Journal of Contemporary Technology and Applied Engineering (JCTAE)
Home Page: https://jctae.journals.ekb.eg/
Published by: Faculty of Engineering, Kafrelsheikh University
Volume: 3 Issue: 2 Year: 2025
Pages: 55 – 62
DOI: 10.21608/jctae.2025.354124.1040
Print ISSN: 2812-5797
Online ISSN: 2812-5800

TOWARDS PAWLAK ROUGH APPROXIMATIONS THEORY WITH APPLICATIONS

Received: 20-01-2025
Accepted: 05-02-2025

Arafa A. Nasef[a], Mona A. Nasef[b,*], S. M. Elkholy[a,*]

[a] Department of Physics and Engineering Mathematics, Faculty of Engineering, Kafrelsheikh University, Kafrelsheikh, Egypt
[b] Department of Tropical Medicine, Infectious Diseases, Faculty of Medicine, Tanta University, Tanta, Egypt
[*]Corresponding author: M.A. Nasef (drmona11@yahoo.com) & S. M. Elkholy (samah.elkhouli1@eng.kfs.edu.eg)

ABSTRACT. When diagnosing a disease, the most difficulty thing doctors face is making an accurate decision to correctly determine the disease due to the similarity of the symptoms of different diseases. Therefore, in this research, using Pawlak's rough set model, upper approximation, lower approximation and by using nano-topology the factors affecting decision-making when afflicted with any disease were identified. Using the nano topology, we reduced the attributes in two real life situations by applying the knowledge as an information system. Here, we have demonstrated by topological reduction of the decision criteria of a recent outbreak of "Hepatitis C" that fever and yellow skin and eyes are the most important indicators of the disease. It became clear from this, that the point of view of mathematical methods is completely consistent with the medical expert's point of view. The effectiveness of rough set theory as a novel mathematical technique for extrapolating inferences from data has been demonstrated. It appears that the rough set concept will soon find very intriguing new applications. These consist of rough control, rough data bases, rough information retrieval, rough neural network and others.

KEYWORDS: Rough set, lower approximation, upper approximation, nano topology, Hepatitis C.
2020 AMS Classification Codes. 54Bo5, 54 Co5.

1. INTRODUCTION

The concept of Rough set was presented by Z. Pawlak [29] in his important paper in 1982 (Pawlak 1982). It is a formal theory based on basic studies of information systems' logical characteristics. Rough set model (RSM) has been methodology of database mining or knowledge discovery in rational databases. In Zadeh, 1965[14], abstract form, it is a new area uncertainty mathematics closely related to fuzzy model. Rough sets can be used to find structural relationships in noisy and imprecise data. Complementary generalizations of classical sets include fuzzy sets and rough sets. Whereas fuzzy sets deal with partial memberships, rough set theory's approximation spaces are sets with multiple memberships. The quick progress of these two methods provides a basis for "soft computing" introduced by Lotfi A. Zadeh [14]. In addition to rough sets, fuzzy logic, neural networks, probabilistic reasoning, belief networks, machine learning, evolutionary computing, and chaos theory are all included in soft computing.

Rough set models are essentially two subsets of a crisp partition established on the universal set involved that approximate a given crisp set. These subsets are referred to as inner and outer approximations, or lower and upper approximations, respectively. We can say that all of the partition's blocks that are part of the represented set make up the lower approximation, and all of the blocks whose intersection with the set is not empty make up the upper approximation. Rough set models are essentially two subsets of a crisp partition established on the universal set in question that approximate a given crisp set. These subsets are referred to as inner and outer approximations, or lower and upper approximations, respectively. We can say that all of the partition's blocks that are part of the represented set make up the lower approximation, and all of the blocks whose intersection with the set is not empty make up the upper approximation. Applications for RSM are numerous and include data mining, pattern recognition, artificial intelligence, knowledge discovery, machine learning, and decision analysis, among other areas.

From practical point of view rough set theory seems to be of fundamental significance to AI and cognitive sciences, especially to machine learning, knowledge discovery, decision analysis, inductive reasoning and pattern recognition. It seems also important to decision support systems and data mining. In fact it is a new mathematical approach to data analysis.

Rudiments of the theory can easily to be understood and applied. Several software systems based on rough set theory have been implemented and many real life, nontrivial applications of tis methodology have been reported, e.g., in medicine, pharmacology, engineering, banking, market analysis, conflict analysis, pattern recognition, environment, linguistics, gene expression and many more.

Rough set theory is based on sound mathematical foundation. The theory is not competitive but complementary to other methods and can also be often used jointly with other approaches (e.g., statistical methods, neural networks, genetic algorithms, fuzzy sets, etc.).

