International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-4, February 2020

A Soft Computing Method for Typical Computation in Medical Data

Janmejay pant, Manoj Kumar Singh, Amit Juyal, Himanshu Pant, Chetan Pandey

***Abstract**: There are several methods of Soft Computing for analyzing the complex data for making decisions and predictions. Rough Set Theory (RST) is one of the best and relatively new intelligent techniques used in different research area for making predictions. RST is used to discover the patterns of data, handle all the redundant objects and attributes. RST is majorly used for extraction the rules from the given data. In this paper, we will use a medical data set example of cancer for retrieving the rule which is useful to make prediction for the unknown class.*

***Keywords**: Rough Set Theory, Information System, Decision class, Indiscernibility, Discernibility matrix, Equivalence class, Rules*

I. INTRODUCTION

As we know the growth of data is increasing rapidly. So it is almost impossible to extract hidden patterns and knowledge manually in the form of rules due to rapid growth of data. To overcome this problem researchers always prefer automatic ways of extracting rules using soft computing. A medical data base contains a huge amount of information about the patient's records and their disease conditions [1].

New medical knowledge can be retrieved through the relationship and patterns within the data. The computation of medical data often requires dealing with inconsistency, incompleteness. It also deals with the manipulation of various levels of representation of data. The existing intelligent techniques [2][3][4] are basically depends on assumptions. These techniques are not much able to retrieve conclusion from the incomplete knowledge. The most commonly used techniques of soft computing in medical data computation are Artificial Neural Network [5], Bayesian Classifier [6], Genetic Algorithms, Decision Tree [2] and Fuzzy Set Theory [7]. In this order RST [8] is introduced by Professor Z. Pawlak in the early 1980.

Rough Set theory is a mathematical tool [8][9][10] for dealing the complex data.

Revised Manuscript Received on February 28, 2020.
* Correspondence Author
Janmejay Pant*, Dept. of Computer Science, Graphic Era Hill University, Bhimtal, India., Email: jpant@gehu.ac.in
Manoj Kumar Singh, Dept. of Computer Science, Graphic Era Hill University, Bhimtal, India., Email: manojthakur1984@rediffmail.com
Amit Juyal , Dept. of Computer Science, Graphic Era Hill University, Dehradun, India., Email: amitjuyal26@gmail.com
Himanshu Pant, Dept. of Computer Science, Graphic Era Hill University, Bhimtal, India, Email: himanshupantbareilly@gmail.com
Chetan Pandey, , Dept. of Computer Science, Graphic Era Hill University, Dehradun, India., Email: schetanpandey@gmail.com

This theory is also used for extracting useful information in terms of knowledge from the complex data having uncertainty and incompleteness.

The RST assumes that the user has much enough information and knowledge about the objects in the universe. These information and knowledge are helpful for dividing the objects into different forms of group. If two objects having the exactly same information then we can say that they are indiscernible or similar in nature [8] [9] [11] or similar. It means we cannot differentiate them with known information or knowledge.

The RST is used to find the dependency among data i.e. how data are dependent to each other , express the importance of attributes, and recognize the hidden patterns, to find decision making rules, reduce redundant attributes, and find the minimum subset of attributes for classification [1]. The algorithms of RST can deal to approximate the decision classes using large patterns [12]. RST has become very popular among researchers in current era. It is now one of the most developed techniques in smart data analyses. One of the important characteristics of RST is it does not require any external parameters unlike other techniques such as Fuzzy set. RST uses only the information given by the data [9][10][11]. In this paper, we will discuss how RST can be used to compute the uncertain data and how we can generate rules from a set of 18 observations of cancer patients.

The whole paper is organized as: The basic concept of rough set theory, description of sample data, and discussion about the experimental results using methodology. After that we will extract reducts from the equivalence classes and Discernibility matrix will be constructed. At last Decision rules are generated based on three extracted reducts.

II. ROUGH SET: BASIC CONCEPT

The RST [8][9][10][11][13] is a mathematical tool for computing the complex data stored in tabular form. An information System [11][13][14] is a table which can be divided in two parts rows and columns. Rows are observations or objects and columns are features or attributes [13]. RST observes the data in terms of equivalence classes [1][13]. An indiscernible group of objects is called equivalence classes and Rough Set Theory treats data in the form of Equivalence classes [8][9][11][13] or similar with respect to the attributes.

