Rough Set Theory-Based Multi-Class Decision Making Framework for Cost Effective Treatment of COVID-19 Suspected Patients

Sandip Majumder

National Institute of Technology Durgapur

Samarjit Kar (✉ dr.samarjitkar@gmail.com)

National Institute of Technology Durgapur <https://orcid.org/0000-0002-5503-9338>

Research Article

Keywords: COVID-19, rough membership function, probabilistic rough set, decision-theoretic rough, Three-way decision, Bayesian decision procedure

Posted Date: November 23rd, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-733073/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Rough Set Theory-Based Multi-Class Decision Making Framework for Cost Effective Treatment of COVID-19 Suspected Patients

Sandip Majumder, Samarjit Kar

Department of Mathematics, National Institute of Technology Durgapur, 713029, West Bengal, India

Corresponding author's email: samarjit.kar@maths.nitdgp.ac.in

ABSTRACT:

Rough set theory approximates a concept by the three regions, namely positive, negative and boundary regions. The three regions enable us to derive three types of decisions, namely acceptance, rejection and deferment. The deferment decision gives us the flexibility to further examine suspicious objects and reduce misclassification. The main objective of this paper is to provide a cost effective treatment of a patient suspect to COVID-19 positive by using multiclass three-way decision making with the help of Rough set theory. The cost-based analysis of three-way decisions brings the theory closer to real-world applications where costs play an indispensable role. In our approach, we extend the three-way decision to three-way multiclass decision, offering a new framework of multiple classes. Different types of misclassification errors are treated separately based on the notation of loss function from Bayesian decision theory. In our cost sensitive classification approach, the cost caused by a different kind of error are not assumed to be equal. Finally, a numerical example for a cost effective treatment of a patient with COVID-19 disease is considered to demonstrate the practicability and efficacy of the developed idea in real-life applications.

KEYWORDS: COVID-19, rough membership function, probabilistic rough set, decision-theoretic rough, Three-way decision, Bayesian decision procedure

1. Introduction

The theory of three-way decision has received much attention in recent years [41, 42, 43]. The theory of three-way decision concerns thinking in threes, working with threes, and processing through threes. It explores the effective uses of triads of three things, for example, three elements, three parts, three perspectives, and so on.

From a new semantic interpretation of the positive, boundary and negative regions, Yao [33] first introduce and study the notion of three-way decisions, consisting of positive, boundary and negative rules. The notion of three-way decisions represents a close relationship between rough set analysis, Bayesian decision analysis, and hypothesis testing in statistics [5, 9, 30, 31]. In a paper “The geometry of three-way decision” by Yao, in a nutshell, he very nicely shown that the three-way decision is about thinking, problem solving, and information processing based on a triad of three things.

The three-way decision rules come closer to the philosophy of the rough set theory, namely, representing a concept using three regions instead of two [19]. By considering three-way decision rules, one may appreciate the true and unique contributions of rough set theory to machine learning and rule induction. With the introduction of the notion of three-way decision, there is a new avenue of research for rough set theory.

The cost-based analysis of three-way decisions brings the theory closer to real-world applications where costs play an indispensable role. In this paper, we extended cost-based analysis of three-way decisions to three-way multiclass decisions to provide a cost effective treatment of a patient suspect to COVID-19 positive. Before going to the main theme of our paper, we take a look at the present scenario of COVID-19.

In late 2019, the world observed the outbreak of the novel coronavirus at Wuhan, China [44], and since then, it has had a rapid spread all over the world from the beginning of the following year. The COVID-19 is a highly infectious disease. After overcoming the first wave, we entered into the second wave. The number of vaccination is quite a few in comparison to their requirement. As the virus mutates continuously, such an available vaccine does not work correctly against some newly upcoming strain. The second wave of covid-19 is affecting most of the world. The scenario is very grim in India, where the daily count on April 15, 2021, itself is double the first peak. India leads the world in the daily average number of new infections reported, accounting for one in every two infections reported worldwide each day. Alarm bells ring as COVID-19 shifts base to villages. A large number of patients in the smaller centers could overwhelm the healthcare facilities, leading to an increase in deaths. The different graphical representations depicted in Fig. 1 – Fig. 2 shows that the present situation is horrible.

Researchers are continuously working to find the possible solutions for this epidemic problem to come out from this unprecedented situation. They are trying to dominate the virus in many ways, like inventing effective treatment, destroying the virus, or protecting it. Majumder et al. [15] describe a decision-making technique for identifying the infected population of COVID-19. Si et al. [29] proposed a decision-making method for selecting preferable medicine for the appropriate treatment of COVID-19 patients in a picture fuzzy environment through the hybrid approach of grey relational analysis and Dempster-Shafer theory. Mishra et al. [14] proposed an extended fuzzy decision-making framework using hesitant fuzzy sets for the drug selection to treat the mild symptoms of COVID-19.

Total COVID-19 tests per 1,000 people, Jun 25, 2021

The most recent figures for selection of countries is shown (you can change the selection using '+ Add country'). Only data relating to the the last 10 days are included.

Source: Official sources collated by Our World in Data

OurWorldInData.org/coronavirus • CC BY

Note: Comparisons of testing data across countries are affected by differences in the way the data are reported. Details can be found at our Testing Dataset page.

One of the major confusion is that some of the symptoms of flu and COVID-19 are similar; it may be hard to tell the difference between them based on symptoms alone. Both COVID-19 and flu can have varying degrees of signs and symptoms, ranging from no symptoms (asymptomatic) to severe symptoms. Only testing may be needed to help confirm a diagnosis. But diagnostic firms testing for coronavirus are nearing breaking point in the metro cities like New Delhi, Kolkata, Bangalore, Chennai and Mumbai as India battles its biggest surge in COVID-19, which may worsen the crisis as many sick people can't get tested fast enough to isolate themselves. Many complaints shows that the test report is negative, but a few days later, the patient condition becomes serious due to covid-19. But we believe that our awareness and proper treatment at the right time can save our life from this pandemic. In the present scenario, our approach provides a cost-sensitive solution to multiclass decision-making for the treatment of covid-19 suspected people.

The new interpretation of a three-way decision is not so critical in the classical rough set model since both positive and negative rules do not involve any uncertainty. It is essential in the probabilistic rough set models, where acceptance and rejection decisions are made with certain levels of tolerance for errors. The probabilistic rough set, which is a generalization of Pawlak's rough set, the same pair of thresholds determine the three regions, namely, positive, negative and boundary regions. Different pairs of thresholds have been discussed in several kinds of literature. However, it is difficult to relate these thresholds with practical notations of real-world applications, such as cost, risk, benefit etc. The required parameters of probabilistic rough set approximations can be systematically determined based on costs of various decisions. The incurred costs of decision rules can be analyzed. Moreover, the Bayesian decision procedure can be used to develop a decision-theoretic rough set model [32, 36, 37, 39, 40]. Based on the decision-theoretic rough set framework, Yao and Zhao [38] suggest changing a $m \times m$ classification problem into $m \times 2$ classification problem. Liu et al. [12] further discussed some practical issues applying Yoa's idea to classify new objects. Their work assumes that the loss incurred for misclassification of an object into any class is the same. This assumption doesn't always hold in real world applications. For example, they are misclassifying a patient with COVID-19 positive to pneumonia costs more than misclassifying this patient to have a general flu.