Rudiments of rough set theory can be found in [28, 30]. For recent development see [4,5 13,21]. Various extension of the theory can be found in [4,26]. Some applications of rough set theory are discussed in [13,21,26]. For more information the reader is advised to consult the internet.

The arrangement of the rest of this work is as follows: Section 2 reviews the literature needed to understand the article's concepts, and results. In Section 3, we first define information systems, Indiscernibility relation, Knowledge system and set of approximations. In Section4, we design the dynamic update reduction of attributes with some applications. Also, we propose an algorithm. Finally, we draw summary and concluding remarks in Section 5.

2. TERMINOLOGY CONCEPTS

This section presents a review of some fundamental notations and basic definitions of Pawlak rough set model.

RSM include the following notations:

U: denotes the universe of objects (states, patients, digits, cars, …, etc.), which can not be empty,

$R \subseteq U \times U$: denotes the indiscernibility relation or equivalence relation defined by an attribute set (i.e. $R = I(A)$ for some attribute set A),

$\boldsymbol{P} = (\boldsymbol{U}, \boldsymbol{R})$: denotes an approximation space or Pawlak approximation space,

$[\boldsymbol{x}]_{\boldsymbol{R}}$: denotes the equivalence class of an element $x$of Uunder the indiscernibility relation R, where $[\boldsymbol{x}]_{\boldsymbol{R}} = \{\boldsymbol{y} \in \boldsymbol{U}: \boldsymbol{xRy}\}$,

$\boldsymbol{U/R} = \{\boldsymbol{Y_1}, \boldsymbol{Y_2}, \ldots, \boldsymbol{Y_m}\}$on Udenotes the partition or the knowledge base $(\boldsymbol{U}, \boldsymbol{R})$of $\boldsymbol{Y_1}, \boldsymbol{Y_2}, \ldots, \boldsymbol{Y_m}$are called the equivalence classes generated by $\boldsymbol{R}$or elementary sets in $\boldsymbol{U/R}$,

$\boldsymbol{A}$: denotes a set consisting of attributes,

$(\boldsymbol{U}, \boldsymbol{A})$: is termed an information system (see, e.g. [28]). Any information system can be characterized by a data table with rows and columns labeled by objects and attributes, respectively.

For example, in Fig(1), the set of objects $\boldsymbol{U}$is divided into equivalence classes. The positive region of a set $\boldsymbol{X}$are all classes in the lower approximation of $\boldsymbol{X}$and the $\boldsymbol{Pos_R}(\boldsymbol{X})$are represented by the latter $\boldsymbol{P}$. Also, the classes in negative region $\boldsymbol{Neg_R}(\boldsymbol{X})$ are denoted with the letter $\boldsymbol{N}$and all other classes belong to the boundary region of the upper approximation.

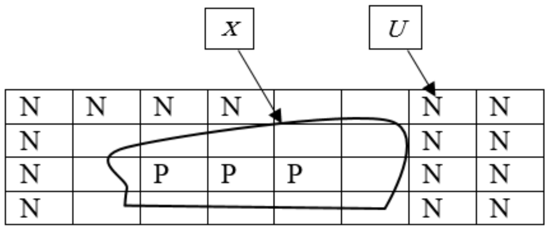

Fig. 1. Example of a positive and negative regions

3. LITERATURE REVIEW

3.1. INFORMATION SYSTEMS

Data model information is kept in tables in Rough Set. Every raw (tuples) denotes a fact or an object. A data table is named an information system in Pawlak model terminology [12, 28]. Thus, the information table expresses input data, collected from any domain of

Table 1. An Information Table

$\boldsymbol{U}$	Set of Attributes (A)			
Cases	**Temp.**	**Headache**	**Nausa**	**Cough**
C_1	high	yes	no	yes
C_2	very high	yes	yes	no
C_3	high	no	no	no
C_4	high	yes	yes	yes
C_5	normal	yes	no	no
C_6	normal	no	yes	yes

Note: For Table 1, rows are named examples (objects, entities). Information system denote as a pair (U, A), where $U$is a non-empty finite set of objects and A is a non-empty finite set of attributes. An information table sometimes termed as decision table (see Table 2) if it contains decision attribute/ attributes. Decision system in pairs of $(U, AU\{d\})$, where $d$is decision attribute (instead of one we can consider more decision attributes).

Table 2, can be shown in relation to function of nominal values of considered attributes, as in Table 3.

3.2. INDISCERNIBILITY RELATION

An indiscernibility relation is a relationship between two or more objects in which all values are the same with respect to a subset of the attributes under consideration. Numerous items with similar characteristics may be found in tables. One way to reduce the size of a table is to save a single representative item for each group of objects that share the same characteristics. These items are referred to as tuples or indiscernible objects. Any P subset of A has a corresponding equivalence relation, represented by IND(P):

$IND(P) = \{(x,y) \in U^2 | \forall a \in P, a(x) = a(y)\}$.

