IMPACT OF BAYESIAN DECISION THEORETIC ROUGH SET-BASED ATTRIBUTE REDUCTION ON CART

Utpal Pal[1], Sharmistha Bhattacharya (Halder)[2]

[1]Research Scholar, [2]Associate Professor

[1, 2]Tripura University (A Central University), Tripura, India

***Abstract:** The classical Decision Tree algorithms have been around for decades and Classification and Regression Tree (CART) is a most powerful technique among these, especially for classification and predictive analysis. But, the accuracy of CART degrades while dealing with high dimensional large data sets. Bayesian Decision Theoretic Rough Set (BDTRS) is a sophisticated rough set based method for attribute reduction. This paper represents the impact of BDTRS-based attribute reduction on CART. This paper also studies CART based on Pawlak rough set model. Experimental result, on R environment, using three different data sets shows that the BDTRS-based CART constructs improved decision tree for better classification efficiency.*

***Keywords:** CART, BDTRS, Pawlak rough set, Attribute Reduction, R.*

I. INTRODUCTION

The CART [3, 14, 21] is a sophisticated Decision Tree [15, 16] algorithm. It is direct viewing, can extract knowledge rules preferably and suitable to deal with classification problems [26]. It outperforms several other popular decision tree algorithms with respect to simplicity, comprehensibility, classification accuracy and being able to handle mixed-type data. Moreover CART efficiently deals with missing values, uses cost-complexity pruning strategy and can handle outliers. But, when the data has large number of attributes and involves impurity, the decision tree constitutive property is poor and difficult to find some information that could have been found and be useful. Another problem with the CART cross validation method is that it can be computationally too expensive, because it requires the growing and pruning of auxiliary trees as well. In order to overcome these drawbacks, rough set based attribute reduction has been introduced. Ever since the introduction of rough set theory by Pawlak [13] in 1982, many extensions have been made [20, 23, 24, 27]. The BDTRS [1, 2, 11, 13] is an excellent rough set based tool for attribute reduction. The indiscernibility relation and set approximation remains unaltered before and after reduction. The positive region computed, involves maximum number of objects. As a result the reduced attribute set involves all the significant attributes eliminating all insignificant attributes. This paper also studies CART based on Pawlak rough set [13] for attribute reduction. In this research work we have used R [5, 9, 17], version 3.2.2, for experimentation and implementation of BDTRS. Data sets have been taken from UCI Machine learning repository.

In Section 2, Pawlak rough set, BDTRS and CART are introduced. Section 3, discusses about processing steps, algorithm and execution process in R environment. Experimental results and concluding remarks are presented in Section 4 and Section 5 respectively.

II. THEORETICAL BACKGROUND

2.1 Information System [13, 22]

An information system is defined as: $S = (U, At, \{V_a \mid a \in At\}, \{I_a \mid a \in At\})$, where, U is a finite nonempty set of objects, At is a finite nonempty set of attributes, V_a is a nonempty set values of $a \in At$ and $I_a: U \to V_a$ is an information function that maps an object in U to exactly one value in V_a.

2.2 Indiscernibility [13, 10]

The indiscernibility relation is intended to express the fact that due to the lack of knowledge we are unable to discern some objects employing the available information. The notion of equivalence is recalled first. A binary relation $R \subseteq X \times X$ which is reflexive (i.e. an object is in relation with itself x R x), symmetric (if x R y then y R x) and transitive(if x R y and y R z then x R z) is called an equivalence relation. The equivalence class of an element x $\in$X consists of all objects y$\in$X such that xRy. For any $B \subseteq A$, there is associated an equivalence relation $IND_S(B)$:

$$IND_S(B) = \{(x, x') \in U \times U \mid \forall a \in B, a(x) = a(x')\}.$$

$IND_S(B)$ is called the B-indiscernibility relation. If $(x, x') \in IND_S(B)$, then objects x and x' are indiscernible from each other by attributes from B. The equivalence classes of the B- indiscernibility relation are denoted by $[x]_B$.