Instead of making an immediate acceptance or rejection in a three-way decision, a third option is deferment to each class. The deferment decision gives users flexibility for further examination of suspicious objects and reduces misclassification, which is cost effective. For example, if the doctor

cannot diagnose a few different types of flu based on a patient's symptoms, a series of medical tests can be performed to gather more information to help the doctor make the decision. Moreover, the losses incurred for misclassifying an object into any classes are treated differently, and the losses incurred for making deferred and rejective decisions to different classes are considered.

In this paper, our main objective is to provide a cost-sensitive solution to multiclass three-way decision-making, which may be difficult to achieve in binary classifications because a forced definite decision may come with a higher cost.

The rest of the paper is organized as follows: The basic concepts of rough sets and related properties are discussed in Section 2. The motivation for three-way decision-making is discussed in Section 3. The Bayesian decision procedure is described in detail in section 4. The proposed classification techniques based on a decision-theoretic rough set is presented in Section 5. In Section 6, an illustrative example of the treatment of a patient suspect to COVID-19 positive is provided to show the effectiveness of the proposed method. Section 7 provides the concluding remarks with future research directions of this study.

2. Background, motivation and research issues

In this section, we review the related knowledge about classical rough set, probabilistic rough set, three-way decision based on probabilistic rough set and decision theoretic rough set model based on Bayesian decision making process [33, 34, 35, 36].

2.1 Classical rough set theory

The theory of rough sets was introduced by Pawlak [18, 19]. In this theory, an approximation space (X, R) is developed given a datum in terms of the attribute-value table. X is the universal set of objects and R is the equivalence relation generated out of the data-table. The relation R , obviously, partitions X into the equivalence classes [1, 2, 3]. The equivalence classes are also called blocks and if R is understood from the context they may be written simply as $[·]$. Given this perspective, for any subset A of X two other subsets of X are defined by

$$\begin{aligned}\underline{A} &= \{x \in X \mid [x]R \subseteq A\} \text{ and} \\ \overline{A} &= \{x \in X \mid [x]R \cap A \neq \emptyset\}.\end{aligned}\tag{1}$$

$\underline{A}$ and $\overline{A}$ are called the lower and upper approximations of A respectively. The set $\overline{A} \setminus \underline{A}$ is called the boundary of A . Elements of $\underline{A}$ and $\overline{A}$ may be interpreted respectively as the objects that are 'definitely' included in A and 'possibly' included in A .

Based on the rough set approximations of A , one can divide the universe X into three pair-wise disjoint regions: the positive region $POS(A)$ is the union of all the equivalence classes that are included in A ; the negative region $NEG(A)$ is the union of all equivalence classes that have an empty intersection with A , and the boundary region $BND(A)$ is the difference between the upper and lower approximations:

$$\begin{aligned}POS(A) &= \underline{A} \\ NEG(A) &= X - \overline{A} \\ BND(A) &= \overline{A} - \underline{A}\end{aligned}\tag{2}$$

Hence, if $x \in POS(A)$, then x surely belongs to the concept A . If $x \in NEG(A)$, then x certainly does not belong to target set A . If an object $x \in BND(A)$, then it may or may not belong to A .

Afterwards, in 1994, Skowron and Pawlak [20,21] introduced the notion of rough membership function. Given an approximation space (X, R) , for any subset $A \subseteq X$, a rough membership function

[4] f_A is defined by $f_A : X \rightarrow [0, 1]$ such that $f_A(x) = \frac{|A \cap [x]|}{|[x]|}$, $|S|$ being the cardinality of the subset S of X . Although in [20], the universe was taken to be finite, the above definition is extended to the case when all the equivalence classes are finite. The following consequences of the above definition are immediate.

- (i) There can be two sets A, B ; $A \neq B$ but $f_A = f_B$.
- (ii) Any element of a block $[x]$ receives the same value under each rough membership function f_A . The value that every element of a block receives is called the block-value and denoted by $f_A[x]$.
- (iii)
$$f_A(x) = \begin{cases} 1 & \text{iff } [x] \subseteq A, \text{ i.e., } x \in \underline{A} \\ 0 & \text{iff } [x] \cap A = \emptyset, \text{ i.e., } x \in \overline{A}^c \\ 0 < f_A(x) < 1 & \text{iff } [x] \cap A \neq \emptyset, [x] \not\subseteq A, \text{ i.e., } x \in (A) \end{cases}$$

2.2 Probabilistic rough set

The positive region, negative region and boundary region defined by Pawlak (1982) are perfect. But the main drawback is that it cannot make decisions for most of the object. With this knowledge, the probabilistic rough set model [10, 11, 20, 36, 37, 46] was proposed. The main intuition of the probabilistic rough set model is to expand the decision region, i.e., to expand positive and negative regions using two parameters α and β . Let (X, E) be the approximation space, then (X, E, P) is a probabilistic approximation space (Yao 2008), where P is a probability measure defined on a subset of universe X . For any $A \subseteq X$, containing instances of a concept, $P(A|[x])$ denotes the conditional probability of an object in A given that the object in $[x]$.

$$P(A|[x]) = \frac{\text{card}(A \cap [x])}{\text{card}([x])} = \frac{|A \cap [x]|}{|[x]|} \quad (3)$$

where $|\cdot|$ denotes the cardinality.

According to the above definitions, the three regions defined in Eq. (2) can be equivalently defined by:

$$\begin{aligned} POS(A) &= \{x \in X \mid P(A|[x]) \geq 1\}, \\ BND(A) &= \{x \in X \mid 0 < P(A|[x]) < 1\}, \\ NEG(A) &= \{x \in X \mid P(A|[x]) \leq 0\}. \end{aligned}$$

2.3 Decision-theoretic rough set model

The decision theoretic rough set (DTRS) model first introduced by Yao et al. [32, 36, 37]. It is a more general probabilistic model in which a threshold pair (α, β) with $0 \leq \beta < \alpha \leq 1$ is used to define three probabilistic regions. Now for $0 \leq \beta < \alpha \leq 1$, upper and lower approximation of $A \subseteq X$ are given by

$$\begin{aligned} \underline{apr}_\alpha(A) &= \{x \in X \mid P(A|[x]) \geq \alpha\} \text{ and} \\ \overline{apr}_\beta(A) &= \{x \in X \mid P(A|[x]) > \beta\} \end{aligned}$$

Using the above two approximation the three probabilistic regions are define by

$$\begin{cases} POS_{(\alpha, \beta)}(A) = \{x \in X \mid P(A|[x]) \geq \alpha\} \\ NEG_{(\alpha, \beta)}(A) = \{x \in X \mid P(A|[x]) \leq \beta\} \\ BND_{(\alpha, \beta)}(A) = \{x \in X \mid \beta < P(A|[x]) < \alpha\} \end{cases} \quad (4)$$

The 0.5-probabilistic rough set model is a special case of the decision-theoretic rough set model, which is formulated based on a particular choice of α and β values, namely, $\alpha = \beta = 0.5$.

