# Rough Set Based Unsupervised Feature Selection Using Relative dependency Measures

**C. Velayutham**
Department of Computer Science,
Aditanar College of Arts and Science,
Tiruchendur,
Thoothukudi, Tamil Nadu 628216, India
cvelayutham22@yahoo.com

**K. Thangavel**
Department of Computer Science,
Periyar University, Salem,
Tamil Nadu 636011, India
drktvelu@yahoo.com

## Abstract

Feature Selection (FS) is a process which attempts to select features which are more informative. It is an important step in knowledge discovery from data. Conventional supervised FS methods evaluate various feature subsets using an evaluation function or metric to select only those features which are related to the decision classes of the data under consideration. However, for many data mining applications, decision class labels are often unknown or incomplete, thus indicating the significance of unsupervised feature selection. However, in unsupervised learning, decision class labels are not provided. The problem is that not all features are important. Some of the features may be redundant, and others may be irrelevant and noisy. In this paper, we propose a new rough set-based unsupervised feature selection using relative dependency measures. The method employs a backward elimination-type search to remove features from the complete set of original features. As with the classification performance is evaluated using WEKA tool. The method is compared with an existing supervised method and demonstrates that it can effectively remove redundant features.

**Keywords**: Data Mining, Rough set, Supervised and Unsupervised Feature Selection.

## 1. Introduction

Feature Selection (FS) [1] is a process which attempts to select features which are more informative. It is an important step in knowledge discovery from data. The high dimensionality of databases can be reduced using suitable techniques, depending on the requirements of the data mining processes. Large dimensionality presents a problem for handling data due to the fact that the complexity of many commonly used operations are highly dependent on the level of dimensionality. The problems associated with such large dimensionality mean that any attempt to use machine learning or data mining tools to extract knowledge results in very poor performance. The main aim of feature selection is to determine a minimal feature subset from a problem domain while retaining a suitably high accuracy in representing the original features.

Conventional supervised FS methods evaluate various feature subsets using an evaluation function or metric to select only those features which are related to, or lead to, the decision classes of the data under consideration. However, for many data mining applications, decision class labels are often unknown or incomplete, thus indicating the significance of unsupervised feature selection. In a broad sense, two different types of approach to unsupervised FS have been adopted: Those which maximise clustering performance using an index function, and those which consider features for selection on the basis of dependency or relevance. The central idea, behind the latter, is that any single feature which carries little or no further information than that subsumed by the remaining features is redundant and can therefore be eliminated [2][3]. The method described in this work is related to these techniques since it involves the removal of features which are considered to be redundant.

The rest of the paper is organized as follows: Section 2 presents an introduction to the Rough Set Theory. Section 3 describes the supervised rough set based relative dependency measures for feature selection. Section 4 describes the proposed unsupervised rough set based relative dependency measures for feature selection. Section 5 describes the classification measures. The experimental results are discussed in section 6 and conclusion is presented in section 7.

## 2. Rough Set Theory

Rough Set Theory (RST) has been used as a tool to discover data dependencies and to reduce the number of attributes contained in a dataset using the data alone, requiring no additional information [4] [5]. Over the past ten years, RST has become a topic of great interest to researchers and has been applied to many domains. Given a dataset with discretized attribute values, it is possible to find a subset (termed a *reduct*) of the original attributes using RST that are the most informative; all other attributes can be removed from the dataset with minimal information loss.

