

Soft-Matrices Computation for Feature Selection on Traditional Malay Musical Instrument Sounds Dataset

Sarjon Defit Tutut Herawan

*Universitas Putra Indonesia "YPTK" Padang
Jl. Raya Lubuk Begalung Padang, Sumatera Barat, Indonesia
sarjond@yahoo.co.uk*

²*Ahmad Dahlan University, Yogyakarta, Indonesia
AMCS Research Center, Yogyakarta, Indonesia
tutut@uad.ac.id*

Abstract

Soft set theory proposed by Molodtsov is a new mathematical tool for dealing with the uncertain data. Soft set-based dimensionality reduction can be considered as a technique for feature selection. In this paper, we present an alternative technique for feature selection of traditional Malay musical instrument sounds dataset. The technique is based on matrices computation of multi-soft sets for dimensionality reduction. The dataset is taken from our previous work which consists of 1116 instances and 37 attributes of Malay Musical Instrument Sounds Dataset. The result shows that the obtained features of proposed technique are 35 out of 37 features.

Keywords: *Soft set theory, Feature selection, Traditional Malay musical instruments sound dataset*

1. Introduction

In making sure data made available for the process of decision making is sufficient enough, then reduction of data ought to be properly addressed. The main objective of reduction is to lessen the number of parameters, and at the same time, attaining the property of information in helping the process of decision making. The theory of soft set [1] proposed by Molodtsov 1999 is a new method for handling uncertain data. Soft sets are called elementary neighborhood systems. Molodtsov pointed out that one of the main advantages of soft set theory is that it is free from the inadequacy of the parameterization tools, like in the theories of fuzzy set, probability and interval mathematics [1]. As for standard soft set, it may be redefined as the classification of objects in two distinct classes (yes/1 and no/0), thus confirming that soft set can deal with a Boolean-valued information system. One of the potential applications of soft set theory is to be used as the dimensionality reduction and decision making method [2–15]. However, most of the current applications are deal with the Boolean information system [2-7, 12, 13, 15]. To date, less effort has been done in handling the non-Boolean dataset especially for the categorical dataset. In previous papers, Herawan and Mustafa proposed the idea of multi-soft sets for representing a categorical-valued information system [16]. Sub sequentially, Herawan, *et al.*, proposed the concept of matrices representation of multi-soft sets and employ such matrices for finding reducts in a categorical-valued information system [17].

In this paper, we employ our early works in [17] as a feature selection technique. It is based on matrices computation of multi-soft sets and further applied it in handling Traditional Malay musical instruments sounds dataset. To accomplish this study, the dataset consists of 1116 instances with 37 attributes from two categories of features schemes which are perception-based and MFCC is exploited. In order to utilize the soft set theory, it is essential to transform the original dataset (continuous values with non-

categorical form) into discretized values with categorical form. For that, equal width and equal frequency binning algorithm in [18] is employed as the discretization method. Afterwards, the matrices computation of multi-soft sets method for soft dimensionality reduction in multi-valued information system in [17] is adopted to select the best feature set from the large number of features available.

The rest of this paper is organized as follows: Section 2 discuss the theory of soft set and multi-soft sets will be explained in Section 3. Section 4 describes the modeling process of this study. A discussion of the result can be found in Section 5 followed by the conclusion in Section 6.

2. Soft Set Theory

Presently, research on soft sets has been very active and many important results have been achieved in the theoretical aspect.

2.1. Information System

Data are often presented as a table, columns of which are labeled by *attributes*, rows by *objects* of interest and entries of the table are *attribute values*. By an *information system*, we mean a 4-tuple (quadruple) $S = (U, A, V, f)$, where U is a non-empty finite set of objects, A is a non-empty finite set of attributes, $V = \bigcup_{a \in A} V_a$, V_a is the domain (value set) of attribute a , $f : U \times A \rightarrow V$ is a total function such that $f(u, a) \in V_a$, for every $(u, a) \in U \times A$, called information (knowledge) function. An information system is also called a knowledge representation systems or an attribute-valued system and can be intuitively expressed in terms of an information table (refer to Table 1). In many applications, there is an outcome of classification that is known. This *a posteriori* knowledge is expressed by one (or more) distinguished attribute called decision attribute; the process is known as supervised learning. An information system of this kind is called a decision system, $S = (U, A \cup \{d\}, V, f)$.

Definition 1. Let $S = (U, A, V_{\{0,1\}}, f)$ Abe an information system. If $V_a = \{0,1\}$, for every $a \in A$, then $S = (U, A, V_{\{0,1\}}, f)$ is called a Boolean-valued information system.