Probabilistic three regions may be interpreted in terms of the costs of different types of classification decisions [33, 34]. One can obtain larger positive and negative regions by introducing classification errors in the trade of a smaller boundary region so that the total classification cost is minimum [35]. Considering the errors introduced, the three regions are semantically interpreted as the following three-way decisions [33, 34, 35]. We accept an object x to be a member of A if the conditional probability is greater than or equal to α , with an understanding that it comes with an $(1 - \alpha)$ -level acceptance error and associated cost. We reject x to be a member of ' A ' if the conditional probability is less than or equal to β , with an understanding that it comes with an β -level of rejection error and associated cost. We neither accept nor reject x to be a member of ' A ' if the conditional probability is between of α and β , instead, we make a decision of deferment. The boundary region does not involve acceptance and rejection errors, but it is associated with the cost of deferment. The three probabilistic regions are obtained by considering a trade-off between various classification costs.

3. Motivation of three way Decision

Binary classification gives two decisions, namely yes or no, acceptance or rejection. In many real-world tasks, it is not easy to make definite decisions. For instance, when a patient's symptoms are not sufficient to support a particular disease, the doctor may not be able to make a correct / positive diagnosis right away. Then instead of making a decision that could give wrong results, a better way to perform diagnostic tests is to collect more evidence. This kind of option we termed as "deferment" decision. Many real-world decision-making problems become more efficient and easier by adding this third option.

One of the other issues is to minimize the number of misclassification examples [16, 22, 23]. Different types of misclassification errors are treated separately based on the notation of loss function from Bayesian decision theory. In the cost-sensitive classification approach, the cost caused by a different kind of error is not assumed to be equal. Cost-sensitive learning is one of the challenging issues in real-life problems. As an example, in medical diagnosis, not treating a blood cancer patient could cause death or injury. On the other hand, unnecessarily treating a patient who doesn't have blood cancer will waste resources and harm the patient.

Three-way decision making has been studied and applied in many fields [13, 25, 26, 30], but our main objective in this study is to provide a cost-sensitive solution to multiclass /three-way decision making, which may be difficult to achieve in binary classifications because forced definite decision may come with a higher cost.

The following example explains the main advantages of adding the third option i.e. deferment decision.

Example 1. Consider a story given in the book by Savage [24]. "Your wife has just broken five good eggs into a bowl when you come in and volunteer to finish making the omelette. A sixth egg, which for some reason must either be used for the omelette or wasted altogether, lies unbroken beside the bowl. You must decide what to do with this unbroken egg."

Consider first a binary-decision scenario, namely, to break it into the bowl containing the other five or to throw it away without inspection. Depending on the state of the egg, each of these two actions will have some consequences and they are given by the following 2×2 pay off matrix:

	Good	Rotten
Break into bowl	Six-egg omelette	No omelette and five good egg destroyed
Thrown away	Five-egg omelette and one good egg Destroyed	Five-egg omelette

If the sixth egg is good and you made the right decision by breaking it into the bowl, your wife will be happy to see a finished six-egg omelette. If the sixth egg is bad and you made the right decision of throwing it away, your wife will still be happy to see a finished five-egg omelette. Wrong decisions cost much more. In one case, if the sixth egg is good, but you throw it away, one good egg will be wasted. In the worst case, the sixth egg is bad, but you break into the bowl, five good eggs will be destroyed, and you had better think about how to explain your decision to your wife. In binary decision-making, each action comes with the cost of ruining some good eggs.

Is there a better way of doing this? A better choice would be not to break the sixth egg into the bowl or throw it away but to break it into a saucer for inspection. By adding this third action, the consequence changed to the following 3×2 pay off matrix:

	Good	Rotten
Break into bowl	Six-egg omelette	No omelette and five good egg destroyed
Thrown away	Five-egg omelette and one good egg Destroyed	Five-egg omelette
Break into saucer	Six-egg omelette, and a saucer to wash	Five-egg omelette, and a saucer to wash

By further examining the state of the sixth egg, the cost is reduced from ruining good eggs to washing a saucer. A similar three-way decision-making process is used in both the above examples, where the doctor needed to decide whether to treat the patient, not treat the patient, or perform diagnostic tests to collect more information [17].

4. The Bayesian decision procedure

In decision-theoretic rough set model, a pair of threshold (α, β) determined three probabilistic regions. The selection of threshold is a major issue. Based on the Bayesian decision procedure, values of α and β are calculated [6, 7, 8, 27]. For a given object let X be the description of the object, $\Omega = \{s_1, s_2, s_3, \dots, s_m\}$ be a finite set of m possible states that x is possibly in and $A = \{a_1, a_2, \dots, a_n\}$ be a finite set of n possible action. Let $P(s_j|X)$ be the conditional probability of x being in state s_j and the loss function $\lambda(a_i|s_j)$ denote the loss (or cost) for taking the action a_i when the state is s_j . Hence, we can construct an $m \times n$ matrix which represents all possible loss function. Table 1, expressed all the values of the loss function where the columns denoting the set Ω of m states and the rows the set A of n actions. Each cell denotes the cost $\lambda(a_i|s_j)$ for taking the action a_i in the state s_j . For simplicity, the cost $\lambda(a_i|s_j)$ can be written as λ_{ij} .

Table 1 (Loss function for m state and n actions):

	s_1	s_2	.....	s_j	.....	s_m
a_1	$\lambda(a_1 s_1) = \lambda_{11}$	$\lambda(a_1 s_2) = \lambda_{12}$	.....	$\lambda(a_1 s_j) = \lambda_{1j}$	.....	$\lambda(a_1 s_m) = \lambda_{1m}$
a_2	$\lambda(a_2 s_1) = \lambda_{21}$	$\lambda(a_2 s_2) = \lambda_{22}$	.....	$\lambda(a_2 s_j) = \lambda_{2j}$	.....	$\lambda(a_2 s_m) = \lambda_{2m}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_i	$\lambda(a_i s_1) = \lambda_{i1}$	$\lambda(a_i s_2) = \lambda_{i2}$	.....	$\lambda(a_i s_j) = \lambda_{ij}$	.....	$\lambda(a_i s_m) = \lambda_{im}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
a_n	$\lambda(a_n s_1) = \lambda_{n1}$	$\lambda(a_n s_2) = \lambda_{n2}$	.....	$\lambda(a_n s_j) = \lambda_{nj}$	.....	$\lambda(a_n s_m) = \lambda_{nm}$

In general, the selection of the values of the loss function can be represented as:

$$\lambda(a_i|s_j) = \begin{cases} 0 & \text{if } i = j \\ \lambda_{ij} & \text{if } i \neq j \end{cases}$$

Where λ_{ij} ($i = 1, 2 \dots n; j = 1, 2 \dots m$) denotes the loss incurred for deciding class s_i when the true class is s_j . There is no cost for making a correct decision, i.e., when $i = j$.