*Basic Rough Set Concepts*

Let $I = (U, A \cup \{d\})$ be an information system, where $U$ is the universe with a non-empty set of finite objects. $A$ is a non-empty finite set of condition attributes, and $d$ is the decision attribute (such a table is also called decision table), $\forall a \in A$ there is a corresponding function $f_a : U \to V_a$, where $V_a$ is the set of values of $a$. If $P \subseteq A$, there is an associated equivalence relation:

$$\mathrm{IND}(P) = \{(x, y) \in U \times U \mid \forall a \in P, f_a(x) = f_a(y)\} \quad (1)$$

The partition of $U$ generated by IND($P$) is denoted $U/P$. If $(x, y) \in \mathrm{IND}(P)$, then $x$ and $y$ are indiscernible by attributes from $P$. The equivalence classes of the $P$-indiscernibility relation are denoted $[x]_P$. Let $X \subseteq U$, the $P$-lower approximation $\underline{P}X$ and P-upper approximation $\overline{P}X$ of set $X$ can be defined as:

$$\underline{P}X = \{x \in U \mid [x]_P \subseteq X\} \quad (2)$$

$$\overline{P}X = \{x \in U \mid [x]_P \cap X \neq \phi\} \quad (3)$$

Let $P, Q \subseteq A$ be equivalence relations over $U$, then the positive, negative and boundary regions can be defined as:

$$\mathrm{POS}_P(Q) = \bigcup_{X \in U/Q} \underline{P}X \quad (4)$$

$$\mathrm{NEG}_P(Q) = U - \bigcup_{X \in U/Q} \overline{P}X \quad (5)$$

$$\mathrm{BND}_P(Q) = \bigcup_{X \in U/Q} \overline{P}X - \bigcup_{X \in U/Q} \underline{P}X \quad (6)$$

The positive region of the partition $U/Q$ with respect to $P$, $\mathrm{POS}_P(Q)$, is the set of all objects of $U$ that can be certainly classified to blocks of the partition $U/Q$ by means of $P$. $Q$ depends on $P$ in a degree k ($0 \le k \le 1$) denoted $P \Rightarrow_k Q$

$$k = \gamma_P(Q) = \frac{|\mathrm{POS}_P(Q)|}{|U|} \quad (7)$$

where $P$ is a set of condition attributes and $Q$ is the decision, $\gamma_P(Q)$ is the quality of classification [4]. If $k = 1$, $Q$ depends totally on $P$, if $0 < k < 1$, $Q$ depends partially on $P$, and if $k = 0$ then $Q$ does not depend on $P$. The goal of attribute reduction is to remove redundant attributes so that the reduced set provides the same quality of classification as the original. The set of all reducts is defined as:

$$\mathrm{Red} = \{R \subseteq C \mid \gamma_R(D) = \gamma_C(D), \forall B \subset R, \gamma_B(D) \neq \gamma_C(D)\} \quad (8)$$

A dataset may have many attribute reducts. The set of all optimal reducts is:

$$\mathrm{Red}_{\min} = \{R \in \mathrm{Red} \mid \forall R' \in \mathrm{Red}, |R| \le |R'|\} \quad (9)$$

## 3. Supervised Feature Selection

The supervised FS methods evaluate various feature subsets using an evaluation function or metric to select only those features which are related to, or lead to, the decision classes of the data under consideration.

In many real world problems FS is a must due to the abundance of noisy, irrelevant or misleading features. For instance, by removing these factors, learning from data techniques can take place very effectively. A detailed review of feature selection techniques devised for classification tasks can be found in [6][7].

The usefulness of a feature or feature subset is determined by both its relevancy and redundancy. A feature is said to be relevant if it is predictive of the decision feature(s), otherwise it is irrelevant. A feature is considered to be redundant if it is highly correlated with other features. Hence, the search for a good feature subset involves finding those features that are highly correlated with the decision feature(s), but are not correlated with each other.

### *3.1 Relative Dependency Measures*

In [8], a feature selection method based on an relative dependency measure is presented. The technique was originally proposed to avoid the calculation of discernability functions or positive regions, which can be computationally expensive without optimizations. The authors replaced the traditional rough set degree of dependency with an alternative measure, the relative dependency of which is defined as follows for an attribute subset $R$:

$$K_R(D) = \frac{|U / \mathrm{IND}(R)|}{|U/\mathrm{IND}(R \cup D)|} \quad (10)$$

Then it was proved that $R$ is a reduct if and only if $K_R(D) = K_C(D)$ and $\forall X \subset R$, $K_X(D) \neq K_C(D)$.

