-processing of all possible reduct generated by SAVGenetic Reducer [20] with an average similarity measure for maintaining diversity among the reducts.

In 2018, A. Trabelsi et al. [30] proposed a GA based belief discernibility matrix [20] approach, where the first set of reduct is chosen randomly from the initial list. Then the reducts added progressively with the highest degree of diversity. The diversity degree is maintained as the cumulative average of the candidate reducts.

Both the works in [7,30] have focused on the computation of diverse multiple reducts for facilitating uncorrelated base classifiers. But the selection of multiple reducts are made out of the reducts computed through SAVGenetic [20] algorithm. Hence the selection of reducts is applied as a postprocessing step of computation of multiple reducts.

In this paper, an alternative approach for multiple reducts computation is proposed by utilizing the search space generated in R.Susmaga's modified reduct generation algorithm(MRGA) [29]. The MRGA approach provides a systematic way for computation of all reducts, working sequentially over clauses in discernibility function. In $i+1^{th}$ iteration, reducts induced by the initial 'i' clauses are used for computations of reducts for initial $i+1$ clauses. This systematic process inspired us to design an approach for arriving at diverse multiple reducts without prior computation of all or many candidate reducts. Thus our goal is to develop an efficient methods for multiple reducts computation using classical rough sets for ensemble classification. The proposed approach is named as $MRGA\_MRC$ (Modified Reduct Generation Algorithm based Multiple Reducts Computations).

The work of the paper is organized as follows: Section-2 presents the required basics of all reducts computation based on discernibility matrix and discernibility function. Section-3 presents the review of R.Susmaga's reduct generation algorithm MRGA [29] followed by $MRGA\_MRC$ approach. Section-4 covers the experimental results along with analysis. Finally, section-5 concluded this work with some motivation for the extension of work towards the further improvement of the proposed approaches.

2 Basics for Multiple Reducts Computation

2.1 Decision System

Let the decision system be represented as $DS = (\mathbb{U}, \mathbb{Q})$, where $\mathbb{U}$ denotes the universe of objects and each object is having a description over the feature set $\mathbb{Q} = \{C \cup \{d\}\}$. Here C is the finite set conditional attributes and $\{d\}$ is decision attribute. Then for any $\mathbb{B} \subseteq C$, there is an associated relation, called discernibility relation, denoted as

Discernibility Relation: $DISC_{DS}(\mathbb{B})$ [19], where

$$DISC_{DS}(\mathbb{B}) = \{(x, y) \in \mathbb{U} \times \mathbb{U} \;/\; \exists\; q \in \mathbb{B},\;\; q(x) \neq q(y)\} \tag{1}$$

Hence two objects are related by discenibilitry relation based on the attribute set $\mathbb{B}$, if and only if they differ atleast on one attribute in $\mathbb{B}$. In a decision system the useful features have the ability to discern objects from different decision classes.

2.2 Discernibility Matrix

For computational representation, the knowledge of discernibility relation is represented in the form of a decision related discernibility matrix

$$DM = [c_{ij}]_{i,j \in \{1,..,|U|\}}, \quad (2)$$

where

$$c_{ij} = \begin{cases} \{q \in C \ / \ q(x_i) \neq q(x_j)\}, & \text{for } \ d(x_i) \neq d(x_j) \\ \lambda & \text{otherwise} \end{cases}$$

The discernibility matrix is a symmetric matrix, and valid entries are provided only for a pair of objects belonging to different decision classes. All the other entries are represented as λ and omitted from further computations.

2.3 Discernibility Function

A boolean discenibility function is constructed from non empty entries of discernibility matrix. A discernibility function [27] f_S is a boolean function of m-boolean variables $b_1^*, ...b_m^*$ (corresponding to the attributes $b_1, ..., b_m$) defined as:

$$f_S(b_1^*, ..., b_m^*) = \bigwedge \left\{ \bigvee c_{ij}^* \ / \ 1 \leq j \leq i \leq |\mathbb{U}|, c_{ij} \neq \phi \right\}, \quad (3)$$

$$\text{where } c_{ij}^* = \{b^* / \ b \in c_{ij}\}$$

Any prime implicant of discernibility function corresponds to a set of attributes which are necessary and sufficient to discern any pair of objects of different decision classes. Hence the attribute corresponding to prime implicant is a reduct for the given decision system. The set of all possible prime implicants of f_S determines the set of all possible reducts of the decision system but computationally is an NP-hard problem [27].