Rough set cannot be defined by a unique or single class as Rough Set is a collection of observations and an observation partially belongs to with at least one of the equivalence class [13]. It may hence only be approximately described by two levels of approximations named as lower and upper approximation.

These classes (equivalence) are completely part of the set and classes (equivalence) having at least one object is part of the set are known as lower approximation and upper approximation respectively [13].The above statements can be justified by the given figure.

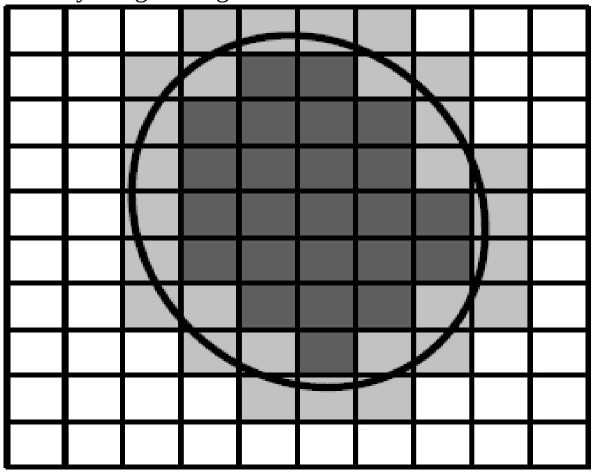

Fig. 1. Roughness in Decision System [13]

It is clear, in Fig.1 that roughness is there to represent the ellipse by the squares which are equivalence classes. It is defined by lower approximation shown as dark grey and upper approximation shown as dark and light grey. Since there is roughness in ellipse so it is known as rough set.

Decision Attribute

An Information System is combination of conditional attributes and decision attributes. The decisions attribute [13] is a target attribute and all the objects of the observations may belong to a single or multiple decision classes. All objects are divided by this target attribute into decision class [13][14].
A training set is formed by combining the information system with the decision attribute. So it is also called decision system.

Indiscernibility Relation

Indiscernibility can be defined as
Let us consider IS = (U, A) is an information system. Where U= {(o_1, o_2, o_3…... o_n)} is defined as the object of finite set of n tuples. This finite set is supposed to be a non empty finite set. A= {A_1, A2, A3…A_n} is defined as a finite set of attributes which is non empty set [13][12]. By the definition of indiscernibility [13] a group of objects in which all the attribute values are with same set then this set can be known as P indiscernible objects[8][12][13]. In the above section we have discussed that in RST data can be utilized as equivalence classes, that means a set of objects are indiscernible with respect to attributes'[8[13][14].

III. AN EXAMPLE OF DECISION TABLE AS A DATA SET

In this paper we are taking a decision table or system to analyze the disease for our computation.
Table 1 is an example of a decision system having the 18 cancer patient data. This data i.e. objects can be divides into two parts: attributes and decision or target class. In our example four properties or attributes of a patient are used.
The recorded attributes about the patients are the expression having three genes with different values and whether patient smokes or not. The value of three genes is: 0 for unchanged, ↓ for down-regulated, ↑ for up-regulated. Decision class having two values which are Lung and Colon. Lung and Colon may be represented by L and C respectively.

Table 1: An Example of Decision Table

Patients	Gene 1	Gene 2	Gene 3	Smoking	Site of Origin
Pa1	↓	↓	0	Yes	Lung
Pa2	0	0	0	Yes	Lung
Pa3	0	↓	↑	No	Colon
Pa4	0	0	0	Yes	Lung
Pa5	0	↓	0	Yes	Lung
Pa6	↓	↓	0	Yes	Lung
Pa7	↓	↑	0	No	Colon
Pa8	↓	↑	0	No	Colon
Pa9	0	↑	0	Yes	Colon
Pa10	↓	↓	↑	No	Lung
Pa11	0	↓	0	Yes	Lung
Pa12	0	↓	0	Yes	Lung
Pa13	0	↓	↑	No	Colon
Pa14	0	↑	↑	No	Colon
Pa15	↓	↑	0	NO	Colon
Pa16	↓	↓	↑	No	Colon
Pa17	0	↓	0	Yes	Lung
Pa18	0	↓	↑	No	Lung

In the above decision table Pa1 to Pa18 are objects or observations. Gene 1, Gene 2, Gene 3 and smoking are conditional attributes and Site of Origin is target attribute.