3.3. KNOWLEDGE SYSTEM

A knowledge expression system can be stated as a four –tuple in rough set theory as $T = (U,R,V,F)$, where

U: is a non-empty finite set (i.e., the universe) with n objects $\{\boldsymbol{u_1}, \boldsymbol{u_2}, \ldots, \boldsymbol{u_n}\}$, also known as the domain of discourse,

R: is a non-empty finite set of characteristics,

$V = \bigcup_{r \in R} V_r$ is the set of attribute values, the attribute value range for the $r \in R$is denoted by V_r,

$F: U \times R \to V$ is an information function, which corresponds to the attribute value of object x(i.e. $\forall x \in U, r \in R$, there is $F(x,r) \in V_r$) .

The decision table is the information knowledge expression system.

The attribute set is $R = C \cup D$, where $C$is the condition attribute set and $D$is the decision attribute set ($D \neq \varphi$).

Consequently, the information system is composed of all conditional and decision attributes.

3.4. SET OF APPROXIMATIONS

For a crisp set, its formal approximation defined by its two approximations are namely Upper approximation and Lower approximation [5,6, 11, 28].

The set of objects which possibly belong to the target set $X$is the Upper approximation, mathematically take the form:

$R^*[X] = \cup \{Y \in U/R: Y \cap X \neq \varphi\} = \{x \in U: [x] \cap X \neq \varphi\}$.

Lower approximation is the set of objects that positively belong to the target set X, and written as follows:

$R_*[X] = \cup \{Y \in U/R: Y \subseteq X\} = \{x \in U: [x] \subset X\}$,

where, $[x]$is the equivalence class of an element x.

The boundary region of the target set Xtake the mathematical form:

$BN_R(X) = R^*(X) - R_*(X)$.

A set is said to be rough if its boundary region is non-empty as shown in Fig. 2, otherwise the set is crisp.

Table 2. Decision Table

U	**Set of Conditional Attributes (*A*)**				**Decision Attribute (*d*)**
Cases	**Temp.**	**Headache**	**Nausa**	**Cough**	**Flu**
C_1	high	yes	no	yes	yes
C_2	very high	yes	yes	no	no
C_3	high	no	no	no	No
C_4	high	yes	yes	yes	yes
C_5	normal	yes	no	no	no
C_6	normal	no	yes	yes	yes

Table 3. Nominal Values

	Attributes	**Nominal Values**
Conditional Attributes	Temperature	Very high, high, normal
	Headache	Yes, No
	Nausa	Yes, No
	Cough	Yes, No
Decision Attributes	Flu	Yes, No

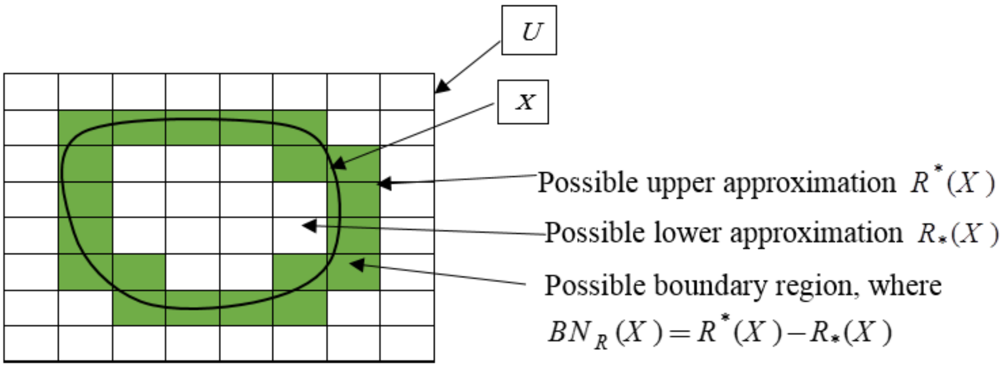

Fig. 2. Example of a rough set

PROPOSITION 3.1 [2,3,4,27]
We have the following properties of the Pawlak 's rough sets, in the following let $-X$be the complement of Xin U,