### *3.2 The Relative Reduct Algorithm*

The algorithm in Fig. 1 is constructed for feature selection based on the measure of backward elimination of features where attributes are removed from the set of considered attributes if the relative dependency equals one upon their removal. Attributes are considered one at a time, starting with the first, evaluating their relative dependency[9][10].

---

**RelativeReduct**($C$, $D$)
$C$, the conditional attributes;
$D$, the decision attributes;
(1) $R \leftarrow C$
(2) $\forall a \in C$
(3) **if** ($K_{R-\{a\}}(D) == 1$)
(4) $R \leftarrow R - \{a\}$
(6) **return** $R$

---

Fig. 1 The Relative Reduct Algorithm

## 4. Unsupervised Feature Selection

In this section, a novel unsupervised reduct algorithm is proposed. The method is based on relative dependency measure using rough set theory. In data mining applications, decision class labels are often unknown or incomplete, during this situation the unsupervised feature selection is play vital role to select features.

### *4.1 Relative Dependency Measure*

The unsupervised relative dependency measure for an attribute subset is defined as follows:

$$K_R(\{a\}) = \frac{|U/IND(R)|}{|U/IND(R \cup \{a\})|}, \forall a \in A \qquad (11)$$

Then show that $R$ is a reduct if and only if $K_R(\{a\}) = K_C(\{a\})$ and $\forall X \subset R$, $K_X(\{a\}) \neq K_C(\{a\})$.

In this case, the decision attribute used in the supervised feature selection, is replaced by the conditional attribute *a*, which is to be eliminated from the current reduct set *R*.

Table.1 Example dataset

| x∈U | a | b | c | d |
|---|---|---|---|---|
| **1** | 1 | 0 | 2 | 1 |
| **2** | 1 | 0 | 2 | 0 |
| **3** | 1 | 2 | 0 | 0 |
| **4** | 1 | 2 | 2 | 1 |
| **5** | 2 | 1 | 0 | 0 |
| **6** | 2 | 1 | 1 | 0 |
| **7** | 2 | 1 | 2 | 1 |

### *4.2 Unsupervised Relative Reduct(USRR) Algorithm*

The new USRR algorithm is shown in Fig. 2. The algorithm starts by considering all of the features contained in the dataset.

---

**USRelativeReduct**(*C*)
*C*, the conditional attributes;
(1) $R \leftarrow C$
(2) $\forall a \in C$
(3) **if** ($k_{R-\{a\}}(\{a\}) == 1$)
(4) $R \leftarrow R - \{a\}$
(6) **return** *R*

---

Fig. 2 The USRelative Reduct Algorithm

Each feature is then examined iteratively, and the relative dependency measure is calculated. If the relative dependency is equal to 1 then that feature can be removed. This process continues until all features have been examined.

### *4.3 Worked Example*

Now consider the example dataset given in Table 1. The backward elimination algorithm initializes *R* to the set of conditional attributes, {*a, b, c, d*}. Next, the attribute *a* is considered for elimination:

$$K_{\{b,c,d\}}(\{a\}) = \frac{|\frac{U}{IND(b,c,d)}|}{|\frac{U}{IND(\{a,b,c,d\})}|}$$

$$= \frac{|\{\{1\}\{2\}\{3\}\{4\}\{5\}\{6\}\{7\}\}|}{|\{\{1\}\{2\}\{3\}\{4\}\{5\}\{6\}\{7\}\}|} = \frac{7}{7}$$

