IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 4, April 2015.
www.ijiset.com
ISSN 2348 – 7968

Soft set, a soft Computing Approach for Dimensionality Reduction

Janmejay Pant[1], Amit Juyal[2] and Shivani Bahuguna[3]

[1] Department of Computer Applications, Graphic Era University, Dehradun, Uttakhand 248002, India

[2] Department of Computer Applications, Graphic Era University, Dehradun, Uttakhand 248002, India

[3] Department of Computer Applications, Graphic Era University, Dehradun, Uttakhand 248002, India

Abstract

Data uncertainty is a main problem of information science to handle the data and information. Many theories handle the uncertainty problem. This paper analysed soft set reduction and described how a data set is converted into binary information system and also analysed how it is better to reduce the dimension of the data. Like rough set, fuzzy set etc., are dealing with uncertainty. Soft set theory also do vital role to handle the uncertainty problem.

1. Introduction

There are three theories to solve the complicated problems in economics, engineering and environment and these theories are theory of probability, theory of fuzzy sets, and the interval mathematics. These are the mathematical tools for dealing with uncertainties. But all there are their own problems.

Theory of probabilities can deal only with stochastically stable phenomena.

Interval mathematics have arisen as a method of taking into account the errors of calculations by constructing an interval estimate for the exact solution of a problem. This is useful in many cases, but the methods of interval mathematics are not sufficiently adaptable for problems with different uncertainties. They cannot appropriately describe a smooth changing of information, unreliable, not adequate, and defective information, partially contradicting aims, and so on [1].

At the present time, the theory of fuzzy sets is progressing rapidly. But there exists a difficulty how to set the membership function in each particular case.

We should not impose only one way to set membership function. The nature of membership function is extremely individual. Very one may understand the notation $\mu_F(_X) = 0.7$ in his own manner. So, the fuzzy set operations based on the arithmetic operations with membership functions do not look natural. It may occur that these operations are similar to the addition of weights and lengths [1]. So the reason for the difficulties is the inadequacy of the parameterization tool of the theory. Consequently, Molodtsov [1] initiated the concept of soft theory as a mathematical tool for dealing with uncertainties which is free from the above mentioned difficulties.

2. Theory of Soft Sets

Molodtsov [1] defined the soft set in the following way. Let U be an initial universe set and E be a set of parameters. Let P (U) denotes the power set of U and$A \subset E$.

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

In other words, a soft set over U is a parameterized family of subsets of the universe. For $\varepsilon \subset A$. $F(\varepsilon)$ May be considered as the set of ε-approximate elements of the soft set (F, A). Clearly a soft set is not a set.

Definition2- The class of all value sets of a soft set (F,E) is called value class of the soft set and is denoted by $C_{(F,E)}$.

$C_{(F, E)}=\{p1=v1,p2=v2,p3=v3,\ldots p_n=v_n\}$[1][2].

Definition3- For two soft sets (F, A) and (G, B) over a common universe U, we say that (F, A) is a soft subset of (G, B) if

- $A \subset B$
- And $\forall \varepsilon \in A$, F($\varepsilon$) and G($\varepsilon$) are identical approximations

We write (F, A) $\tilde{\subseteq}$ (G, B).

(F, A) is said to be a soft super set of (G, B), if (G, B) is a soft subset of (F, A). We denote it by (F, A) $\tilde{\supseteq}$ (G, B) [1] [2].

Definition4- Two soft sets (F, A) and (G, B) over a common universe U are said to be soft equal if (F, A) is a soft subset of (G, B) and (G, B) is a soft subset of (F, A) [1] [2].

Soft set is also used finding the reducts for dimensionality reduction. As we know that reduct is a subset of attributes that are jointly sufficient and individually necessary for preserving a particular property of a given information system [3]. The reduct approaches under the soft set theory are still based on Boolean valued information system. The soft set theory has been applied to data analysis and decision support system.

The idea of dimensionality reductions under soft set theory have been proposed and compared, including the works [4-7]. The restriction of those techniques is that they are applicable only for Boolean valued information systems. But in the real applications a given parameter may have different values depending on the set of parameters. For example, the mathematics degree of student can be classified into three values, high, median and low. In this situation, every parameter determines a partition of the universe which is contains more than two disjoint subsets [3].

In Boolean valued information system we cannot directly define the standard soft set. To this the idea of multi soft sets [4] has been proposed. This approach deals with multi valued information systems.

