Research Article

Experimental Analysis of Soft Set Based Parameter Reduction Algorithms for Decision Making

Priyanka D. Lanjewar[a]* , B. F. Momin[b]

[a] Computer Science and Engineering Department, Walchand College of Engineering, Sangli 416415, Shivaji Univeristy, Kolhapur,India
email:[a]*priyanka.lanjewar@walchandsangli.ac.in,[b]bfmomin@yahoo.com

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 4 June 2021

Abstract: In the field of data mining, the parameter reduction method solves the decision making problems for the knowledge discovery process. Big data faces many problems which can be solved with the help of parameter reduction. Now a day's reduction of data is extremely significant to make the optimal decision on the basis of some parameters. In this paper, the literature survey shows the various methods of parameter reduction which are based on the Soft Set theory. Soft set theory is based on the parameterized reduction property. This paper mainly focuses on the analysis of existing parameter based reduction methods using the soft set concept which are practically implemented with machine learning. The new soft set based approach for parameter reduction is also proposed called as ranked based parameter reduction method for the optimal selection of object to take the correct decision. For a better understanding, a comparison of various implemented algorithms is also presented.

Keywords: Soft set theory, Soft set, parameter reduction, optimal selection, decision making, etc.

1. Introduction

These Days decision making process is very important for various field of study. It is very difficult to handle the uncertain and imprecise data in field of medical, social science, engineering and economics etc. Basically there are some mathematical theories such as Fuzzy Set, Rough Set which handles the vague data. However such theories comprise of individual problems. Fuzzy set theory requires defining the membership function for each different case and in case of Rough set, no direct relationship among the decision and conditional parameters is given. In these theories there is essential to design the mathematical model for the precise solution. However it is very difficult to design such model for the exact solution.

So to avoid this problem, Russian scientist D. Molodtsov proposed the new mathematical model called as Soft set Theory (SST) in 1999[1]. This theory deals with the uncertain data such as unknown and missing data. The soft computing approaches handle such type of uncertain data and provide the optimal decision or discover the knowledge over the big data. A researcher is concentrating on new area as "SST" as parameter reduction philosophy specifically depends on parameterization property that selects the parameters for the optimal solution by reducing the core parameters. So, this type of concept overcomes the problems of existing theories and finally gives the approximate solution to the problem without defining any precise solution. The soft set theory is very convenient which is easy to applicable for practice because this theory does not require designing of the exact model because it is approximate in nature. This theory is useful in decision making process to give optimal selection using various parameter reduction algorithms.

Parameter reduction is the process of reducing the parameters which identify the core or dispensable parameter for the meaningful reduction of data to give the optimal selection over data. Soft set theory is basically parameter reduction theory that handles the problems of parameter reduction and proposed many algorithms for reduction of parameter as discussed in literature survey and methodology. This paper presents new algorithm for the optimal selection based on the ranking concept.

2.Literature Survey

Soft Set Theory gives the new property of parameterization which is very useful for decision making process to take the optimal decision. This theory is applicable to many fields such as medical, economics and in engineering etc. There are various parameter reduction algorithms that are being developed using the concept of soft set theory which are useful for better and correct decision making process.

2.1 Soft Set based Parameter Reduction Algorithms:

P. K. Maji [2] has developed the very first soft set based parameter reduction algorithm to solve the problem of decision making where need to select a house from many houses. At last, this algorithm gives the optimal selection of house using few parameters only. Soft set can be defined as follows:

$$F: E \rightarrow P(U) \quad (1)$$

Here, the power set of *U* is *P(U)* and soft set is *(F,E)*and, *F* is the mapping of *E* where *U* is the parameterized soft set.

