Mathematical Innovations of a Modern Topology in Medical Events

M. Lellis Thivagar[1], Carmel Richard[2,*], Nirmala Rebecca Paul[2]

[1]School of Mathematics, Madurai Kamaraj University, Madurai-625021, TamilNadu, India
[2]Department of Mathematics, Lady Doak College, Madurai - 625002, Tamil Nadu, India
mlthivagar@yahoo.co.in, carmel09richard@gmail.com, nimmi_ rebecca@yahoo.com

Abstract   The purpose of this paper is to introduce a new topology called Rough Topology in terms of rough sets and prove that rough topology can be used to analyze many practical/real life problems. Using this concept, we find the deciding factors for the most common diseases chikungunya and diabetes.

Keywords   Rough Sets, Rough Topology, Lower Approximation, Upper Approximation, Core

1. Introduction

Rough set theory, introduced by Zdzislaw Pawlak, is a mathematical tool for representing, reasoning and decision making in the case of uncertain information. This theory deals with the approximation of sets or concepts by means of equivalence relations and is considered as one of the first non-statistical approaches in data analysis. Several interesting applications of the theory have come up, in particular, in Artificial Intelligence and Cognitive Sciences. The main advantage of rough set theory in data analysis is that, it does not require any preliminary or additional information of the data. The main difference between rough sets and fuzzy sets is that the rough sets have precise boundaries whereas fuzzy set theory is generally based on ill-defined sets of data, where the bounds are not precise and hence fuzzy predictions tend to deviate from exact values. The lower and upper approximations of a set are analogous to the interior and closure operations in a topology generated by data. In this paper, we have introduced a new topology called rough topology in terms of lower and upper approximations of a rough set and we have applied the concept of topological basis to find the deciding factors for chikungunya and diabetes.

2. Preliminaries

Definition 2.1 [6]: Let U be a non-empty finite set of objects called the universe and R be an equivalence relation on U named as the indiscernibility relation. The pair (U,R) is called the approximation space. Let X be a subset of U.

i) The lower approximation of X with respect to R is the set of all objects, which can be for certain classified as X with respect to R and it is denoted by $R_*(X)$. That is, $R_*(X) = \bigcup_{x \in U} \{R(x) : R(x) \subseteq X\}$ where $R(x)$ denotes the equivalence class determined by x.

ii) The upper approximation of X with respect to R is the set of all objects, which can be possibly classified as X with respect to R and it is denoted by $R^*(X)$. That is, $R^*(X) = \bigcup_{x \in U} \{R(x) : R(x) \cap X \neq \phi\}$.

iii) The boundary region of X with respect to R is the set of all objects, which can be classified neither as X nor as not-X with respect to R and it is denoted by $B_R X$. That is, $B_R(X) = R^*(X) - R_*(X)$.

The set X is said to be rough with respect to R if $R^*(X) \neq R_*(X)$. That is, if $B_R(X) \neq \phi$.

Proposition 2.2[6]: If (U, R) is an approximation space and X and Y are subsets of U, then

i) $R_*(X) \subseteq X \subseteq R^*(X)$
ii) $R_*(\phi) = R^*(\phi) = \phi$ and $R_*(U) = R^*(U) = U$
iii) $R^*(X \cup Y) = R^*(X) \cup R^*(X)$
iv) $R_*(X \cup Y) \supseteq R_*(X) \cup R_*(Y)$
v) $R_*(X \cap Y) = R_*(X) \cap R_*(Y)$
vi) $R^*(X \cap Y) \subseteq R^*(X) \cap R^*(Y)$
vii) $R_*(X) \subseteq R_*(Y)$ and $R^*(X) \subseteq R^*(Y)$ whenever $X \subseteq Y$
viii) $R_*(X^C) = [R^*(X)]^C$ and $R^*(X^C) = [R_*(X)]^C$
ix) $R_* R_*(X) = R^* R_*(X) = R_*(X)$
x) $R^* R^*(X) = R_* R^*(X) = R^*(X)$