IV. METHODOLOGY

In this paper we will use RST for finding the rules in medical data that is describe in table. In order to achieve decision rules we first build equivalence classes having the similar attributes with similar decision class. Total eight equivalence classes are build names as Eq1 to Eq8. Table 2 defines the equivalence classes. Based on the equivalence class lower and upper approximations are defined. How equivalence classes may represent in the terms of lower and upper approximations is shown in Fig.2. A Discernibility matrix is constructed in table3 to calculate reducts. At last IF-THEN rules are constructed using Boolean reasoning and Rough Set Theory.

V. EXPERIMENTAL RESULTS

In this Section we discuss the implementation of results using the concept of RST. All the experiments are described in the following sections and results are shown in Table2, Fig 2, Table 3, Table 4 and Table 5.

A.Generation of Equivalence Classes

In our decision system in table 1 –
Patients Pa5, Pa11, Pa12 and Pa17 are similar or indiscernible about the recorded attributes and decision class. This statement implies that these patients having the similar conditional attributes and all these patients belong to the same decision class. Therefore these patients belong to the same class. Patients Pa3, Pa13 and Pa18 also form an equivalence class as the attributes values of these patients are similar.

Retrieval Number: D9078019420/2020©BEIESP
DOI: 10.35940/ijitee.D9078.029420
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

However these patients belong to different decision classes.Generalized decision is expressed by set of all decision classes which are related to the equivalence class. In this case {C, L} is a set of decision classes.

So this practice of making equivalence classes continue for the given decision table and finally we may summaries the given decision system in terms of equivalence classes. The summarized Decision System in terms of equivalence classes may be considered as follows:

Table 2: Decision Classes in terms of Equivalence Classes

Equivalence Classes	Gene 1	Gene 2	Gene 3	Smoking	Site of origin
Eq1={Pa1,Pa6}	↓	↓	0	Yes	{L}
Eq2={Pa2,Pa4}	0	0	0	Yes	{L}
Eq3={Pa3,Pa13,Pa18}	0	↓	↑	No	{C,L}
Eq4={Pa5,Pa11,Pa12,Pa17}	0	↓	0	Yes	{L}
Eq5={Pa7,Pa8,Pa15}	↓	↑	0	No	{C}
Eq6={Pa9}	0	↑	0	Yes	{C}
Eq7={Pa10,Pa16}	↓	↓	↑	No	{C,L}
Eq8={Pa14}	0	↑	↑	No	{C}

It is clear from the table 2 that the patients having both Lung and Colon as target values of class site of origin can be defined as rough because these patients cannot be represented in terms of equivalence classes in a unique way. These classes can be expressed in terms of lower and upper approximations [13][14].

For example, we can define decision class {L} by all equivalence classes in which patients have the same target value L in the decision class. In the above table 2 equivalence classes {Eq_1, Eq_2, and Eq_4} have the same target value L in the decision class. These equivalence classes can be defined as lower approximation. We can also define decision class {L} by all equivalence classes in which at least one patient belongs to the decision class Lung L. Theses equivalence classes can be defined as upper approximations. In the above table equivalence classes {Eq_1, Eq_2, Eq_3, Eq_4 and Eq_7} can be represented as upper approximations.

The concept of lower and upper approximation [13][22][26] of the given table can be easily understood by the figure 2.

In this section we have discussed how roughness may occur in the decision table and how we can make equivalence classes based on the decision attribute. How equivalence classes may represent in the terms of lower and upper approximations. Now by doing this exercise we are able to define the decision class of each patient.

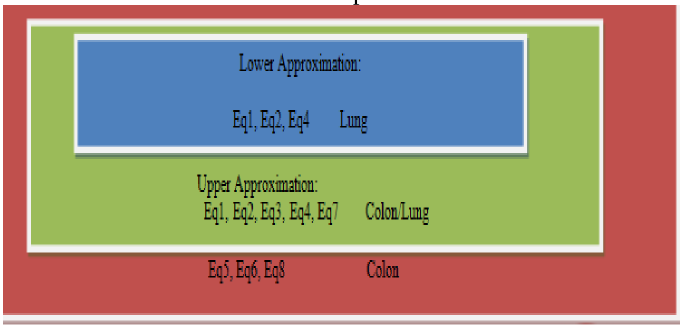

Fig.2. Lower and Upper Approximations

B. Reduct

In this section we will discuss how data can be reduced by Boolean reasoning [14][15][16] in a decision system