D. Chen *et al.*[3]there is inaccurate computation in [2]. This is enhanced by reduction of parameters by means of soft set parameterization reduction concept. The same method of [2] is provided here but gives the different solution because of new parameterization property. In soft set, decision value is computed using the parametric values of each object but in rough set it does not provide any relation which becomes main diversity of these two sets and finally presented alternative method using weighted soft set. Z. Kong *et al*. [4] has presented a new definition of Normal Parameter Reduction (NPR) algorithm based on soft set and proposed one algorithm for the same. In that parameter importance degree is calculated using that value one feasible parameter reduction set is created which satisfy the some equality condition which resulted in parameter reduction set as output. Complexity of this algorithm is O (n^3). Also proposed one algorithm called as NPR for fuzzy soft set (NPRFSS). Herawan *et al.*[5]has developed the alternative soft set based method for reduction of attribute over multi valued information system using AND and OR operations. X. Ma *et al.*[6] has proposed a new efficient normal parameter reduction algorithm (NENPR) for soft stet. In NPR algorithm there is need to calculate the parameter importance degree which requires more amount of time. In order to avoid this time, NENPR algorithm is developed which calculate the total of oriented parameter only. X Ma *et al.*[7] has designed a novel definition of parameterization value reduction for soft set theory in which only one parameter is kept that has maximum value which is denoted by value "1" and others parameter values are deleted. On the basis of only one parameter with least parameter values optimal selection is made for decision making.

M. I. Ali [8] introduce the concept of soft equivalence relation and also presented the new method to select the house using dispensable parameter without distorting classification ability. Z. Kong *et al.*[9] has proposed new idea for reduction of parameter using two cases that are based on soft set, as first is done by changing entries and second is done by adding objects after validate the results. D. A. Kumar and R. Rengasamy[10] have proposed method that reduced the dispensable parameters in which sample data in converted into binary data for better decision making process. B. Han and X. Li [11] have proposed a method to compile the various normal parameter reduction algorithms of soft set using three decision based rules. In which these three rules are combine in one algorithm. Z. Li *et al.*[12] has presented new algorithm of parameter reduction with the concept of soft covering where the discernibility matrix is calculated for the reduction of parameters. T. Bakshi *et al.*[13] has presented seven different correlated algorithms for parameterization reduction of soft set that gives polynomial time of computation. S. Danjuma *et al.*[14] has developed new algorithm called as alternative approach to NPR (ANPR) algorithm which reduced the problem of NPR algorithm in the view of complexity. The complexity of this algorithm is reduced as O (n^2). Comparison of NPR, NENPR and ANPR algorithms is also presented in this paper. How these three algorithms works is also shown with the help of one common example of six patients who have different symptoms of thrombocythemia disease as shown in following table 3. S. Danjuma *et al.*[15] has presented the review of various soft set based parameter reduction algorithms that are being useful for decision making process and also provide the comparison of each algorithm with its advantages and disadvantages. X. Ma and H. Qin [16] has proposed the parameter value reduction algorithm for soft set after that one more algorithm for maximal parameter value reduction and same algorithm which is based on normal parameter reduction are proposed.

3.Methodology

The new parameter reduction theory which is very useful to get the best choice called as Soft Set Theory which takes better decision for many areas. This theory is applicable in many fields such as medical, science, economic and engineering. Data analysis is extremely essential to obtain correct decision using optimal data for many realistic applications. It means that rather to consider all the parameters while taking certain decision, selection of some parameters will give the optimal selection with less parameter. Here various parameter reduction algorithms are discussed with the help of examples.

3.1.Parameter reduction algorithm for selection of the house [2] [3]:

In the above algorithm consider one example for the selection of house from the total six numbers of houses, as U= {h_1, h_2, h_3, h_4, h_5, h_6} be a set of six houses, and E is the collection of parameters which provides the "characteristics of house", given by (F, E) as shown in following figure 1.

{Expensive houses = Ø,
(e_1) Beautiful houses= $\{h_1, h_2, h_3, h_4, h_5, h_6\}$,
(e_2) Wooden houses = $\{h_1, h_2, h_6\}$,
Modern houses = $\{h_1, h_2, h_6\}$,
Houses in bad repair = $\{h_2, h_4, h_5\}$,
(e_3) Cheap houses = $\{h_1, h_2, h_3, h_4, h_5, h_6\}$,
(e_4) Houses in green surroundings = $\{h_1, h_2, h_3, h_4, h_6\}$
(e_5) Houses in good repair = $\{h_1, h_3, h_6\}$}

Figure 1: Soft Set (F, E) Example for selection of house

If someone wants to buy a house using some parameters such as beautiful, wooden, cheap, in green surroundings, in good repair which form the set E = $\{e_1,e_2,e_3,e_4,e_5\}$. According to the above example shown in figure 1 following table 1 is constructed.

