ISSN 2278-3091
Volume 10, No.2, March - April 2021
International Journal of Advanced Trends in Computer Science and Engineering
Available Online at http://www.warse.org/IJATCSE/static/pdf/file/ijatcse921022021.pdf
https://doi.org/10.30534/ijatcse/2021/921022021

Cognitive Properties of MADM and Hybrid Rough Sets for Efficient Healthcare Test Diagnosis

Indrani Kumari Sahu[1], Susant Kumar Das[2]
[1]Khallikote (A) College, Odisha, India, indranisahu@gmail.com
[2]Berhampur University, Odisha, India, dr.s.k.das.1965@gmail.com

ABSTRACT

Healthcare test diagnostic comprises the methods of finding test results in accordance with the prescribed symptoms of disease. The logic used in terms of attribute values are not sufficient to quantify the actual conditions of diagnostics, tests, or prognostics. In other words, a patient's diagnosis-test measurement is assumed to be very severe, severe, or starting to be severe. Torationalization the quantified value of severity is not always discrete. To solve this problem, we use multi attribute decision making methods to identify alternative attributes through a hybrid methodology of rough and fuzzy relations in apessimistic and optimistic parameter of covering RST in the domain of healthcare test diagnostic domain.

Key words: Rough set, fuzzy rough set, variable precession RST, MADM, test-diagnostics, healthcare decision making.

1. INTRODUCTION

Medical informatics is an important service in healthcare system which has been stemming with innovations in medicine and information science since 1970. As, systematic diagnosis is an inherent property of successful treatment, modern healthcare emphasizes on some proactive cognitive properties like recommending for a treatment or identifying case-history of patient in addition to routine tasks and diagnosing. The advancement in computing technology in healthcare also encourages the practitioners for effective decision-making capabilities and reduces patient-favored attributes like time, cost, and diagnostic errors [1].

To date, significant contributions and investigations have been made in healthcare-based decision-making systems. Clinical decision support systems (CDSS) [2,3] are evolving in an aim to assist the choices of medical practitioners, to achieve patient satisfaction in terms of quality of care, uniform structure of medical interventions with an optimum associated cost. Most of the time, the decision-making capability purely depends on the medical practitioners' autonomy and authority in discharging duty. In some cases, the CDSS found to be challenging and assumed as a threat to some practitioners for their professional authority because of personal and biased insecure feeling on adopting the technology would lose their personal credentials and they may lose the control over the authorized health care work [4].

At the same time, CDSS found to be effective in the domain of knowledge acquisition and helpful to medical experts and researchers in identifying hidden relations among different medical indicators. The minimum set of indicators are very much helpful in identifying or predicting diseases, to make analysis for preventing the false diagnosis results, affect of parameters for prognosis, diagnostic tests, and treatment. The clinical knowledge acquisition tools help and treated as an integral part of medical decision-making system.

Rough set theory [5,6] modeled to describe a given set or concept through two determined sets, namely "*lower and upper approximation sets*", which divides the universe into positive and negative region with a difference between them as boundary region. The model is used in many applications to resolve several problems, such as to deal some key applications including, the "*information uncertainty*", "*feature selection*", "*knowledge reduction*", and "*rule extraction*". To curb the inexact or vague data, several extensions of RST have been evolved in updating the requirements of applications; "*Decision-Theoretic RST model*" [7], "*Variable Precision RST*"[8]are an effect on this regard that serves to induce decision rules from incomplete information systems [9]. Covering rough set [10] is a special model and applicable to deal real data sets in the cases of availability of overlapping multiple knowledge(s)in the dataset.

Fuzzy sets introduced in [11] came up with numerous fruitful real-life applications .It is to note that, in comparison to crisp relations, the fuzzy relation identifies relations from set of elements with natural proximity. Two extensions in this effect ("*fuzzy equivalence relation*" and "*fuzzy proximity relations*")are used in special cases of identifying fuzzy relations. In healthcare, test-diagnostic datasets also contain with continuous numerical attributes and often encounter as a special information system due to availability of heterogenous data from different sources. Hence, it is appropriate to use fuzzy-RST in CDSS. Many researchers brought multiple applications in this regard. Attribute

reduction and rule induction are some of the novel approaches on this regard. In [12], fuzzy-RST model is used for brain segmentation using MRI images. The method also handles the uncertainty and vagueness of information in MRI images.