(1L) $R_*(U) = U$ (Co-normality)
(1H) $R^*(U) = U$ Co-normality)
(2L) $R_*(\theta) = \varphi$ (Normality)
(2H) $R^*(\theta) = \varphi$ (Normality)
(3L) $R_*(X) \subseteq X$ (Contraction)
(3H) $X \subseteq R^*(X)$ (Extraction)
(4L) $R_*(X \cap Y) = R_*(X) \cap R_*(Y)$ (Multiplication)
(4H) $R^*(X \cup Y) = R^*(X) \cup R^*(Y)$ (Addition)
(5L) $R_*(R_*(X)) = R_*(X)$ (Idempotency)
(5H) $R^*(R^*(X)) = R^*(X)$ (Idempotency)
(6L) $X \subseteq Y \Rightarrow R_*(X) \subseteq R_*(Y)$ (Monotone)
(6H) $X \subseteq Y \Rightarrow R^*(X) \subseteq R^*(Y)$ (Monotone)
(7L) $R_*(-R_*(X)) = -R_*(X)$ (Lower complement relation)
(7H) $R^*(-R^*(X)) = -R^*(X)$ (Upper complement relation)
(8LH) $R_*(-X) = -R^*(X)$ (Duality)
(9LH) $R^*(-X) = -R_*(X)$ (Duality)
(10L) $\forall K \in U/R, R_*(K) = K$ (Granularity)
(10H) $\forall K \in U/R, R^*(K) = K$ (Granularity)

From papers [1,11,23,24, 30], the above properties contain all essential properties of lower and upper approximations, since all the other properties of Pawlak 's lower and upper approximations can be inferred from the above properties.

DEFINITION 3.1[15-19]:
Let Ube the universe, Rbe an equivalence relation on $U$and $\tau_R(X) = \{U, \varphi, R_*(X), R^*(X), BN_R(X)\}$, where $X \subseteq U$. Then by Proposition 3.1, the class $\tau_R(X)$ satisfies the following axioms:

(i) $U$and$\varphi \in \tau_R(X)$,
(ii) The union of the elements of any subcollection of $\tau_R(X)$ is in $\tau_R(X)$,
(iii) The intersection of the elements of any finite subcollection of $\tau_R(X)$is in $\tau_R(X)$.

That is, $\tau_R(X)$is a topology on Unamed the nano topology on $U$with respect to $X$and the pair $(U, \tau_R(X))$ is named a nano topological space. The elements of $\tau_R(X)$are called nano open sets in $U$and the complement of nano open set is called a nano closed set. Elements of $[\tau_R(X)]^c$being called dual nano topology of $\tau_R(X)$.

REMARK 3.1:
If $\tau_R(X)$is a nano topology on $U$with respect to X, then Thivagar and Richared [13] observed that the family $\beta = \{U, R_*(X), BN_R(X)\}$is the basis for $\tau_R(X)$.

REMARK 3.2:
Let $(U, \tau_R(X))$be a nano topological space with respect to $X$where $X \subseteq U$and Rbe an equivalence relation on U. Then U/R denotes the family of equivalence classes of $U$by R.

DEFINITION 3.2 [1,7,16]:
If $(U, \tau_R(X))$is a nano topological space with respect to $X$where $X \subseteq U$and if $A \subseteq U$, then:

(i) The nano interior of the set $A$is defined as the union of all nano open subsets contained in $A$and denoted by $n\,Int(A)$. That is $n\,Int(A)$is the largest nano open subset of A.
(ii) The nano closure of the set $A$is defined as the intersection of all nano closed sets containing $A$and is denoted by $n\,Cl(A)$. That is, $n\,Cl(A)$is the smallest nano closed set containing A.

4. Reduction of Attributes

The challenge of whether some condition features may be removed without affecting the system's fundamental characteristics—that is, whether there is any unnecessary data—occurs when researching information systems. Rough set models can be used to reduce the number of attributes. This procedure is referred as attribute reduction or in context of machine learning as feature selection [7,10,20,25].

Getting the minimum possible subset of attributes that preserves the information of interest is the main idea of reducts. The procedure assumed is shown as in Fig.3.

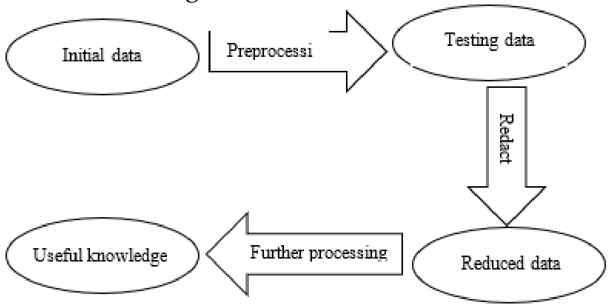

Fig. 3. Attribute Reduction Framework

This idea can be more exactly states as: Suppose $C, D \subseteq A$be subsets of conditions and decision attributes. We can say that $T \subseteq C$is a D- reduct (reduct with respect to D) of C, if $T$is a minimal subset of Csuch that:

$$\gamma(C, D) = \gamma(T, D)$$

The CORE is the intersection of all the reducts. This represents that the attributes are existing in all the reducts. Thus, the notion of the CORE is the most important set of attributes in the decision system and forms a basis of classification or decision power of attributes.