Table 1. An Information System

U/A	a_1	a_2	...	a_k	...	$a_{ A }$
u_1	$f(u_1, a_1)$	$f(u_1, a_2)$...	$f(u_1, a_k)$...	$f(u_1, a_{ A })$
u_2	$f(u_2, a_1)$	$f(u_2, a_2)$...	$f(u_2, a_k)$...	$f(u_2, a_{ A })$
u_3	$f(u_3, a_1)$	$f(u_3, a_2)$...	$f(u_3, a_k)$...	$f(u_3, a_{ A })$
\vdots	\vdots	\vdots	\ddots	\vdots	\ddots	\vdots
$u_{ U }$	$f(u_{ U }, a_1)$	$f(u_{ U }, a_2)$...	$f(u_{ U }, a_k)$...	$f(u_{ U }, a_{ A })$

2.2. Soft Set

Throughout this section U refers to an initial universe, E is a set of parameters in relation to objects in U , $P(U)$ is the power set of U .

Definition 2. (See [1].) A pair (F, E) is called a soft set over U , where F is a mapping given by $F : E \rightarrow P(U)$.

In other words, a soft set over U is a parameterized family of subsets of the universe U . For $e \in E$, $F(e)$ may be considered as a set of e -elements of the soft set (F, E) or as the set of e -approximate elements of the soft set. Clearly, a soft set is not a (crisp) set. Based on the definition of an information system and a soft set, in this section we show that a soft set is a special type of information systems, *i.e.*, a Boolean-valued information system.

Example 1. Based on [13], let consider a soft set (F, E) which describes the “conditions of patients suspected influenza” that a hospital is considering to make a decision. The six influenza symptoms, *i.e.*, fever, respiratory, nasal discharges, cough, headache and lethargic. Suppose that there are six patients in the hospital under consideration, $U = \{p_1, p_2, p_3, p_4, p_5, p_6\}$, and E is a set of decision parameters $E = \{e_1, e_2, e_3, e_4, e_5, e_6, e_7\}$, where e_1 is the parameter “fever”, e_2 is the parameter “respiratory”, e_3 is the parameter “nasal discharges”, e_4 is the parameter “cough”, e_5 is the parameter “headache”, e_6 is the parameter “sore throat”, e_7 is the parameter “lethargic”

Consider the mapping $F : E \rightarrow P(U)$ given by “patients(\cdot)”, where (\cdot) is to be filled in by one of parameters $e \in E$. Suppose that $F(e_1) = \{p_1, p_3, p_4, p_5, p_6\}$, $F(e_2) = \{p_1, p_2\}$, $F(e_3) = \{p_1, p_2, p_4\}$, $F(e_4) = \{p_1\}$, $F(e_5) = \{p_3, p_4\}$, $F(e_6) = \{p_2, p_4\}$, and $F(e_7) = \{p_1, p_3, p_5, p_6\}$. Therefore, $F(e_1)$ means “patients suffer fever”, whose functional value is the set $\{p_1, p_3, p_4, p_5, p_6\}$. Thus, we can view the soft set (F, E) as a collection of approximations as below:

$$(F, E) = \left\{ \begin{array}{l} \text{fever} = \{p_1, p_3, p_4, p_5, p_6\}, \\ \text{respiratory} = \{p_1, p_2\}, \\ \text{nasal discharges} = \{p_1, p_2, p_4\}, \\ \text{cough} = \{p_1\}, \\ \text{head ache} = \{p_3, p_4\}, \\ \text{sore throat} = \{p_2, p_4\}, \\ \text{lethargic} = \{p_1, p_3, p_5, p_6\} \end{array} \right\}$$

The relation between a soft set and a Boolean-valued information system is given in the following proposition.

Proposition 1. If (F, A) is a soft set over the universe U , then (F, A) is a Boolean-valued information system $S = (U, A, V_{\{0,1\}}, f)$.

Proof. Let (F, A) be a soft set over the universe U , we define a mapping $F = \{f_1, f_2, \dots, f_n\}$, where $f_i : U \rightarrow V_i$ and $f_i(x) = \begin{cases} 1, & x \in F(a_i), \text{ for } 1 \leq i \leq |A| \\ 0, & x \notin F(a_i) \end{cases}$.

Hence, if $V = \bigcup_{a_i \in A} V_{a_i}$, where $V_{a_i} = \{0,1\}$, then a soft set (F, A) can be considered as a Boolean-valued information system $S = (U, A, V_{\{0,1\}}, f)$.

From Proposition 1, it is easily to understand that a Boolean-valued information system can be represented as a soft set. Thus, we can make a one-to-one correspondence between (F, E) over U and $S = (U, A, V_{\{0,1\}}, f)$. The tabular representation of the above soft set is given in the Table 2. From Table 2, each approximation has two parts, a predicate e and an approximate value set p . For example, for the approximation “fever = $\{p_1, p_3, p_4, p_5, p_6\}$ ”, we have the predicate name of patients with fever and its value set is $\{p_1, p_3, p_4, p_5, p_6\}$.