Table 1: Soft Set

U/E	e_1	e_2	e_3	e_4	e_5	Choice value c_i
h_1	1	1	1	1	1	5
h_2	1	1	1	1	0	4
h_3	1	0	1	1	1	4
h_4	1	0	1	1	0	3
h_5	1	0	1	0	0	2
h_6	1	1	1	1	1	5

Algorithm for selection of house:

1. Input the soft set (F, E)
2. Input the set P of choice parameters which is a subset of E
3. Identify all reduct-soft-sets of (F, P)
4. Select one reduct-soft-set say (F, Q) of (F, P)
5. Obtain k, for which Ck = max c_i.

Figure 2: Parameterization Algorithm [2][3]

The parameterization algorithm shown in figure 2 is work as follows. First input to above algorithm is (F, E) and the collection of parameters as $\{e_1 \ldots e_5\}$is the second input as shown in figure 6, it is obvious that h_1 and h_6 both have maximum choice value so h_1 and h_6 are becomes the optimal selection. For huge data it is complicated to obtain solution, so third step becomes finding of the reduct soft set by deleting e_1 and e_3 because of similar data which is resultant in reduction thus soft set (F, P) becomes as $\{e_2 ,e_4 ,e_5\}$. After deleting $e_4$from (F, P) then only e_2 and e_5 are remained which is called as (F,Q). From this reduct soft set, optimal choices for best house selection are h_1 or h_6 which has highest choice value. So the suboptimal choices are h_2 and h_3. Output of this algorithm is shown in figure 7.

3.2.Parameterization Value Reduction algorithm (PVR) [7]:

(1) Input the soft set (F, E) and the parameter set E;

(2) Obliterate the parameter values denoted by 0.

(3) keep the one parameter value denoted by 1 of the object which has the maximum choice value and delete the remainder parameter values.

Figure 3: Parameterization Value Reduction algorithm (PVR) [7]

The above algorithm shown in figure 3 first take the soft set as input with the parameter values for each houses as shown in table 2. For the reduction of parameter, this algorithm delete all the parameter values which is indicated by “0” and write “1” for only those house/object who has maximum sum value *f()*as given in figure 8 .

Table 2: Soft set (F, E)

U/E	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	*f()*
h_1	1	0	0	0	0	1	1	0	3
h_2	0	0	1	1	1	1	0	0	4
h_3	1	0	0	1	1	1	1	1	6
h_4	1	1	0	1	0	0	0	1	4
h_5	0	0	1	0	1	0	0	1	3
h_6	1	0	0	1	1	0	0	1	4

3.3.New algorithm for selection of house [8]:

(1) Input the soft set (F, E).
(2) Input the choice (condition) parameters C
(3) Input decision parameter $d = \sum h_{ij}$ as the last column in table obtained by choice parameters.
(4) Rearrange the Input by placing the objects having the same value for the parameter d adjacent to each other.
(5) Distinguish the objects with different values of d by double line.
(6) Identify core parameters
(7) Output by eliminating the entire dispensable parameters one by one, resulting a table with minimum number of condition parameters having the same classification ability for *d* as the original table with *d*.

Figure 4: New Algorithm for selection of house [8]

Table 3: Soft set (F, E)

U/E	e_1	e_2	e_3	e_4	e_5	e_6	e_7	*D*
h_1	1	0	1	1	1	0	0	4
h_2	0	1	1	1	0	1	1	5
h_3	0	0	1	0	1	0	1	3
h_4	1	0	1	1	0	0	0	3
h_5	1	0	1	0	0	1	0	3
h_6	0	1	1	1	1	0	0	4

The algorithm shown in figure 4 provide the new concept of the decision parameter as shown in table 3, *d* is the decision parameter which is the sum of values of choice parameter. After that rearrange the objects according to the *d* values as classification of objects is done according to value *d*. Then eliminate the dispensable parameter and check whether classification pattern change or not. Here in above example e_3 is the dispensable parameter because after deleting the e_3 classification pattern does not changed. Classification pattern before and after deleting e_3 is same ash_2, h_1, h_6, h_3, $h_4$and h_5. The output of the above algorithm is shown in figure 9 that gives h_2 as the first selection of house because of the maximum *d* value.