Example 4.1 In the following information table, we have a record of an exam in three different languages, German (G), English (E) and French (F), respectively, for five students in some school.

Table 4. Information system

Students U	**Condition attributes** G	E	F	**Decision attributes (result)**
St_1	false	true	false	fail
St_2	true	false	true	fail
St_3	true	true	true	pass
St_4	false	true	false	pass
St_5	true	false	false	fail

From the above table we have,
$U = \{St_1, St_2, St_3, St_4, St_5\}$and the knowledge base is $U/R = \{\{St_1, St_4\}, \{St_2\}, \{St_3\}, \{St_5\}\}$.

Case 1: Let $X = \{St_3, St_4\}$ be the set of students pass in the exam, then we can get that
$R_*(X) = \{St_3\}$, $R^*(X) = \{St_1, St_3, St_4\}$ and $Bnd_R(X) = \{St_1, St_4\}$.
Hence, $\beta(\tau_R(X)) = \{U, \{St_3\}, \{St_1, St_4\}\}$.

Step 1: If the attribute German (G) is removed from the set of conditions attributes (C)
, then the knowledge base is:
$U/R - G = \{\{St_1, St_4\}, \{St_2\}, \{St_3\}, \{St_5\}\}$.
Hence,$R_*(X) = \{St_3\}$, $R^*(X) = \{St_1, St_3, St_4\}$.
So, $\beta(\tau_{R-G}(X)) = \{U, \{St_3\}, \{St_1, St_4\}\} = \beta(\tau_R(X))$.
Thus $G \notin CORE(C)$.

Step 2: If we eliminate the attribute English (E) from the set of conditions attributes (C)
, then the knowledge base is:
$U/R - E = \{\{St_1, St_4\}, \{St_2, St_3\}, \{St_5\}\}$.
Consequently, $R_*(X) = \varphi$,
$R^*(X) = \{St_1, St_2, St_3, St_4\} = Bnd_R(X)$.
Hence, $\beta(\tau_{R-E}(X)) = \{U, \varphi, \{St_1, St_2, St_3, St_4\}\} \neq \beta(\tau_R(X))$.
Subsequently, $E \in CORE(C)$.

Step 3: If we remove the attribute French (F) from the set of conditions attributes (C)
, then the knowledge base is:
$U/R - F = \{\{St_1, St_4\}, \{St_2, St_5\}, \{St_3\}\}$ and we can get that $R_*(X) = \{St_3\}$, $R^*(X) = \{St_1, St_3, St_4\}$and $Bnd_R(X) = \{St_1, St_4\}$. hence,
$\beta(\tau_{R-F}(X)) = \{U, \{St_3\}, \{St_1, St_4\}\} = \beta(\tau_R(X))$.
Thus, $F \notin CORE(C)$.

Observation: From the above, we accomplish that English is the main attribute essential to decide whether the student has pass the exam or not.

Example 4.2: Consider Table 5. Which contains information about patients having Hepatitis C. The objects (the patients) are represented in rows and the columns of the table represent the attributes of Hepatitis C, where $A = \{Y, D, J, F\}$and described as follows:
Ystands for yellow skin and eyes,
Dstands for dark urine,
Jstands for joint and abdominal pains and
Fstands for fever.

Table 5. Infection information about some patients

Patients U	Set of attributes (A) Y	D	J	F	**Decision (C)**
Pa_1	1	1	1	+	Yes
Pa_2	1	0	0	+	No
Pa_3	1	0	0	+	Yes
Pa_4	0	0	0	++	No
Pa_5	0	1	1	+	No
Pa_6	1	1	0	++	Yes
Pa_7	1	1	0	-	No
Pa_8	1	1	0	++	Yes

where the symbol "1" means the patient has the symptom, "0" otherwise. Also, the symbol "+" means the patient has high fever, "++" means the patient has very high fever and the symbol "-" means the patient has no fever.

Set

$$U = \{Pa_1, Pa_2, Pa_3, Pa_4, Pa_5, Pa_6, Pa_7, Pa_8\}$$

represents the set of eight patients.

Case 1: Let $\boldsymbol{X} = \{\boldsymbol{Pa_1}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}$ be the set of patients having Hepatitis C. The family of all knowledge base from Table 4 is given by:

$\boldsymbol{U/R(A)} =$
$\{\{\boldsymbol{Pa_1}\}, \{\boldsymbol{Pa_4}\}, \{\boldsymbol{Pa_5}\}, \{\boldsymbol{Pa_7}\}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}\}$, where R represents equivalence relation on the universe U with respect to the condition attributes A.