3 Proposed Work

3.1 All Reducts Generation

To generate all exact reducts, R.Susmaga proposed a two phase approach MRGA[29]. The algorithm starts with creating a discenibility matrix [27]. Absorbed discernibility list(ADL) is computed from the clauses of dicernibility matrix, by removal of redundant entries for prime implicants computation in the repeated application of absorption law and arranged in ascending order of the cardinality of its elements.
$ADL = \{C_1, C_2...C_k\}$, where $k \in \{[1, N(N-1)/2]\}$ and is defined as,

$$ADL = \{c_{i,j} \in DM \ / \ c_{i,j} \neq \phi \wedge \nexists c \in DM : c \subset c_{i,j}\} \quad (4)$$

The second phase of MRGA computes the reducts by minimal subsets having non-empty intersection with the entries of the ADL. Let R_i denotes the set of reducts for $C_1, ..., C_i$ clauses. i.e., R_i is having non-trivial intersection with $C_1..C_i$ along with being minimal. In computation of reducts from C_1 to C_{i+1} clauses, S_{i+1}, T_{i+1} are computed. S_{i+1} contains the reducts in R_i having non-trivial intersection with C_{i+1} and hence becomes reducts for C_1 to C_{i+1} clauses. For each of the reducts in $R_i - S_{i+1}$, reducts are generated in T_{i+1} by augmenting every attributes in C_{i+1}. The entries of T_{i+1} are checked against a minimal criteria for removal of non minimal entries. The resulting $S_{i+1} \cup T_{i+1}$ becomes all reducts set R_{i+1} for C_1 to C_{i+1}. After completion of k-iterations algorithm results in all reducts for the given decision system.

3.2 Proposed Multiple Reducts Computation Algorithm($MRGA\_MRC$)

The objective of MRGA_MRC is to compute "n" diverse multiple reducts, where "n" is user supplied parameter based on the intended number of the base classifier for ensemble classifier construction. $MRGA\_MRC$ algorithm restricts MRGA algorithm to generate the desired n-multiple reducts which satisfy the diverse reducts criteria. The algorithm $MRGA\_MRC$ is given in algorithm-1. In algorithm-1 RED denotes R_i. In i^{th} iteration after computation of RED following MRGA procedure, if the size of the RED exceeds "n", then a selection criteria is employed in $MRGA\_MRC$. Initially, the entries of RED are sorted in increasing order of their length, and the first entry(shortest reduct) is entered into multiple reducts list SRED. The selection of remaining "n-1" multiple reducts is done in sequence. The next reduct is selected based on the proposed heuristic for achieving diversity. For each of the candidate reducts, B in $RED - SRED$ the following measure is computed.

$$L_B = \max_{j \in SRED} |B \cap SRED(j)| \tag{5}$$

The next reduct B^* is selected having the least value for heuristic measure L. This ensures the selected B^* has a lesser overlap with the entries of SRED. As the elements of RED are sorted in the increasing order of the length, in case of any tie, the lesser length reduct is selected into SRED. The selected B^* is then added to the SRED, and the procedure continues until the required number of reducts are computed. In algorithm-1, Phase-I and phase-II correspond to R.Susmaga's MRGA [29] and the Phase-III correspond to multiple reducts computation. In this manner, $MRGA\_MRC$ controls the search space of the MRGA algorithm and arrives at computations of diverse multiple reducts without generating all candidate reducts.

3.3 Ensemble of classifiers using multiple reducts

The rough set reducts of any data set characterize the system in distinct subspaces and capture different knowledge about the classification [7]. Thus an

Algorithm 1 MRGA based Multiple Reduct Computation (MRGA_MRC)

Input 1: Finite Set of objects $\mathbb{U}(|\mathbb{U}| = N)$ characterized by values of the features set $Q = \{C \cup \{d\}\}$, and n (cardinality of multiple reduct).

Output: The set of n-diverse reducts as samples from $\mathbb{Q}$.

PHASE-I ***Generating Discernibility Matrix and create ADL .***

Step-1

Create discernibility matrix and transform to absorbed discernibility list: $ADL = \{C_1, C_2 ... C_k\}$ where $k \in [1, N(N-1)/2]$ with non empty and non-minimal elements.