3.4.Normal Parameter Reduction (NPR) Algorithm [4]:

Table 4: Soft set

U/E	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8	*f()*
p_1	1	0	1	1	1	1	0	1	6
p_2	0	0	1	1	0	0	1	1	4
p_3	1	0	1	0	1	1	1	0	5
p_4	1	0	1	1	1	1	0	1	6
p_5	1	0	1	1	1	1	0	1	6

p_6	0	0	1	0	0	1	1	1	4

Above table 4 gives the soft set (F, E) for the example where a doctor need to check- ups the various patients who are identified with Thrombocythemia symptoms as shown in figure 5. The collection of six patients is the set U = $\{p_1, p_2, p_3, p_4, p_5, p_6\}$. Here parameters are the set of 8 symptoms can be represented as set E = $\{e_1, e_2, e_3, e_4, e_5, e_6, e_7, e_8\}$ as shown in figure 10. Here, doctor need to identify the patients who has thrombocythemia disease with the few symptoms/parameters.

$$(F,E) = \left\{ \begin{array}{c} \text{Headache} = \{p_1, p_3, p_4, p_5, \} \\ \text{Dizziness} = \{\ \} \\ \text{Weakness} = \{p_1, p_2, p_3, p_4, p_5, p_6\} \\ \text{Fainting} = \{p_1, p_2, p_4, p_5\} \\ \text{Numbness} = \{p_1, p_3, p_4, p_5\} \\ \text{Throbbing} = \{p_1, p_3, p_4, p_5, p_6\} \\ \text{Change in vision} = \{p_2, p_3, p_6\} \\ \text{Chest pain} = \{p_1, p_2, p_4, p_5, p_6\} \end{array} \right\}$$

Figure 5: Soft Set (F, E) Example for selection of patients of thrombocythemia disease [14]

NPR Algorithm:

Step1: Compute the parameter importance degree using this formula:

$re_i = \frac{1}{|U|}(\propto_1, e_i, \propto_2, e_i + \cdots \propto_s, e_i)$

$re_1 = 1/|6| *(4) = 2/3$

$re_2 = 1/|6| *(0) = 0$

$re_3 = 1/|6| *(6) = 1$

$re_4 = 1/|6| *(4) = 2/3$

$re_5 = 1/|6| *(4) = 2/3$

$re_6 = 1/|6| *(5) = 5/6$

$re_7 = 1/|6| *(3) = 1/2$

$re_8 = 1/|6| *(5) = 5/6$

Step 2: Using above calculation, Find out the maximal subset *A* which satisfy that sum of $re_i = (1 \le i \le p)$ which has to be non negative integer that provide the any of the following set.

A = $\{e_2, e_3, e_4, e_5, e_6, e_8\}$ *A*= $\{e_1, e_3, e_4, e_5\}$

A = $\{e_1, e_2, e_3, e_4, e_5\}$ *A*= $\{e_1, e_7, e_8\}$

$re_{2+} re_{3+} re_{4+} re_{5+} re_{6+} re_8 = 4$

$re_{1+} re_{2+} re_{3+} re_{4+} re_5$ or $re_{1+} re_{3+} re_{4+} re_5 = 3$

$re_{1+} re_{7+} re_8 = 2$

Step 3: Form the above sets *A* = $\{e_2, e_3, e_4, e_5, e_6, e_8\}$ *A*= $\{e_1, e_3, e_4, e_5\}$ *A* =$\{e_1, e_2, e_3, e_4, e_5\}$ and *A*= $\{e_1, e_7, e_8\}$ only *A* = $\{e_1, e_7, e_8\}$ satisfy $f_A(p_1) = f_A(p_2) = \ldots = f_A(p_n)$ here, $f_A(p_1) = f_A(p_2) = f_A(p_3) = f_A(p_4) = f_A(p_5) = f_A(p_6) = 2$ so this condition is satisfied for only one set *A* = $\{e_1, e_7, e_8\}$ and ignore other sets.