The medical lower and upper approximations of the set patients having Hepatitis C with respect to R is given by:

$\boldsymbol{R_*}(\boldsymbol{X}) = \{\boldsymbol{Pa_1}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}$,
$\boldsymbol{R^*}(\boldsymbol{X}) = \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}$
and $\boldsymbol{BN_R}(\boldsymbol{X}) = \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}$.

Therefore, the base of nano topology is given by $\boldsymbol{\beta}(\boldsymbol{\tau_R}(\boldsymbol{X})) = \{\boldsymbol{U}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}\}$.

Hence,

$\boldsymbol{\tau_R}(\boldsymbol{X}) = \{\boldsymbol{U}, \boldsymbol{\varphi}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}\}$.

Step 1: If we remove the attribute "Y" from the set of condition attributes, then the family of knowledge base associated to the resulting set of attributes is given by:

$\boldsymbol{U/R-Y} =$
$\{\{\boldsymbol{Pa_4}\}, \{\boldsymbol{Pa_7}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_5}\}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}\}$.

Consequently, the corresponding lower and upper approximations are given by:

$\boldsymbol{R_*}(\boldsymbol{X}) = \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}$,
$\boldsymbol{R^*}(\boldsymbol{X}) = \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_5}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}$
and $\boldsymbol{BN_R}(\boldsymbol{X}) = \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_5}\}$.

Therefore, the base of nano topology is given by
$\boldsymbol{\beta}(\boldsymbol{\tau_{R-Y}}(\boldsymbol{X})) = \{\boldsymbol{U}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_5}\}\}$.

Hence,

$\boldsymbol{\tau_{R-Y}}(\boldsymbol{X}) = \{\boldsymbol{U}, \boldsymbol{\varphi}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_5}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_5}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}\} \neq \boldsymbol{\tau_R}(\boldsymbol{X})$.

Consequently, yellow skin and eyes is a factor from the core of Hepatitis C.

Step 2: If we remove the attribute "D" from the set of condition attributes, then the elementary set is given by:

$\boldsymbol{U/R-D} =$
$\{\{\boldsymbol{Pa_1}\}, \{\boldsymbol{Pa_4}\}, \{\boldsymbol{Pa_5}\}, \{\boldsymbol{Pa_7}\}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}\}$.

Which is the same knowledge base $U/R(A)$ and hence $\boldsymbol{\tau_{R-D}}(\boldsymbol{X}) = \boldsymbol{\tau_R}(\boldsymbol{X})$.

So, Dark urine does not a factor from the core of Hepatitis C.

Step 3: When the attribute "J" is removed from Table 5, then

$\boldsymbol{U/R-J} =$
$\{\{\boldsymbol{Pa_1}\}, \{\boldsymbol{Pa_4}\}, \{\boldsymbol{Pa_5}\}, \{\boldsymbol{Pa_7}\}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}\}$.

Which is the same knowledge base $U/R(A)$ and hence each of the basis and topologies are coincides. Hence "Joint and abdominal pains" does not a factor from the core of Hepatitis C.

Step 4: When attribute "F" is omitted from Table 5. Then one deduce that:

$\boldsymbol{U/R-F} =$
$\{\{\boldsymbol{Pa_1}\}, \{\boldsymbol{Pa_4}\}, \{\boldsymbol{Pa_5}\}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_7}, \boldsymbol{Pa_8}\}\}$.

and the related approximations are:

$\boldsymbol{R_*}(\boldsymbol{X}) = \{\{\boldsymbol{P_1}\}\}$,
$\boldsymbol{R^*}(\boldsymbol{X}) = \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_7}, \boldsymbol{Pa_8}\}$
and $\boldsymbol{BN_R}(\boldsymbol{X}) = \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_7}, \boldsymbol{Pa_8}\}$.

Therefore,

$\boldsymbol{\beta}(\boldsymbol{\tau_{R-F}}(\boldsymbol{X})) = \{\boldsymbol{U}, \{\boldsymbol{Pa_1}\}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_7}, \boldsymbol{Pa_8}\}\}$.

Also the corresponding nona topology is given by:

$\boldsymbol{\tau_{R-F}}(\boldsymbol{X}) = \{\boldsymbol{U}, \boldsymbol{\varphi}, \{\boldsymbol{Pa_1}\}, \{\boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_7}, \boldsymbol{Pa_8}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_7}, \boldsymbol{Pa_8}\}\} \neq \boldsymbol{\tau_R}(\boldsymbol{X})$.

Consequently, Fever is a factor from the core of Hepatitis C.

Step 5: If $\boldsymbol{M} = \{\boldsymbol{Y}, \boldsymbol{F}\}$which is a subset of the set of attributes, then the corresponding knowledge base with respect to M (i.e. when attributes D, J are omitted) is given by:

$\boldsymbol{U/R(M)} =$
$\{\{\boldsymbol{Pa_4}\}, \{\boldsymbol{Pa_5}\}, \{\boldsymbol{Pa_7}\}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}\}$.