To maintain the indiscernibility relations among the objects of a data sample or decision table, a minimum group of attributes are required. This minimal group of attributes is known as reduct [13].Let A= (A_1, A_2, A_3…) is finite set of attributes and U= (a_1, a_2, a_3…) is a set of objects. Then the decision table or Information System can be defined as *I (U.A)* [13].Based on the definition of reduct it defined as a minimal set of attributes $B \subseteq A$ such that $IND_I(B) = IND_I(A)$[13].A Boolean function can be constructed for each object. A Boolean function always evaluates to true or false. It is known as Discernibility function. This function returns true for all attributes that are dissimilar from other objects having different target value. Solutions interpreted as so-called reducts [14][15][17][22].The minimum group of attributes that discerns one objects from other with a different decision or target value. There are various algorithms based on approximation through which reducts can be extracted. These approximation algorithms include greedy algorithm [11] and genetic algorithms.For extracting the reducts, let us consider table 2. Where equivalence class of the decision system is are mentioned. A Discernibility function is defined by creating the Discernibility matrix by using all the equivalence classes [18][19][20][21]. This matrix specifies which attribute that discerns the different equivalence classes. Table 3 describes the Discernibility characteristics among all the attributes.

Retrieval Number: D9078019420/2020©BEIESP
DOI: 10.35940/ijitee.D9078.029420
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Table 3. Discernibility Matrix

	Eq1	**Eq2**	**Eq3**	**Eq4**	**Eq5**	**Eq6**	**Eq7**	**Eq8**
Eq1	ϕ							
Eq2	ϕ	ϕ						
Eq3	Ge1,Ge3,Sm	Ge2,Ge3,Sm	ϕ					
Eq4	ϕ	ϕ	Ge3,Sm	ϕ				
Eq5	Ge2,Sm	Ge1,Ge2,Sm	Ge1,Ge2,Ge 3	Ge1,Ge2,Sm	ϕ			
Eq6	Ge1,Ge2	Ge2	G2,G3,Sm	Ge2	ϕ	ϕ		
Eq7	Ge3,Sm	Ge1,Ge2,Ge3,Sm	ϕ	Ge1,Ge3,Sm	Ge2,Ge3	Ge1,Ge2,Ge3,Sm	ϕ	
Eq8	Ge1,Ge2,Ge3,Sm	Ge2,Ge3,Sm	Ge2	Ge2,Ge3,Sm	ϕ	ϕ	Ge1,Ge2	ϕ

The following points may be noticed from the table 2:

- The entry Eq1-Eq2 is empty (ϕ), because the equivalence classes E1 cannot be discerned from itself.
- Eq1-Eq2 is also empty because both have the same decision class. So no need to discern equivalence classes with the same generalized decision.
- The entry Eq1-Eq3 has different decision class (in table 2). The values of attributes of Eq1 and Eq3 are different. The value of Ge1 (Gene 1) is down regulated, value of Ge3 is unchanged and value of Smoking is yes for Eq1. While Eq3 has the different values for the same attributes. The value of Ge1 (Gene 1) is unchanged, value of Ge3 is up-regulated and value of Smoking is No for Eq3.
- Based on the attribute values we can develop Discernibility matrix. This matrix is symmetric. i.e. Entries Eq1-Eq4 and Eq4-Eq1 are identical. Entries Eq1-Eq5 and Eq5-Eq1 are identical and so on. So we need to construct on half of the matrix.

After getting the Discernibility matrix, minimal information are required to discern E1 from all other objects with different decision. This information can be represented by Discernibility function that is defined here as:

f_{E1} (Ge_1, Ge_2, Ge_3, S_m) = (Ge_1 OR Ge_3 OR S_m) AND (Ge_2 OR S_m) AND (Ge_1 OR Ge_2) AND (Ge_3 OR S_m) AND (Ge_1 OR Ge_2 OR Ge_3 OR S_m).

In order for this function to be true, at least one of the attributes from each Eq1-related entry (blue) in the Discernibility matrix need to be included. Thus, the function can be simplified to:

f_{E1} (Ge_1, Ge_2, Ge_3, S_m) = (Ge_1 AND S_m) OR (Ge_1 AND S_m) OR (Ge_2 AND S_m). The reducts reflected by the Discernibility function are as below: {Ge_2, Ge_3}, {Ge_1, S_m} and {Ge_2, S_m}. It should be noticed that each reduct having the minimal set of attributes that are available in all the ANDs in the Discernibility function.