Step 4: Finally do *E – A*. Here *A*= $\{e_1, e_7, e_8\}$ so that normal parameter reduction set is $\{e_4, e_5, e_6\}$ which is shown in figure 11.

3.5.New Efficient Normal Parameter Reduction (NENPR) Algorithm [6]:

Step 1: There exists e^o_j and e^1_j it means that e_2 has value "0" and e_3 has value "1" for all patients. So that e_2 and e_3 are deleted and put them in set *C*.

Step 2: Computed oriented parameter sum *f ()* for the reaming parameter.

Step 3: Obtain subsets A⊂ E where, S_A is a multiple of |U| =6. Which results in many subsets among few are {e_1,e_2,e_3,e4,e_5}, {e_1,e_3,e_4,e_5},{e_4,e_5,e_6,e_8} that are called as candidate parameter reduction set. Here for each of these three set S_A = 18 which is multiple of |6|. S_A is the sum of all *f ()* of each patient.

Step 4: Sort out the candidate parameter set which satisfy $f_A(p_1) = f_A(p_2) = \ldots = f_A(p_n)$ and delete the remainders. In this case, set A= {e_5, e_7, e_8} satisfied the above condition in which $f_A(p_1) = f_A(p_2) = f_A(p_3) = f_A(p_4) = f_A(p_5) = f_A(p_6) = 2$ for the set *A* = {e_5, e_7, e_8} so delete that e_5,e_7 and e_8 parameters and put them in set *A*.

Step 5: Finally does *E-A-C* = {e_1, e_4, e_6} as parameter reduction set which is shown in figure 12.

3.6.Alternative Approach to Normal Parameter Reduction (ANPR) Algorithm [14]:

Step 1: Soft set with its parameter is the input.

Step 2: If $e_i = e_j$ is exists then choose only one of them and here $e_1 = e_5$ have same parameters so delete one of them and here delete e_5 and put it in *Q* set. Also check if there exists e^0_j and e^1_j so that here e_2 and e_3 are also reduced here and put them in set *C*.

Step 3: Calculate oriented parameter sum as *f ()*.

Step 4: Check if $f_A(p_1) = f_A(p_2) = \ldots = f_A(p_n)$ and also check $f_B(p_1) = f_B(p_2) = \ldots = f_B(p_n)$ so set {e_1, e_7, e_8} and {e_4, e_6, e_7} satisfy above condition

Step 5: The intersection of (A U B) & (A ∩ B) is calculated such as *A* = {e_1, e_7, e_8} and *B*= {e_4, e_6, e_7} which result in (A ∩ B) = {e_7} and put it in *D* set for reduction.

Step 6: Finally do *E – C – D – Q* which gives following reduction of parameter as {e_1, e_4, e_6, e_8} which is shown in figure 13.

4.Proposed method:

New approach to select the patients using different optimal parameters is **Ranked Based Parameter Reduction Algorithm** (RBPR) using the concept of soft set:

Step 1: Soft set (*F, E*) with the choice parameters *E* is the input as shown in table 4.

Step2: Calculate $f() = \sum h_{ij}$ for each object.

Step 3: Rearrange all objects according to the *f* () with the highest rank and compute the ranking order all objects.

Step 4: Compute the ranked distance d_r values for each consequent objects.

Step 5: Check if $d_r(p_i, p_{i+1}) \neq 0$ then $f_A(p_i) - f_A(p_{i+1}) < d_r(p_i, p_{i+1})$

Check if $d_r(p_i, p_{i+1}) = 0$ then $f_A(p_i) - f_A(p_{i+1}) = d_r(p_i, p_{i+1})$

If above conditions is satisfied then put those parameter in reduction set *A*.

Step 6: Check if there exists $e_i = e_j$ then put them in reduction set *C*

Step 7: Finally do *E- A-C* and choose the objects with the highest ranked for optimal selection.