Then:

$\boldsymbol{R_*}(\boldsymbol{X}) = \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}$,
$\boldsymbol{R^*}(\boldsymbol{X}) = \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}$
and $\boldsymbol{BN_R}(\boldsymbol{X}) = \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}$.

Therefore, the base for the nano topology corresponding to M is given by:

$\boldsymbol{\beta}(\boldsymbol{\tau_{R(M)}}(\boldsymbol{X})) = \{\boldsymbol{U}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}\}$.

So, the associated nano topology is

$\boldsymbol{\tau_{R(M)}}(\boldsymbol{X}) = \{\boldsymbol{U}, \boldsymbol{\varphi}, \{\boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}\}, \{\boldsymbol{Pa_1}, \boldsymbol{Pa_2}, \boldsymbol{Pa_3}, \boldsymbol{Pa_6}, \boldsymbol{Pa_8}\}\} \neq \boldsymbol{\tau_R}(\boldsymbol{X})$.

From the above steps we conclude that, CORE (Hepatitis C)={ Yellow skin and eyes, Fever}.

Case 2: Similarly, if X is taken as the set of patients not having "Hepatitis C", then by the same above technique, we conclude that CORE (Hepatitis C) = {Y,F}. So, Yellow skin and eyes and fever are the key attributes that has close connection to the disease of Hepatitis C.

Finally, we propose the following algorithm (Table 6) to describe how to use the nano topology in decision making for any information systems, through the Pawlak approximations.

Table 6. Algorithm 1: Algorithm on nano topology

Step no.	A decision making via nano topology
Step 1:	Input the information about patients suffering from Hepatitis C using the finite set U and a finite set A of attributes.
Step 2:	Compute the lower and upper approximation and boundary region for decision set $\boldsymbol{X} \subseteq \boldsymbol{U}$according to Definition 3.1.
Step 3:	Generate the nano topology $\boldsymbol{\tau_R}(\boldsymbol{X})$on U.
Step 4:	Remove the attribute $\boldsymbol{e_i}$from the condition attributes (C) and find the lower, upper approximations and boundary region of target set X on $\boldsymbol{C} - \boldsymbol{e_i}$for each $\boldsymbol{i} \in \boldsymbol{N}$.
Step 5:	Generate the associated nano topology $\boldsymbol{\tau_{R-\{e_i\}}}(\boldsymbol{X})$on U.
Step 6:	Repeat steps 4 and 5 for all attributes in C.
Step 7:	Those attributes in C for which $\boldsymbol{\tau_{R-\{e_i\}}}(\boldsymbol{X}) \neq \boldsymbol{\tau_R}(\boldsymbol{X})$forms the CORE (Hepatitis C).

5. Summary and Concluding Remark

As a novel mathematical tool for deriving conclusions from data, rough set theory has shown to be beneficial. The rough set concept appears to have very intriguing new applications coming soon. They consist of rough neural networks, rough data bases, rough information retrieval, rough control, and more.

Furthermore, other researchers have examined the algebraic and logical underpinnings of rough sets, providing a more comprehensive comprehension of the theoretical foundations of rough sets. The relationship between rough sets and statistical reasoning techniques has also attracted the attention of numerous researchers. Numerous additions to the "basic" paradigm of rough sets have also been developed and analyzed.

We have used the knowledge as an information system in characteristics reduction utilizing the nano topology in two real-life situations. We have identified a recent "Hepatitis C" outbreak using topological reduction in decision factors. Fever and yellow skin and eyes were recognized to be key factors for Hepatitis C.

Finally, a fascinating study that links the foundations of mathematics, quantum physics, and rough sets has recently been published [4,5,6,7].

Rough set model has also been used for knowledge representation, data mining, reducing knowledge representation, dealing with imperfect data and for analyzing attribute dependencies.

Finally, let us say that rough set theory is neither new set theory nor its improvement and it can be embedded in classical set theory

Acknowledgment

The authors would like to thank the editors and the anonymous reviewers for their comments and suggestions for improving the quality of the paper.

REFERENCES

[1] A. A. Nasef, A. I. Aggour, S. M. Darwesh, On some classes of nearly open sets in nano topological spaces, Journal of Egyptian Mathematical Society, 24, 4, (2016), 585-589.

[2] A.A. Nasef, A.E.F.A. El Atik, Topological kernel of sets and application on fractals, Journal of Contemporary Technology and Applied Engineering, 1(1) (2022) 1-8.

[3] A.A. Nasef, A.E.F.A. El Atik, Nearly open sets in infra topology and related infra continuity, Journal of Contemporary Technology and Applied Engineering, 2(1) (2023) 12-18.