C.Decision Rules

The decision rules [22][23][24] may be constructed once reducts are extracted. The rules can be generated by expressing each attribute by its value of a reduct.

In IF- THEN rule, if part is known as antecedent i.e. a_1=v_1 AND a_2=v_2, where a_1 and a_2 are attributes in reduct and v_1 and v_2 are attribute values. IF part is associate with THEN in decision rule. THEN part is known as consequent i.e. d_1=d_2 OR d=d_2 where d is a target attribute and d_1 and d_2 are target values. Rules can be measured by their accuracy and coverage.

Basically accuracy interpret the rules in the terms of specific way i.e. fraction of similar objects that are from the target class of the THEN part. And coverage is how general they are i.e. fraction of target class objects in THEN part matching same in the IF part.

The three minimal set of attributes known as reduct {Ge_2, Ge_3}, {Ge_1, S_m} and {Ge_2, S_m} discerning equivalence class Eq1 in the above section in table 2.

The following rules can be generated through the reducts for Eq1.

First we construct the rule later on we will describe about the rules.

Table 4. Rule Generation

		Support	Accuracy	Coverage
R1	IF Gene2(↓) AND Gene3(0) **THEN** Site (Lung)	6	1.0(6/6)	0.60(6/10)
R2	IF Gene1(↓) AND Smoking(Yes) **THEN** Site (Lung)	2	1.0(2/2)	0.20(2/10)
R3	IF Gene2(↓) AND Smoking(Yes) **THEN** Site (Lung)	6	1.0(6/6)	0.60(6/10)

Here we can see Table 2, 6 objects match in rule R1. So the support count is 6 i.e. patients Pa1, Pa5, Pa6, Pa11, Pa12 and Pa17.

These objects have matching values in IF part and matching values in THEN part of Rule 1. The THEN part (site of origin) for all objects is Lung.

That's why the accuracy is 100% or 1.0. THEN parts of 10 objects that is decision class are similar as LUNG of rule R1 in Table 2.

Retrieval Number: D9078019420/2020©BEIESP
DOI: 10.35940/ijitee.D9078.029420
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
and Sciences Publication (BEIESP)

Out of 10 6 objects have the matching IF part of rule R1. So we define the coverage for rule R1 as 10/6=0.6. Thus the very first rule R1 elaborates the decision class Lung having the 60% of matching objects.Same procedure is applied to calculate the accuracy and coverage of all the other rules i.e. R2 and R3. The rules having high accuracy and high coverage is preferred. So from table 4 rule R1 is better than R2 Rest of rules can be generated by applying the same procedure which we have used to find rule R1, R2 and R3.

Table 5.Rule Generation

		Support	Accuracy	Coverage
R4	**IF** Gene 2(0) **THEN** Site (Lung)	2	1.0	0.20
R5	**IF** Gene2(↓) AND Gene3(↑) **THEN** Site (Lung) **OR** Site (Colon)	2,3	0.4,0.6	0.20,0.38
R6	**IF** Gene2(↓) AND Smoking(No) **THEN** Site (Lung) **OR** Site (Colon)	2,3	0.4,0.6	0.20,0.38
R7	**IF** Gene3(0) AND Smoking(No) **THEN** Site (Colon)	3	1.0	0.38
R8	**IF** Gene2(↑) **THEN** Site (Colon)	5	1.0	0.60
R9	**IF** Gene1(↓) AND Gene3(↑) **THEN** Site (Lung) OR Site (Colon)	1,1	0.5,0.5	0.10,0.13

Thus these above rules construct a model for classification. . Nine rules have been generated for this model. This model can be used for the two main purposes. First on is predictive purpose and another one is descriptive purpose.
Rules are beneficial for prediction. Predictive purpose can be used to make predictions for the unknown patients. Descriptive purpose can be used to separate patients with different site of origin.

VI. CONCLUSION

In this paper we have used soft computing for computation in cancer data example of 18 tuples. We have used RST to generate the rules. For this first equivalence classes are constructed then reducts are extracted based on Discernibility matrix and finally 9 IF- THEN rules are constructed along their support, accuracy and coverage. These rules can be useful for making predictions for the new observation.