2.3 Process of Reduction

Discernibility matrix [19]: Given an information system the discernibility matrix of S is a symmetric $n \times n$ matrix with entries C_{ij} as given below:

$$C_{ij} = \{a \in A \mid I_a(x_i) \neq I_a(x_j)\} \text{ for } i,\ j = 1, 2, ..., n.$$

Discernibility function [19]: Each entry thus consists of the set of attributes upon which objects x_i and x_j differ. A discernibility function f_A for an information system S is a Boolean function of m Boolean variables $a^*_1,\ a^*_2,\ ..,\ a^*_m$ (corresponding to the attributes $a_1,\ a_2,\ ..,\ a_m$) defined as follows:

$f_A(a^*_1,\ a^*_2,\ ..,\ a^*_m) = \wedge\{\vee\, c^*_{ij} \mid 1 \leq j \leq i \leq n, c_{ij} \neq \phi\}$, where $c^*_{ij} = \{a^* \mid a \in c_{ij}\}$.

2.4 Attribute Significance [13]

The consistency factor is defined as $\gamma(C, D) = |POS_C(D)|/|U|$. The decision table is consistent if $\gamma(C, D) = 1$. The significance, $\sigma(a)$, of any attributes a, can be defined as:

$$\sigma(C, D)(a) = \big(\gamma(C, D) - \gamma(C - \{a\}, D)\big)/\gamma(C, D) = 1 - (\gamma(C - \{a\}, D)/\gamma(C, D)).$$

Where, $0 \leq \sigma(a) \leq 1$.

2.5 Pawlak's Rough Set Model [13, 22]

The upper and lower approximation of X with respect to equivalence relation R with condition probability are denoted as $\overline{R}X$ and $\underline{R}X$, respectively, and defined as follows:

$$\overline{R}X = \cup\{[x]_R \mid P(X/[x]_R) > 0,\ [x]_R \in \pi_R\}.$$

$$\underline{R}X = \cup\{[x]_R \mid P(X/[x]_R) = 1,\ [x]_R \in \pi_R\}.$$

Three kinds of approximation regions of X with respect to A can be defined according to its upper and lower approximation, respectively.

Positive region: $POS_R(X) = \underline{R}X = \cup\left\{[x]_R | P\left({}^{X}/_{[x]_R}\right) = 1, [x]_R \in \pi_R\right\}.$

Negative region: $NEG_R(X) = U - \overline{R}X = \cup\left\{[x]_R | P\left({}^{X}/_{[x]_R}\right) = 0, [x]_R \in \pi_R\right\}.$

Boundary region: $BND_R(X) = \overline{R}X - \underline{R}X = \cup\left\{[x]_R | \ 0 < P\left({}^{X}/_{[x]_R}\right) < 1, [x]_R \in \pi_R\right\}.$

2.6 Bayesian Decision Theoretic Rough Set [1, 2, 12, 19]

Let, D_{POS} denotes the positive region in BDTRS model. For an equivalence class, $[x]_c \in \pi_A$,

$$D_{POS}([x]_c) = \{D_i \in \pi_D : P(D_i/[x]_c) > P(D_i)\}.$$

For equivalence classes $[x]_c$ and $[y]_c$ the elements of a positive decision-based discernibility matrix, $M_{D_{pos}}$ is defined as follows.

$$M_{D_{POS}}([x]_c, [y]_c) = \{a \in C :\ I_a(x)\ \neq I_a(y) \wedge D_{POS}([x]_c) \neq D_{POS}([y]_c)\}.$$

A positive decision reduct is a prime implicant of the reduced disjunctive form of the discernibility function.

$$f\left(M_{D_{POS}}\right) = \wedge \left\{\vee\left(M_{D_{POS}}([x]_c, [y]_c)\right) :\ \forall\, x, y \in U\left(M_{D_{POS}}([x]_c, [y]_c) \neq \emptyset\right)\right\}.$$

In order to derive the reduced disjunctive form, the discernibility function $f\left(M_{D_{POS}}\right)$ is transformed by using the absorption and distributive laws. Accordingly, finding the set of reducts can be modeled based on the manipulation of a Boolean function.

2.7 Decision Tree: CART [3, 14, 15, 21]

The CART algorithm starts with the initial decision table(data set) D, attribute set At and *gini-index*, as attribute selection method. Initially, it creates a node, N, which incorporates D. If all the objects of D belong to same class, node N is returned as leaf node. Otherwise, select a condition attribute (that maximizes the reduction in impurity of D) that divides D, in a manner such that height of the tree is as small as possible. The node N is labeled with the selected attribute. After that, branches are grown from N for each of the outcomes of the splitting attributes. The algorithm works recursively on each subset of D. Recursion may stop in one of these cases:

- All the objects of a node belong to a particular class (same class).
- There are no attributes left on which the objects may be further be separated.
- There are no objects for a particular branch and a partition is empty.