Table 2. Conditions of Patients Suspected Influenza

U/E	e_1	e_2	e_3	e_4	e_5	e_6	e_7
p_1	1	1	1	1	0	0	1
p_2	0	1	1	0	0	1	0
p_3	1	0	0	0	1	0	1
p_4	1	0	1	0	1	1	0
p_5	1	0	0	0	0	0	1
p_6	1	0	0	0	0	0	1

Definition 3. (See [19].) *The class of all value sets of a soft set (F, E) is called value-class of the soft set and is denoted by $C_{(F,E)}$.*

3. Method and Main Result

The “standard” soft set deals with a Boolean-valued information system. For a multi-valued information system $S = (U, A, V, f)$, where $V = \bigcup_{a \in A} V_a$, V_a is the domain (value set) of attribute a which has multi value, a decomposition can be made from S into $|A|$ number of Boolean-valued information systems $S = (U, A, V_{\{0,1\}}, f)$. In this sub-section, we present the construction of *multi-soft sets* representing $S = (U, A, V, f)$ based on such decomposition [16].

3.1. Multi-soft Sets

In the following sub-section, we propose an idea of decomposition of a multi-valued information system $S = (U, A, V, f)$ into $|A|$ -numbers of Boolean-valued information system $S^i = (U, a_i, V_{\{0,1\}}, f)$, where $|A|$ is the cardinality of A .

3.2. Decomposition of a Multi-valued Information System

The decomposition of $S = (U, A, V, f)$ is based on decomposition of $A = \{a_1, a_2, \dots, a_{|A|}\}$ into the disjoint-singleton attribute $\{a_1\}, \{a_2\}, \dots, \{a_{|A|}\}$. Here, we only consider for complete information systems. Let $S = (U, A, V, f)$ be an information system such that for every $a \in A$, $V_a = f(U, A)$ is a finite non-empty set and for every $u \in U$, $|f(u, a)| = 1$. For every a_i under i^{th} -attribute consideration, $a_i \in A$ and $v \in V_a$, we define the map $a_v^i : U \rightarrow \{0,1\}$ such that $a_v^i(u) = 1$ if $f(u, a) = v$, otherwise

$a_v^i(u) = 0$. The next result, we define a Boolean-valued information system as a quadruple $S^i = (U, a_i, V_{\{0,1\}}, f)$. The information systems $S^i = (U, a_i, V_{\{0,1\}}, f)$, $1 \leq i \leq |A|$ is referred to as a decomposition of a multi-valued information system $S = (U, A, V, f)$ into $|A|$ Boolean-valued information systems, as depicted in Figure 1.