Remark 2.3: $R^*: P(U) \to P(U)$ satisfies the Kuratowski closure axioms that

i) $R^*(\phi) = \phi$
ii) $X \subseteq R^*(X)$
iii) $R^*(X \cup Y) = R^*(X) \cup R^*(X)$
iv) $R^* R^*(X) = R^*(X)$ for all subsets X and Y of U

If $F = \{X \subseteq U / R^*(X) = X\}$,using conditions (i) to (iv), we see that ϕ and U are in F; $X \bigcup Y \in F$ whenever X and Y are in F and $\bigcap X_\alpha \in F$ for all X_α in F. Therefore, the family

T, of complements of members of F is a topology on U. Thus, F is the family of T-closed sets. Also, $Cl(X) = R^*(X)$. Therefore, R^* is the Kuratowski's closure operator.

Remark 2.4: Since $R_*: P(U) \to P(U)$ satisfies the following properties that

i) $R_*(U) = U$

ii) $R_*(X) \subseteq X$

iii) $R_*(X \cap Y) = R_*(X) \cap R_*(Y)$

iv) $R_*R_*(X) = R_*(X)$ for all subsets X and Y of U,

the operator R_* is the interior operator.

3. Rough Topology

In this section we introduce a new topology called rough topology in terms of the lower and upper approximations.

Remark 3.1: Let U be the universe of objects and R be an equivalence relation on U. For $X \subseteq U$, we define $\tau_R = \{U, \phi, R_*X, R^*(X), B_RX\}$, where $R^*(X)$, $R_*(X)$ and $B_R(X)$ are respectively the upper approximation, the lower approximation and the boundary region of X with respect to R. We note that U and $\phi \in \tau_R$. Since $R_*(X) \subseteq R^*(X)$, $R_*(X) \cup R^*(X) = R^*(X) \in \tau_R$. Also, $R^*(X) \cup B_R(X) = R^*(X) \in \tau_R$ and $R_*(X) \cup B_R(X) = R^*(X) \in \tau_R$. Also, $R_*(X) \cap R^*(X) = R_*(X) \in \tau_R$; $R^*(X) \cap B_R(X) = B_R(X) \in \tau_R$ and $R_*(X) \cap B_R(X) = \phi \in \tau_R$.

Definition 3.2: Let U be the universe, R be an equivalence relation on U and $\tau_R = \{U, \phi, R_*(X), R^*(X), B_R(X)\}$ where $X \subseteq U$. τ_R satisfies the following axioms:

i) U and $\phi \in \tau_R$.

ii) The union of the elements of any subcollection of τ_R is in τ_R.

iii) The intersection of the elements of any finite subcollection of τ_R is in τ_R.

τ_R forms a topology on U called as the rough topology on U with respect to X. We call (U, τ_R, X) as the rough topological space.

Example 3.3: Let $U = \{a,b,c,d,e\}$, $U/R = \{\{a,b\},\{c,d\},\{e\}\}$, the family of equivalence classes of U by the equivalence relation R and $X = \{a,c,d\}$. Then $R^*(X) = \{a,b,c,d\}$, $R_*(X) = \{c,d\}$ and $B_R(X) = \{a,b\}$. Therefore the rough topology $\tau_R = \{U, \phi, \{a,b,c,d\}, \{c,d\},\{a,b\}\}$.

Proposition 3.4: If τ_R is the rough topology on U with respect to X, then the set $B = \{U, R_*(X), B_R(X)\}$ is the basis for τ_R.

Proof:

i) $\bigcup_{A \in B} A = U$.

ii) Consider U and $R_*(X)$ from B. Let $W = R_*(X)$. Since $U \cap R_*(X) = R_*(X)$, $W \subset U \cap R_*(X)$ and every x in $U \cap R_*(X)$ belongs to W. If we consider U and $B_R(X)$ from B, taking $W = B_R(X)$, $W \subset U \cap B_R(X)$ and every x in $U \cap B_R(X)$ belongs to W, since $U \cap B_R(X) = B_R(X)$. And when we consider $R_*(X)$ and $B_R(X)$, $R_*(X) \cap B_R(X) = \phi$. Thus, B is a basis for τ_R.