III. IMPLEMENTATION OF BDTRS-BASED CART

3.1 Processing Steps

This section shows the implementation and execution procedure for BDTRS-based CART. BDTRS-based decision tree induction is performed in two basic steps. First, attribute reduction and second, decision tree induction using the reduced information. Procedure for attribute reduction is shown in Algorithm-1(AttReduction ()). Case: 1 and Case: 2 represent BDTRS and Pawlak rough set respectively for attribute reduction. Based on rough set theory equivalent classes [13] are computed. Procedure for computation of positive region, discernibility matrix [19], discernibility function [19] and reduced attribute set are shown in Algorithm-1. The discernibility function is a conjunction over the disjunction of the matrix elements. The function can be transformed in to a reduced attribute set using absorption and distributive laws of Boolean algebra. The classical CART is implemented in package 'rpart' of R. Induction of decision tree using R commands has been shown in section 3.3.

3.2 Algorithm

Algorithm-1. AttReduction (*DT*):
Input: Decision Table (*DT*), Decision column (*D*), Data Matrix (D_m).
Output: Reduced date set, D_r.
Variables D_{pos} , represents Positive Region. *X*, *objIdx*, *objIdx1*, *objIdx2* represents intermediate variables.

1. Compute [IND], the Indiscernible relation/Equivalent classes from *DT*.
2. Repeat step 3 for [IND$_i$], i=1 to n, n being the number of Equivalent classes.
3. Repeat 3.1 to 3.2 for each decision (D_j), j=1 to k, k is number of unique decision.
   Case: 1 [For Bayesian Decision Theoretic Rough Set]
   3.1. X = (Objects in [IND$_i$] having D_j as decision ÷ Objects in [IND$_i$])
   3.2. If ($X > P(D_j)$) [$P(D_j)$ is the probability of D_{jth} Decision]
   Assign Decision D_j, to D_{pos}, corresponding to each object of [IND$_i$]
   Case: 2 [For Pawlak Rough Set]
   3.1. X = (Objects in [IND$_i$] having D_j as decision ÷ Objects in [IND$_i$])
   3.2. If (X = =1)
   Assign Decision D_j, to D_{pos}, corresponding to each object of [IND$_i$].
4. Repeat for i^{th} object in *DT*, where i = 1 to (n-1), n is the number of objects in *DT*
   4.1. Find the objects (in *DT*) whose decision in *D* doesn't match with the decision of i^{th} object. [i.e. *objIdx1* = (decVector[i]!=(decVector[i+1] to decVector[n]))]
   4.2. Find objects (in *DT*) whose decision in D_{pos} doesn't match with the decision of i^{th} object.
   [i.e. *objIdx2* = (D_{pos}[i] != (D_{pos}[i+1] to D_{pos}[n]))]
   4.3. *objIdx* = { *objIdx1* ∩ *objIdx2* } [Common objects].
   4.4. Repeat for j^{th} object in *objIdx*, where j = 1 to no. of objects in *objIdx*
   4.4.1 A_l = Column names of D_m where Column name of i^{th} object != Column name of j^{th} object..
   4.4.2 $D_{mat} = A_l$ [Assign the attribute list A_l to Discernibility Matrix D_{mat}].
5. Compute discernibility function, D_f from D_{mat}
6. Transform D_f into reduced attribute set, A_r (using laws of Boolean Algebra)
7. D_r = Data set corresponding to the attributes of A_r.
8. Return (D_r).