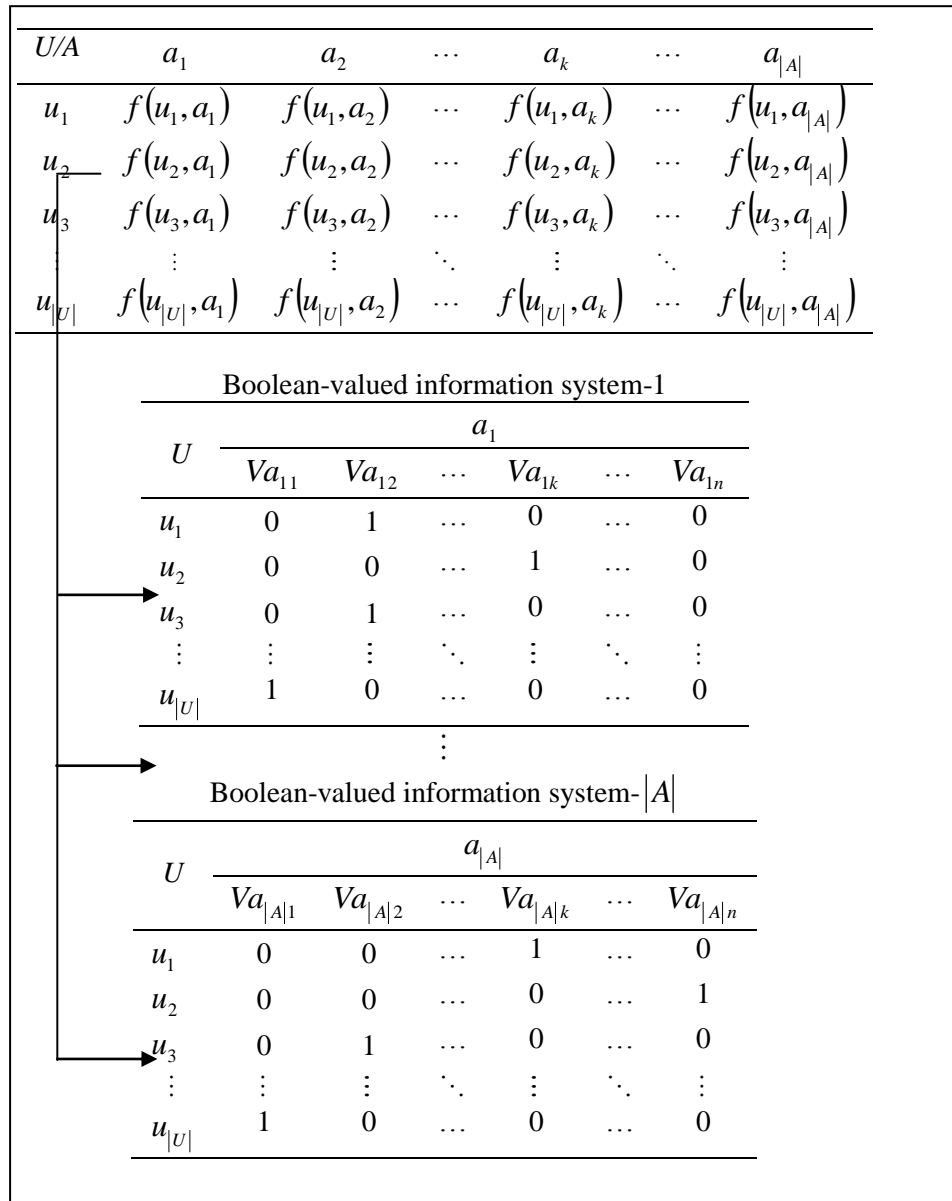


Figure 1. A Decomposition of Information Systems

Every information system $S^i = (U, a_i, V_{a_i}, f)$, $1 \leq i \leq |A|$ is a deterministic information system since for every $a \in A$ and for every $u \in U$, $|f(u, a)| = 1$ such that the structure of a multi-valued information system and $|A|$ number of Boolean-valued information systems give the same value of attribute related to objects.

3.3. Multi-soft Sets Construction

The idea of multi-soft sets is based on a decomposition of a multi-valued information system $S = (U, A, V, f)$, into $|A|$ number of Boolean-valued information systems $S = (U, A, V_{\{0,1\}}, f)$, where $|A|$ denotes the cardinality of A . Consequently, the $|A|$ Boolean-valued information systems define *multi-soft sets* $(F, A) = \{(F, a_i) : 1 \leq i \leq |A|\}$.

Based on the notion of a decomposition of a multi-valued information system in the previous section, in this sub-section we present the notion of multi-soft set representing multi-valued information systems. Let $S = (U, A, V, f)$ be a multi-valued information system and $S^i = (U, a_i, V_{a_i}, f)$, $1 \leq i \leq |A|$ be the $|A|$ Boolean-valued information systems. From Proposition 1, we have

$$S = (U, A, V, f) = \begin{cases} S^1 = (U, a_1, V_{\{0,1\}}, f) & \Leftrightarrow (F, a_1) \\ S^2 = (U, a_2, V_{\{0,1\}}, f) & \Leftrightarrow (F, a_2) \\ \vdots & \vdots \\ S^{|A|} = (U, a_{|A|}, V_{\{0,1\}}, f) & \Leftrightarrow (F, a_{|A|}) \end{cases} \\ = ((F, a_1), (F, a_2), \dots, (F, a_{|A|}))$$

We define $(F, A) = ((F, a_1), (F, a_2), \dots, (F, a_{|A|}))$ as a *multi-soft sets* over universe U representing a multi-valued information system $S = (U, A, V, f)$.

3.4. Matrices Representation of Multi-Soft Sets

The concept of matrix representing multi-soft sets is described here. In this sub-section, the notation (F, A) represents a multi-soft set over universe U of a multi-valued information system $S = (U, A, V, f)$.

$$(F, A) \Leftrightarrow \begin{cases} (F, a_1) \Leftrightarrow S^1 = (U, a_1, V_{\{0,1\}}, f) \Leftrightarrow M_{a_1} \\ (F, a_2) \Leftrightarrow S^2 = (U, a_2, V_{\{0,1\}}, f) \Leftrightarrow M_{a_2} \\ \vdots \\ (F, a_{|A|}) \Leftrightarrow S^{|A|} = (U, a_{|A|}, V_{\{0,1\}}, f) \Leftrightarrow M_{|A|} \end{cases}$$

Definition 4. Matrix M_{a_i} , $1 \leq i \leq |A|$ is called matrix representation of soft set (F, a_i) over universe U . The dimension of matrices is defined by $\dim(M_{a_i}) = |U| \times |Va_i|$. All entries of $M_{a_i} = [a_{ij}]$ is belong to a set $\{0,1\}$, where

$$a_{ij} = \begin{cases} 0, & \text{if } |f(u, \alpha)| = 0, \quad 1 \leq i \leq |U|, \quad 1 \leq j \leq |Va_i|, \quad u \in U \text{ and } \alpha \in Va_i. \\ 1, & \text{if } |f(u, \alpha)| = 1 \end{cases}$$

The collection of all matrices representing (F, A) is denoted by \mathcal{M}_A , i.e.,

$$\mathcal{M}_A = \{M_{a_i} : 1 \leq i \leq |A|\}.$$

Definition 5. Let $M_{a_i} \in \mathcal{M}_A$ be a matrix representation of a soft set (F, a_i) over U . The value-class of M_{a_i} , i.e., class of all value sets of M_{a_i} , denoted $C_{M_{a_i}}$ is defined by

$$C_{M_{a_i}} = \left\{ \{u : |f(u, \alpha_1)| = 1\}, \dots, \{u : |f(u, \alpha_{|Va_i|})| = 1\} \right\},$$

where $1 \leq i \leq |Va_i|$, $u \in U$ and $\alpha \in Va_i$.