In this part of dissertation work We used the idea of dimensionality reduction for multi valued information systems under the soft set theory. The main contributions of this work are as follows: firstly, I apply the idea of multi soft sets construction from a multi valued information System, and apply AND and OR operations on multi soft sets. Then I will present the applicability of the soft set theory for data reduction under multi valued IS using multi valued sets and AND operation.

AND and OR operations in multi soft sets

The notions of AND and OR operations in multi soft sets are given as-

AND Operation

Let (F,E) = ((F,a_i): i=1,2,3,…|A|) be a multi soft set over U representing a multi valued information system S=(U,A,V,f). The AND operation between (F, a_i) and (F,a_j) is defined as [3]

(F,a_i) AND (F,a_j) = (F, $a_i \times a_j$),

Where G (Va_i, Va_j) = F(Va_i) ∩ F(Va_j) ,∀ (Va_i, Va_j) ∈ $a_i \times a_j$,for $1 \le i, j \le |A|$.

OR Operation

Let (F,E) = ((F,a_i): i=1,2,3,…|A|) be a multi soft set over U representing a multi valued information system S=(U,A,V,f). The OR operation between (F, a_i) and (F, a_j) is defined as [3]

(F,a_i) OR (F,a_j) = (F, $a_i \times a_j$),

Where G (Va_i, Va_j) = F(Va_i) ∪ F(Va_j) ,∀ (Va_i, Va_j) ∈ $a_i \times a_j$,for $1 \le i, j \le |A|$.

Attribute Reduction based on AND operation

Let (F,E) = ((F,a_i): i=1,2,3,…|A|) be a 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 $C_{F(b1\times b2\ldots\times b\,|B|)} = C_{F(a1\times a2\ldots\times b\,|A|)}$ [3].

Experimental Results

In this section I present the applicability of soft set for finding reducts. As for first step, I need a transformation from a multi valued information system to multi soft sets. In the multi soft sets, I present the notion of AND and OR operations. For attribute reduction, I employ an AND operation to show that the reducts. The features are same as I used in rough set theory. The features I choose for dimensionality reduction are[8].

protocol, service, flage, Src_bytes, dst_byte, srv_cnt, serror_rate, srv_serrpr_rate, rerror_rate, srv_rerror_rate, same_srv_rate, diff_srv_rate,srv_diff_host_rate, dst_host_cnt,dst_host_srv_cnt,dst_host_diff_host,dst_srv_error_rate, dst_same_srv_rate and class

3. Multi-soft sets construction from multi-information systems

In this section We are going to decompose of a multi valued information system S= (U, A, V, f) into |A| number of binary valued information systems. The decomposition of S= (U, A, V, f) is based on decomposition of A = {a1, a2, a3…}[3][7][8].

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∈V_a, we define the map $a^i_v : U \longrightarrow \{0,1\}$ such that $a^i_v(u)$=1 if f(u,a)=v, otherwise a^i_v =0[3][4][5].

Using the above methodology I decompose the multi valued data set into binary valued information system. The features are mentioned in section 3.6.

Each feature is multi valued so I will make a separate IS for each feature. In my IS feature one is named as a1, second feature is named as a2 and so on. The decomposition of data set in the form of binary valued Information System (IS) is as shown-

Table 1 Decomposition of multi valued Information System

U	a1	
	udp	**tcp**
1	1	0
2	0	1
3	0	1
4	0	1
5	0	1
6	0	1
7	0	1
8	0	1
9	0	1
10	0	1

U	a2				
	other	**Private**	**http**	**remote**	**net_bios**
1	1	0	0	0	0
2	0	1	0	0	0
3	0	0	1	0	0
4	0	1	0	0	0
5	0	1	0	0	0
6	0	0	0	1	0
7	0	1	0	0	0
8	0	1	0	0	0
9	0	0	0	0	1
10	0	0	1	0	0

U	a3		
	SF	**S0**	**rej**
1	1	0	0
2	0	1	0
3	1	0	0
4	0	0	1
5	0	1	0
6	0	1	0
7	0	0	1
8	0	1	0
9	0	1	0
10	1	0	0

U	a4		
	Low	**medium**	**high**
1	0	1	0
2	1	0	0
3	0	1	0
4	1	0	0
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	1	0	0
10	0	0	1