In the first step of the above algorithm, it takes the table 4 as input to solve the problem where need to select the patients/objects who are suspected with Thrombocythemia disease with less symptoms. In the second step compute $f() = \sum h_{ij}$ for each patients which is the sum of parameter values for each patients. In third step ranked table need to generate according to the highest rank. In step fourth ranked distance d_r is need to compute for each patients as follows.

$d_r p_1\, p_4) = fE(p_1) - fE(p_4) = 6 - 6 = 0$

$d_r(p_4, p_5) = fE(p_4) - fE(p_5) = 6 - 6 = 0$

$d_r(p_5, p_3) = fE(p_5) - fE(p_3) = 6 - 5 = 1$

$d_r(p_3, p_2) = fE(p_3) - fE(p_2) = 5 - 4 = 1$

$d_r(p_2, p_6) = fE(p_2) - fE(p_6) = 4 - 4 = 0$

In fifth step two conditions have to check,

Check if $d_r(p_i, p_{i+1}) \neq 0$ then $f_A(p_i) - f_A(p_{i+1}) < d_r(p_i, p_{i+1})$

Check if d_r (p_i, p_{i+1}) = 0 then f_A (p_i)-f_A (p_{i+1}) = d_r (p_i, p_{i+1})

If both conditions are satisfied by the parameter then put them in parameter reduction set *A*. In this example parameter {e_2, e_3, and e_7} satisfied the condition which is considered as reduction of parameter and put them in set *A*. In step sixth similarity condition is to be checked using e_i = e_j. Here parameters e_1 and e_5 satisfy the condition which is to be deleted and put in reduction set *C*. In last step only do the *E- A-C* which gives the {e_4, e_6, e_8} as the output of the algorithm. So using these e_4, e_6, e_8 three parameters, patient's p_1, p_4 and $p_5$can be the optimal solution that has thrombocythemia disease because of the highest choice values as shown in table 5.

5.Experimental analysis of existing methods

Implementation of all above algorithm is done in machine learning platform using python version 3.7.1, on Intel® core™ i3-5005U CPU @ 2.00GHzwith 8.00 GB RAM and 64-bit operating system. Result of various algorithms of parameter reduction for selection of houses and selection of patients for Thrombocythemia disease are shown here. Those algorithms are parameterization algorithm, parameterization value reduction (PVR), new algorithm for selection of houses, NPR, NENPR, ANPR and RBPR.

	U	e1	e2	e3	e4	e5	sum
0	h1	1	1	1	1	1	5
1	h2	1	1	1	1	0	4
2	h3	1	0	1	1	1	4
3	h4	1	0	1	1	0	3
4	h5	1	0	1	0	0	2
5	h6	1	1	1	1	1	5

	U	e2	e5
0	h1	1	1
1	h2	1	0
2	h3	0	1
3	h4	0	0
4	h5	0	0
5	h6	1	1

Figure 6: Soft set **Figure** 7: Output of parameterization algorithm

Above figure 6 is the input data given to the parameterization algorithm that consist of six houses and produced the output as shown in figure 7.So using only two parameters e_2 and e_5, houses h_1 and h_6 is the optimal selection using the parameterization reduction concept.

	e1	e2	e3	e4	e5	e6	e7	e8	sum
0	0	0	0	0	0	0	0	0	0
1	0	0	0	0	0	0	0	0	0
2	0	1	0	0	0	0	0	0	1
3	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0	0	0	0

Figure 8: Output of Parameterization value reduction (PVR)

Above figure 8 shows the output of PVR algorithm in which h_3 house is the optimal choice selection which has the maximum choice value and writes "0" for all remaining objects.

	houses	e1	e2	e4	e5	e6	e7	sum
1	2	0	1	1	0	1	1	4
0	1	1	0	1	1	0	0	3
5	6	0	1	1	1	0	0	3
2	3	0	0	0	1	0	1	2
3	4	1	0	1	0	0	0	2
4	5	1	0	0	0	1	0	2

Figure 9: Output of new algorithm for selection of house

Above figure 9 shows the output of the new method to select the best house using the decision parameter *d*. In this algorithm, the core parameters need to identify for the reduction of data here e_3 becomes dispensable parameter. The classification patterns is same as after deleting e_3 is $h_{2,}$ h_1, $h_{6,}$ $h_{3,}$ h_4 and $h_5$which gives h_2 as the highest choice value selection for the house.