Clearly $C_{M_{a_i}} \subseteq P(U)$.

3.5. AND and OR Operations in a Collection of Matrices

The AND and OR operation of the matrices are described in this sub-section.

Definition 6. Let $M_{a_i} = [a_{kl}]$, $1 \leq k \leq |U|$, for $1 \leq l \leq |Va_i|$ and $M_{a_j} = [a_{mn}]$, for $1 \leq m \leq |U|$, $1 \leq n \leq |Va_j|$ be two matrices in \mathcal{M}_A . The "AND" operation $M_{a_i} \text{ AND } M_{a_j}$ of matrices M_{a_i} and M_{a_j} is defined as follows

$$M_{a_i} \text{ AND } M_{a_j} = M_{a_{ij}} = [a_{pq}] \text{ with } \dim(M_{a_{ij}}) = |U| \times (|Va_i| \times |Va_j|),$$

where

$$a_{p1} = \min\{a_{k1}, a_{m1}\}, a_{p2} = \min\{a_{k2}, a_{m2}\}, \dots, a_{p(|Va_i| \times |Va_j|)} = \min\{a_{k|Va_i|}, a_{m|Va_j|}\}.$$

Definition 6. Let $M_{a_i} = [a_{kl}]$ and $M_{a_j} = [b_{mn}]$ be two matrices in \mathcal{M}_A . The "OR" operation $M_{a_i} \text{ OR } M_{a_j}$ of matrices M_{a_i} and M_{a_j} is defined as follows

$$M_{a_i} \text{ OR } M_{a_j} = M_{a_{ij}} = [c_{pq}] \text{ with } \dim(M_{a_{ij}}) = |U| \times (|Va_i| \times |Va_j|),$$

where

$$a_{p1} = \max\{a_{k1}, a_{m1}\}, a_{p2} = \max\{a_{k2}, a_{m2}\}, \dots, a_{p(|Va_i| \times |Va_j|)} = \max\{a_{k|Va_i|}, a_{m|Va_j|}\}.$$

3.6. Reduction Method

A *reduct* is a minimal set of attributes that preserve the classification. In [17], we presented the application of multi-soft sets for finding reducts. The proposed approach is based on AND operation in multi-soft sets [16].

Definition 7. Let $(F, A) = ((F, a_i): 1 \leq i \leq |A|)$ be multi-soft set over U representing a multi-valued information system $S = (U, A, V, f)$. A set of attributes $B \subseteq A$ is called a *reduct* for A if

$$C_{F(b_1 \times \dots \times b_B)} = C_{F(a_1 \times \dots \times a_A)} \text{ and } C_{F(b_1 \times \dots \times b_{B^*})} \neq C_{F(a_1 \times \dots \times a_A)}, \quad B^* \subset B,$$

where

$$(F, a_i \times a_j) = (F, a_i) \text{ AND } (F, a_j).$$

In [17], we explore the applicability of matrices representation for finding reducts. We show that the reducts obtained are equivalent with that in [20].

Proposition 2. Let \mathcal{M}_A be a matrices over U representing multi-soft sets (F, A) . A set of attributes $B \subseteq A$ is a *reduct* for A if only if $C_{\text{AND}_{b_i: b \in B}} = C_{\text{AND}_{a_i: a \in B}}$.

Proof. It is clear, since every (F, a_i) can be represented as M_{a_i} , $1 \leq i \leq |A|$ and vice versa.

Example 2. Let given a small dataset as follow.

Table 3. An Information System

U/A	a_1	a_2	a_3	a_4
1	low	bad	Loss	small
2	low	good	Loss	large
3	high	good	Loss	medium
4	high	good	Loss	medium
5	low	good	Profit	large

Let $A = \{a_1, a_2, a_3, a_4\}$, the multi-soft sets representing Table 3 is given as

$$(F, A) = \left(\left\{ \left\{ \text{low} = \{1, 2, 5\}, \text{high} = \{3, 4\} \right\}, \left\{ \text{bad} = \{1\}, \text{good} = \{2, 3, 4, 5\} \right\}, \left\{ \text{loss} = \{1, 2, 3, 4\}, \text{profit} = \{5\} \right\}, \left\{ \text{small} = \{1\}, \text{large} = \{2, 5\}, \text{medium} = \{3, 4\} \right\} \right), \right)$$

and the matrices representing (F, A) is $\mathcal{M}_A = \{M_{a_1}, M_{a_2}, M_{a_3}, M_{a_4}\}$, where

$$M_{a_1} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}, M_{a_2} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, M_{a_3} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \text{ and } M_{a_4} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Since $A = \{a_1, a_2, a_3, a_4\}$, then

$$M_{AND_{a_i:1 \leq i \leq 4}} = \begin{bmatrix} 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where