VII. FUTURE SCOPE

In our work we have used a technique of Soft Computing for rule generation. In the current research era soft computing is the first interest of researchers in medical science. We can also extend our work by using other soft computing techniques like ANN, Fuzzy, neuro fuzzy, Soft Set [15][19] etc. We also can extend this work by doing classification for the large data set with better accuracy

ACKNOWLEDGMENT

The first author is very much thankful to Prof R P Pant, Dept. of Mathematics, DSB Campus, Kumaun University Nainital and Prof. Durgesh Pant, UOU, Haldwani and Dr. Ashutosh Bhatt, Associate Prof. BIAS, Bhimtal for their continuous guidance and valuable suggestations.

REFERENCES

1. Aboul-Ella Hassanien, "Rough Set Approach for Attribute Reduction and Rule Generation: A Case of Patients With Suspected Breast Cancer", JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY, 55(11):954–962, 2004.
2. Hassanien, A.-E. (2003a),"Classification and feature selection of breast cancer data based on decision tree algorithm", International Journal of Studies in Informatics and Control Journal, 12, 33–39.
3. Lavraic, N., Keravnou, E., & Zupan, B. (1997)," Intelligent data analysis in medicine and pharmacology", Boston/Dordrecht: Kluwer.
4. Wolf, S., Oliver, H., Herbert, S., & Michael, M. (2000, August), "Intelligent data mining for medical quality management", Proceedings of the Fifth International Workshop on Intelligent Data Analysis in Medicine and Pharmacology (IDAMAP-2000), Berlin, Germany European Society on Artificial Intelligence.
5. Choi, S.-E., & Rockett, P.," The training of neural classifiers with Condensed datasets",IEEE Transactions on Systems, Man, and Cybernetics, Part B, 32(2), 202–206,2002.
6. Cheeseman, P., & Stutz, J.," Bayesian classification (autoclass): Theory and results", In U.M. Fayyad, G. Piatetsky-Shapiro, P. Smyth, & R. Uthurusamy (Eds.), Advances in knowledge discovery and data Mining. Boston: AAAI Press/MIT Press, 1996.
7. Parido, A., & Bonelli, P,". A new approach to fuzzy classifier systems", In Lori Costello (Ed.), Proceedings of the Fifth International Conference on Genetic Algorithms (ICGA '93) (pp. 223–230), University of Illinois at Urbana-Champaign, (1993, July 17–21).
8. Pawlak, Z. ,"Rough sets", International Journal of Computer and Information Science", 11, 341–356, 1982.
9. Priyanka Suyal, Janmejay Pant, Akhilesh Dwivedi, Manoj Chandra Lohani, "Performance evaluation of rough set-based classification models to intrusion detection system", Advances in Computing, Communication, & Automation (ICACCA)(Fall), IEEE International Conference, 2016/9/30, 1-6.
10. Janmejay pant, Kamlesh Padaliya, Himanshu Pant,"Rough Set Approach for Feature Selection in IDS, International Journal of Innovations & Advancement in Computer Science, Volume 4,8-11,2015.
11. Janmejay Pant, R P Pant, Amit Juyal, Himanshu Pant, "Rule Generation Method for Medical Data (Heart disease) using Rough Set Approach (March 14,2019)",International Conference on Advances in Engineering Science Management & Technology(ICAESMT)-2019, Uttranchal University, Dehradun, India, Available at SSRN: https://ssrn.com/abstract=3384652.
12. Grzymala-Busse, J., Pawlak, Z., Slowinski, R., & Ziarko, W.,"Rough sets", Communications of the ACM, 38(11) 89–95, 1999.
13. Samir Roy,Udit Chakraborty, "Introduction to Soft Computing", Pearson Education India, 2013 - 604 pages .
14. Janmejay Pant,Puspha Pant, Ashutosh Bhatt, Himanshu Pant,Nirmal Pandey," Feature Selection towards Soil Classification in the context of Fertility classes using Machine Learning", International Journal of Innovative Technology and Exploring Engineering (IJITEE),Volume 8,issue 12,pages 4000-4004,2019.
15. Janmejay Pant, Amit Juyal, Himanshu Pant, Akhilesh Dwivedi, "A Real Time Application of Soft Set in Parameterization Reduction for Decision Making Problem ".International Journal of Electrical and Computer Engineering (IJECE), 2017, Vol. 7, No. 1, pp. 324~329 ISSN: 2088-8708, DOI: 10.11591/ijece.v7i1.pp324-329.