3.3 Execution of CART and BDTRS-Based CART in R Environment

Decision Tree Induction Using CART. This section explains decision tree induction by taking *housing* [8] as example data set. Installation procedure of R and relevant packages ("RoughSets"[18], "rpart" etc.) are available in Comprehensive R Archive Network (CRAN). In order to perform attribute reduction, raw data (in .txt, .csv, .xlsx format) is first converted into *data frame* object which is then converted into *DecisionTable* format. For this, functions like *read.table*(), *SF.asDecisionTable*()[18] may be used. The package "rpart" is installed using the R command: > library(rpart). Similarly other packages like "rattle", "caret", "RoughSets" etc. are also installed. The Decision Tree, T_1 obtained from the following commands is shown in Figure 1.

```
>HData = RoughSetData$housing.dt#Using housing DataSet
>train.flag = createDataPartition(y = HData $MEDV,p=0.7,list=FALSE) #Training and test sample
>training = HData [train.flag,]
>modfit = train(MEDV~.,method="rpart",data=training) #Building the Mmodel
>fancyRpartPlot(modfit$finalModel)    #DrawCART Decision Tree
```

Decision Tree Induction Using BDTRS**-based CART.** BDTRS-based CART is inducted in two steps. Step 1: Performs data reduction by BDTRS. Step2: Decision Tree induction based on the output of Step 1. Following R commands are executed on *housing* data set (*housing.csv* format). Reduced attribute set is shown in Table 1.

```
>HousingFrm = read.table ("housing.csv", header = TRUE, sep = ",") #Computation of DataFrame
>HousingDecTable = SF.asDecisionTable (HousingFrm, decision.attr = 14, indx.nominal = c(1:13) #to Decision table
>BDTRS_PosRegion = BayesDtPos (HousingDecTable, c(1:13))
```

>BDTRS_DisMatrix = BayesDtDisMat (HousingDecTable, BDTRS_PosRegion, range.object = NULL, return.matrix =TRUE)# Computation of Discernibility Matrix
>ReducedDataSet=FS.one.reduct.computation(BDTRS_DisMatrix)

The procedure for decision tree induction is same as CART and hence not repeated. The decision tree (T_2) thus obtained is shown in Figure 1.

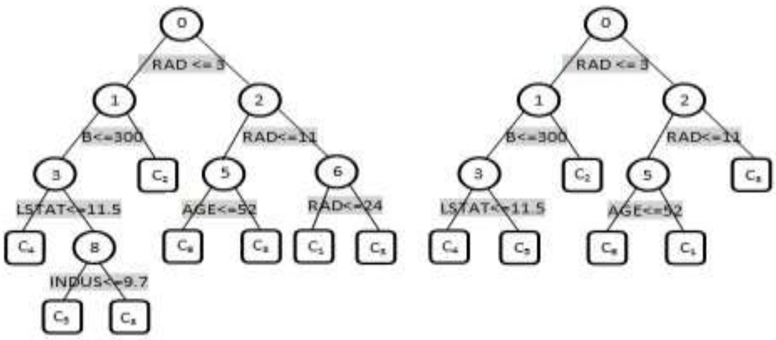

Fig. 1. Decision Tree T_1 using CART method (Left hand side), before attribute reduction and Decision Tree T_2 using BDTRS-CART() method (Right hand side) after attribute reduction.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

The data sets used for experimentation are: *Audiology* (226 objects with 69 attributes) [28], *Cervical-Cancer* (858 objects with 36 attributes) [29] and *housing* (506 objects with 14 attributes). The housing data set is already introduced in the previous section. Each of these data sets has been studied by BDTRS and Pawlak rough set. For housing data, at first, we pre-processed the sample data (*D*) and filled in missing values using built in functions available in R. Next, we run the original algorithm CART, on *D*, to construct a Decision Tree T_1 as shown in Figure 1.

Table 1. Attribute reduction using BDTRS model on *housing* dataset.

Attribute set of *D* before reduction	**Total**	**Attribute set of D_{red} after reduction**	**Total**
{RAD, RM, DIS, B, AGE, CRIM, CHAS, LSTAT, TAX, ZN, NOX, PTRATIO, INDUS}	13	{RAD, RM, DIS, B, AGE, CRIM, CHAS, LSTAT, TAX}	09
Computed attribute significance in ascending order: B: 0.75, RAD: 0.74, RM: 0.74, LSTAT: 0.70, AGE: 0.68, CHAS: 0.50, CRIM: 0.50, TAX: 0.49 DIS: 0.42, ZN: 0.12, NOX: 0.08, PTRATIO: 0.06, INDUS: 0.02.			