As the relative dependency is equal to 1, attribute *a* can be removed from the reduct candidate becomes *R* ={a, *b, c, d*}. Hence the current reduct candidate *R*={*b, c, d*}. The algorithm then considers the elimination of attribute *b* from *R*:

$$K_{\{c,d\}}(\{b\}) = \frac{|\frac{U}{IND(c,d)}|}{|\frac{U}{IND(\{b,c,d\})}|}$$

$$= \frac{|\{\{1\,4\,7\}\{2\}\{3\,5\}\{6\}\}|}{|\{\{1\}\{2\}\{3\}\{4\}\{5\}\{6\}\{7\}\}|} = \frac{4}{7}$$

As this does not equal 1, attribute *b* is not removed from *R*. The algorithm then evaluates the elimination of attribute *c* from *R* :

$$K_{\{b,d\}}(\{c\}) = \frac{|\frac{U}{IND(b,d)}|}{|\frac{U}{IND(\{b,c,d\})}|}$$

$$= \frac{|\{\{1\}\{2\}\{3\}\{4\}\{5\,6\}\{7\}\}|}{|\{\{1\}\{2\}\{3\}\{4\}\{5\}\{6\}\{7\}\}|} = \frac{6}{7}$$

Again, the relative dependency does not evaluate to 1. Hence attribute *c* is retained in the reduct candidate. The next step evaluates the removal of *d* from the reduct candidate *R:*

$$K_{\{b,c\}}(\{d\}) = \frac{|\frac{U}{IND(b,c)}|}{|\frac{U}{IND(\{b,c,d\})}|}$$

$$= \frac{|\{\{1\ 2\}\{3\}\{4\}\{5\}\{6\}\{7\}\}|}{|\{\{1\}\{2\}\{3\}\{4\}\{5\}\{6\}\{7\}\}|} = \frac{6}{7}$$

Again, the relative dependency does not evaluate to 1. Hence attribute *d* is retained in the reduct candidate and the current reduct candidate *R={b, c, d}*. As there are no further attributes to consider, the algorithm terminates and outputs the reduct *{b, c, d}*.

## 5. Classification

The classifier tool WEKA [11] an open source java based machine –learning workbench that can be run on any computer that has a java run time environment installed. It brings together many machine learning algorithm and tools under a common frame work. The WEKA is a well known package of data mining tools which provides a variety of known, well maintained classifying algorithms. This allows us to experiment with several kinds of classifiers quickly and easily. The tool is used to perform benchmark experiment. Four classifier learners were employed for the classification of the data, DTNB, JRip, J48, and LMT**.**

## 6. Experimental Results

This section presents the results of experimental studies using both crisp-valued and real-valued data sets. The USRR method is compared with the SRR method. All data sets have been obtained from the UCI Repository Machine Learning Database [12]. A comparison of the USRR method, and SRR method is made based on the subset size, time taken to discover subsets, and classification accuracy. A short experimental evaluation for 7 benchmark datasets is presented. The information of the data sets contains names of dataset, number of objects,

number of classes and number of attributes which are given in Table 2:

Table.2 Dataset Information

| Index | Data Set | Objects | Class | Attr_size |
|---|---|---|---|---|
| **1** | Iris | 150 | 3 | 4 |
| **2** | WBCD | 699 | 2 | 9 |
| **3** | Car | 1728 | 4 | 6 |
| **4** | ECOLI | 336 | 8 | 7 |
| **5** | Heart_s | 270 | 2 | 15 |
| **6** | BUPALiver | 345 | 2 | 6 |
| **7** | PimaIn Diabetes | 768 | 2 | 8 |

### *5.1 Feature Selection by SRR and USRR*

The features are reduced by the the Supervised Relative Reduct (SRR) Algorithm and the Unsupervised Relative Reduct(USRR) Algorithm and time(in seconds) taken to find reduct are tabulated in Table 4

Table 4. The Features Selected by SRR Algorithm and USRR Algorithm.