Indu Yekkala, Sunanda Dixit," Prediction of Heart Disease Using Random Forest and Rough Set Based Feature Selection", International Journal of Big Data and Analytics in Healthcare Volume 3 • Issue 1 • January-June 2018.
16. Siri Krishan Wasan, Vasudha Bhatnagar," THE IMPACT OF DATA MINING TECHNIQUES ON MEDICAL DIAGNOSTICS", Data Science Journal, Volume 5, 19 October 2006.
17. Aboul Ella Hassanien, Ajith Abraham, James F. Peters, Gerald Schaefer," Rough Sets in Medical Informatics Applications," Applications of Soft Computing, AISC 58, pp. 23–30, Springer-Verlag Berlin Heidelberg 2009.
18. Janmejay Pant, Amit Juyal , Shivani Bahuguna,"Soft set, a soft Computing Approach for Dimensionality Reduction", International Journal of Innovative Science, Engineering & Technology,2015,Vol. 2, Issue 4.
19. Aboul Ella Hassanien, Mohamed E. Abdelhafez," Rough Sets Data Analysis in Knowledge Discovery: A Case of Kuwaiti Diabetic Children Patients", Hindawi Publishing Corporation, Advances in Fuzzy Systems Volume 2008, Article ID 528461, 13 pages
20. Rahman Ali,Muhammad Hameed Siddiqi," Rough set-based approaches for discretization: a compact review", Springer Science+Business Media Dordrecht, 2015.
21. P K Srimani, Manjula Sanjay Koti" Rough set (RS) approach for optimal rule generation in medical data", International Journal of Conceptions on Computing and Information Technology Vol.2, Issue 2, April' 2014; ISSN: 2345 – 9808.
22. Thanh-Trung Nguyen, Phi-Khu Nguyen," Reducing Attributes in Rough Set Theory with the Viewpoint of Mining Frequent Patterns", (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 4, 2013
23. P.K Maji,R. Biswas,"Soft Set Theory",International Journal of Computers & Mathematics with Applications",555-562,2003.
24. G Ilczuk, R Mlynarski," Rough Set Techniques for Medical Diagnosis Systems", Computers in Cardiology 2005; 32:837−840, IEEE..

AUTHORS PROFILE

Janmejay Pant is pursuing his Doctorate in Information Technology from Kumaun University Nainital and done M.tech in Information Technology from Graphic Era deemed to be University, Dehradun. Currently he is working as Assistant Professor in Dept. of computer science at Graphic Era Hill University, Bhimtal Campus. He has total Academic teaching experience of more than 10 years with more than 25 publications in reputed National and International SCOPUS, UGC Approve Journals and conferences. His research area includes Machine Learning, Soft Computing, and Data Mining and Deep Learning. He is an active member of CSI. He secured First Rank in Uttarakhand State Eligibility Test (USET) in Computer Science and Applications in 2018. He also received Research Award at University Level in September 2019.

Manoj Kumar Singh is pursuing Doctorate in Computer Science from Graphic Era Deemed to be University, Dehradun. He received his M.tech in Computer Science from Graphic Era Deemed to be University. His areas of interest are Machine Learning, Soft computing and Internet of Things. He has published various papers in the different journals. Currently he is working as Assistant professor in Graphic Era Hill University Bhimtal. He has an experience of around 11 years of teaching.

Amit Juyal is pursuing Doctorate in Computer Science from Graphic Era Deemed to be University, Dehradun. He received his M.tech in Computer Science from Graphic Era Deemed to be University. His areas of interest are Machine Learning; Soft computing .He has published various papers in the different journals with good impact factor. Currently he is working as Assistant professor in Graphic Era Hill University Dehradun. He has an experience of around 11 years of teaching.

Himanshu Pant is pursuing Doctorate from Graphic Era Hill University. He received his MCA from UPTU Lucknow. His areas of interest are Machine Learning, Soft computing and data mining. He has published various papers in the different journals. Currently he is working as Assistant professor in Graphic Era Hill University Bhimtal. He has an experience of around 07 years of teaching

Chetan Pandey is an Assistant Professor in the School of Computing, Graphic Era Hill University, Dehradun (UK), India. He has received Master in Technology in CSE from Graphic Era Deemed to be University, Dehradun (UK), India. His research interests are Data Processing and Machine Learning.