After that, we run Case: 1 of Algorithm-1 (attribute reduction by BDTRS) to reduce the insignificant attributes of *D*. We reduced the number of attributes down to 9, saved the new sample data set as D_{red}. The deleted attributes are: 'ZN', 'NOX', 'PTRATIO' and 'INDUS'. We computed attribute significance of all the attributes to show that deleted attributes have little effect on decision making. This is shown in Table 1. We further computed consistency factor (C.F) (shown in Table 2) which remains same (one) before and after attribute reduction. This ensures that the integrity of the data set remains unchanged after attribute reduction. Finally, we run CART on the reduced data set D_{red}, to construct a simplified decision tree, T_2 (Figure 1). The above mentioned experimental procedure is repeated using the functions AttReduction() (Algorithm-1) and CART method for *Audiology* and *Cervical-Cancer* data sets. The results obtained from the computations are shown in Table 2.

Table 2. Comparison of CART, BDTRS-CART and Pawlak-CART using *Audiology*, *Cervical-Cancer* and *housing* data sets.

Data Set used	Model used	No. of conditi-on Attribu -tees	Depth of Tree	No. of Leav -es	Average Length of Rules	Averag-e Exe. Time in Seconds	Classifi-cation Accurac y	C.F
Audiol-gy	CART	69	19	18	18.42	0.98	86.33	1
	BDTRS-CART	51	17	16	16.03	1.28	92.79	1
	Pawlak-CART	56	18	17	17.11	1.12	90.16	1
Cervic-al Cancer	CART	36	10	09	9.22	0.72	88.56	1
	BDTRS-CART	27	09	08	8.03	1.11	91.68	1
	Pawlak-CART	32	10	09	8.15	1.06	90.05	1
housing	CART	13	05	04	3.45	0.30	88.65	1
	BDTRS-CART	09	04	03	2.88	0.65	88.86	1
	Pawlak-CART	12	05	04	2.98	0.66	88.86	1

For *housing* data, tree T_1, shows that the classification tree has 15 nodes (7 internal and 8 leaf nodes). On the other hand, Tree T_2, of Figure 1 shows that the classification tree has 11 nodes (5 internal and 6 leaf nodes). The domain of decision column (Minimum: 5 and Maximum: 50) is divided in to six equivalent classes: C_1, C_2, C_3, C_4, C_5 and C_6. There are eight (8) classification rules corresponding to the leaf nodes of T_1 and six (6) rules corresponds to T_2. For example, the classification rule: "If (RAD <= 3) and (B <= 300) and (LSTAT >11.5) and (INDUS<=9.7) then MEDV = C_5", corresponds to leaf node C_5 of tree T_1. Similar results obtained from the other two data sets and the comparison of CART, BDTRS-CART and Pawlak-CART methods are shown in the Table 2. It can be observed from Table 2 that CART model deals with the original unreduced data set, whereas, other three approaches work on the reduced data set. The BDTRS method gives the best attribute reduction. As a result, number of nodes, number of leaf nodes, depth and average length of the classification rules has been decreased for all the three approaches except classical CART Model. This ultimately improves the classification accuracy of the decision trees as shown in Figure 2.

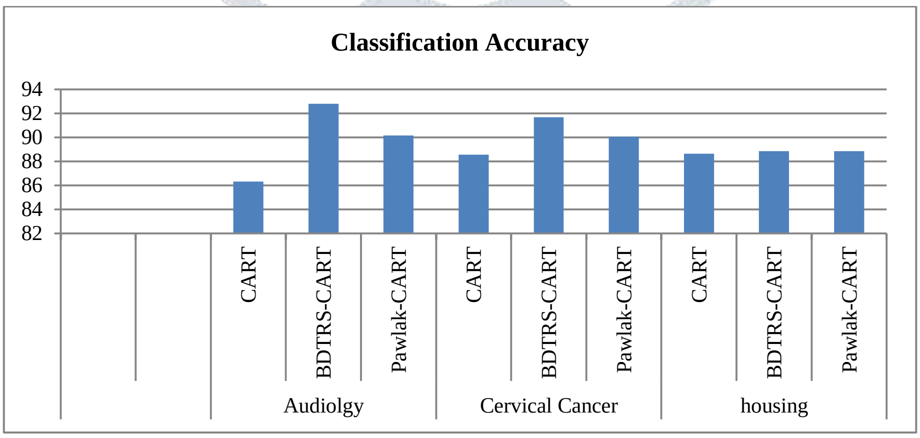

Fig. 2. Classification accuracy of CART, BDTRS-CART and Pawlak based CART on *Cervical Cancer*, *Audiolgy* and *housing* data sets.