For an object x with description X , suppose action a_i is taken. The expected cost associated with action a_i is given by:

$$R(a_i|X) = \sum_{j=1}^m \lambda(a_i|s_j) P(s_j|X) \quad (5)$$

The $n \times m$ matrix in table-1 has two important applications. First, given the loss functions and the probabilities, one can compute the expected cost of a certain action. Furthermore, by comparing the expected costs of all the actions, one can decide on a particular action with the minimum cost. Second, according to the loss functions, one can determine the condition or probability for taking a particular action.

Example 2. The idea of the Bayesian decision procedure can be demonstrated by the following example. Suppose there are two states: s_1 indicates that a meeting will be over in less than or equal to 4 hours, and s_2 indicates that the meeting will be over in more than 4 hours. Two states are complement to each other. Suppose, the probability for having the state s_1 is 0.70, then the probability for having the state s_2 is 0.30, i.e., $P(s_1|X) = 0.70$ and $P(s_2|X) = 1 - 0.70 = 0.30$. There are two actions: a_1 means to park the car on meter, and a_2 means to park the car in the parking lot. The loss functions for taking different actions in different states can be expressed as the following matrix:

	a_1 (Parking on meter)	a_2 (parking in a parking lot)
$(s_1 \leq 4 \text{ hr})$	Rs. 40	Rs. 90
$(s_2 > 4 \text{ hr})$	Rs. 160	Rs. 90

In this case, the cost for each action can be calculated as follows:

$$R(a_1|X) = 40 \times 0.70 + 160 \times 0.30 = 76$$

$$R(a_2|X) = 90 \times 0.70 + 90 \times 0.30 = 90$$

As, Rs. 76 < Rs. 90, according to the minimum cost, one may decide to park the car on meter, instead of in a parking lot. Suppose, a person wants to decide where to park the car. It is interesting to know if parking on meter is more suitable. According to Eq. (5), one can obtain:

$$40 \times P(s_1|X) + 160 \times (1 - P(s_1|X)) \leq 90$$

That is, $P(s_1|X) \geq 0.50$. Thus, if the probability that the meeting is over within 4 hours is greater than or equal to 0.50, then it is more profitable to park the car on meter, otherwise, park in a parking lot.

From equation (5) the classifier is said to assign an object to the class s_i if the expected cost of making action a_i , is less than the cost of making action a_j i.e.

$$R(a_i|[x]) < R(a_j|[x]), \text{ for all } i \neq j.$$

One has the option of assigning the object to one of the classes. In other words, each class s_j is associated with an action a_i of accepting the object to be a member of that class. However, an

immediate decision has to be made to either accept or reject the object to be a member of one of the classes.

Yao and Zhao [38] suggested a change in this class-classification problem and made a three-way decision for each class. For the finite set of m classes $\Omega = \{s_1, s_2, s_3, \dots, s_m\}$ the s_j 's form a family of pairwise disjoint subsets of X , namely, $s_i \cap s_j = \emptyset$ for $i \neq j$, and $\cup s_j = X$. For each s_j , a two-class classification $\{s_j, s_j^c\}$ can be defined. The loss function for each s_i is given by the following table.

Table 2: (Loss function for m state and 3 actions):

	s_1	s_2	...	s_j	...	s_m
a_{P_i}	$\lambda(a_{P_i} s_1) = \lambda_{P1}$	$\lambda(a_{P_i} s_2) = \lambda_{P2}$	...	$\lambda(a_{P_i} s_j) = \lambda_{Pj}$	...	$\lambda(a_{P_i} s_m) = \lambda_{Pm}$
a_{N_i}	$\lambda(a_{N_i} s_1) = \lambda_{N1}$	$\lambda(a_{N_i} s_2) = \lambda_{N2}$	...	$\lambda(a_{N_i} s_j) = \lambda_{Nj}$	...	$\lambda(a_{N_i} s_m) = \lambda_{Nm}$
a_{B_i}	$\lambda(a_{B_i} s_1) = \lambda_{B1}$	$\lambda(a_{B_i} s_2) = \lambda_{B2}$	...	$\lambda(a_{B_i} s_j) = \lambda_{Bj}$	...	$\lambda(a_{B_i} s_m) = \lambda_{Bm}$

The m pairs of thresholds can be systematically calculated based on the given loss functions. The results from DTRS can be immediately applied. However, the losses incurred for making any wrong decisions are considered as the same, that is:

$$\begin{aligned}
\lambda(a_{P_i}|s_j) &= \lambda_{P1} = \lambda_{P2} = \dots = \lambda_{Pm}, \text{ for all } i \neq j, \\
\lambda(a_{N_i}|s_j) &= \lambda_{N1} = \lambda_{N2} = \dots = \lambda_{Nm}, \text{ for all } i \neq j, \\
\lambda(a_{B_i}|s_j) &= \lambda_{B1} = \lambda_{B2} = \dots = \lambda_{Bm}, \text{ for all } i \neq j
\end{aligned} \tag{6}$$

This assumption does not always hold in many real world scenario, which makes it impractical for real applications.

5. Class classification based on decision-theoretic rough set

Here we proposed a new formulation of decision-theoretic rough set model. Let $\Omega = \{s_1, s_2, s_3, \dots, s_m\}$ be a given set of m class. Each class is associated with three action $A = \{a_{P_i}, a_{B_i}, a_{N_i}\}$. where a_{P_i} is the action for deciding an object x belongs to positive region of the i^{th} class. Similarly, a_{B_i}, a_{N_i} are action for deciding an object x belongs to boundary and negative region of the i^{th} class respectively. Now similar to table 2, here we also get a loss function table. We define $\lambda(a_{P_i}|s_j)$ as the loss incurred for accepting an object x as a member of s_i when it's real class is s_j , $\lambda(a_{N_i}|s_j)$ represent as the loss incurred for rejecting an object x as a member of s_i when it's real class is s_j , and $\lambda(a_{B_i}|s_j)$ is defined as the loss incurred for neither accepting nor rejecting an object x as a member of s_i when it's real class is s_j .

The main distinction between Yao's approach and our approach is that in Yao assumed the losses incurred for making any wrong decisions are the same. This assumption is not true for a lot of real world situations.