| Index | Algorithm's Selected Features | | Runtime(S) | |
|---|---|---|---|---|
| | SRR | USRR | SRR | USRR |
| **1** | (2,3,4) | (1, 2, 3, 4) | 0.8602 | 2.1234 |
| **2** | (5, 6, 7, 8, 9) | (1, 2, 3, 4, 5, 6, 7, 8, 9) | 4.4469 | 13.1817 |
| **3** | (1, 2, 3, 4, 5, 6) | (1, 2, 3, 4, 5, 6) | 1.9830 | 1.8431 |
| **4** | (2, 5, 7) | (2, 5, 7) | 16.3145 | 21.8152 |
| **5** | (6, 11, 12, 13) | (6, 9, 11, 12) | 23.3529 | 36.2886 |
| **6** | (3, 4, 5) | (2, 4, 5) | 9.3458 | 17.7308 |
| **7** | (6, 7, 8) | (6, 7, 8) | 150.9474 | 201.6087 |

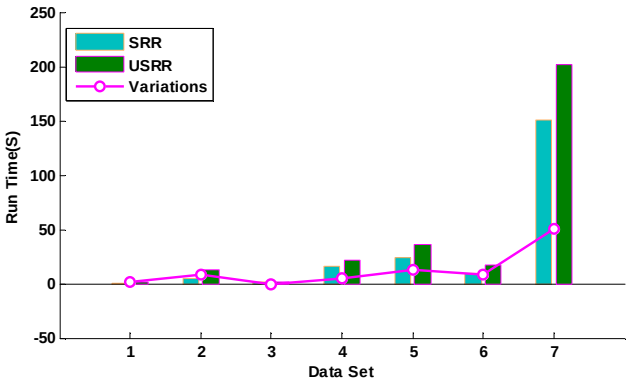



Fig. 3 Runtime for SRR vs USRR

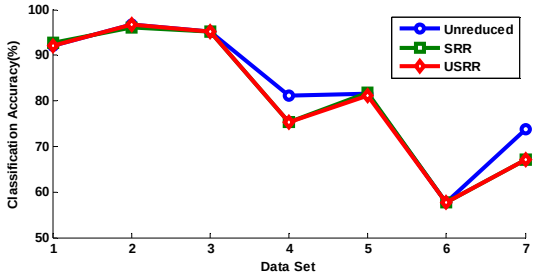
In an attempt to compare the complexity of both the SRR and USRR methods from an application viewpoint, both FS methods were applied to the data sets and the time taken to find a reduct was recorded in each case. The results show that there is only a marginal increase in runtime for the USRR Method. There is even a decrease in car dataset. However, Fig. 3 demonstrates that for increased dimensionality there is little overall difference in runtime between the methods.

### *5.2 Comparison of SRR and USRR*

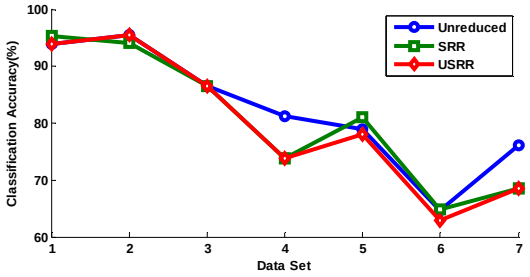
In this section, the USRR method is compared with the SRR method. The classification was initially performed on the unreduced data set, followed by the reduced data sets which were obtained by using the SRR and USRR dimensionality reduction techniques, respectively. Results are presented both in terms of classification accuracy and classification mean absolute error. The data presented in Table 5 and Table 6 shows the classification accuracy values and classification mean absolute error values respectively.