$ADL = \{c_{i,j} \in DM \ / \ c_{i,j} \neq \phi \wedge \nexists \ \ c \in DM : c \subset c_{i,j}\}$

Where $DM = [c_{ij}]_{i,j \in \{1,..,|\mathbb{U}|\}}$ and $c_{i,j} = \{q \in \mathbb{Q} \ / \ q(x_i) \neq q(x_j)\}$ for $i, j = 1, 2..., |\mathbb{U}|$

Step-2

Sort ADL in the ascending order of the cardinality of its elements.

PHASE-II **:*Computing ALL Reducts***

Step-1

$RED = \{\phi\}$

Step-2

for i= 1,2...$|ADL|$, then compute.

$SL_i = \{B \in RED : B \cap C_i = \phi\}$

$TR_i = \bigcup_{a \in C_i} \bigcup_{B \in RED \ \wedge \ B \cap C_i \neq \phi} \{B \cup \{a\}\}$

$MIN_i = \{B \in TR_i \ / \ Min(B, ADL, i) = TRUE\}$

$RED = SL_i \cup MIN_i$

PHASE-III **:*MRC(Multiple Reducts Computation)***

Step-1

$S\_RED = \{\phi\}$

Step-2

if $|RED| > n$,then

$S\_RED = Sort(RED)$: Sort the RED ascending order of the reduct size.

$SRED = \{S\_RED(1)\}$

Step-3

for j = 2 to n:

$BestCount = length(ADL)$

$CurrentCount = \phi$

for B in $S\_RED - SRED$

$L = \max_{j \in SERD} |B \cap SRED(j)|$

if $(L < BestCount)$

$BestCount = l$

$CurrentBest = B$

$SRED = SRED \cup CurrentBest$

$RED = SRED$

$Return$ $RED.$

attempt to train multiple classifiers using multiple reducts generated by either feature partitioning or random subspace approach will induces an ensemble, which expected to achieve better classification accuracy. Multiple reducts based ensemble can be constructed in three different phases: first creates sample by feature partitioning, (i.e., generating multiple rough set reducts using proposed $MRGA\_MRC$ algorithm as the sample for the ensemble). In the second phase, construct base classifiers with selected reducts. Finally, integrate all the predictions of the participated classifiers as a predictive ensemble model. The outline of the above-described ensemble is depicted in Fig-1.

4 Experimentation

For validating the relevance of selected reducts through diversity criterion and efficiency of ensemble using them, the experimental computations are carried out for performance analysis of the ensemble with JRip [5] classifiers based on diverse-multiple reducts generated from the proposed MRGA_MRC approach. The Results of the $MRGA\_MRC$ based multiple reducts ensemble is compared with the random reducts ensemble: R2-JRip, all attribute JRip (All-JRip:Single classifier with all attributes) and REUCS [7] (i.e., Reduct based ensemble of a supervised classifier system, is an accuracy-based hybrid model to solve the real-valued classification problem using approximate reducts) based ensemble classifier.

In R2-JRip multiple reducts ensemble, the multiple reducts are computed through $MRGA\_MRC$ approach but using a random selection in R_i without using the proposed heuristic. The comparison with R2-JRip is done for the purpose of validating the proposed heuristic for diverse reducts computation in contrast to random selection.

All-JRip classifier is constructed using all attributes. A comparison is done with REUCS ensemble classifier proposed in [7] which is based on diverse multiple reducts computed from candidate reducts generated through a genetic algorithm. The results reported for REUCS are taken from the [7] and reported along with our results.

4.1 Experimental Setting

The experimental setup includes, Processor: Intel(R) Core i5-7500 CPU, Clock Speed: 3.40 GHz $\times$ 4, RAM: 8GB DDR4, Operating System: Ubuntu-18.04.1 LTS 64 bit, and Implemented in software: R Studio Version 1.1.463 with R version 3.5.2

Experiments are conducted, on six benchmark categorical datasets, are taken from UCI-machine learning repository [8]. To generate training samples as well as test samples, a 10-fold cross validation strategy is used. The precise description of data sets is given in Table-1. Table -1 provides the number of all possible