U	a5		
	low	**medium**	**high**
1	1	0	0
2	1	0	0
3	0	1	0
4	1	0	0
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	1	0	0
10	0	0	1

U	a6		
	Low	**medium**	**high**
1	1	0	0
2	0	1	0
3	1	0	0
4	0	1	0
5	0	1	0
6	0	0	1
7	0	0	1
8	0	1	0
9	1	0	0
10	1	0	0

U	a7		
	low	medium	high
1	1	0	0
2	1	0	0
3	0	0	1
4	0	1	0
5	0	1	0
6	0	1	0
7	1	0	0
8	1	0	0
9	0	1	0
10	1	0	0

U	a8		
	Low	medium	high
1	1	0	0
2	0	0	1
3	1	0	0
4	1	0	0
5	0	0	1
6	0	0	1
7	1	0	0
8	0	0	1
9	0	0	1
10	1	0	0

U	a9		
	low	medium	high
1	1	0	0
2	0	0	1
3	1	0	0
4	1	0	0
5	0	0	1
6	0	0	1
7	1	0	0
8	0	0	1
9	0	0	1
10	1	0	0

U	a10		
	Low	medium	high
1	1	0	0
2	1	0	0
3	1	0	0
4	0	0	1
5	1	0	0
6	1	0	0
7	0	0	1
8	1	0	0
9	1	0	0
10	1	0	0

U	a11		
	low	medium	high
1	1	0	0
2	1	0	0
3	1	0	0
4	0	0	1
5	1	0	0
6	1	0	0
7	0	0	1
8	1	0	0
9	1	0	0
10	1	0	0

U	a12		
	Low	medium	high
1	0	0	1
2	0	1	0
3	0	0	1
4	1	0	0
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	1	0	0
10	0	0	1

U	a13		
	low	medium	high
1	0	0	1
2	0	0	0
3	1	0	0
4	0	0	0
5	0	0	0
6	1	0	0
7	0	0	0
8	0	0	0
9	1	0	0
10	1	0	0

	a14		
U	low	medium	high
1	1	0	0
2	1	0	0
3	0	0	1
4	1	0	1
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	1	0	0
10	0	0	1

U	a15		
	low	medium	high
1	0	0	1
2	0	0	1
3	0	0	1
4	0	0	1
5	0	0	1
6	0	0	1
7	0	0	1
8	0	0	1
9	0	1	1
10	0	1	0

U	a16		
	Low	medium	high
1	1	0	0
2	1	0	0
3	0	0	1
4	1	0	0
5	1	0	0
6	1	0	0
7	1	0	0
8	1	0	0
9	1	0	0
10	1	0	1

U	a17		
	low	medium	high
1	1	0	0
2	0	0	1
3	0	0	1
4	0	1	0
5	0	1	0
6	0	0	1
7	0	1	0
8	0	1	0
9	1	0	0
10	0	0	1

U	a18		
	low	medium	high
1	0	0	1
2	1	0	0
3	1	0	0
4	0	0	0
5	0	0	0
6	1	0	0
7	0	0	0
8	0	0	0
9	1	0	0
10	1	0	0

So the decomposition of multi valued data set which I have used is. There are 18 features I used then 18 information systems are generated as above.
From the above tables the corresponding soft sets are given below.
(F, a1)= {{udp=1}, {tcp=2, 3, 4, 5, 6, 7, 8, 9.10}},
(F, a2)={{other=1},{private=2,4,5,7,8},{http=3,10},{remote_job=6},{netbios_ns}},
(F, a3)= {{sf=1, 3, 10}, {s0=2, 5, 6, 8, 9}, {rej=4, 7}},
(F, a4)= {{low=2, 4, 5, 6, 7, 8, 9}, {medium=1, 3}, {high=10}},
(F, a5)= {{low=1, 2, 4, 5, 6, 7, 8, 9}, {medium=3},{high=10}},
(F, a6)= {{low=1, 3, 9, 10},{medium=2,4,5.8},{high=6,7}},
(F, a7)= {{low=1, 2,7, 8,10},{medium=4,5,6,9},{high=3}},
(F, a8)= {{low=1, 3, 4,7, 10},{medium=∅},{high=2,5.6,8,9}},
(F, a9)= {{low=1, 3, 4,7, 10},{medium=∅},{high=2,5.6,8,9}},
(F, a10)= {{low=1, 2, 3, 5,6,8,9, 10},{medium=∅},{high=4,7}},
(F, a11)= {{low=1, 2, 3, 5,6,8,9, 10},{medium=∅},{high=4,7}},
(F, a12)= {{low=4, 5, 6, 7, 8, 9}, {medium=2}, {high=1, 3, 10}},
(F, a13)= {{low=3, 6, 9, 10}, {medium=2, 4, 5, 7, 8}, {high=1}},
(F, a14)= {{low=1, 2, 4, 5, 6, 7, 8, 9}, {medium=0}, {high=3, 10}},
(F, a15)= {{low=4, 5, 6, 7, 8, 9}, {medium=2}, {high=1, 3, 10}},
(F, a16)= {{low=1, 2, 4, 5, 6, 7, 8, 9}, {medium=0}, {high= 3, 10}},
(F, a17)= {{low=1, 9}, {medium=4, 5, 7, 8}, {high=2, 3, 6,10}},
(F, a18)= {{low=2, 3, 6, 10}, {medium=4, 5, 7, 8, 9}, {high=1}}.

