JOURNAL OF NORTHEASTERN UNIVERSITY

Volume 25 Issue 04, 2022 ISSN: 1005-3026 https://dbdxxb.cn/ Original Research Paper

REDUCT BASED CLASSIFICATION OF PATIENTS USING ROUGH SET THEORY-AN APPLICATION IN HOMEOPATHY

Gayatri Bhupalrao Patil[1], Dr. Ajit More[2]

[1] Research Scholar, Department of Computer Application, Bharati Vidyapeeth Deemed to be University, Pune, Maharashtra, India.

[2] Department of Computer Application, Bharati Vidyapeeth Deemed to be University, Pune, Maharashtra, India.

Abstract:

Rough Set Theory a mathematical tool to handle ambiguous, incomplete and vague data has proved to be effective in the field of medical science. The process of remedy diagnosis is a tedious job for a homeopath but concept of reduct and rule generation in the Rough Set Theory can effectively used. This research aims on classification of patient in respective group from among the existing groups.The mental state of the patient is of more importance when the study is in homoepathy. By the reduction of attributes and application of selected rules a model can be generated that can easily classify the accurate group of the patient.

Keywords: Rough Set Theory, Classification, Homeopathy.

Introduction:

Diagnosis of disease is very essential and important task in the field of medicine. If the diagnosis is accurate the medicine can be accurate otherwise it may cause inconvenience to the patients as well as reputation and expertise of the doctor can be questionable. Progress of computer science and technology development has a great impact on the field of medicine. Many computer aided diagnostic systems (CAD) help the physicians. The most essential part of such CAD is classification.

Classification is the task of assigning labels to the unlabelled data instances and a classifier is used to perform such a task. The classifier has to be efficient enough to classify the object in the respective class correctly. To achieve the accuracy, extract data, and reduce the data we can implement the Rough Set Theory (RST). RST is a mathematical tool for extracting knowledge from uncertain and incomplete databased information. This theory became very popular among scientist around the world and the RS is now one of the most intelligent data analysis tools. This is the reason RST to become more impactful in the field of medicine. Homeopathy is a medical science with a holistic approach and self-healing. The remedy just boosts the patients' immunity against the disease. 100% results can be achieved if exact remedy match is found. There is a specific method or approach through which one can reach to an exact remedy. The attribute reduction and rule generation are important features of Rough Set Theory. By applying these features and getting the reduct set we can reach till accurate/ matching remedy. Most commonly neural networks, Bayesian classifier, genetic algorithms, decision trees, fuzzy theory are intelligent techniques used in medical data analytics. In this contribution the Rough Set Theory is introduced; RST was introduced by polish professor Z. Pawlak early eighties. This theory became very popular among scientist and this is now one of the most developing

Submitted: 17/10/2022 **Accepted**: 14/11/2022

intelligent data analysis theory. Attribute reduction is one of the important features in RST. This feature allows selecting only those features, which are essential for the study, but at the same time, the result does not affect. Basic concepts of RST:

1. Indescernibility relation
2. Approximations
3. Rough membership
4. Dependency of attributes
5. Reduct and core

Objective of the study:

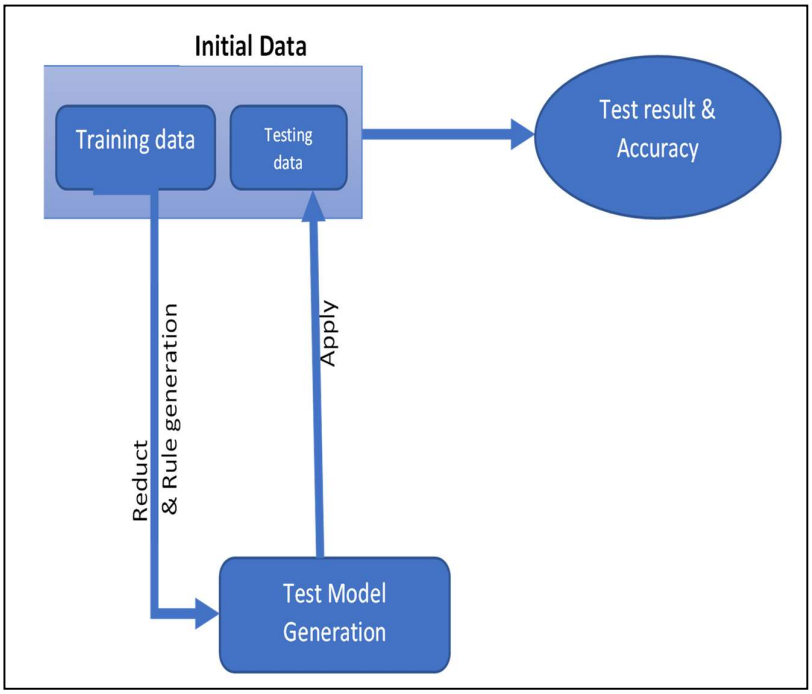
1. To classify the patient's symptoms and provide a reliable and more accurate solution for remedy/ group of the patient using homeopathy.
2. To design a system/ algorithm for automation of classification based on rough sets.
3. To test the system with training and test data sets and validate this model with real life data in homeopathy.
4. To find out most important or valuable attributes which can help the homeopath to come to an accurate remedy group.