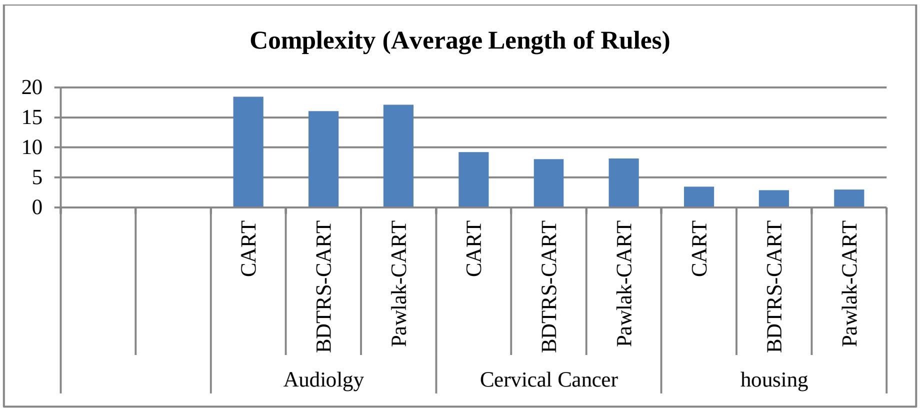

Fig. 3. Complexity representation of CART, BDTRS-CART and Pawlak based CART on *Cervical Cancer*, *Audiolgy* and *housing* data sets.

On the other hand, the rough set based decision tree approach suffers in terms of total execution time as it involves the attribute reduction phase also. The minimal increase of execution time is acceptable, keeping the classification accuracy and reduction of tree complexity (shown in Fig. 3) in mind. Tree complexity mainly depends on the average length of the decision rules which is lowest in case of BDTRS-based CART method for all data sets.

V. Conclusion

Efficiency of CART-based decision tree becomes an issue of concern to deal with high-dimensional large data sets. This study focuses on reducing number of insignificant attributes from the original data set before induction of CART-based decision tree. The reduced attribute set preserves the indiscernibility relation and set approximation. This is ensured by computing attribute significance and consistency factor. In this research work we have also implemented the BDTRS using R language in order to study the classical CART. The experimental results show that the decision tree induced by BDTRS-based CART is the simplest and most efficient in terms of depth, number of nodes, average rule length and classification accuracy compared to the other methods mentioned in this work.