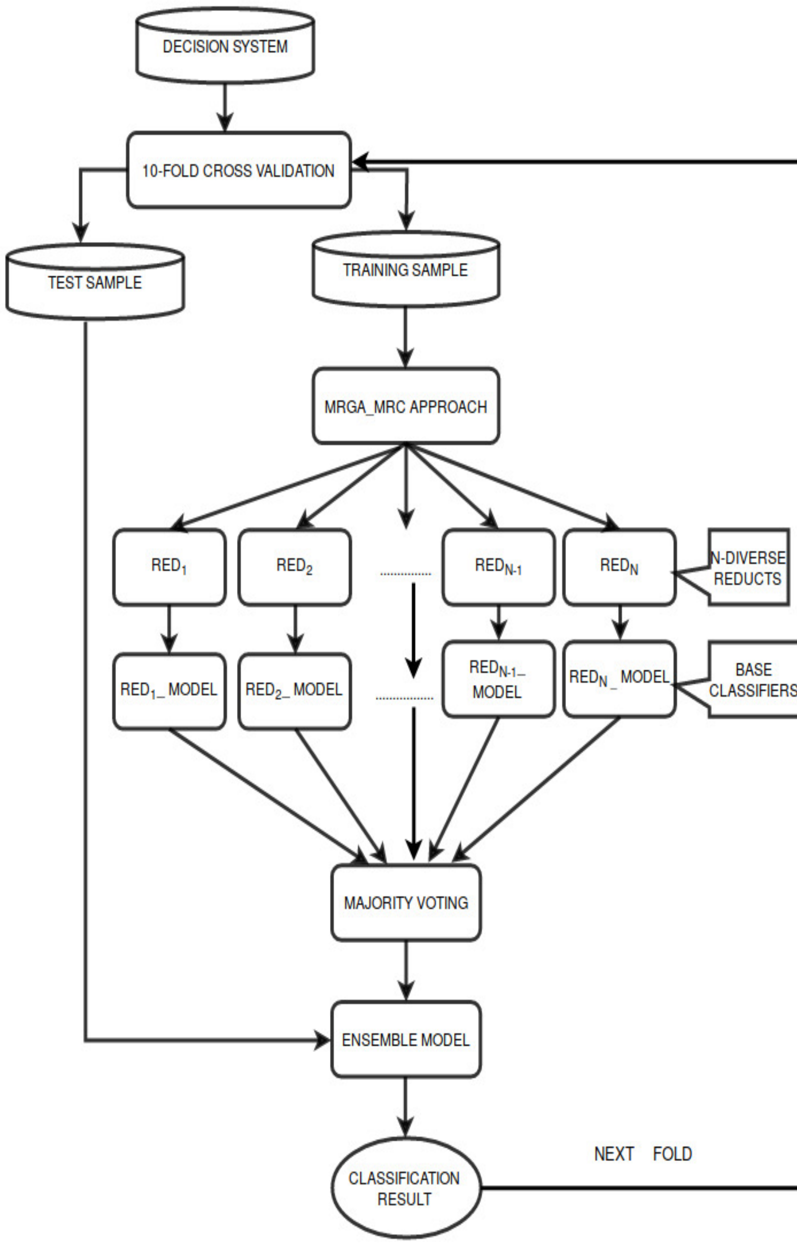

Fig. 1: Architecture of $MRGA\_MRC$ ensemble classification through 10-fold cross validation.

reducts computed using MRGA algorithm along with basic details of data sets. The computation could not be completed for Sonar, Spectf, Wpbc data sets, due to system limitations but based on the processing that was feasible; minimum bound cardinality is provided.

For example, the sonar data set is established to have more than 90000 reducts, illustrating the huge search space within which $MRGA\_MRC$ arrives at “n” diverse multiple reducts. The “n” value used in the experiment is 23.