$$C_{M_{AND_{a_i:1 \leq i \leq 4}}} = \{\{1\}, \{2\}, \{3,4\}, \{5\}\}.$$

Let, given $B = \{a_1, a_2, a_3\}$ and $C = \{a_3, a_4\}$, then we have

$$M_{a_1} \text{ AND } M_{a_2} \text{ AND } M_{a_3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$M_{a_3} \text{ AND } M_{a_4} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

where

$$C_{M_{a_1} \text{ AND } M_{a_2} \text{ AND } M_{a_3}} = \{\{1\}, \{2\}, \{3,4\}, \{5\}\} \quad \text{and} \quad C_{M_{a_1} \text{ AND } M_{a_2} \text{ AND } M_{a_3}} = \{\{1\}, \{2\}, \{3,4\}, \{5\}\},$$

respectively. Thus, we have $\{a_1, a_2, a_3\}$ and $\{a_3, a_4\}$ are reducts of A. Furthermore, core is

$$\{a_1, a_2, a_3\} \cap \{a_3, a_4\} = \{a_3\}.$$

Thus a reduct is a set of attributes that preserves classes. It means that a reduct is the minimal subset of attributes that enables the same classification of elements of the universe as the whole set of attributes. In other words, attributes that do not belong to a reduct are superfluous with regard to classification of elements of the universe.

3.7. Dataset

The second part of this article starts from a description of the traditional Malay musical instrument sounds dataset and the feature selection problem for which the data was collected. This work will identify which attributes are dispensable and further can be removed from dataset without affected the original classification power. Then, we make a decision of the reduced dataset based on the attributes dispensed. From this experiment, the ideal dataset can be represented in decision system form as $S = \{U, A \cup \{d\}, V, f\}$. There are 1116 instances in the universe U , with 37 attributes of A as the set of condition attributes (as shown in Table 4) and the family of the instruments as the decision attribute d . The distribution of all instances in each class is uniform with no missing values in the data [21].

Table 4. Features Descriptions

Number	Description
1	Zero Crossing
2-3	Mean and Standard Deviation of Zero Crossings Rate
4-5	Mean and Standard Deviation of Root-Mean-Square
6-7	Mean and Standard Deviation of Spectral Centroid
8-9	Mean and Standard Deviation of Bandwidth
10-11	Mean and Standard Deviation of Flux
12-37	Mean and Standard Deviation of the First 13 MFCCs

The mean and standard deviation were then calculated for each of these features. The problem of feature selection through the above dataset is how to delete irrelevant attributes. It can be done, among all attributes we want to remove dispensable attributes. To select the candidate of the dispensable attributes, we employ the method proposed by [17].

3.8. Result and Discussion

The decision table used in this study is in the form of continuous value with non-categorical features (attributes). In order to employ the proposed method, it is essential to transform the dataset into categorical ones. For that, the equal width binning in [18] is used and the dataset is first discretized into 3 categorical values as shown in Table 5.

Table 5. Summary of Data Distribution

Attribute	Description	#1	#2	#3
ZC	Zero Crossing	456	438	222
MEANZCR	Mean of Zero Crossings Rate	134	638	344
STDZCR	Standard Deviation of Zero Crossings Rate	1092	16	8
MEANRMS	Mean of Root-Mean-Square	724	386	6
STDRMS	Standard Deviation of Root-Mean-Square	350	498	268
MEANC	Mean of Spectral Centroid	198	528	390
STDC	Standard Deviation of Spectral Centroid	166	430	520
MEANB	Mean of Bandwidth	206	336	574
STDB	Standard Deviation of Bandwidth	480	474	162
MEANFLUX	Mean of Flux	484	494	138

STDFLUX	Standard Deviation of Flux	692	304	120
MMFCC1	Mean of the MFCCs #1	1116	0	0
MMFCC2	Mean of the MFCCs #2	984	96	36
MMFCC3	Mean of the MFCCs #3	302	516	298
MMFCC4	Mean of the MFCCs #4	692	352	72
MMFCC5	Mean of the MFCCs #5	596	392	128
MMFCC6	Mean of the MFCCs #6	822	246	48
MMFCC7	Mean of the MFCCs #7	840	222	54
MMFCC8	Mean of the MFCCs #8	924	142	50
MMFCC9	Mean of the MFCCs #9	730	380	6
MMFCC10	Mean of the MFCCs #10	938	136	42
MMFCC11	Mean of the MFCCs #11	950	122	44
MMFCC12	Mean of the MFCCs #12	950	144	22
MMFCC13	Mean of the MFCCs #13	1010	42	64
SMFCC1	Standard Deviation of the MFCCs #1	1116	0	0
SMFCC2	Standard Deviation of the MFCCs #2	856	218	42
SMFCC3	Standard Deviation of the MFCCs #3	772	196	148
SMFCC4	Standard Deviation of the MFCCs #4	778	256	82
SMFCC5	Standard Deviation of the MFCCs #5	724	264	128
SMFCC6	Standard Deviation of the MFCCs #6	786	176	154
SMFCC7	Standard Deviation of the MFCCs #7	846	172	98
SMFCC8	Standard Deviation of the MFCCs #8	864	162	90
SMFCC9	Standard Deviation of the MFCCs #9	808	220	88
SMFCC10	Standard Deviation of the MFCCs #10	752	292	72
SMFCC11	Standard Deviation of the MFCCs #11	898	156	62
SMFCC12	Standard Deviation of the MFCCs #12	986	98	32
MFCC13	Standard Deviation of the MFCCs #13	840	148	128