Thus the multi soft set representing in the above table is
(F, A)= ((F, a1), (F, a2), (F, a3)… (Fa18))
Now, the next assignment is finding the reducts from the above soft sets. For attribute reduction soft set theory use AND and OR operations.

4. **Attribute Selection** This section describes the idea of attributes reduction under soft set theory. AND operation is used for the task. For example- The AND operation between (F,a_i) and (F,a_j) is defined as [3][4][6]

(F,a_i) AND $(F,a_j)=(F, a_i \times a_j)$

So we have to find reducts using AND operation. Then I applied AND operation to each object.

Like (F,a1) AND(F,a2), (F,a1) AND F,a3), (F,a1) AND(F,a4),……………(F, a1) AND(F, a18).

(F, a2) AND (F, a3), (F, a2) AND F, a4)... (F, a2) AND (F, a18), and so on. As we know that (F, a1) AND (F, a2) = (F, a2) AND (F, a1) then one combination we choose for the reduct.

For example, let two multi soft sets

(F {a1, a4}) and (F {a1, a5})

- For(F{a1,a4}), where (F,a1)={{udp=1},{tcp=2,3,4,5,6,7,8,9,10}}

(F,a4)={{low=2,4,5,6,7,8,9},{medium=1,3},{high=10}}

Then we have

(F, a1) AND (F, a4) = (F, a1×a4)

$$= \left\{ \begin{array}{l} (udp, low) = \{\emptyset\},\quad (udp,medium)=\{1\},\quad (udp,high)=\{\emptyset\} \\ (tcp, low) = \{2,4,5,6,7,8,9\},\ (tcp,medium)=\{3\},(tcp,high)=\{10\} \end{array} \right\}$$

So that

$C_{F(a1\times a4)}$ = {{1},{2,4,5,6,7,8,9},{3},{10}} **…………….(1)**

- For(F{a1,a5}), where (F,a1)={{udp=1},{tcp=2,3,4,5,6,7,8,9,10}}

(F, a5)={{low= 1,2,4,5,6,7,8,9},{medium=3},{high=10}}

Then we have

(F, a1) AND (F, a5) = (F, a1×a5)

$$= \left\{ \begin{array}{l} (udp, low) = \{1\},\quad (udp,medium)=\{\emptyset\}\},\quad (udp,high)=\{\emptyset\} \\ (tcp, low) = \{2,4,5,6,7,8,9\},\ (tcp,medium)=\{3\},(tcp,high)=\{10\} \end{array} \right\}$$

So that

$C_{F(a1\times a5)}$ = {{1},{2,4,5,6,7,8,9},{3},{10}} **…………….(2)**

From (1) and (2) we have {a1, a4} and {a1, a5} have the same set of features.

So {a1} is reduct of the above multi soft set.

Now we can find all the reducts from the given data set of multi valued features. I found the different reducts from the network data set using the AND operator described in the above.