	e1	e2	e3	e4	e5	e6	e7	e8	sum
0	1	0	1	1	1	1	0	1	6
1	0	0	1	1	0	0	1	1	4
2	1	0	1	0	1	1	1	0	5
3	1	0	1	1	1	1	0	1	6
4	1	0	1	1	1	1	0	1	6
5	0	0	1	0	0	1	1	1	4

Figure 10: Soft set

	e4	e5	e6	sum
0	1	1	1	3
1	1	0	0	1
2	0	1	1	2
3	1	1	1	3
4	1	1	1	3
5	0	0	1	1

Figure 11: Output of NPR

Figure 10 is the input soft set to NPR algorithm. Here with the help of only three parameter or symptoms e_4, e_5 and e_6, only three patient p_1 , p_4 and $p_5$which has maximum f() value makes the optimal solution.It means that with the help of only three symptoms e_4= fainting, e_5= numbness and e_6= throbbing doctor can detect only three patient having Thrombocythemia disease with the maximum value of f()= 3. It means that paitent p_1, p_4 and p_5 having maximum chance of thrombocythemia disease with the help of only three symptoms as e_4, e_5 and e_6 as shown in figure 11.

	e4	e1	e6	sum
0	1	1	1	3
1	1	0	0	1
2	0	1	1	2
3	1	1	1	3
4	1	1	1	3
5	0	0	1	1

Figure 12: Output of NENPR Algorithm

Figure 12 shows the output of NENPR algorithm in which with help of three symptoms e_1, e_4 and e_6, three patients' p_1, p_4 and $p_5$has thrombocythemia disease which has maximum value of 3.

	e4	e8	e1	e6	sum
0	1	1	1	1	4
1	1	1	0	0	2
2	0	0	1	1	2
3	1	1	1	1	4
4	1	1	1	1	4
5	0	1	0	1	2

Figure 13: Output of ANPR Algorithm

Here also with help of four parameters e_1, e_4, e_6 and $e_{8,}$ three patients p_1, p_4 and p_5 have thrombocythemia disease which has maximum value of 4 as shown in figure 13.

Table 5: Output of RBPR Algorithm

U/E	e_4	e_6	e_8	*F()*
p_1	1	1	1	3
p_2	1	0	1	2
p_3	0	1	0	1
p_4	1	1	1	3
p_5	1	1	1	3
p_6	0	1	1	2

Above table shows the output of ranked based parameter reduction in which with the help of only three parameters optimal selection of three patients are done as p_1, p_4 and p_5 are selected who has high chances of thrombocythemia disease with the high *f()* value as 3

Table 6: Comparison of parameter reduction methods for the selection of house

Sr. No	Name of Algorithm	Advantage	Disadvantage
1	Parameterization reduction algorithm [2]	Proposed first soft set algorithm for parameter reduction to solve decision making problem.	Reduction of parameters of this algorithm is not correctly calculated.
2	Parameterization Value Reduction Algorithm[7]	First time proposed the concept of parameterization value reduction for least parameter values.	There is no any suboptimal choice selection.
3	New Algorithm for parameter reduction[8]	With the help of classification pattern the selection of object is very easy.	Less number of parameters reduced using this algorithm.

Table 7: Comparison of parameter reduction methods for the selection of patients for the Thrombocythemia disease

Sr. No	Name of Algorithm	Advantage	Disadvantage
1	Normal parameter Reduction algorithm(NPR) [4]	The problem of suboptimal choice and added parameter is solved to give the exact optimal choice selection.	Requires more computation for parameter reduction of data as O (n^3)

2	New Efficient Normal parameter Reduction algorithm(NENPR) [6]	Requires low computation as compared to NPR for parameter reduction of data as O (n^2)	This algorithm does not consider the same value parameter (such as $e_i = e_j$) for the reduction purpose.
3	Alternative approach to Normal parameter Reduction algorithm (ANPR) [14]	Requires low computation for parameter reduction of data as O (n^2) and solve the problem of both NPR and NENPR.	Does not always consider last choices for selection.
4	Ranked Based Parameter Reduction Algorithm	Easy to implement as compare to existing parameter reduction algorithms.	Does not always consider last choices for selection