Table 5. Classification Accuracy Values

| Index | Unreduced data | | | | Reduced data by SRR | | | | Reduced data by USRR | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | DTNB | JRip | J48 | LMT | DTNB | JRip | J48 | LMT | DTNB | JRip | J48 | LMT |
| **1** | 92.00 | 94.00 | 96.00 | 94.00 | 92.67 | 95.33 | 96.00 | 95.33 | 92.00 | 94.00 | 96.00 | 94.00 |
| **2** | 96.85 | 95.42 | 94.56 | 95.99 | 96.13 | 94.13 | 94.27 | 95.42 | 96.85 | 95.42 | 94.56 | 95.99 |
| **3** | 95.25 | 86.45 | 92.36 | 98.78 | 95.25 | 86.45 | 92.36 | 98.78 | 95.25 | 86.45 | 92.36 | 98.78 |
| **4** | 81.25 | 81.25 | 84.22 | 87.20 | 75.30 | 73.80 | 78.27 | 76.48 | 75.30 | 73.80 | 78.27 | 76.48 |
| **5** | 81.48 | 78.89 | 76.67 | 83.33 | 81.85 | 81.11 | 82.96 | 80.00 | 81.11 | 78.14 | 81.11 | 79.62 |
| **6** | 57.68 | 64.63 | 68.69 | 66.37 | 57.68 | 64.92 | 63.76 | 64.63 | 57.68 | 62.89 | 61.15 | 59.13 |
| **7** | 73.82 | 76.04 | 73.82 | 77.47 | 67.18 | 68.61 | 67.70 | 69.92 | 67.18 | 68.61 | 67.70 | 69.92 |




(a)




(b)

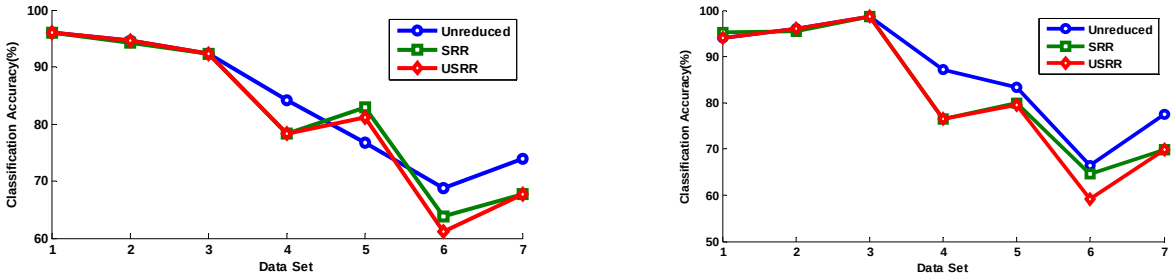



Fig. 4: Classification Accuracy values (a) DTNB classifier (b) JRip classifier (c) J48 classifier (d) LMT classifier

It is interesting to note that where an increase in classification accuracy is recorded for both the USSR and the SRR methods, with respect to the unreduced data in some cases, this increase in classification accuracy is little bit high when comparing both the SRR and the USRR methods to the unreduced data. Also, when comparing classification results, where the USRR method shows a high in classification accuracy, which is demonstrated in Fig. 4.

It should also be noted that a decrease in classification error is recorded for both the USSR and the SRR methods in some cases, with respect to the unreduced data. This is demonstrated in Fig. 5.

Table 6
Classification Mean Absolute Error Values

| Index | Unreduced data | | | | Reduced data by SRR | | | | Reduced data by USRR | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | DTNB | JRip | J48 | LMT | DTNB | JRip | J48 | LMT | DTNB | JRip | J48 | LMT |
| 1 | 0.07 | 0.05 | 0.03 | 0.04 | 0.06 | 0.04 | 0.03 | 0.04 | 0.07 | 0.05 | 0.03 | 0.04 |
| 2 | 0.03 | 0.06 | 0.06 | 0.05 | 0.04 | 0.08 | 0.08 | 0.07 | 0.03 | 0.06 | 0.06 | 0.05 |
| 3 | 0.14 | 0.08 | 0.04 | 0.01 | 0.14 | 0.08 | 0.04 | 0.01 | 0.14 | 0.08 | 0.04 | 0.01 |
| 4 | 0.06 | 0.06 | 0.05 | 0.05 | 0.09 | 0.08 | 0.08 | 0.09 | 0.09 | 0.08 | 0.08 | 0.09 |
| 5 | 0.23 | 0.28 | 0.24 | 0.23 | 0.27 | 0.29 | 0.26 | 0.27 | 0.29 | 0.31 | 0.30 | 0.30 |
| 6 | 0.47 | 0.41 | 0.36 | 0.40 | 0.47 | 0.43 | 0.23 | 0.44 | 0.47 | 0.45 | 0.45 | 0.45 |
| 7 | 0.31 | 0.34 | 0.31 | 0.31 | 0.38 | 0.40 | 0.38 | 0.37 | 0.38 | 0.40 | 0.38 | 0.37 |