From the above theory I found different multi soft sets having the same set of features or values. These sets are given as

{a1, a4}, {a1, a5},{a1,a8},{a1,a9},{a1,a14},{a1,a16},{a3,a8},{a3,a9},{a3,a10},
{a3,a11},{a3,a4},{a3,a15},{a3,a14},{a3,a16},{a4,a10},{a4,a11},{a4,a14},{a4,a16},{a5,a8},{a5,a9},{a5,a10},{a5,a11},{a4,a13},{a5,a12},{a5,a14},{a5,a15},{a5,a16},
{a6,a8},{a6,a9},{a6,a10},{a6,a11},{a7,a8},{a7,a9},{a7,a10},{a7,a11},{a7,a14},
{a7,a15},{a7,a16},{a8,a10},{a8,a11},{a8,a14},{a8,a16},{a9,a10},{a9,a11},{a9,a14},{a9,a16},{a10,a14},{a10,a16},{a11,a14},{a11,a16},{a10,a18},{a11,a18},{a12,a14},{a12,a16},{a13,a14},{a13,a18},{a14,a17},{a17,a18}.

From the above features set there are some important features which preserve the indiscernibility relations are reducts.

{a1, a4}, {a1, a5} are reducts so we can choose only one set for dimensionality reduction. Same as

{a1, a8}, {a1, a9} are reducts,

{a1, a14}, {a1, a16} are reducts,

{a3, a4}, {a3, a15}, {a3, a14} and {a3, a16} are reducts,

{a3, a8}, {a3, a9}, {a3, a10} and {a3, a11} are reducts,

{a4, a10}, {a4, a11} are reducts, {a4, a14}, {a4, a16} are reducts,

{a5, a8}, {a5, a9} are reducts, {a5, a10}, {a5, a11} are reducts,
{a4, a13}, {a5, a12} are reducts, {a5, a14}, {a5, a15} and {a5,a16} are reducts,
{a6, a8}, {a6, a9}, {a6, a10} and {a6, a11} are reducts,
{a7, a8}, {a7, a9} are reducts, {a7, a10}, {a7, a11} are reducts,
{a7, a14}, {a7, a15} and {a7, a16} are reducts,
{a8, a10}, {a8, a11} are reducts, {a8, a14}, {a8, a16} are reducts,
{a9, a10}, {a9, a11} are reducts, {a9, a14}, {a9, a16} are reducts,
{a10, a14}, {a10, a16}, {a11, a14} and {a11, a16} are reducts,
{a10, a18}, {a11, a8} are reducts, {a12, a14}, {a12, a16} are reducts,
{a13, a14}, {a13, a16} are reducts, {a13, a17}, {a13, a18} are reducts,
{a14, a17}, {a17, a18} are reducts,

The reducts of my data set are
{protocol, src_byte }, { protocol, dst_byte } are reducts,
{ protocol, srv_serrpr_rate}, { protocol, rerror_rate} are reducts,
{ protocol, dst_host_cnt}, { protocol, dst_same_srv_rate} are reducts,
{flage,src_byte},{flage,dst_host_srv_cnt},{flage,dst_host_cnt}and{flage, dst_same_srv_rate}
{flage,srv_serrpr_rate},{flage,rerror_rate},{flage,srv_rerror_rate}and{flage, same_srv_rate}
{src_byte,srv_rerror_rate},{src_byte,same_srv_rate}are reducts,{src_byte,dst_host_cnt},
{src_byte, dst_same_srv_rate} are reducts,
{dst_byte, srv_serrpr_rate}, {dst_byte,rerror_rate}are reducts, {dst_byte, srv_rerror_rate },
{ dst_byte, same_srv_rate} are reducts,
{src_byte,srv_diff_host_rate},{dst_byte,diff_srv_rate} are reducts, {dst_byte, dst_host_cnt}, { dst_byte, dst_host_srv_cnt} and { dst_byte,dst_same_srv_rate} are reducts,
{srv_cnt,srv_serrpr_rate},{srv_cnt,rerror_rate},{srv_cnt,srv_rerror_rate}and{srv_cnt,same_srv_rate} are reducts,
{serror_rate,srv_serrpr_rate},{serror_rate,rerror_rate} are reducts, { serror_rate, srv_rerror_rate }, { serror_rate, same_srv_rate} are reducts,
{serror_rate,dst_host_cnt},{serror_rate,dst_host_srv_cnt}and{serror_rate,dst_same_srv_rate} are reducts,{ srv_serrpr_rate, srv_rerror_rate }, {srv_serrpr_rate, same_srv_rate} are reduct, { srv_serrpr_rate, dst_host_cnt}, { srv_serrpr_rate, dst_same_srv_rate} are reducts,
{rerror_rate,srv_rerror_rate},{rerror_rate,same_srv_rate}arereducts,
{ rerror_rate, dst_host_cnt}, { rerror_rate, dst_same_srv_rate} are reducts,
{srv_rerror_rate,dst_host_cnt},{srv_rerror_rate,dst_same_srv_rate},{same_srv_rate,dst_host_cnt} and {same_srv_rate, dst_same_srv_rate} are reducts,
{srv_rerror_rate,dst_srv_error_rate},{same_srv_rate,srv_serrpr_rate}arereducts, {diff_srv_rate, dst_host_cnt}, {diff_srv_rate, dst_same_srv_rate} are reducts,
{srv_diff_host_rate,dst_host_cnt},{srv_diff_host_rate,dst_same_srv_rate}arereducts, {srv_diff_host_rate, dst_host_diff_host}, {srv_diff_host_rate, dst_srv_error_rate} are reducts,{dst_host_cnt, dst_host_diff_host}, {dst_host_diff_host, dst_srv_error_rate} are reducts.