[4] A.A. Nasef, M. Shokry, A.E. Shalabi and S.E. Kozaa, An application of information system via modern topology, Journal of Engineering Research, Vol. 6, No. 1 , (2022), 53-58.

[5] A.A. Nasef, M. Shokry, A.E. Shalabi and S.E. Kozaa, Reduction of information systems via Rough set model, Journal of Advanced Studies in Topology, (2022). 1-17.

[6] A.A.Nasef and M.K.El-Sayed, Rough set theory based on two relations and its applications, Asain Journal of Mathematics and applications, (2017),1-7.

[7] A. A. Nasef, R. Mareay, N. Youns, M. A. Nasef, M. Badr, Some Issues On Topological Operators Via Ideals, Journal of Contemporary Technology and Applied Engineering, 3(1) (2024) 25-31.

[8] A.Jayalakshni and C. Janaki, A new class of sets in nano topological spaces with an application in medical diagnoses, International Journal of Applied Engineering Research, Vol. 12. Nu. 16, (2017), 5894-5899.

[9] Armsteong Gl, Wasley A. Simard EP, McQuillan GM, Kuhnrt Wl and Alter MJ: The prevalence og

Hepatitis C viruss infection in the United States, 1999 through 2002, An International Med., (2006), 144: 705-714.

[10] B. F. Chellas, Modal Logic: An Iintroduction, Cambridge University Press, (1980).

[11] F. Jiang and G. Liu, Rough relational operators and rough entropy in rough relational database, In Second Int. Workshop on Knowledge Discovery and Data Mining, (2009), 761-764.

[12] G.Y. Weng, Rough set theory and knowledge discovery, Xi'an Jieotong Press, Xi'an, China, 2001.

[13] J. Sanabria , K. Rojo and F. Abad, A new approach of soft rough sets and a medical application for the diagnosis of Coronavirus disease, Mathematics, 8(2),(2022), 2686-2707.

[14] L.A. Zadeh, Fuzzy sets, Information Control, 8(1965), 338-353.

[15] Lellis Thivagar, M. and Priyalatha, S.P.R., Medical diagnosis in an indiscernibility matrix based on nano topology, Cogent Mathematics, Vol. 4, Iss. 1, (2017),1-9.

[16] Lellis Thivagar, M. and Richard, C., On nano continuity, Journal of Mathematical Theory and Modeling, 3, (2013b), 32-37.

[17] Lellis Thivagar, M. and Richard, C.: Nutrition Modeling Through Nano Topology, Int. Journal of Engineering Research and Applications, Vol. 4 Issue 1,(2014), 327- 334.

[18] Lellis Thivagar, M. and Richard, On nano forms of weakly open sets, International Journal of Mathematics and Statistics Inventions, Vol. 1, Iss. 1, (2013), 31-37.

[19] Lellis Thivagar, M., Paul Manuel and Sutha Devi, V.: A detection for patent infringement suit via nano topology induced by graph, Cogent Mathematics, 3, (2016), 1-10.

[20] Messina, J. P, et. al, Global distribution and prevalence of Hepatitis C virus genotypes, Hepatology, 61, (2015), 77-87.

[21] M. Riaz, F. Karaaslan, I. Nawaz and M. Sohail, soft multi-rough set topology with applications multi-criteria decision making problems, Soft Computing 25(2021), 799-815.

[22] NCD Risk Factor Collaboration, Worldwide trends in blood pressure from 1975 to 2015: a pooled analysis of 1479 population-based measurement studies with 19.1 million participants, Lance 389, (2017), 37-55.

[23] Q. Shen and R. Jensen, Rough sets, their extensions and applications, Int. J. of Automation and Comp., 4(3) (2007). 217-228.

[24] W Zhu, F.Y. Wang, Reduction and ariomization of covering generalized rough sets, Info. Sci. 152, (2003), 217-230.

[25] Xuekui Yu et al., Cryo-electron microscopy and three-dimensional reconstructions of Hepatitis C virus particles, Virology, 367, (2007), 126-134.

[26] Y. Goo, D. Chen and H. Wang, Optimal granularity selection based on algorithm stability with application to attribute reduction in rough set theory, Information Science, 654(2024),1-20.

[27] Y. Yao, Constructive and algebraic methods of theory of rough sets, Inf. Sci. 109 (1), (1998), 21-47.

[28] Z. Pawlak and Skowran, Rudiments of rough sets, Information Science, 177 (2007), 3-27.

[29] Z. Pawlak, Rough Sets, Int. J. Inf. Copmut. Sci.II, (1982), 341-356.

[30] Z. Pawlak, Rough sets, Theoretical Aspects of Reasoning About Data, Kluwer Academic Publisher Dordrecht, 1991.