Real-life decision-making systems often consists of high dimensional attributes and RST found to be a competitive methodology to reduce attributes using the popular "*attribute reduction*" methodology. Further, in some cases, there is still need of identifying the best alternative from the attribute-optimized decision system where the "Multi Attribute Decision Making (MADM)" outperforms to other classical methods.

This is also helpful in finding alternatives in a rearrangement order. MADM method found to be dominant and most selective decision-making systems, particularly in the domain of management science, economy, engineering, and product selection. MADM has been evolving with distinct methods such as, Aggregation operator method [13], ELECTRE method [17], TOPSIS method [18], PROMETHEE method [20] and its variances to find optimum alternatives from allied datasets.

The rest-structure of this paper is organized as follows. Section 2 shortly reviews the literature regarding RST and VPRST, fuzzy covering and fuzzy neighborhood properties and the hybrid usage of pessimistic and optimistic covering in VPRS. In section 3, we present test-diagnostic scenarios in healthcare system and present some identified tests, their purpose, and master parameters (discrete, sensitivity specificity) to validate our discussions along with classical MADM methodologies. Section 4 presents the usage of above methods (hybrid rough fuzzy and MADM) and discuss with an algorithm. Finally, section 5 lands with conclusion remarks and plan to improve some work using UCI medical datasets.

2. PRELIMINARIES

This section aims in recalling fundamental and relevant methods, concepts and processes which shall be used as a background knowledge for designing the scope of the work. Subsequent sections shall include fundamental concepts of classical rough set (RST), variable precision rough set (VPRST), operators associating fuzzy logic, covering on fuzzy approximation space and derive their reflexive, transitive, symmetric properties.

The structure of a dataset articulated like a table having rows represent data values whereas columns represent variables of data (attributes). An information system is an organized format of dataset with universe of objects and allied attribute sets. The decision system is a classified information system where the object occupies in a row and makes a distinction according to its decision attribute value (true or false) with set of conditional attribute values.

2.1 RST and VPRST

Pawlak's (1982) rough set model (RST) [6,7] offers an approximation space $K = (U, F)$ where $U$is the universe (non-empty set) and$F$ is the family of equivalence relation on U. The equivalence class of $F$is known as the quotient set of U by the relation F, $U/F = \{O_{i\in\{1,2,\dots,n\}}\}$.The elementary sets (equivalence classes of F) where objects of each equivalence class are indistinguishable to other classes. In other words, for every $(x, y) \in R, if\ (x, y) \in U$ then x and y are indistinguishable in K. R is represented as an indiscernibility relation (IND). If any attribute set P where, $P \neq \phi\ and\ P \subseteq R$ then $IND(P) = \cap P$; in other words, it is the intersection of all equivalence relations in P. If any attribute $a_i \in P$ and $IND(P)=IND(P-\{a_i\})$ then the attribute a_i is not necessary and permitted to omit from P, otherwise a_i is necessary element in the attribute set P. In the process of approximation of the objects and to classify using RST, the lower approximation (LA) is based on the concept of certainty $\underline{R}(X) = \{x \in U \mid [x]_A \subseteq X\}$, the upper approximation (UA) $\overline{R}(X) = \{x \in U \mid [x]_A \cap X \neq \phi\}$, has the members that can likely to be classified as members and the difference of them (boundary) is the area of uncertainty.

The above classical RST model has been experimented in numerous applications in data analysis but found to be sensitive for noisy data where both approximations do not evolve with promising normalization. Ziarko (1993) proposed an extensive model (VPRST) with partial classification for functional data patterns, known as variable precision parameter. According to the model, The model is an ordered pair $c(X, Y)$ with two dependent parameters, where X is a set and Y is the non-empty subsets of the universe. The calculation is based on the relative degree of misclassification of the first parameter with respect to each second parameter.

$$c(X,Y) = \begin{cases} 1 - \frac{|X\cap Y|}{|X|}, & |X| > 0 \\ 0, & |X| = 0 \end{cases} where\ 0 \le c(X,Y) \le 1.$$

In using $c(X,Y)$, the LA and UA are generalized with majority inclusion threshold $\beta\ (0 \le \beta < 0.5)$ as, $\underline{R_\beta}X = \cup\{E \in U/R \mid c(E,X) \le \beta\}$ and $\overline{R_\beta}(X) = \cup\{E \in U/R \mid c(E,X) < 1-\beta\}$ and known as β-*LA* and β-*UA* respectively. In accordance to the RST properties, the region specific relations (*positive, negative* and *boundary region)* are, $POS_\beta X = \cup\{E \in U/R \mid c(E,X) \le \beta\}, NEG_\beta X = \cup\{E \in U/R \mid c(E,X) \ge 1-\beta\}$ and $BNR_\beta X = \cup\{E \in U/R \mid \beta < c(E,X) < 1-\beta\}$ respectively.