5. Conclusions

Soft set theory is a general method for solving problems of uncertainty. Soft Sets represent a powerful tool for decision making about information systems, data mining and drawing conclusions from data, especially in those cases where some uncertainty exists in the data. Its efficiency in dealing with uncertainty problems is as a result of its parameterized concept. Recently, various researches had been done various works in theory and in practices. In this paper, a study of the concept of soft set which deals with uncertainty. A concept of soft equivalence relation is introduced. Soft equivalence relations give rise to the concept of soft approximation space and soft rough sets. It is shown that a soft set give rise to an equivalence relation, so a soft set itself has classification ability. In my work, it is studied that how we can reduce the number of parameters for a soft set to the minimum without distorting its original classification ability. It helps us to reach a conclusion with minimum number of parameters in the soft set. KDD data set is used for dimension detection using SST the main purpose was to find the reducts. For this decomposed the multi valued information system and picked 18 features for this and then found the various reducts. There is no tool is available for SST so all the experiments were achieved by manual.

The existing reduct approaches under soft set theory are still based on boolean valued IS. For the real applications the data usually contain non boolean valued. An alternative approach for attribute reductions in multi valued IS under SST has been presented. In the applied approach, notion of multi soft set is used to represent multi valued information systems. The AND operation is used in multi soft sets to present the notion of attribute reduction

REFERENCES

[1] D. Molodtsov, "Soft Set Theory First Results", Computers and Mathematics with Applications, 1999.

[2] P. K. Maji, R.Biswas ,R.Roy," Soft Set Theory", Computers and Mathematics with Applications,2003.

[3] Tutut Herawan,Rozaida Ghazalil,"Soft Sets theoritic approach for dimensionality reduction ",International Journal of Database Theory and Application, Vol.3,June 2010.

[4] P. K. Maji, R.Biswas ,R.Roy," An Application of Soft Sets in a decision making problem", Computers and Mathematics with Applications,2002.

[5] D.Chen,E.C.C Tsang,"Some notes on the parameterizations reduction of soft sets", Proceeding of International Conference on Machine learning and Cybernetics ,IEEE,2003.

[6] D.Chen, E.C.C Tsang, D.S Yeung, "The parameterization Reduction of Soft Sets and its Applications", Computers and Mathematics with Applications, 2005.

[7] Z.Kong,L.Gao,L.wang,"The normal parameter reduction of soft sets and its algorithms", Computers and Mathematics with Applications,2008.

[8] Y. Zou, Z.Xiao,"Data Analysis approaches of soft sets under incomplete informationn", 2008.

First Author- Mr. Janmejay Pant received his M.Tech degree in Information Technology from Graphic Era University. His areas of interest are soft computing and machine learning. He has published various papers in the different journals those have good impact factor and has presented paper in international conference that was sponsored by IEEE. Currently he is working as Assistant professor in Graphic Era University Deharadun.

Second Author- Mr. Amit Juyal received his M.Tech degree in computer science and engineering from Graphic Era University. His areas of interest are big data and data mining. He has published various papers in the different journals. Currently he is working as Assistant professor in Graphic Era University Dehradun.

Third Author- Ms. Shivani Bahuguna received her M.Tech degree in Information Technology from Graphic Era University. Her areas of interest are big data and data mining. She has published various papers in the different journals. Currently she is working as Assistant professor in Graphic Era University Dehradun.