Definition 3.5: Let U be the universe and R be an equivalence relation on U. Let τ_R be the rough topology on U and β_R be the basis for τ_R. A subset M of A, the set of at-

tributes is called the core of R if $\beta_M \neq \beta_{R-(r)}$ for every r in M. That is, a core of R is a subset of attributes which is such that none of its elements can be removed without affecting the classification power of attributes.

4. Rough Topology in Chikungunya

Here we consider the problem of Chikungunya, a disease that is transmitted to humans by virus- carrying Aedes mosquitoes. There have been recent breakouts of CHIKV associated with severe illness. It causes fever and severe joint pain. Other symptoms include muscle pain, headache and nausea. Initial symptoms are similar to dengue fever. It is usually not life threatening. But the joint pain can last for a long time and full recovery may take months. Usually patient gets lifelong immunity from infection and hence re-infection is very rare. In recent decades the disease has spread to Africa and Asia, in particular, the Indian subcontinent.

Consider the following information table giving data about 8 patients.

Table 1.

Patients	Joint pain (J)	Head-ache(H)	Nausea (N)	Temperature	Chikungunya
P_1	Yes	Yes	Yes	High	Yes
P_2	Yes	No	No	High	No
P_3	Yes	No	No	High	Yes
P_4	No	No	No	Very high	No
P_5	No	Yes	Yes	High	No
P_6	Yes	Yes	No	Very high	Yes
P_7	Yes	Yes	No	Normal	No
P_8	Yes	Yes	No	Very high	Yes

The columns of the table represent the attributes (the symptoms for chikungunya) and the rows represent the objects (the patients). The entries in the table are the attribute values. The patient P_5 is characterized by the value set (Joint pain, No), (Headache, Yes), (Nausea, Yes), (Temperature, High) and (Chikungunya, No), which gives information about the patient P_5. In the table, the patients P_1, P_2, P_3, P_6, P_7 and P_8 are indiscernible with respect to the attribute 'Joint pain'. The attribute 'Joint pain' generates two equivalence classes, namely, $\{P_1,P_2,P_3,P_6,P_7,P_8\}$ and $\{P_4,P_5\}$, whereas the attributes 'Joint pain' and 'Headache' generate the equivalence classes $\{P_1, P_6, P_7, P_8\}$, $\{P_2,P_3\},\{P_4\}$ and $\{P_5\}$. The equivalence classes for the attributes Joint pain, Headache, Nausea and Temperature are $\{P_1\}$, $\{P_2,P_3\},\{P_4\},\{P_5\},\{P_6,P_8\}$ and $\{P_7\}$. For the set of patients having chikungunya, the lower approximation $= \{P_1,P_6,P_8\}$ and the upper approximation $= \{P_1, P_2, P_3, P_6, P_8\}$ and hence the boundary region $= \{P_2, P_3\}$. Hence the patients P_2 and P_3 cannot be uniquely classified in view of the available knowledge. The patients P_1, P_6 and P_8 display symptoms which enable us to classify them with certainty as having chikungunya. In our case, the symptoms Jointpain, Headache, Nausea and Temperature are considered as the condition attributes and the disease chikungunya is considered as the decision attribute. Not all

condition attributes in an information system are necessary to depict the decision attribute before decision rules are generated. It may happen that the decision attribute depends not on the whole set of condition attributes but on a subset of it and hence we are interested to find this subset which is given by the core. Here U = {P_1,P_2,...,P_8}.