6.Conclusion

The most recent mathematical tool which is being developed to solve the problem of existing theory is called as soft set theory. The problem of knowledge discovery in decision making process is also handled with the help of the concept of this soft set. Literature survey shows the various parameter reduction methods based on the Soft Set theory to give the optimal selection. Methodology of various parameter reduction algorithms based on the soft set theory is also discussed with the help of practical implementation in machine learning platform using python. New approach to select the optimal parameters based on ranking called as ranked based soft set algorithm for parameter reduction is proposed that gives the better performance than existing algorithms

References

1. D. Molodtsov, “Soft set theory—First results,” Comput. Math. with Appl., vol. 37, no. 4–5, pp. 19–31, 1999.
2. P. K. Maji, A. R. Roy, and R. Biswas, “An application of soft sets in a decision making problem,” Comput. Math. with Appl., vol. 44, no. 8–9, pp. 1077–1083, 2002.
3. C. Dgang, T. E.C.C, Y. D. S., and W. Xizhao, “The Parameterization Reduction of Soft Sets and its Applications,” Int. J. Comput. Math. with Appl., vol. 49, pp. 757–763, 2005.
4. Z. Kong, G. Liqun, W. Lifu, and L. Steven, “The normal parameter reduction of soft sets and its algorithm,” Comput. Math. with Appl., vol. 56, pp. 3029–3037, 2008.
5. T. Herawan, A. N. M. Rose, and M. Mat Deris, “Soft set theoretic approach for dimensionality reduction,” in Communications in Computer and Information Science, 2009, vol. 64, pp. 171–178.
6. X. Ma, N. Sulaiman, H. Qin, T. Herawan, and J. M. Zain, “A new efficient normal parameter reduction algorithm of soft sets,” Comput. Math. with Appl., vol. 62, no. 2, pp. 588–598, 2011.
7. X. Ma, N. Sulaiman, and H. Qin, “Parameterization Value Reduction of Soft Sets and its algorithm,” 2011 IEEE Colloq. Humanit. Sci. Eng., no. Chuser, pp. 261–264, 2011.
8. M. I. Ali, “Another view on reduction of parameters in soft set,” Appl. Soft Comput., vol. 212, pp. 1814–1821, 2012.
9. Z. Kong, L. Wang, and Z. Wu, “Two cases based on normal parameter reduction in soft sets,” in Proceedings - 2012 International Conference on Computer Science and Electronics Engineering, ICCSEE 2012, 2012, vol. 3, pp. 577–581.
10. D. A. Kumar and R. Rengasamy, “Parameterization reduction using soft set theory for better decision making,” in Proceedings of the 2013 International Conference on Pattern Recognition, Informatics and Mobile Engineering, PRIME 2013, 2013, pp. 365–367.
11. B. Han and X. Li, “Propositional Compilation for All Normal Parameter Reduction of a Soft Sets,” springer, pp. 184–193, 2014.
12. Z. Li, N. Xie, and G. Wen, “Soft coverings and their parameter reductions,” Appl. Soft Comput. J., vol. 31, pp. 48–60, 2015.
13. T. Bakshi, A. Sinharay, and T. Som, “An introduction towards automated parameterization reduction of soft set,” 2016 3rd Int. Conf. Recent Adv. Inf. Technol. RAIT 2016, pp. 164–171, 2016.
14. S. Danjuma, M. A. Ismail, and T. Herawan, “An alternative approach to normal parameter reduction algorithm for soft set theory,” IEEE Access, vol. 5, pp. 4732–4746, 2017.
15. S. Danjuma, T. Herawan, M. A. Ismail, H. Chiroma, A. I. Abubakar, and A. M. Zeki, “A Review on Soft Set-Based Parameter Reduction and Decision Making,” IEEE Access, vol. 5, pp. 4671–4689, 2017.

16. X. Ma and H. Qin, "Soft Set Based Parameter Value Reduction for Decision Making Application," IEEE Access, vol. 7, no. c, pp. 35499–35511, 2019.