Table 1: The description of experimental data sets with possible number of reducts

Sl.No.	**Data Sets**	**No.of Objects**	**No.of Variables**	**No.of Classes**	**No.of Reducts**
1	Breast-Cancer	569	30	2	55
2	Hepatitis	155	19	2	238
3	Horse-Colic	368	22	2	9150
4	Sonar	208	60	2	> 90000
5	Spectf	267	44	2	> 1000
6	Wpbc	198	33	2	> 1000

4.2 Experimental Results

To analyze the results of $MRGA\_MRC$ based ensemble and other compared methods, a 10-fold cross validation strategy is used in the above-mentioned data sets. The results are reported in Table-2. For each dataset, the results are reported in the form of a matrix, where the rows and columns correspond to the used approaches. The principle diagonal provides $\mu \pm \sigma$ (mean $\pm$ standard deviation) for the accuracies obtained with the corresponding approach in the 10-fold experiment. Between two approaches, a t-test is conducted using the online tool GraphPad t-test calculator [28], and the resulting p-values are summarized in the non-diagonal entries. If an approach of a row is statistically significant than the approach of the column, is indicated by placing an appropriate number of “*s” in the matrix. Here we use “*” for $p \leqslant 0.05$- statistically significant, “** ” for $p \leqslant 0.01$ -very statistically significant, “***” for $p \leqslant 0.001$-extremely statistically significant. In the corresponding complement, the position obtaining statistical inferior performance is indicated through “#”. If both approaches are similar and results are not statistically significant ($p > 0.05$) is indicated by “-”. For example, in Breast -cancer data set R1-JRip performance is very statistically significant than R2-JRip and REUCS, but it performs similarly to All-JRip.

Table 2: 10-fold experimental results of categorical data sets

Breast-Cancer				
APPROACHES	**R1-JRip**	**R2-JRip**	**All-JRip**	**REUCS**
R1-JRip	93.69505 ± 2.848007	$**$	$-$	$**$
R2-JRip	#	86.44780 ± 7.447370	#	$-$
All-JRip	$-$	$**$	94.05220 ± 3.975904	$**$
REUCS	#	$-$	#	90.7 ± 1.2

Hepatitis				
APPROACHES	**R1-JRip**	**R2-JRip**	**All-JRip**	**REUCS**
R1-JRip	67.00000 ± 11.38008	$-$	$-$	#
R2-JRip	$-$	62.83333 ± 12.52282	$-$	#
All-JRip	$-$	$-$	63.16667 ± 10.49544	#
REUCS	$***$	$***$	$***$	81.8 ± 0.0

Horse-Colic				
APPROACHES	**R1-JRip**	**R2-JRip**	**All-JRip**	**REUCS**
R1-JRip	75.40404 ± 8.314351	$-$	$-$	#
R2-JRip	$-$	68.35859 ± 12.625730	$-$	#
All-JRip	$-$	$-$	70.20202 ± 7.684406	#
REUCS	$**$	$***$	$***$	83.3 ± 2.1

Sonar				
APPROACHES	**R1-JRip**	**R2-JRip**	**All-JRip**	**REUCS**
R1-JRip	75.57143 ± 8.072798	$**$	$-$	$**$
R2-JRip	#	63.35714 ± 10.312850	$-$	$-$
All-JRip	$-$	$-$	71.50000 ± 8.834906	$-$
REUCS	#	$-$	$-$	64.8 ± 5.9

Spectf				
APPROACHES	**R1-JRip**	**R2-JRip**	**All-JRip**	**REUCS**
R1-JRip	79.25275 ± 8.913195	$-$	$-$	$-$
R2-JRip	$-$	76.94505 ± 8.108702	$-$	$-$
All-JRip	$-$	$-$	77.71429 ± 7.647313	$-$
REUCS	$-$	$-$	$-$	81.3 ± 4.0

Wpbc				
APPROACHES	**R1-JRip**	**R2-JRip**	**All-JRip**	**REUCS**
R1-JRip	79.29825 ± 8.666027	$-$	$-$	$-$
R2-JRip	$-$	74.77583 ± 12.599290	$-$	$-$
All-JRip	$-$	$-$	76.14035 ± 10.757691	$-$
REUCS	$-$	$-$	$-$	81.1 ± 3.2

4.3 Analysis of Results

Based on the results reported in Table-2, the following observations are made. R1-JRip has obtained better average accuracies than R2-JRip in all datasets.

Based on the t-test results, R1-JRip has achieved very statistically significant performance over R2-JRip in breast cancer and sonar data sets and performed similarly in the remaining data sets. This result validates the importance of proposed heuristic-based diverse multiple reducts by achieving better ensemble performance than randomly selected multiple reducts, in the same search space of MRGA algorithm.