To tackle the missing values or unknown values in incomplete information systems, VPRST model is found to be promising to derive embedded rules using its properties of $\beta - reducts$. A case study of VPRST, as follows:

A sample healthcare test-diagnose dataset (see, Table-1), with six rows and five columns to prescribed tests, column-D represents the expert decision and intermediate columns represent logical observation identifiers (LOINC)of prescribed symptoms of a probable disease type. The attribute

values are populated with discrete values obtained from sample healthcare practitioners, whereas U/N represents for unknown/not-relevant data value.

Table 1: Sample Healthcare test-diagnose dataset

Tests	$LOINC_1$	$LOINC_2$	$LOINC_3$	$LOINC_4$	D
t_1	1	1	1	0	1
t_2	0	U/N	1	0	1
t_3	U/N	U/N	0	1	0
t_4	1	U/N	1	1	1
t_5	U/N	1	1	1	2
t_6	0	1	1	U/N	1

As the values of the above table are vague, incomplete, or unknown, it is hard to derive conclusions and hence to be treated as a type of incomplete decision table. In using properties of VPRST as discussed above, the model identifies the set of *β-reducts* for incomplete decision table, as $\{S_3,S_4\}$ and the incomplete discernibility matrix for the β-reducts$\{S_3,S_4\}$, relative discernibility functions. The model derives conclusive generalized decision rules as, $\{(S_4,0) \rightarrow (D,1)\}$, $\{(S_3,1)\rightarrow(D,1)\vee(D,2)\}$, $\{(S_3,0)\rightarrow(D,0)\}$

2.2 Fuzzy coverings and fuzzy neighborhood properties

Binary logic is associated with two discrete logic values (true|1, false|0). In healthcare, these logic values are not sufficient to quantify the actual conditions of diagnostics, tests, or prognostics. In other words, a patient's diagnosis-test measurement is assumed to be very severe, severe or starting to be severe and updating the truth value with intensity range (between 0 and 1); the patient is very severe could have the degree of severe around 0.9, whereas the degree of 0.1 could quantify the sickness that the patients is recovered from the severity.

A binary function (*T*) satisfying commutative, associative properties and non-decreasing in both arguments where $T(1,e) = e \; \forall \, e \in [0,1]$ is termed as *t-norm*. The function is *left continuoust*-norm if $T(e, \vee_{i\in\prod} f_i) = \vee_{i\in\prod} T(e, f_i)$, where $\prod$ is an index and$i \in \prod$; $e_i, f \in [0,1]$.

A function (*N*) satisfying *N(1) =0* , *N(0)=1* and *N(N(e)) = e* and non-increasing for each $e \in [0,1]$ is termed as an *involutive-negator*.

A binary function (*V*) satisfying *V(1, 0) = 0, V(0, 1)=1*, *V(0, 0) =1*, *V(1, 1) =1* and *V(. , e)* is non-increasing, *V(e, .)* is non-decreasing for each $e \in [0,1]$is termed as an *implicator*. The In using the *left continuous t-norm*, the residual implicator can be defined as $V_T(e,f) = \vee\{g \in [0,1] \,|T(e,g) \leq f\} \, for \; each \; e,f \in [0,1]$. It is to be noted that, $\forall \, e,f,g \in [0,1], T(e,g) \leq f \leftrightarrow g \leq V_T(e,f)$

The concept of "covering" is identified through a covering approximation space (*U,C*), where *C* be family of subsets of *U* and *C* is a covering if $U = \cup_{C\in C\wedge C\neq 0} C$. Let *F(X)* is a collection of all fuzzy sets on *X* where *F:X*$\rightarrow$[0,1] be a fuzzy set. The variance of collection $C = \{F_i \in F(X) | F_i \neq 0, i \in \prod\}$is termed as a fuzzy coveringif $\forall \, a \in X$ there is a fuzzy set $F \in C$ such that $F(x) = 1$. The pair *(X,C)* is treated as fuzzy covering approximation space. In[14], the fuzzy minimal description (f_{minD}) and the fuzzy maximal description (f_{maxD}) were defined for each $x \in X$.