To find the reducts (relevant features), the proposed model is implemented in MATLAB version 7.6.0.324 (R2008a). It is executed on a processor Intel Core 2 Duo CPUs. The total main memory is 1 gigabyte and the operating system is Windows XP Professional SP3. The results of the selected features are presented in Table 6 as follow:

Table 6. The Selected Attributes

Attribute	Description
ZC	Zero Crossing
MEANZCR	Mean of Zero Crossings Rate
STDZCR	Standard Deviation of Zero Crossings Rate
MEANRMS	Mean of Root-Mean-Square
STDRMS	Standard Deviation of Root-Mean-Square
MEANC	Mean of Spectral Centroid
STDC	Standard Deviation of Spectral Centroid
MEANB	Mean of Bandwidth
STDB	Standard Deviation of Bandwidth
MEANFLUX	Mean of Flux
STDFLUX	Standard Deviation of Flux
MMFCC1	Mean of the MFCCs #1
MMFCC2	Mean of the MFCCs #2
MMFCC3	Mean of the MFCCs #3
MMFCC4	Mean of the MFCCs #4
MMFCC5	Mean of the MFCCs #5

MMFCC6	Mean of the MFCCs #6
MMFCC7	Mean of the MFCCs #7
MMFCC8	Mean of the MFCCs #8
MMFCC9	Mean of the MFCCs #9
MMFCC10	Mean of the MFCCs #10
MMFCC11	Mean of the MFCCs #11
MMFCC12	Mean of the MFCCs #12
MMFCC13	Mean of the MFCCs #13
SMFCC1	Standard Deviation of the MFCCs #1
SMFCC2	Standard Deviation of the MFCCs #2
SMFCC3	Standard Deviation of the MFCCs #3
SMFCC4	Standard Deviation of the MFCCs #4
SMFCC5	Standard Deviation of the MFCCs #5
SMFCC6	Standard Deviation of the MFCCs #6
SMFCC7	Standard Deviation of the MFCCs #7
SMFCC8	Standard Deviation of the MFCCs #8
SMFCC9	Standard Deviation of the MFCCs #9
SMFCC10	Standard Deviation of the MFCCs #10
SMFCC11	Standard Deviation of the MFCCs #11
SMFCC12	Standard Deviation of the MFCCs #12
MFCC13	Standard Deviation of the MFCCs #13

From Table 6, the result shows that the proposed model successfully generates the best features set with 35 attributes from 37 of original full attributes. It is found that {MMFCC1, SMFCC1} is the dispensable set of attributes. Thus, the relevant features (reduct) is $A - \{MMFCC1, SMFCC1\}$.

4. Conclusion

In this paper, the matrices computation of multi-soft sets has been used as feature selection (attribute reduction) technique to identify the best features of Traditional Malay musical instrument sounds. To perform this task, two categories of features schemes which are perception-based and MFCC which consist of 37 attributes are extracted. The equal width and equal frequency binning discretization technique is then employed to transform this continuous-value with non-categorical features (attributes) into categorical form. Afterward, the multi-soft sets approach is adopted for feature selection through dimensionality reduction in multi-valued domain. It is found that 35 best features are successfully generated by the proposed model. For that, our future work will examine others feature selection technique and examine the effectiveness of the selected features towards the classification performance.

Acknowledgement

This work is supported by UPI-YPTK Research Grant ref number 065/UPI-YPTK/RG/II/2014.

References

- [1] D. Molodtsov, "Soft Set Theory-First Results", Computer and Mathematics with Applications, vol. 37, (1999), pp. 19–31.
- [2] M. Xiuqin, N. Sulaiman, Q. Hongwu and T. Herawan, "A Novel Normal Parameter Reduction Algorithm of Soft Sets", In Gaspar Cunha et al. (Eds.): AISC 2011, Advances in Intelligent and Soft Computing, vol. 96, (2011), pp. 221–229.