REFERENCE

[1] Bhattacharya(Halder), S.: A Study on Bayesian Decision Theoretic Rough Set, International Journal of Rough Sets and Data Analysis (IJRSDA) 1.1, 1-14 (2014).
[2] Bhattacharya(Halder), S., Debnath, K.: Attribute Reduction Using Bayesian Decision Theoretic Rough Set Models. Int. Journal of Rough Sets and Data Analysis (IJRSDA) 1.1, 15-31 (2014).
[3] Breiman, L., Friedman, J. H., Olshen, R. A., Stone, C. J.: Classification and regression trees. Monterey, CA: Wadsworth & Brooks/Cole Advanced Books & Software, (1984). ISBN 978-0-412-04841-8.
[4] Chang, C. D., Chien-Chih, W., Bernard, C. J.: Using data mining techniques for multi-diseases prediction modeling of hypertension and hyperlipidemia by common risk factors, Expert Systems with Applications 38.5, 5507-5513 (2011).
[5] Ferraro, M. B., Giordani, P.: A toolbox for fuzzy clustering using the R programming language, Fuzzy Sets and Systems(ELSEVIER), 279, 1–16 (2015).
[6] Gelfand, S.B., Ravishankar, C.S., Delp, E. J.: An iterative growing and pruning algorithm for classification tree design, Systems Man and Cybernetics. Conference Proceedings. IEEE International Conference, 818-823 (1989).
[7] Han, J., Kamber, M.: Data Mining: Concepts and Techniques, The Morgan Kaufmann Series in Data Management Systems, 2nd Edition, 291-309 (2006).
[8] Harrison, D., Rubinfeld, D. L.: Hedonic Prices and the Demand for Clean Air, Journal of Environmental Economics and Management, 5, 81-102 (1978).
[9] Ihaka, R., Gentleman, R.: R: a language for data analysis and graphics, Journal of Computational and Graphical Statistics, 5, 299–314 (1996).
[10] Orlowska, E., Pawlak, Z.: Measurement and indiscernibility, Bulletin of Polish Academy of Sciences, Mathematics, 32, 617-624 (1984).
[11] Pal, U., Bhattacharya(Halder), S., Debnath, K.: A Study on Maximum Probabilistic Based Rough Set (MPBRS), LNAI-Springer, MIKE-2017, 10682, pp. 1–12 (2017). https://doi.org/10.1007/978-3-319-71928-3_39.
[12] Pal, U., Bhattacharya(Halder), S., Debnath, K.: R Implementation of Bayesian Decision Theoretic Rough Set Model for Attribute Reduction, LNNS-Springer, I3SET, 11, (2017). doi.org/10.1007/978-981-10-3953-9_44.
[13] Pawlak, Z.: Rough Sets, International Journal of Computer and Information Sciences, 11, 341-356 (1982).
[14] Questier, F., Put, R., Coomans, D., Walczak, B., Vander Heyden, Y.: The use of CART and multivariate regression trees for supervised and unsupervised feature selection, Chemometrics and Intelligent Laboratory Systems, 76(1), 45-54 (2005), ISSN 0169-7439, http://dx.doi.org/10.1016/j.chemolab.2004.09.003.
[15] Quinlan, J. R.: Induction of Decision Trees, Machine Learning, 1, 81-106 (1986).
[16] Quinlan, J. R.: Simplifying Decision Trees, Int. J. Man-Machine Studies, 27, 221-234 (1987).
[17] R Development Core Team, R: A Language and Env. for Statistical Computing. Vienna, Austria, 2011: the R Foundation for Statistical Computing. ISBN: 3-900051-07-0. Available online at: http://www.R-project.org/. Last accessed on: 08-06-2016.
[18] Riza, L. S., Janusz, A., Bergmeir, C., Cornelis, C., Herrera, F., Slezak, D., Benitez, J. M.: Implementing algorithms of rough set theory and fuzzy rough set theory in the R package ‘‘RoughSets’’ , Information Sciences(ELSEVIER), 287, 68–89 (2014).

[19] Skowron, A., Rauszer, C.: The Discernibility Matrices and Functions in Information Systems, in: Slowiski R (ed), Intelligent Decision Support. Handbook of Applications and Advances of the Rough Set Theory, Kluwer Academic Publishers, Dordrecht, 311-362 (1992).

[20] Slezak, D., Ziarko, W.: Bayesian Rough set model, In: Proc. of the International Workshop on Foundation of Data mining, Japan, 131-135 (2002).

[21] Stuart, L. C.: Extensions to the CART algorithm, International Journal of Man-Machine Studies, 31(2) (1989) 197-217, ISSN 0020-7373, http://dx.doi.org/10.1016/0020-7373 (89)

[22] Yao, Y. Y.: Generalized rough set models, In: Polkowski, L., Skowron, A. (eds.), Rough Sets in Knowledge Discovery, Physica-Verlag, Heidelberg, 286-318 (1998).

[23] Yao, Y. Y.: Probabilistic approaches on rough sets, Expert Systems, 20, 287-297 (2003).

[24] Yao, Y. Y., Wong, S.K., Lingras, P.: A decision theoretic rough set model, In: Ras Z. W., Zemankova M., Emrich M. L.(Eds), Methodologies for intelligent systems, North Holland, New York, 5, 17-24 (1990).

[25] Zhengpin, X., Hua, D., Weixing, Y.: Based on decision tree and chart level superposition accurate agricultural output chart analysis method [J], Agricultural engineering journal, 22(8), 140-144 (2006).

[26] Zhiling, C., Qingmin, Z., Qinglian, Y.: A method based on rough set to construct decision tree, Journal of Nanjing University of Technology, 27, 80-83 (2005).

[27] Ziarko, W.: Variable precision rough set model, Journal of computer and System Sciences, 46, 39-59 (1993).

[28] https://archive.ics.uci.edu/ml/datasets/Audiology+%28Standardized%29.

[29] https://archive.ics.uci.edu/ml/datasets/Cervical+cancer+%28Risk+Factors%29.