Moreover, for each $x \in X$, a neighborhood of x is,$N(x) = \cap\{C \in C : x \in f_{minD}(C,x)\}$. The mapping of $N: X \rightarrow F(X)$is a fuzzy neighborhood operator.For each $x_1, x_2 \in X$,there associate 4-types fuzzy neighborhood operators [14]. It is to note that, type-1 and type-3 satisfy "*reflexive and T-transitive*" with "*R-implicator V*", type-4 is "*reflexive and symmetric*" and if *C* is finite, type-2 is "*reflexive*".

$$N_1^C(x_1)(x_2) = \bigwedge_{F\in C} V\big(F(x_1), F(x_2)\big),$$

$$N_2^C(x_1)(x_2) = \bigvee_{F\in f_{minD}(C,x_1)} T(F(x_1), F(x_2)),$$

$$N_3^C(x_1)(x_2) = \bigwedge_{F\in f_{maxD}(C,a)} V(F(a), F(b)) \,, \; N_4^C(x_1)(x_2) = \bigvee_{F\in C} T(F(x_1), F(x_2))$$

2.3Covering,VPRST and Fuzzy Rough Set

Let $F \in F(X)$ with a fuzzy relation *R* on *X* having a variable precision $\gamma \in [0,1)$, both approximations (lower and upper) of variable precision fuzzy rough set [15] for any $x_1, x_2, x_3 \in X$ are,

$$\underline{Apr}^v_{R,\gamma}(F)(x_1) = \bigwedge_{x_1\in X} V(R(x_1,x_2), \gamma \vee F(x_2))$$

$$\overline{Apr}^T_{R,\gamma}(F)(x_1) = \bigvee_{x_2\in X} T(R(x_1,x_2), N(\gamma) \wedge F(x_2))$$

Similarly, a (*V, T*)-covering variable precision fuzzy RST on a fuzzy neighborhood having variable precision $\gamma \in [0,1)$ is defined through the following pairs of *F* on *X*.

$$\underline{C}^v_{N,\gamma}(F)(x_1) = \bigwedge_{x_2\in X} V(N(x_1,x_2), \gamma \vee F(x_2))$$

$$\overline{C}^p_{N,\gamma}(F)(x_1) = \bigvee_{x_2\in X} T(N(x_1,x_2), N(\gamma) \wedge F(x_2))$$

Similarly, on fuzzy covering approximations with fuzzy neighborhood operator[15], the covering -variable precession -fuzzy-rough set (CVPFR) of *F* on *X* is defined with following four equations. The first two pairs are pessimistic, and rest two pairs are optimistic covering based variable precession fuzzy rough sets. The role of variable precision (γ) identifies through the membership degree of *F* (the fuzzy set) and to enhance the representation of knowledge on fuzzy approximation space, the idle representation of $\gamma \in [0,1)$.

$$\underline{C}^{p}_{N,\gamma}(F)(x_1) = \bigvee_{x_2 \in X} T(N(x_2)(x_1), \bigwedge_{x_3 \in X} V(N(x_2)(x_3), \gamma \vee F(x_3)))$$

$$\overline{C}^{p}_{N,\gamma}(F)(x_1) = \bigwedge_{x_2 \in X} V(N(x_2)(x_1), \bigvee_{x_3 \in X} T(N(x_2),(x_3), N(\gamma) \wedge F(x_3)))$$

$$\underline{C}^{O}_{N,\gamma}(F)(x_1) = \bigwedge_{x_2 \in X} V(N(x_2)(x_1), \bigwedge_{x_3 \in X} V(N(x_2)(x_3), \gamma \vee F(x_3)))$$

$$\overline{C}^{O}_{N,\gamma}(F)(x_1) = \bigwedge_{x_2 \in X} T(N(x_2)(x_1), \bigvee_{x_3 \in X} T(N(x_2)(x_3), N(\gamma) \wedge F(x_3)))$$

Moreover, for any $0 \leq \gamma_1 \leq \gamma_2 < 1$ on fuzzyapproximation space, following properties hold:

1. $\underline{C}^{p}_{N,\gamma 1}(A) \subseteq \underline{C}^{p}_{N,\gamma 2}(A)$
2. $\overline{C}^{p}_{N,\gamma 2}(A) \subseteq \overline{C}^{p}_{N,\gamma 1}(A)$
3. $\underline{C}^{o}_{N,\gamma 1}(A) \subseteq \underline{C}^{o}_{N,\gamma 2}(A)$
4. $\overline{C}^{o}_{N,\gamma 2}(A) \subseteq \overline{C}^{o}_{N,\gamma 1}(A)$
5. $\underline{C}^{p}_{N,\gamma}(A)(X) = X$
6. $\overline{C}^{p}_{N,\gamma}(0) = 0$
7. $\underline{C}^{p}_{N,\gamma}(A) \subseteq \gamma_X \cup A$
8. $\underline{C}^{p}_{N,\gamma}(\mu_x) \subseteq (\gamma \vee \mu)_x$
9. $N(\gamma)_x \cap A \subseteq \overline{C}^{p}_{N,\gamma}(A)$
10. $\underline{C}^{p}_{N,\gamma}(\mu_x) = (\gamma \vee \mu)_x$
11. $\overline{C}^{p}_{N,\gamma}(\mu_x) \subseteq (N(\gamma) \wedge \mu)_x$
12. $\underline{C}^{p}_{N,\gamma}(A \cap B) \subseteq \underline{C}^{p}_{N,\gamma}(A) \cap \underline{C}^{p}_{N,\gamma}(B)$
13. $\overline{C}^{p}_{N,\gamma}(A \cup B) \subseteq \overline{C}^{p}_{N,\gamma}(A) \cup \overline{C}^{p}_{N,\gamma}(B)$
14. $\overline{C}^{p}_{N,\gamma}(A \cap B) \subseteq \overline{C}^{p}_{N,\gamma}(A) \cap \overline{C}^{p}_{N,\gamma}(B)$
15. $\underline{C}^{p}_{N,\gamma}(A \cup B) \subseteq \underline{C}^{p}_{N,\gamma}(A) \cup \underline{C}^{p}_{N,\gamma}(B)$
16. $\underline{C}^{p}_{N,\gamma}(\gamma_x \cup A) = \subseteq \underline{C}^{p}_{N,\gamma}(A)$
17. $\underline{C}^{p}_{N,\gamma}(\underline{C}^{p}_{N,\gamma}(A)) = \gamma_x \cup \underline{C}^{p}_{N,\gamma}(A)$
18. $IF\ A \subseteq B, then\ \underline{C}^{p}_{N,\gamma}(A) \subseteq \underline{C}^{p}_{N,\gamma}(B)$
19. $IF\ A \subseteq B, then\ \overline{C}^{p}_{N,\gamma}(A) \subseteq \overline{C}^{p}_{N,\gamma}(B)$

3. CVPFR FOR HEALTHCARE TEST DIAGNOSIS

In an approach to our earlier study [16], it is identified that, diagnosis error is a major flaw in present clinical healthcare systems. Some of the observed sources of such errors are due to, unintentional delay of diagnosis for lack of sufficient information, wrong or another diagnosis made before the correct one and no diagnosis is exhibited. The finding of evidence for studies of medical test performance on disease diagnosis is thus crucial.

3.1 Tests in Healthcare

A medical diagnostic test happens to be a systematic process of identifying, diagnosing, monitoring activities of disease(s) and most of the time prevention of disease(s). According to a healthcare market research firm ("*Kalorama*"), most of the physician's (approximately 80%) diagnoses process purely depend on the results of laboratory tests.

The patient-care test-setting is normally associated with two sections, *Anatomic pathology* (broad areas of histopathology, cytopathology, and electron microscopy) and *Medical Laboratory* (broad areas of, clinical chemistry/microbiology, molecular diagnostics, reproductive biology, hematology and blood bank). There are some common medical-tests and their purpose in context of our study is shown in Tabl-2. Table-3 and Table-4 represent reference Diagnosis-test master parameters of five test types (samples), used to identify the sensitivity of disease with discrete and sensitivity values.