Consider an example where we only have three types of disease where $s_1 = \text{covid-19}$, $s_2 = \text{cancer}$, and $s_3 = \text{pneumonia}$. $\lambda(a_{P_1}|s_2)$ represents the loss function of treating a cancer patient as he/she has covid-19, $\lambda(a_{N_3}|s_3)$ represents the loss function of not treating a pneumonia patient, and $\lambda(a_{B_3}|s_2)$ represents the loss function of not making an immediate diagnosis to a cancer patient, i.e., further tests are needed (e.g., perform a CT scan). As we know that not treating a cancer patient could cause death or injury. On the other hand, resources will be wasted for unnecessary treatment to a patient who does not have

cancer. These two types of loss functions are not equal and should be treated differently. One can easily relate these loss functions with the actual cost of making a different diagnosis.

Hence in our approach, the losses incurred for making different wrong decisions are considered differently.

The expected losses for taking different actions namely acceptance, rejection and deferment for an objects $x \in [x]$ can be expressed by equation (5) as:

$$\begin{cases} R(a_{P_i}||[x]) = \sum_{j=1}^m \lambda(a_{P_i}|s_j)P(s_j|[x]), \\ R(a_{N_i}||[x]) = \sum_{j=1}^m \lambda(a_{N_i}|s_j)P(s_j|[x]), \\ R(a_{B_i}||[x]) = \sum_{j=1}^m \lambda(a_{B_i}|s_j)P(s_j|[x]) \end{cases} \quad (7)$$

The Bayesian decision procedure leads to the following minimum-risk decision rules:

$$\begin{aligned} (P) \text{ If } R(a_{P_i}||[x]) \leq R(a_{B_i}||[x]) \text{ and } R(a_{P_i}||[x]) \leq R(a_{N_i}||[x]), \text{ decide } x \in POS(s_i); \\ (N) \text{ If } R(a_{N_i}||[x]) \leq R(a_{B_i}||[x]) \text{ and } R(a_{N_i}||[x]) \leq R(a_{P_i}||[x]), \text{ decide } x \in NEG(s_i); \\ (B) \text{ If } R(a_{B_i}||[x]) \leq R(a_{P_i}||[x]) \text{ and } R(a_{B_i}||[x]) \leq R(a_{N_i}||[x]), \text{ decide } x \in BND(s_i); \end{aligned} \quad (8)$$

Now consider a special kind of loss function where

$$\begin{cases} \lambda(a_{P_i}|s_i) \leq \lambda(a_{B_i}|s_i) < \lambda(a_{N_i}|s_i) \\ \lambda(a_{N_i}|s_j) \leq \lambda(a_{B_i}|s_j) < \lambda(a_{P_i}|s_j) \text{ for all } j \text{ where } j \neq i \end{cases} \quad (9)$$

That is, the loss of classifying an object x belonging to s_i into the positive region $POS(s_i)$ is less than or equal to the loss of classifying x into the boundary region $BND(s_i)$, and both of these losses are strictly less than the loss of classifying x into the negative region $NEG(s_i)$. The reverse order of losses is used for classifying an object not in s_i . Under these special kind of loss function we can simplify decision rules (P), (N) and (B) are as follows.

For rule (P), $R(a_{P_i}||[x]) \leq R(a_{B_i}||[x])$ and $R(a_{P_i}||[x]) \leq R(a_{N_i}||[x])$,

Now, $R(a_{P_i}||[x]) \leq R(a_{B_i}||[x])$

$$\begin{aligned} &\Leftrightarrow \sum_{j=1}^m \lambda(a_{P_i}|s_j)P(s_j|[x]) \leq \sum_{j=1}^m \lambda(a_{B_i}|s_j)P(s_j|[x]). \\ &\Leftrightarrow \lambda(a_{P_i}|s_i)P(s_i|[x]) + \sum_{j=1, j \neq i}^m \lambda(a_{P_i}|s_j)P(s_j|[x]) \leq \lambda(a_{B_i}|s_i)P(s_i|[x]) \\ &\quad + \sum_{j=1, j \neq i}^m \lambda(a_{B_i}|s_j)P(s_j|[x]). \\ &\Leftrightarrow P(s_i|[x]) \geq \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{ \lambda(a_{P_i}|s_j) - \lambda(a_{B_i}|s_j) \}}{\{ \lambda(a_{B_i}|s_i) - \lambda(a_{P_i}|s_i) \}} \end{aligned}$$

And $R(a_{P_i}||[x]) \leq R(a_{N_i}||[x])$

$$\Leftrightarrow P(s_i|[x]) \geq \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{ \lambda(a_{P_i}|s_j) - \lambda(a_{N_i}|s_j) \}}{\{ \lambda(a_{N_i}|s_i) - \lambda(a_{P_i}|s_i) \}}.$$

Similarly, for rule (N)

$R(a_{N_i}||[x]) \leq R(a_{B_i}||[x])$

$$\Leftrightarrow P(s_i|[x]) \geq \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{ \lambda(a_{B_i}|s_j) - \lambda(a_{N_i}|s_j) \}}{\{ \lambda(a_{N_i}|s_i) - \lambda(a_{B_i}|s_i) \}}$$

And $R(a_{N_i}|[x]) \leq R(a_{P_i}|[x])$

$$\Leftrightarrow P(s_i|[x]) \geq \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{\lambda(a_{P_i}|s_j) - \lambda(a_{N_i}|s_j)\}}{\{\lambda(a_{N_i}|s_i) - \lambda(a_{P_i}|s_i)\}}$$

Likewise for rule (B)

$R(a_{B_i}|[x]) \leq R(a_{P_i}|[x])$

$$\Leftrightarrow P(s_i|[x]) \geq \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{\lambda(a_{P_i}|s_j) - \lambda(a_{B_i}|s_j)\}}{\{\lambda(a_{B_i}|s_i) - \lambda(a_{P_i}|s_i)\}}$$

And $R(a_{B_i}|[x]) \leq R(a_{N_i}|[x])$

$$\Leftrightarrow P(s_i|[x]) \geq \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{\lambda(a_{B_i}|s_j) - \lambda(a_{N_i}|s_j)\}}{\{\lambda(a_{N_i}|s_i) - \lambda(a_{B_i}|s_i)\}}$$

Now introducing parameter as

$$\begin{cases} \alpha_i = \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{\lambda(a_{P_i}|s_j) - \lambda(a_{B_i}|s_j)\}}{\{\lambda(a_{B_i}|s_i) - \lambda(a_{P_i}|s_i)\}} \\ \beta_i = \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{\lambda(a_{B_i}|s_j) - \lambda(a_{N_i}|s_j)\}}{\{\lambda(a_{N_i}|s_i) - \lambda(a_{B_i}|s_i)\}} \\ \gamma_i = \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{\lambda(a_{P_i}|s_j) - \lambda(a_{N_i}|s_j)\}}{\{\lambda(a_{N_i}|s_i) - \lambda(a_{P_i}|s_i)\}} \end{cases} \quad (10)$$