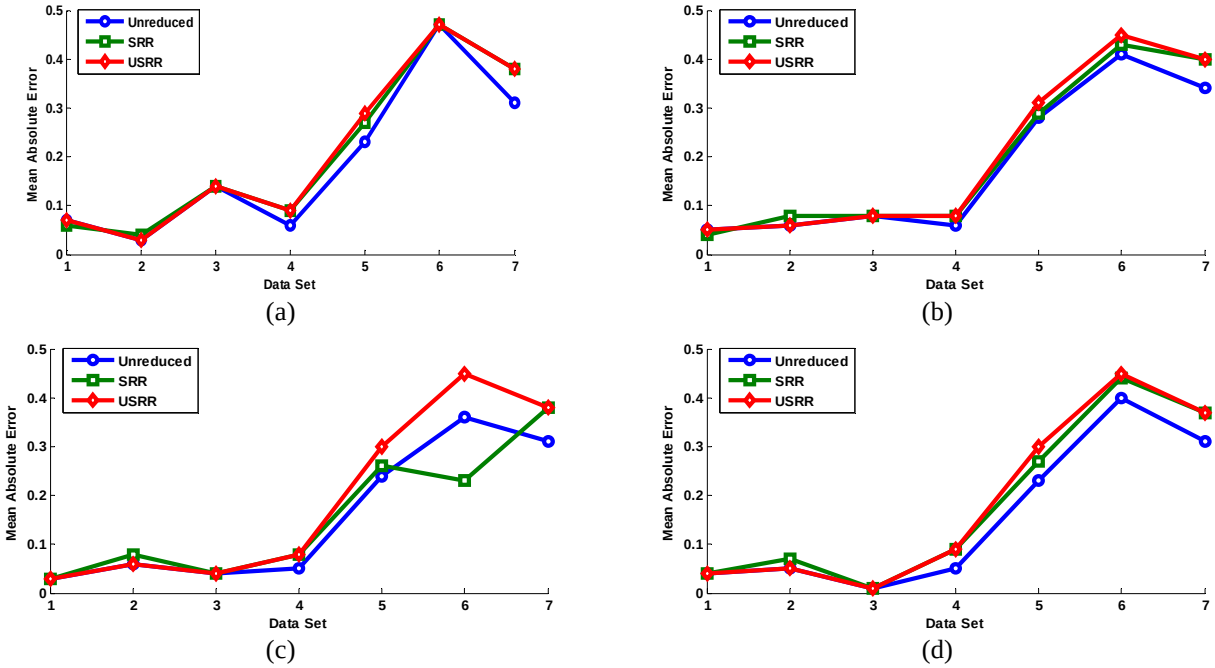



Fig. 5 Classification mean absolute error values (a) DTNB classifier (b) JRip classifier (c) J48 classifier (d) LMT classifier

## 7. Conclusion

In this work, the rough set based unsupervised feature selection method using relative dependency measures is proposed. The method employs a backward elimination-type search to remove features from the complete set of original features. As with the WEKA tool is used to classify the data and the classification performance is evaluated using classification accuracy and mean absolute error, the method is compared with an existing supervised method and it demonstrates that it can effectively remove redundant features. The subsets returned by this unsupervised method are of similar size to that of the supervised method and classification of the reduced data shows that the method selects useful features which are of comparable quality. In future, the same approach can be extended to mammogram image datasets for breast cancer diagnosis.