- [3] A. N. M. Rose, H. Hassan, M. I. Awang, T. Herawan and M. M. Deris, "Solving Incomplete Datasets in Soft Set Using Parity Bits of Supported Sets", In T. H. Kim, *et al.*, (Eds.): UCMA 2011, Communication of Computer and Information Sciences, vol. 151, (2011), pp. 33–43.
- [4] M. Xiuqin, N. Sulaiman, Q. Hongwu, J. M. Zain and T. Herawan, "A New Efficient Normal Parameter Reduction Algorithm of Soft Sets", *Computer and Mathematics with Applications*, vol. 62, no. 2, (2011), pp. 588–598.
- [5] T. Herawan and M. M. Deris, "A soft set approach for association rules mining", *Knowledge Based Systems*, vol. 24, no. 1, (2011), pp. 186–195.
- [6] A. N. M. Rose, M. I. Awang, T. Herawan and M. M. Deris, "Hybrid Reduction in Soft Set Decision Making", In De-Shuang Huang, *et al.*, (Eds.): ICIC 2011, Lecture Notes in Computer Sciences, vol. 6838, (2011), pp. 108–115.
- [7] Q. Hongwu, M. Xiuqin, T. Herawan and J. M. Zain, "Data filling approach of soft sets under incomplete information", In N. T. Nguyen, C. G. Kim, A. Janiak (Eds.): ACIIDS 2011, Lecture Notes in Computer Science, vol. 6592, (2011), pp. 302–311.
- [8] Q. Hongwu, M. Xiuqin, T. Herawan and J. M. Zain, "An adjustable approach to interval-valued intuitionistic fuzzy soft sets based decision making", In N. T. Nguyen, C. G. Kim, A. Janiak (Eds.): ACIIDS 2011, Lecture Notes in Computer Science, vol. 6592, (2011), pp. 80–89.
- [9] R. Mamat, T. Herawan and M. M. Deris, "Super Attribute Representative for Decision Attribute Selection", In J. M. Zain, *et al.*, (Eds.): ICSECS 2011, Communication of Computer and Information Sciences, vol. 188, no. II, (2011), pp. 516–529.
- [10] Q. Hongwu, M. Xiuqin, J. M. Zain, N. Sulaiman and T. Herawan, "A Soft Set Model on Information System and Its Application in Clustering Attribute Selection", In J. M. Zain, *et al.*, (Eds.): ICSECS 2011, Communication of Computer and Information Sciences, vol. 188, no. II, (2011), pp. 16–27.
- [11] M. I. Awang, A. N. M. Rose, T. Herawan and M. M. Deris, "Soft set approach for selecting a clustering attribute", In L. Chao, J. Zhong and Y. Feng (Eds.): ADMA 2010, Lecture Notes in Computer Science, vol. 6441, (2010), pp. 87–98.
- [12] A. N. M. Rose, T. Herawan and M. M. Deris, "A framework of decision making based on maximal supported sets", In L. Zhang, J. Kwok and B.-L. Lu (Eds.): ISNN 2010, Part I, Lecture Notes in Computer Science, vol. 6063, (2010), pp. 473–482.
- [13] T. Herawan and M. M. Deris, "Soft decision making for patients suspected influenza", In D. Taniar, *et al.*, (Eds.): ICCSA 2010, Part III, Lecture Notes in Computer Science, vol. 6018, (2010), pp. 405–418.
- [14] T. Herawan, A. N. M. Rose and M. D. Mustafa, "Soft set theoretic approach for dimensionality reduction", In D. Ślęzak *et al.* (Eds.): DTA 2009, Communication of Computer and Information Sciences, vol. 64, (2009), pp. 180–187.
- [15] T. Herawan, I. T. R. Yanto and M. M. Deris, "Soft set approach for maximal association rules mining, In D. Ślęzak, *et al.*, (Eds.): DTA 2009, Communication of Computer and Information Sciences, vol. 64, (2009), pp. 163–170.
- [16] T. Herawan and M. D. Mustafa, "On multi-soft sets construction in information systems", In D. S. Huang, *et al.*, (Eds.): ICIC 2009, Lecture Notes in Artificial Intelligence, vol. 5755, (2009), pp. 101–110.
- [17] T. Herawan, M. D. Mustafa and J. H. Abawayjy, "Matrices representation of multi soft-sets and its application", In D. Taniar *et al.* (Eds.): ICCSA 2010, Part III, Lecture Notes in Computer Science, vol. 6018, (2010), pp. 201–214.
- [18] S. Palaniappan and T. K. Hong, "Discretization of Continuous Valued Dimensions in OLAP Data Cubes", *IJCSNS, International Journal of Computer Science and Network Security*, vol. 8, (2008), pp. 116–126.
- [19] P. K. Maji and A. R. Roy, "An application of soft sets in a decision making problem", *Computers and Mathematics with Applications*, vol. 44, (2002), pp. 1077–1083.
- [20] Z. Pawlak and A. Skowron, "Rudiments of rough sets", *Information Sciences*, vol. 177, no. 1, (2007), pp. 3–27.
- [21] N. Senan, R. Ibrahim, N. M. Nawi, M. M. Mokji and T. Herawan, "The Ideal Data Representation for Feature Extraction of Traditional Malay Musical Instrument Sounds Classification", In De-Shuang Huang, *et al.*, (Eds.): ICIC 2010, Lecture Notes in Computer Science, vol. 6215, (2010), pp. 345–353.