R1-JRip achieved better average accuracies than that of All-JRip in all data sets except breast cancer. Based on the t-test results both performed similarly in all data sets. Mixed results are noticed among R1-JRip and REUCS based ensembles. R1-JRip has achieved very statistically significant results than REUCS, in breast cancer and sonar datasets, while REUCS achieved very statistically significant results in hepatites and horse-colic data sets. Both approaches perform similarly in spectf and wpbc data sets. In spectf and wpbc datasets, REUCS achieved better average accuracies.

Summing of, these results establish the relevance of the proposed MRGA_MRC approach for induction of multiple diverse reducts and their utility in the construction of ensemble classifier. The research will be continued in MRGA_MRC in the form of alternative heuristics for diverse multiple reducts computation and parallelization of computation for obtaining improved accuracies over existing approaches.

5 Conclusion

Rough sets provide a unique approach for computation of all possible reducts in the given decision system. As computing all possible reduct is an NP-hard problem, the researchers have investigated the multiple reducts computation aiding in ensemble classification. This work proposed a novel approach for multiple reducts computation as MRGA_MRC algorithm. MRGA_MRC works in the search space of all reducts computation of MRGA algorithm using a proposed heuristic for selection of diverse multiple reducts. The experimental results amply validated the relevance of the proposed approach. Alternative heuristic and control strategies will be investigated in the future for further improvements in the algorithm.

6 Acknowledgment

This work is supported by Ministry of Human Resource Development(MHRD) Government of India under the scheme, Council of Scientific and Industrial Research - Junior Research Fellow(CSIR-JRF) with the grant number - [File No:09/414(1117)/2016-EMR-I].

References

1. Bazan, J.G., Nguyen, H.S., Nguyen, S.H., Synak, P., Wróblewski, J.: Rough set algorithms in classification problem. In: Rough set methods and applications, pp. 49–88. Springer (2000)

2. Beaubouef, T., Petry, F.E., Arora, G.: Information-theoretic measures of uncertainty for rough sets and rough relational databases. Inf. Sci. **109**(1-4), 185–195 (1998)
3. Bhatt, R.B., Gopal, M.: On the compact computational domain of fuzzy-rough sets. Pattern Recognition Letters **26**(11), 1632–1640 (2005)
4. Brown, G., Wyatt, J.L., Harris, R., Yao, X.: Diversity creation methods: a survey and categorisation. Information Fusion **6**(1), 5–20 (2005)
5. Cohen, W.W.: Fast effective rule induction. In: In Proceedings of the Twelfth International Conference on Machine Learning. pp. 115–123. Morgan Kaufmann (1995)
6. Das, A.K., Sil, J.: An efficient classifier design integrating rough set and set oriented database operations. Applied Soft Computing **11**(2), 2279–2285 (2011)
7. Debie, E.S., Shafi, K., Lokan, C., Merrick, K.E.: Reduct based ensemble of learning classifier system for real-valued classification problems. In: CIEL. pp. 66–73. IEEE (2013)
8. Dua, D., Karra Taniskidou, E.: UCI machine learning repository (2017), http://archive.ics.uci.edu/ml
9. Elkan, C.: Boosting and naıve bayesian learning. Proceedings of KDD-97, New Port Beach, CA (1997)
10. Gao, S., Dai, J., Shi, H.: Discernibility matrix-based ensemble learning. In: 2018 24th International Conference on Pattern Recognition (ICPR). pp. 952–957. IEEE Computer Society (2018)
11. Guan, J., Bell, D.A.: Rough computational methods for information systems. Artificial intelligence **105**(1-2), 77–103 (1998)
12. Hu, X., Cercone, N., Ziarko, W.: Generation of multiple knowledge from databases based on rough sets theory. In: Rough Sets and Data Mining, pp. 109–121. Springer (1997)
13. Jr., M.P.P.: Combining classifiers: From the creation of ensembles to the decision fusion. In: 2011 24th SIBGRAPI Conference on Graphics, Patterns, and Images Tutorials. pp. 1–10. IEEE (Aug 2011)
14. Kaneiwa, K., Kudo, Y.: A sequential pattern mining algorithm using rough set theory. International Journal of Approximate Reasoning **52**(6), 881 – 893 (2011)
15. Kuncheva, L.I., Skurichina, M., Duin, R.P.: An experimental study on diversity for bagging and boosting with linear classifiers. Information fusion **3**(4), 245–258 (2002)
16. LIANG, J., SHI, Z.: The information entropy, rough entropy and knowledge granulation in rough set theory. International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems **12**(01), 37–46 (2004)
17. Lin, T., Cercone, N.: Rough sets and data mining: Analysis of imprecise data (1996)
18. Nguyen, H.S.: Approximate boolean reasoning: Foundations and applications in data mining. In: Peters, J.F., Skowron, A. (eds.) Transactions on Rough Sets V. pp. 334–506. Springer Berlin Heidelberg, Berlin, Heidelberg (2006)
19. Nguyen, H.S.: Rough sets and boolean reasoning. In: Rough Computing: Theories, Technologies and Applications, pp. 38–69. IGI Global (2008)
20. Øhrn, A., Komorowski, J., et al.: Rosetta–a rough set toolkit for analysis of data. In: Proc. Third International Joint Conference on Information Sciences. Citeseer (1997)
21. Pan, X., Zhang, S.: Ensemble remote sensing classifier based on rough set theory and genetic algorithm. In: 2010 18th International Conference on Geoinformatics. pp. 1–5 (June 2010)