Data collection:

This study is based on the real-life cases of patients suffering from arthritis. The patients are from almost every i.e. 30-70 and above age group. The sample comprises of 500 patients. The data covers all the personal information on their site of pain, type if pain and the mental state of the patient. Among the 500 records 300 are used as a training data set, to classify the given patient in an appropriate class. The remaining 200 records are used to validate or test the developed model. Weka, a collection of state of art ML algorithms and data processing tool is used for pre-processing of data. Decision rules and reduct are developed by using RSES. RSES is a software tool with an easy-to-use interface and at the same time featuring a bunch of methods that make it possible to perform compound, non-trivial experiments in data exploration with use of RS methods.

Proposed rough set based reduct and decision rule generation (RSD) method:

Diagram 1: Research design



The procedure for computing reduct and decision rules is as follows:

1. A data set is considered containing information about patients after taking the case (detailed information). This is the raw data of the patient.
2. The information system (IS) or table is constructed by considering each patient as an object and their information as an attribute. Information system is represented as {P, A} where set of objects P= {P1, P2, P3……., P500} say P is the patient 1 for classification and A= {A1, A2, A3……, An} representing set of attributes for a patient.
3. The data pre-processing is done in WEKA, for checking missing values, noisy data and other inconsistences before executing it to the algorithm.
4. The Reduct set R is generated by RSES which is a reduced subset of the original attribute set which retains the accuracy of the original set.
5. Final part in RS analysis is generating decision rules; based on reduct R. A decision rule consists of IF (condition) THEN element belongs to 'DECISION CLASS'. Using these decision rules we can classify the patient in his/her respective and correct class.

For example, consider the following table containing information of patients as objects in decision systems to generate decision rules. Here patient ID is unique (serial number). Reduct set has 30 different attributes and class is the decision attribute.

Table 1: Information system (IS)

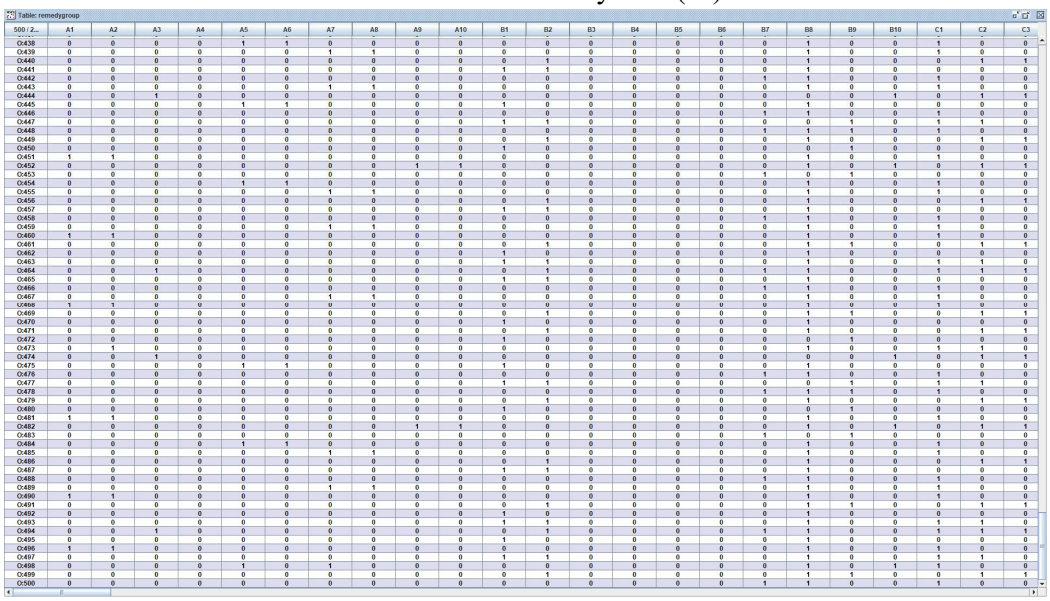


Table 2: Reduct Set

The 10 reduct set is found which consists of 30 attributes, each set contains 3 attributes. The very important and mandatory attributes when decision rules are generated. The data is divided as testing data and training data.