If we impose

$$\frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{\lambda(a_{P_i}|s_j) - \lambda(a_{B_i}|s_j)\}}{\{\lambda(a_{B_i}|s_i) - \lambda(a_{P_i}|s_i)\}} > \frac{\sum_{j=1, j \neq i}^m P(s_j|[x]) \{\lambda(a_{B_i}|s_j) - \lambda(a_{N_i}|s_j)\}}{\{\lambda(a_{N_i}|s_i) - \lambda(a_{B_i}|s_i)\}}$$

$$\Rightarrow \frac{\{\lambda(a_{P_i}|s_j) - \lambda(a_{B_i}|s_j)\}}{\{\lambda(a_{B_i}|s_i) - \lambda(a_{P_i}|s_i)\}} > \frac{\{\lambda(a_{B_i}|s_j) - \lambda(a_{N_i}|s_j)\}}{\{\lambda(a_{N_i}|s_i) - \lambda(a_{B_i}|s_i)\}}$$

$$\text{i.e.} \quad \alpha_i > \beta_i \quad (11)$$

From (9) and (11) it implies that $\alpha_i > \gamma_i > \beta_i$ and we get simplified rule as follows:

$$\begin{aligned} (P) \quad & P(s_i|[x]) \geq \alpha_i, \text{ decide } x \in POS(s_i); \\ (N) \quad & P(s_i|[x]) \leq \beta_i, \text{ decide } x \in NEG(s_i); \\ (B) \quad & \beta_i < P(s_i|[x]) < \alpha_i, \text{ decide } x \in BND(s_i) \end{aligned} \quad (12)$$

With this simplified (P)-(B) rule, the three probabilistic region for any class s_i can be derived from the equation (3) as:

$$\begin{cases} POS_{(\alpha_i, \beta_i)}(s_i) = \{x \in X | P(s_i|[x]) \geq \alpha_i\} \\ NEG_{(\alpha_i, \beta_i)}(s_i) = \{x \in X | P(s_i|[x]) \leq \beta_i\} \\ BND_{(\alpha_i, \beta_i)}(s_i) = \{x \in X | \beta_i < P(s_i|[x]) < \alpha_i\} \end{cases} \quad (13)$$

Definition 1:

Let $\Omega = \{s_1, s_2, s_3, \dots, s_m\}$ be a partition of the universe X , Then three regions of partition are given by

$$\begin{cases} POS_{(\alpha_i, \beta_i)}(\Omega) = \cup_{i=1}^m POS_{(\alpha_i, \beta_i)}(s_i) \\ BND_{(\alpha_i, \beta_i)}(\Omega) = \cup_{i=1}^m BND_{(\alpha_i, \beta_i)}(s_i) \\ NEG_{(\alpha_i, \beta_i)}(\Omega) = X - \{POS_{(\alpha_i, \beta_i)}(\Omega) \cup BND_{(\alpha_i, \beta_i)}(\Omega)\} \end{cases} \quad (14)$$

Where s_i class represents an arbitrary class from the set of decision classes Ω . It is necessary to have a further study on the probabilistic three regions of a classification, as well as the associated rules. In general, one has to consider the problem of rule conflict resolution in order to make effective acceptance, rejection, and abstaining decisions.

Furthermore, for an object $x \in X$, it is desirable that the decision made with more attributes should stay the same as the decision made with fewer attributes. The decision-monotonicity properties for DTRS are discussed by Yao and Zhao [38]. Slezak and Ziarko [28] and Zhang et al. [45] also introduced criteria of decision-monotonicity for their generalized rough set models.

6. Numerical Illustration

Here, we incorporated cost into the classification process and vary costs for different types of misclassification errors. The following medical diagnostic example illustrates how our approach works and shows that the classification results based on the same training set could be different based on the different loss functions provided.

Let us consider four types of diseases $D = \{s_1, s_2, s_3, s_4\}$. Three actions $A = \{a_{P_i}, a_{B_i}, a_{N_i}\}$ are associated with corresponding to each disease. Where a_{P_i} indicates that the doctor decides to treat the i^{th} patient a_{N_i} indicates that the doctor decides not to treat the i^{th} patient and a_{B_i} indicates that the doctor neither treats nor not treats the patient, further tests are needed. The loss function of the two medical centres is given in table 3 and 4 respectively. As the two tables provide the loss function of two different medical centres, so they are significantly different. Also, it is clear that the cost for medical centre -1 is more than medical centre -2 for neglecting a sick patient, i.e. for delayed treatment and cost of unnecessary treatment, i.e. misdiagnosing a healthy patient. Our approach gives the flexibility of tailoring the classification results to meet individual requirements in terms of the minimum overall cost.

A part of the training set acquired from statistical data of a medical journal that shows the relationships between patients' symptoms to a certain disease is represented by table-5. where the rows represent the patients $\{P_1, P_2, P_3, P_4, \dots\}$ and columns represent the symptoms $\{O_1, O_2, O_3, O_4, \dots\}$. We use binary values for the sake of simplicity where "Y" indicate symptoms is present and "N" denotes symptoms not present.

Now for a new patient P_n , let probabilities of having each disease be calculated from the training set using the equation (3) as

$$P(s_1|[x]) = 0.38, P(s_2|[x]) = 0.22, P(s_3|[x]) = 0.14, P(s_4|[x]) = 0.26.$$

Table 3 (Loss function of Medical centre-1):

	s_1	s_2	s_3	s_4
a_{P_1}	0	5	6	10
a_{B_1}	5	3	3	7
a_{N_1}	18	0	0	0
a_{P_2}	18	0	36	46
a_{B_2}	9	14	20	30
a_{N_2}	0	35	0	0
a_{P_3}	55	50	0	65
a_{B_3}	24	22	30	26
a_{N_3}	0	0	70	0
a_{P_4}	16	10	8	0
a_{B_4}	10	3	2	15
a_{N_4}	0	0	0	55

Table 4 (Loss function of Medical centre-2):

	s_1	s_2	s_3	s_4
a_{P_1}	0	4	3	6
a_{B_1}	4	2	2	3
a_{N_1}	10	0	0	0
a_{P_2}	8	0	16	26
a_{B_2}	6	6	12	15
a_{N_2}	0	18	0	0
a_{P_3}	20	20	0	25
a_{B_3}	10	9	12	12
a_{N_3}	0	0	25	0
a_{P_4}	8	4	2	0
a_{B_4}	3	2	1	12
a_{N_4}	0	0	0	25

Table 5 (statistical data):

	O_1	O_2	O_3	O_4	O_5	.....	D
P_1	Y	Y	N	N	Y	.....	s_2
P_2	Y	N	Y	Y	N	.....	s_3
P_3	N	Y	Y	Y	Y	.....	s_4
P_4	Y	N	N	N	N	.....	s_1
P_5	N	Y	N	N	Y	.....	s_3
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	.....	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	.....	$\vdots$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	.....	$\vdots$

Table 6 (Testing set):