## Acknowledgement

The first author wholeheartedly thanks the help rendered by the *UGC, SERO, Hyderabad* to carry out this research under FDP during XI plan period.

The second author immensely thanks the *UGC, New Delhi* for financial assistance under major research project grant No. F-34-105/2008.

## References

[1] C. Velayutham, K. Thangavel, "Improved Rough Set Algorithms for Optimal Attribute Reduct", Journal of Electronic Science and Technology (JEST)(International), Vol. 9, No. 2, June 2011., pp 108-117.

[2] R. Roselin, K. Thangavel, C. Velayutham, "Fuzzy Rough Feature Selection for Mammogram Classification", Journal of Electronic Science and Technology (JEST)(International), Vol. 9, No. 2, June 2011., pp 124-132.

[3] C. Velayutham, K. Thangavel, "Unsupervised Quick Reduct Algorithm Using Rough Set Theory", Journal of Electronic Science and Technology (JEST)(International), Vol. 9, No. 3, Sep. 2011., pp 193-201.

[4] Z. Pawlak, Rough Sets: Theoretical aspects of reasoning about data. Kluwer Academic Publishers, Dordrecht,( 1991).

[5] Z. Pawlak, Rough set approach to knowledge-based decision support, European Journal of Operational Research 99 (1997) 48-57.

[6] R. Jensen, Q. Shen, Semantics-Preserving Dimensionality Reduction: Rough and Fuzzy-Rough based Approaches, *IEEE Transactions onKnowledge and Data Engineering* 16(12) (2004) 1457-1471.

[7] K.Thangavel, A. Pethalakshmi, Dimensionality reduction based on Rough set theory: A review, *Applied Soft computing*,9(2009), 1-12.

[8] R. Jensen, Q. Shen, Fuzzy-rough attribute reduction with application to web categorization, *Fuzzy Sets and Systems* 141 (2004) 469-485.

[9] R. Jensen, Combining rough and fuzzy sets for feature selection, *PhD thesis*. Doctor of Philosophy, School of Informatics, University of Edinburgh, (2004).

[10] M. Dash, & H. Liu, Feature Selection for Classification. *Intelligent Data Analysis*, 1(3) (1997), pp. 131-156.

[11] J. R. Quinlan, (1993). C4.5: Programs for Machine Learning. The Morgan Kaufmann Series in Machine Learning. Morgan Kaufmann Publishers, San Mateo, CA.

[12] C. L. Blake and C. J. Merz. UCI Repository of machine learning databases. Irvine, University of California,(1998), http://www.ics.uci.edu/~mlearn/.

**C.Velayutham** was born in 1965 at Thanjavur, Tamilnadu, India. He is received the Master of Science in Applied Mathematics in 1989, and Post Gratuate Diploma in Computer Application in 1990, from Bharathidasan University, Trichy, India. He obtained his M.Phil (Computer Science) Degree from Manonmaniam Sundaranar University, Trunelveli, India in 2002. Currently he is working as Associate Professor, Department of Computer Science, Aditanar College of Arts and Science, Tiruchendur, Tamil Nadu, India. His area of interests includes Medical Image Processing, Data Mining Neural Network, Fuzzy logic, and Rough Set.

**K.Thangavel** was born in 1964 at Namakkal, Tamilnadu, India. He received his Master of Science from the Department of Mathematics, Bharathidasan University in 1986, and Master of Computer Applications Degree from Madurai Kamaraj University, India in 2001. He obtained his Ph.D. Degree from the Department of Mathematics, Gandhigram Rural Institute-Deemed University, Gandhigram, India in 1999. Currently he is working as Professor and Head, Department of Computer Science, Periyar University, Salem. He is a recipient of Tamilnadu Scientist award for the year 2009. His area of interests includes Medical Image Processing, Artificial Intelligence, Neural Network, Fuzzy logic, Data Mining and Rough Set.