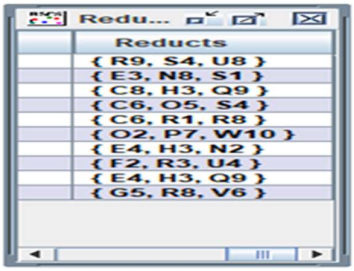


Table 3: Confusion matrix

	POTASSIUM KALI GROUP(PKG)	CALCAREA GROUP(CG)	NATRUM GROUP(NG)	MAGNESIUM GROUP(MG)	
POTASSIUM KALI GROUP(PKG)	47	02	01	00	50
CALCAREA GROUP(CG)	00	49	00	01	50
NATRUM GROUP(NG)	02	02	46	00	50
MAGNESIUM GROUP(MG)	00	01	00	49	50
	49	54	47	51	200

Among 500 records, 300 records are training data and 200 are testing data. A model developed with the help of reduct and decision rules generated. The model was exposed to the training data set. For the verification of results confusion matrix was used. The result is as shown is table 2. These results show that the rough set methodology gives efficient result in the classification of the data set.

Results:

TP:191

Overall Accuracy: 95.5%

CLASS	n(truth)	n(classified)	Accuracy	Precision	Recall	F1Score
POTASSIUM KALI GROUP(PKG)	49	50	97.5%	0.94	0.96	0.95
CALCAREA GROUP(CG)	54	50	97%	0.98	0.91	0.94
NATRUM GROUP(NG)	47	50	97.5%	0.92	0.98	0.95
MAGNESIUM GROUP(MG)	51	50	99%	0.98	0.98	0.98

Table 4: Results

Conclusion:

In this study we presented a rough set methodology determining decision rules and reduct helpful for the correct classification of patient in his respective class/ group. The study shows that 191 instances were correctly classified with the overall accuracy of 95.5% This system will help homeopaths reach to a correct remedy once the patient's group is correctly identified. The model reduces the time and skip the unnecessary data or redundant data which is unimportant for class-identification. For this study 4 major groups are chosen which commonly appear in the population; other groups can be considered as the further scope of study.

References:

1. A. Manimaran1,*, v. M. Chandrasekaran2, aishwarya asesh Rough set approach for an efficient medical diagnosis system School of advanced sciences,vit university, vellore-6320141,2
2. M.Durairaj1*, T.Sathyavathi21*,2Applying Rough Set Theory for Medical Informatics Data Analysis Department of Computer Science, Engineering andTechnology, Bharathidasan University,Tamilnadu, India.
3. zbigniew suraj h. Sucharskiego str rzeszów, polandZsuraj An introduction to rough set theory and its applications.
4. Raj K Manchanda Building quality research evidence in Homoeopathy Editor http://www.ijrh.org on October 11, 2018, IP: 42.106.207.131
5. K PrasannaJyothi Dr R SivaRanjani DrTusarKanti Mishra S Ranjan Mishra A Study of Classification Techniques of Data Mining Techniques in Health Related Research
6. International Journal of Innovative Research in Computerand Communication Engineering ISSN(Online) 2320-9801ISSN (Print): 2320-9798 Vol. 5, Issue 7, July 2017.
7. 1.Qinghua Zhang 2. Qin Xie 3.Guoyin Wang A Survey On Rough Set Theory And Its Applications Science Direct CAAI Transactions on Intelligence Technology 1 (2016) 323e333
8. A.Manimaran1,*, V. M. Chandrasekaran2, Aishwarya Asesh3 School of Advanced Sciences, VIT University, ,2School of Computing Sciences and Engineering Rough set approach for an efficient medical diagnosis system
9. S. Udhaya Kumara, H.HannahInbaranib A Novel Neighborhood Rough set Based Classification Approach for Medical Diagnosis 1877-0509 © 2015 The Authors. Published by Elsevier B.V. doi: 10.1016/j.procs.2015.03.216 ScienceDirect
10. 1.Kareem Kamal A.Ghany 2.Heba Ayeldeen 3. Hossam M. Zawbaa 4. Olfat Shaker Rough set based reasoner for medical diagnosis 978-1-4673-7910-6/15/$31.00 c 2015 IEEE 2015 InternationalConference on Green Computing and Internet of Things (ICGCIoT)
11. MuktaMajumder, Utsav Barman, Rahul Prasad, Kumar Saurabh, Sujan Kumar Saha A Novel Technique for Name Identification from Homeopathy Diagnosis Discussion Forum Sciverse Science Direct Elsevier Ltd.2nd International Conference on Communication, Computing & Security [ICCCS-2012]

12. JyotiSoni Ujma Ansari Dipesh Sharma SunitaSoni Predictive Data Mining for Medical Diagnosis: An Overviewof Heart Disease Prediction International Journal of Computer Applications Volume 17– No.8, March 2011
13. Mrs.Yogita Bhapkar Bvdu, Yashwantrao Mohite College, Pune, India ,Dr.Ajit More Mca Director, Bvdu, India Volume-Iv, Issue-Vi Issn Comparative Analysis Of Data Mining Algorithms For Classification Of Home Loan Applications : 2350-0476 Issn (Print): 2394-207x Impact Factor: 4.205 International Journal Of Multifaceted And Multilingual Studies.