	O_1	O_2	O_3	O_4	O_5	.....	D
P_n	Y	N	Y	N	Y	.....	$?$

Now Based on Eq. (10) and Table-3, we can compute the pairs of thresholds (α_1, β_1) for diseases S_1 of medical centre-1 as:

$$\begin{aligned}
\alpha_1 &= \frac{\sum_{j=1, j \neq 1}^m P(S_j|[x]) \{ \lambda(a_{P_1}|s_j) - \lambda(a_{B_1}|s_j) \}}{\{ \lambda(a_{B_1}|s_1) - \lambda(a_{P_1}|s_1) \}} \\
&= \frac{P(S_2|[x]) \{ \lambda(a_{P_1}|s_2) - \lambda(a_{B_1}|s_2) \} + P(S_3|[x]) \{ \lambda(a_{P_1}|s_3) - \lambda(a_{B_1}|s_3) \} + P(S_4|[x]) \{ \lambda(a_{P_1}|s_4) - \lambda(a_{B_1}|s_4) \}}{\{ \lambda(a_{B_1}|s_1) - \lambda(a_{P_1}|s_1) \}} \\
&= \frac{0.22(5-3) + 0.14(6-3) + 0.26(10-7)}{5-0} = \frac{1.64}{5} = 0.32. \\
\beta_1 &= \frac{\sum_{j=1, j \neq 1}^m P(S_j|[x]) \{ \lambda(a_{B_1}|s_j) - \lambda(a_{N_1}|s_j) \}}{\{ \lambda(a_{N_1}|s_1) - \lambda(a_{B_1}|s_1) \}} \\
&= \frac{P(S_2|[x]) \{ \lambda(a_{B_1}|s_2) - \lambda(a_{N_1}|s_2) \} + P(S_3|[x]) \{ \lambda(a_{B_1}|s_3) - \lambda(a_{N_1}|s_3) \} + P(S_4|[x]) \{ \lambda(a_{B_1}|s_4) - \lambda(a_{N_1}|s_4) \}}{\{ \lambda(a_{N_1}|s_1) - \lambda(a_{B_1}|s_1) \}} \\
&= \frac{0.22(3-0) + 0.14(3-0) + 0.26(7-0)}{18-5} = \frac{2.9}{13} = 0.22.
\end{aligned}$$

Similarly, for the disease S_2, S_3, S_4 we can get $\alpha_2 = 0.70$, $\beta_2 = 0.66$; $\alpha_3 = 0.93$, $\beta_3 = 0.52$ and $\alpha_4 = 0.31$, $\beta_4 = 0.11$ respectively. Also thresholds pair for table-2 (medical centre-2) corresponding to the diseases S_1, S_2, S_3, S_4 are $\alpha_1 = 0.34$, $\beta_1 = 0.25$, $\alpha_2 = 0.69$, $\beta_2 = 0.65$; $\alpha_3 = 0.80$, $\beta_3 = 0.68$ and $\alpha_4 = 0.22$, $\beta_4 = 0.13$ respectively.

Comparison of thresholds values:

For medical centre -1.

$P(s_1|[x]) = 0.38 > \alpha_1 = 0.32$, so class S_1 is in the positive region. The resulting diagnosis is that the new patient has disease S_1 .

$P(s_2|[x]) = 0.22 < \beta_2 = 0.66$, so class S_2 is in the negative region. The resulting diagnosis is that the new patient has no disease S_2 .

$P(s_3|[x]) = 0.14 < \beta_3 = 0.52$, so class S_3 is in the negative region. The resulting diagnosis is that the new patient has no disease S_3 .

$$\beta_4 = 0.11 < P(s_4|[x]) = 0.26 < \alpha_4 = 0.31.$$

So class S_4 is in the boundary region. The resulting diagnosis is that the new patient need further test regarding the disease S_4 .

For medical centre-2.

$P(s_1|[x]) = 0.38 > \alpha_1 = 0.34$, hence class S_1 is in the positive region. The resulting diagnosis is that the new patient has disease S_1 .

$P(s_2|[x]) = 0.22 < \beta_2 = 0.65$, so class S_2 is in the negative region. The resulting diagnosis is that the new patient has no disease S_2

$P(s_3|[x]) = 0.14 < \beta_3 = 0.68$, so class S_3 is in the negative region. The resulting diagnosis is that the new patient has no disease S_3 .

$P(s_4|[x]) = 0.26 > \alpha_4 = 0.22$, hence class S_4 is in the positive region. The resulting diagnosis is that the new patient has disease S_4 .

As for the medical centre-2, class S_1 and S_4 are both in the positive region. The new patient could have both disease S_1 and S_4 , and either one of them. Rule conflict resolution should be added to further distinguish which disease the patient is more likely to have.

From the above example, we can see that different cost setting could lead to different decisions. Therefore, instead of classifying objects based on misclassification error rate, in our approach, classification decisions are made based on minimum overall cost.

7. Concluding Remarks:

There are so many techniques for medical or clinical decision making under uncertainty. We mainly consider a rough set based decision-making technique which typically involves evaluation criteria to construct a decision region. This paper aims to provide a cost-sensitive solution of a COVID-19 suspected people to multi-class decision making. Instead of pair of lower and upper approximation in this paper, we use three pair-wise disjoint positive, negative and boundary regions. This approach can be considered a straightforward generalization of the three-way classification introduced in decision-theoretic rough set models. One may further study three-way decision rules generated from different classes and the associated rule conflict resolutions for real classification applications.

Compliance with ethical standards

Conflict of interest: Both authors of this research paper declare that, there is no conflict of interest.

Ethical approval: This article does not contain any studies with human participants or animals performed by any of the authors.

References:

1. M. Banerjee, M. K. Chakraborty, Rough sets through algebraic logic, *Fundamental Informaticae*. 28 (1996) 211-221.
2. G. Cattaneo, D. Ciucci, Heyting-Wajsberg algebras as an abstract environment linking fuzzy and rough sets, in *Proceedings of the 3rd International conference on Rough Sets and Current Trends in Computing (RSCTC 2002)*, Malvern, PA, USA, *SLNAI2475*, 2004, pp 77-84.
3. G. Cattaneo, Abstract approximation spaces for rough theories in *Polkowski & Skowron (Eds.), Rough Sets in Knowledge Discovery, applications, case studies and software systems*, Physica-Verlag, Heidelberg, (1998) 59 –106.
4. M. K. Chakraborty, Membership function based rough set, *International Journal of Approximate Reasoning*, 55 (2014) 402 – 411.
5. M. R. Foster, Key concepts in model selection: performance and generalizability, *Journal of Mathematical Psychology*, 44 (2000) 205-231.
6. S. Greco, B. Matarazzo, R. Słowiński, Parameterized rough set model using rough membership and Bayesian confirmation measures, *International Journal of Approximate Reasoning*, 49 (2009) 285–300.
7. S. Greco, Z. Pawlak, R. Słowiński, Bayesian confirmation measures within rough set approach, in: *Rough Sets and Current Trends in Computing*, (2004) 264–273.