22. Pawlak, Z.: Rough sets. International journal of computer & information sciences **11**(5), 341–356 (1982)
23. Pawlak, Z., Sets, R.: Theoretical aspects of reasoning about data. Kluwer, Netherlands (1991)
24. Sai Prasad, P.S.V.S., Rao, C.R.: Iquickreduct: An improvement to quick reduct algorithm. In: Sakai, H., Chakraborty, M.K., Hassanien, A.E., Slezak, D., Zhu, W. (eds.) RSFDGrC. Lecture Notes in Computer Science, vol. 5908, pp. 152–159. Springer (2009)
25. Sai Prasad, P.S.V.S., Rao, C.R.: Extensions to iquickreduct. In: Sombattheera, C., Agarwal, A., Udgata, S.K., Lavangnananda, K. (eds.) MIWAI. Lecture Notes in Computer Science, vol. 7080, pp. 351–362. Springer (2011)
26. Shi, L., Duan, Q., Zhang, J., Xi, L., Ma, X.: Rough set based ensemble learning algorithm for agricultural data classification. Filomat **32**(5) (2018)
27. Skowron, A., Rauszer, C.: The Discernibility Matrices and Functions in Information Systems, pp. 331–362. Springer Netherlands, Dordrecht (1992)
28. Software, G.: Graphpad t-test calculator (2018), https://www.graphpad.com/quickcalcs/ttest1/l
29. Susmaga, R.: Parallel computation of reducts. In: Polkowski, L., Skowron, A. (eds.) Rough Sets and Current Trends in Computing. pp. 450–458. Springer Berlin Heidelberg, Berlin, Heidelberg (1998)
30. Trabelsi, A., Elouedi, Z., Lefevre, E.: Ensemble enhanced evidential k-nn classifier through rough set reducts. In: Medina, J., Ojeda-Aciego, M., Galdeano, J.L.V., Pelta, D.A., Cabrera, I.P., Bouchon-Meunier, B., Yager, R.R. (eds.) IPMU (1). Communications in Computer and Information Science, vol. 853, pp. 383–394. Springer (2018)
31. Wang, J., Wang, J.: Reduction algorithms based on discernibility matrix: The ordered attributes method. Journal of Computer Science and Technology **16**(6), 489–504 (Nov 2001)
32. Wong, S., Ziarko, W.: On optimal decision rules in decision tables. University of Regina, Computer Science Department (1985)
33. Wroblewski, J.: Finding minimal reducts using genetic algorithms. In: Proccedings of the second annual join conference on infromation science. vol. 2, pp. 186–189 (1995)
34. Wu, Q., Bell, D.A., McGinnity, T.M.: Multiknowledge for decision making. Knowl. Inf. Syst. **7**(2), 246–266 (2005)
35. Yang, X., Yao, Y.: Ensemble selector for attribute reduction. Appl. Soft Comput. **70**, 1–11 (2018)
36. Zhou, Z.H., Wu, J., Tang, W.: Ensembling neural networks: many could be better than all. Artificial intelligence **137**(1-2), 239–263 (2002)