Table-2: Common Diagnostic Tests and their purposes

TEST NAME	PURPOSE	TEST NAME	PURPOSE
Adam Test	*Diagnosis Of Scoliosis*	Ischiara Test	*Colour Vision*
Aids-Cd4 Count	*Predictor OfHiv (<200)*	Jegars Type Card Test	*Near Vision*
Aids-Elisa Test	*Screening ForHiv*	Knee Kiss Test	*Meningitis in children*
Aids-Western Blot Test	*Confirmative Test ForHiv*	Kramer's Index	*Neonatal Jaundice*
Aldehyde Test	*Leishmaniasis*	Kveim Test	*Sarcoidosis*
Allen Test	*Abg Analysis*	Nerve Condcn Test	*Gbs*
Arthrocentesis Test	*Joint Inflamation And Infection*	Orthotolidin Test	*Check Chlorine in water*
Bangle Test	*Protein Energy Malnutrition*	Pap's Smear	*Cancer of cervix*
Bender Gestalt Test	*Organic Mental Disorder*	Patch Test	*Allergic reaction*
Benedict Test	*Urine Glucose*	Paul Bunnel Test	*Epstein barr virus*
Benzidine Test	*Detection of blood in urine , stool & stomach contents*	PhallenManeuver /Tinnel Test	*Carpal tunnel syndrome*
Binnet Test	*Intelligence Quotient*	Pulmonary Functn Test	*Measuring Lung volume/capacity*
Bonny Test/ Marsall Test	*Stress Incontinence*	Rinee Test	*Conductive Hearing loss*
Braden Scale Test	*Measure integrity of bed sore*	Roll Over Test	*Pre-Eclampsia*
Burrow Ink Test	*Scabies /The Itch*	Romberg Test	*Neurological Function*
Cancer-Ca125 Test	*Ovarian Cancer*	Rorschach Test	*Schizophrenia*
Cancer Ca15-3 Test	*Breast Cancer*	Rothera Test	*Acetone In Urine*
Cancer Ca19-9	*Git (Pancreatic,*	Rubin Test	*Patency of*

Test	*Liver Cancer)*		*fallopian tube*
Complement Fixation Test	*Serological diagn-ones (Gonorrhoea, Syphlish etc.)*	Schick Test	*Diptheria*
Confrontation Test	*Central & Peripheral Visions*	Schiling Test	*Vitamin B12 Absorption*
Creatinine Clearance Test	*Estimation OfGfr*	Sgot And Sgpt,Ast /Alt	*Liver Function*
D-Dimer Test	*Measuring Clot Formation (Dvt,Pe,Dic)*	Shake /Bubble Test	*Surfactant/Fetal Lungs Maturity*
Dexa Scan	*Bone Metabollic Dis. (Osteoporosis)*	Smith Test	*Bile Pigment Urine*
Dexamethasone Suppression Test	*Acth/Pituitary /Ad-renal Gland Tumor*	Snellen Chart Test	*Distance Visions*
Dick Test	*Scarlet Fever*	Sweat Chloride Test	*Cystic Fibrosis*
Direct Coomb Test	*HemolyticAnemia*	Tape Test	*Pinworms*
Fem/Nitrizine Test	*Leakageof Amniotic-fluid& Anovulation*	Tensilon Test	*Myasthenia Gravis*
Fouchet's Test	*Bilirubin in Urine*	Tourniquet Test	*Dengue*
Frie Test	*Lymphogranuloma Inguinale*	Treadmill/Stres s Test	*Heart Function*
Glucose Tolerance Test	*Diabetes Mellitus*	Trendelen-burg Test	*Vericose Vein*
Glycosylated Hemoglobin /Hb1ac Test	*Diabetes Mellitus*	Triple Test	*Down Syndrome*
Gold Quantiferon Test	*Tuberculosis*	Tzanck Test	*Herpes Genital or Vericella*
Guaic Test	*Occult blood in stool*	Urea Breath Test	*H.Pylori*
Guthrie Test	*Phenzlketonuria*	Vdrl Test	*Syphlish*
Halo Test	*Csf Leakage*	Vma (Vanel-lylMan dellic Acid)	*Pheochromo-cyt oma*
Hanging Drop Test	*Cholera*	Water Depri-vation Test	*Diabetes Insipidus*
Hay's Test	*Bile Salt Urine*	Weber Test	*Localization Hearing Loss*
Heel to Shin Test	*Body coordination*	Weil Felix Test	*Typhoid Fever*
Histamine Test	*Leprosy*	Widal Test	*Typhoid (On 2nd Week)*
Holter Monitory Test	*4hr EcgMonitoiry In Dysrhythmia*	Wood's Light Examination	*Superficial infection of skin*

Today's healthcare is well integrated with Hospital Information System (HIS), Electronic Health Record (EHR), Laboratory instruments having standard test processing and reporting terminologies like, Logical Observation Identifiers (LOINC) and Nomenclature for Properties and Units (NPU) terminology.