8. S. Greco, Z. Pawlak, R. Slowiński, Can Bayesian confirmation measures be useful for rough set decision rules?, *Engineering Applications of Artificial Intelligence* 17 (2004) 345–361.
9. R. Goudey, Do statistical inferences allowing three alternative decision give better feedback for environmentally precautionary decision-making, *Journal of Environmental Management*, 85 (2007) 338–344.
10. J. P. Herbert, J. T. Yao, Game-theoretic risk analysis in decision-theoretic rough sets, in: *Proceedings of RSKT'08, Lecture Notes in Computer Science*, 5009 (2008) 132–139.
11. J. P. Herbert, J. T. Yao, Game-theoretic rough sets, *Fundamental Informaticae - Advances in Rough Set Theory* 108 (3–4) (2011) 267–286.
12. D. Liu, T. R. Li, P. Hu, H. X. Li, Multiple-category classification with decision-theoretic rough sets, in: *RSKT*, (2010) 703–710.
13. D. Liu, T. R. Li, D. C. Liang, Incorporating logistic regression to decision-theoretic rough sets for classifications, *International Journal of Approximate Reasoning*, 2013.
14. A.R. Mishra, P. Rani, R. Krishankumar, K. S. Ravichandran, S. Kar, An extended fuzzy decision-making framework using hesitant fuzzy sets for the drug selection to treat the mild symptoms of Coronavirus Disease 2019 (COVID-19), *Applied Soft Computing*, 107155, 2021.
15. S. Majumder, S. Kar, E. Samanta, A fuzzy rough hybrid decision making technique for identifying the infected population of COVID-19, *Soft Computing*, 2020.
16. T. M. Mitchell, *Machine Learning*, McGraw-Hill, New York, 1997.
17. S. G. Pauker, J. P. Kassirer, The threshold approach to clinical decision making, *New England Journal of Medicine* 302 (1980) 1109–1117.
18. Z. Pawlak, Rough sets, *International Journal of Computer and Information Sciences* 11 (1982) 341–356.
19. Z. Pawlak, *Rough Sets, Theoretical Aspects of Reasoning about Data*, Kluwer Academic Publishers, Dordrecht, 1991.
20. Z. Pawlak, S. K. M. Wong, W. Ziarko, Rough sets: probabilistic versus deterministic approach, *International Journal of Man-Machine Studies* 29 (1988) 81–95.
21. Z. Pawlak, A. Skowron, Rough membership function. in: *R. Yager, M. Fedrizzi, J. Kacprzyk (Eds.) Advances in Dempster-shafer Theory of Evidence*. Wiley, New York, (1994) 251–271.
22. J. R. Quinlan, Learning efficient classification procedures and their application to chess endgames, in: J. S. Michalski, J. G. Carbonell, T. M. Mitchell (Eds.), *Machine Learning: An Artificial Intelligence Approach*, vol. 1, Morgan Kaufmann, Palo Alto, CA, 1983, pp. 463–482.
23. J. R. Quinlan, Induction of decision trees, *Machine Learning* 1 (1) (1986) 81–106.
24. L. J. Savage, *The Foundation of Statistics*, second ed., Dover, 1972.
25. C. B. Schechter, Sequential analysis in a Bayesian model of diastolic blood pressure measurement, *Medical Decision Making* 8 (1988) 191–196.
26. M. Sherif, C. I. Hovland, *Social Judgment: Assimilation and Contrast Effects in Communication and Attitude Change*, Yale University Press, New Haven, 1961.
27. D. Slezak, W. Ziarko, Bayesian rough set model, in: *Proceedings of FDM'2002*, (2002) 131–135.
28. D. Slezak, W. Ziarko, The investigation of the Bayesian rough set model, *International Journal of Approximate Reasoning* 40 (2005) 81–91.
29. A. Si, S. Das, S. Kar, Picture fuzzy set-based decision-making approach using Dempster–Shafer theory of evidence and grey relation analysis and its application in COVID-19 medicine selection, *Soft Computing*, doi: 10.1007/s00500-021-05909-9, 2021.
30. A. Wald, Sequential tests of statistical hypotheses, *The Annals of Mathematical Statistics* 16 (1945) 117–186.

31. P.W. Woodward and Naylor, J.C. An application of Bayesian methods in SPC, *The Statistician*, 42, 461-469, 1993.
32. Y. Y. Yao, Decision-theoretic rough set models, in: *Rough Sets and Knowledge Technology, Lecture Notes in Computer Science*, vol. 4481, 2007, pp. 1–12.
33. Y. Y. Yao, Three-way decision: an interpretation of rules in rough set theory, in: *Rough Sets and Knowledge Technology, Lecture Notes in Computer Science*, vol. 5589, 2009, pp. 642–649.
34. Y. Y. Yao, Three-way decisions with probabilistic rough sets, *Information Sciences* 180 (2010) 341–353.
35. Y. Y. Yao, The superiority of three-way decisions in probabilistic rough set models, *Information Sciences* 181 (2011) 1080–1096.
36. Y. Y. Yao, S. K. M. Wong, A decision theoretic framework for approximating concepts, *International Journal of Man-machine Studies* 37 (1992) 793–809.
37. Y. Y. Yao, S. K. M. Wong, P. Lingras, A decision-theoretic rough set model, in: Z.W. Ras, M. Zemankova, M.L. Emrich (Eds.), *Methodologies for Intelligent Systems*, vol. 5, North-Holland, New York, (1990) 17–24.
38. Y. Y. Yao, Y. Zhao, Attribute reduction in decision-theoretic rough set models, *Information Sciences* 178 (17) (2008) 3356–3373.
39. Y.Y. Yao, Probabilistic approaches to rough sets, *Expert Systems*, 20 (2003) 287-297.
40. Y.Y. Yao, Probabilistic rough set approximations, *International Journal of Approximation Reasoning*, 49 (2008) 255-271.
41. B Yang, J. H. Li, Complex network analysis of three-way decision researches. *Int J Mach Learn Cybern* 11 (2020) 973–987.
42. D. J. Yu, Z. S. Xu, W. Pedrycz, Bibliometric analysis of rough sets research. *Applied Soft Computing* 94 (2020) 106467.
43. H. Yu , X. C. Wang, G. Y. Wang, X. H. Zeng, An active three-way clustering method via low-rank matrices for multi-view data. *Information Sciences* 507 (2020) 823–839.
44. C. Yang, and J. Wang, A mathematical model for the novel coronavirus epidemic in Wuhan, China, *Mathematical Biosciences and Engineering* 17 (3) (2020) 2708–2724.
45. W. X. Zhang, J. S. Mi, W. Z. Wu, Knowledge reduction in inconsistent information systems, *Chinese Journal of Computers* 1 (2003) 12–18.
46. W. Ziarko, Variable precision rough sets model, *Journal of Computer and Systems Sciences* 46 (1993) 39–59.