Case1: Let X = {P_1,P_3,P_6,P_8}, the set of patients having chikungunya. Let R be the equivalence relation on U with respect to the condition attributes. The family of equivalence classes corresponding to R is given by U/I(R) = {{P_1}, {P_2,P_3}, {P_4}, {P_5}, {P_6,P_8}, {P_7}}. The lower and upper approximations of X with respect to R are given by $R_*(X) =$ {P_1,P_6,P_8} and $R^*(X)=$ {P_1,P_2,P_3,P_6,P_8}. Therefore, the rough topology on U with respect to X is given by τ_R = {U, φ, {P_1,P_6,P_8}, {P_1,P_2,P_3,P_6,P_8}, {P_2,P_3}}.The basis for this topology τ_R is given by β_R = {U, {P_1,P_6,P_8},{P_2,P_3}}. If we remove the attribute 'Joint pain' from the set of condition attributes, the family of equivalence classes corresponding to the resulting set of attributes is given by U/I(R-(J)) = {{P_1,P_5}, {P_2,P_3}, {P_4}, {P_6,P_8}, {P_7}}. The corresponding lower and upper approximations are given by $(R-(J))_*(X)=$ {P_6,P_8}; $(R-(J))^*(X)=$ {P_1,P_2,P_3,P_5,P_6,P_8} and hence $\tau_{R-(J)}$ ={U,φ, {P_6,P_8}, {P_1,P_2,P_3,P_5,P_6,P_8}, {P_1,P_2,P_3,P_5}} and its basis $\beta_{R-(J)}$={U,{P_6,P_8}, {P_1,P_2,P_3,P_5}}≠ β_R. If we remove the attribute 'Headache' from the set of condition attributes, the family of equivalence classes corresponding to the resulting set of attributes is given by U/I(R-(H))={{P_1},{P_2,P_3}, {P_4},{P_5}, {P_6,P_8},{P_7}} which is the same as U/I(R) and hence $\tau_{R-(H)} = \tau_R$ and $\beta_{R-(H)} = \beta_R$. On removal of the attribute 'Nausea', we get U/I(R-(N))={{P_1},{P_2,P_3}, {P_4},{P_5},{P_6,P_8}, {P_7}}= U/I(R) and hence $\tau_{R-(N)} = \tau_R$ and $\beta_{R-(N)} = \beta_R$. When the attribute 'Temperature' is omitted, U/I(R-(T)) = {{P_1}, {P_2,P_3}, {P_4}, {P_5}, {P_6,P_7,P_8}}. $(R-(T))_*(X)=${P_1}; $(R-(T))^*(X)=${P_1,P_2,P_3,P_6,P_7,P_8}. Therefore,$\tau_{(R-(T))}$={U,φ, {P_1}, {P_1,P_2,P_3,P_6,P_7,P_8},{P_2,P_3,P_6,P_7,P_8}} and its basis $\beta_{R-(T)}$ = {U, {P_1}, {P_2,P_3,P_6,P_7,P_8}} ≠ β_R. If M = {J,T}, then the basis for the rough topology corresponding to M is given by β_M = {U,{P_6,P_8}, {P_1,P_2,P_3}}. Also $\beta_M ≠ \beta_{R-(x)}$ for all x in M. Therefore, **CORE(R) = {J,T}**.

Case2: Let X ={P_2,P_4,P_5,P_7},the set of patients not having chikungunya. Then U/I(R) = {{P_1}, {P_2,P_3}, {P_4}, {P_5}, {P_6,P_8}, {P_7}}. $R_*(X)$ = {P_4,P_5,P_7} and $R^*(X)$ = {P_2,P_3,P_4, P_5,P_7}. Therefore τ_R = {U,φ, {P_4,P_5,P_7}, {P_2,P_3,P_4,P_5,P_7}, {P_2,P_3}} and β_R = {U, {P_4,P_5,P_7},{P_2,P_3}}. Omitting the attribute 'Jointpain', U/I(R-(J)) = {{P_1,P_5}, {P_2,P_3}, {P_4}, {P_6,P_8}, {P_7}}. $(R-(J))_*(X)$ = {P_4,P_7}; $(R-(J))^*(X)$ = {P_1,P_2,P_3,P_4,P_5,P_7} and hence $\tau_{R-(J)}$ = {U, φ, {P_4,P_7}, {P_1,P_2,P_3,P_4,P_5,P_7}, {P_1,P_2,P_3,P_5}} and its basis $\beta_{R-(J)}$ {U,{P_4,P_7}, {P_1,P_2,P_3,P_5}} ≠ β_R. If the attribute 'Headache' is removed, U/I(R-(H)) = {{P_1}, {P_2,P_3}, {P_4}, {P_5},{P_6,P_8}, {P_7}} which is the same as U/I(R) and hence $\beta_{R-(H)} = \beta_R$. On removal of the attribute 'Nausea', we get U/I(R-(N)) = {{P_1}, {P_2,P_3}, {P_4}, {P_5}, {P_6,P_8}, {P_7}},which is the same as U/I(R) and hence $\tau_{R-(N)} = \tau_R$ and $\beta_{R-(N)} = \beta_R$. When the attribute 'Temperature' is omitted, U/I(R-(T)) = {{P_1}, {P_2,P_3}, {P_4}, {P_5}, {P_6,P_7,P_8}}. $(R-(T))_*(X)$ = {P_4,P_5}; $(R-(T))^*(X)$ = {P_2,P_3,P_4,P_5,P_6,P_7, P_8}.Therefore, $\tau_{R-(T)}$ = {U, φ, {P_4,P_5},

{P_2, P_3, P_4, P_5, P_6, P_7, P_8}, {P_2,P_3,P_6,P_7,P_8}} and its basis $\beta_{R-(T)}$ = {U, {P_4,P_5}, {P_2,P_3,P_6,P_7,P_8}} ≠ β_R. If M = {J,T}, U/I(r)= {{P_1,P_2,P_3}, {P_4,P_5}, {P_6,P_8}, {P_7}}, $r_*(X)$= {P_4,P_5,P_7} and $r^*(X)$= {P_1, P_2, P_3, P_4, P_5, P_7} where r is the equivalence relation on U with respect to M. Therefore, β_M = {U, {P_4,P_5,P_7},{P_1,P_2,P_3}} ≠ $\beta_{R-(x)}$ for every x in M. Therefore, here again, **CORE(R) = {J,T}**.

Observation: From both cases we conclude that 'Joint-pain' and 'Temperature' are the key attributes necessary to decide whether a patient has chikungunya or not.

5. Rough Topology in Diabetes

Diabetes is a group of metabolic diseases in which a person has high blood sugar, either because the body does not produce enough insulin, or because cells do not respond to the insulin that is produced. In diabetes, glucose in the blood cannot move into cells, so it stays in the blood. This not only harms the cells that need the glucose for fuel, but also harms certain organs and tissues exposed to the high glucose levels. This high blood sugar produces the classical symptoms of polyuria (frequent urination), weight loss and polyphagia (increased hunger).

Consider the following table giving information about six patients

Table 2.

Patients	Frequent Urination (F)	Weight Loss (W)	Increased Hunger (H)	Diabetes
P_1	Yes	Yes	No	Yes
P_2	Yes	No	Yes	Yes
P_3	Yes	No	No	Yes
P_4	No	Yes	Yes	No
P_5	No	Yes	No	No
P_6	No	No	Yes	No

Here, U = {P_1,P_2,P_3,P_4,P_5,P_6} and Frequent Urination (F), Weight Loss (W) and Increased Hunger (H) form the condition attributes. Let X ={P_1,P_2,P_3}, the set of patients having diabetes.

U/I(R) ={{P_1}, {P_2}, {P_3}, {P_4}, {P_5}, {P_6}}. The lower and upper approximations of X with respect to R are given by $R_*(X)$ = {P_1,P_2,P_3} and $R^*(X)$ = {P_1,P_2,P_3}. Therefore $B_R(X)$ = φ. The rough topology on U is τ_R = {U, φ, {P_1,P_2,P_3}} and its basis, β_R = {U, {P_1,P_2,P_3}}. If the attribute 'Frequent Urination' is removed from the set of condition attributes, then U/I(R-(F)) ={{P_1,P_5},{P_3},{P_4},{P_2,P_6}}. $(R-(F))_*(X)=${P_3}; $(R-(F))^*(X)$ = {P_1,P_2,P_3,P_5, P_6} and hence $\tau_{R-(F)}$ ={U, φ, {P_3}, {P_1,P_2,P_3,P_5,P_6},{P_1,P_2,P_5,P_6}} and its basis $\beta_{R-(F)}$= {U, {P_3}, {P_1,P_2,P_5,P_6}} ≠ β_R. If the attribute 'Weight Loss' is removed, then U/I(R-(W)) = {{P_1,P_3}, {P_2}, {P_4}, {P_6}, {P_5}}. $(R-(W))_*(X)$ = {P_1,P_2,P_3} and $(R-(W))^*(X)$={P_1,P_2,P_3} and hence $\tau_{R-(W)}$ ={U,φ,{P_1,P_2,P_3}} and its basis $\beta_{R-(W)}$ = {U,{P_1,P_2,P_3}} = β_R. If the attribute ' Increased Hunger' is removed, U/I(R-(H)) = {{P_1}, {P_2,P_3}, {P_4,P_5}, {P_6}}. $(R-(H))_*(X)$ = {P_1,P_2,P_3} and $(R-(H))^*(X)=$ {P_1,P_2,P_3}. Therefore $\tau_{R-(H)} = \tau_R$ and its basis $\beta_{R-(H)} = \beta_R$. If M ={F}, we see that $\beta_M ≠ \beta_{R-(F)}$. Therefore, CORE(R) = {F}. Similarly if

X is taken as the set of patients not having diabetes, then again **CORE(R) = {F}.**

Observation: Since the core of R has F as its only element, 'Frequent Urination' is the key attribute that has close connection to the disease diabetes .

The procedure applied in the above two cases can be put in the form of an algorithm as follows:

Algorithm:

Step 1: Given a finite universe U, a finite set A of attributes that is divided into two classes, C of condition attributes and D of decision attribute, an equivalence relation R on U corresponding to C and a subset X of U, represent the data as an information table, columns of which are labeled by attributes, rows by objects and entries of the table are attribute values.

Step 2: Find the lower approximation, upper approximation and the boundary region of X with respect to R.

Step 3: Generate the rough topology τR on U and its basis β_R.

Step 4: Remove an attribute x from C and find the lower and upper approximations and the boundary region of X with respect to the equivalence relation on C -(x).

Step 5: Generate the rough topology $\tau_{R-(x)}$ on U and its basis $\beta_{R-(x)}$.

Step 6: Repeat steps 3 and 4 for all attributes in C.

Step 7: Those attributes in C for which $\beta_{R-(x)} \neq \beta_R$ form the core (R).

6. Conclusions

In this work, we have shown that real world problems can be dealt with the rough topology. The concept of basis has been applied to find the deciding factors of a recent outbreak 'Chikungunya' which had been reported especially, in South India and a chronic disease 'Diabetes'. We could find that Joint pain and Temperature are the deciding factors for chikungunya and frequent urination is the only deciding symptom for diabetes. It is also seen that from a clinical point of view, the rough topological model is on par with the medical experts with respect to the diseases analyzed here. The proposed rough topology can be applied to more general and complex information systems for future research. The rough set model is based on the original data only and does not need any external information, unlike probability in statistics or grade of membership in the fuzzy set theory. It is also a tool suitable for analyzing not only quantitative attributes but also qualitative ones. The results of the rough set model are easy to understand, while the results from other methods need an interpretation of the technical parameters. Thus it is advantageous to use rough topology in real life situations.

REFERENCES

[1]   Jansi Rani,P.G., and Bhaskaran,R., Computation of reducts using topology and measure of significance of attributes, Journal of Computing 2 (2010)50-55.

[2]   Lashin,E.F., and Medhat,T., Topological reduction of information systems, Chaos, Solitons and Fractals 25 (2005) 277-286.

[3]   Rady,E.A., Kozae,A.M. and Abd El-Monsef,M.M.E., Generalized rough sets, Chaos, Solitons and Fractals 21 (2004) 49-53.

[4]   Salama,A.S., Some topological properties of rough sets with tools for data mining, IJCSI International Journal of Computer Science issues 8 (2011) 588-595.

[5]   Yuhuaqian, Chuangyin Dang, Jiye Liang, and Dawei Tang, Set- valued information systems, Information Sciences 179 (2009) 2809-2832.

[6]   Zdzislaw Pawlak, Rough set theory and its applications, Journal of Telecommunications and Information Technology 3(2002)7-10

[7]   Zuqiang Meng, and Zhongzhi Shi, A fast approach to attribute reduction in incomplete decision systems with tolerance relation - based rough sets, Information Sciences 179 (2009) 2774-2793.