At the same time, most clinical practitioners believe that, it is most difficult and challenging task to identify an appropriate or optimum test(s), out of allied numerous available probable tests. This brings towards an incorrect process of diagnoses (delayed, incorrect, or erroneous) through irrelevant test data and leads to failure in interpreting test results. Hence, the source of clinical errors associates with three major activities, first the challenge of prescribing most relevant diagnostic tests, second, difficulty in interpreting the diagnosis test reports appropriately and dealing with large volume of diagnostic-test cases. to establish patient specific and expert driven interpretations there is a vital need of CDSS, which would assist proactively, how to order for correct tests and help in making an appropriate analysis and prognosis.

3.1 Tests in Differential Diagnosis

Differential diagnosis happens to be a selective healthcare process of identifying one or multiple diseases of patient (multimorbidity)and considered to process through three interrelated phases. In phase-1, observed signs and symptoms of patient under study is interpreted into a "*diagnostic hypothesis*" (set of feasible diseases that could affect the patient) and passes to phase-2"*diagnostic tests (DT)*" (possibility of finding new facts/cases to accept or discard any set of hypotheses). Phase-3"*diagnostic refinement*" refines(adds/discards) the current set of diagnostic hypotheses in accordance with the results of phase-2. There may need of repetition of last two phases to obtain a final diagnosis. This is applicable when the results of phase-2 are not sufficient or more DTs are required to make a conclusion.

Table-3: Diagnosis-Test Master parameters (Discrete)

Test Name	MST_obs_Normal	MST_obs_Sensitive
Allen Test	(<3 Second) for Good	(>5 Second) for NotGood
Bangle Test	>80% for Normal	71-80% for Mild 61-70% for Moderate 51-60% for Severe
Glucose Tolerance Test	(100-125 mg/dL) fasting for Normal	(>=126 mg/dL) for Diabetes
Glycosylated Hemoglobin /Hb1ac Test	(< 5.7%) for Normal (5.7%-6.4%) for prediabetics	(>=6.5%) for diabetics
Schick Test	(=0.01) for bord_line (>0.01) for protected	(<0.01) for Risk

The selection of DTs to validate or discard a disease is in accordance to set of clinical scrutiny [21]. We discuss some of the following cases where the process of DT is invoked multiple times or infer to other DTs and is due to the extensive diagnostic observations during the phase-2,3 of Differential diagnosis process. Moreover, it also carries with parametric evaluations like, "*availability of DTs in the system"*, "*feasibility of DTs*", "*patient comfort and safety*". and "*medical costs*". Under such circumstances, the normal process of obtaining decision with such classical systems undermines the ranking properties, on test results. The

construct of covering based fuzzy RST with MADM approaches lead to promising rankings.

A. Repetition of same DT into number of times for observational conformity like high blood pressure symptom needs to be validated and repetition followed for second time, third time for observation conformity.

B. Few (not all) signs of symptoms can infer to a diagnostic hypothesis.

C. Selection of DT depends on present set of symptoms which infer to additional set of diagnostics, like a patient with "*dysthermia feeling*" can be checked to measure temperature to confirm for presence of fever.

D. The acceptance factor or discard factor of diagnostic hypothesis depends on observed findings.

E. Acceptance or discard of a hypothesis is dependent on absence of negation of some DT.

Table-4: Diagnosis-Test Master parameters(Sensitivity,Specificity)

Test Name	Sensitivity	Specificity
Adam Test	Sensitivity: 0.92	Specificity: 0.60
Aldehyde Test	Sensitivity 34.7% fever< 3 months Sensitivity 90.90% fever = 3 to< 6 months sensitivity 100% fever >= 6 months	Specificity is +ve at 96% Specificity is -ve at 94.9%
Urea Breath Test	Sensitivity 59% at 5 min & 100% at all other times of breath collection	Specificity 96% except at 10 and 20 min
Histamine Test	sensitivity: 90.7% [95% CI, 81.7 to 96.1]	specificity: 91.7% [73.0 to 98.9]
Roll Over Test	For the 20 mm Hg cut-off point, sensitivity: 20%	specificity: 93%

3.2 Multi Attribute Decision Making Methods

The selection of relevant test for specific disease on criteria-based symptoms is the concern in clinical expert system which is classified by MADM and FCAS approach. In this study, we discuss four well known methods, ELECTRE, TOPSIS, PROMETHEE and VIKOR method.

A ELECTRE method

The ELECTRE method ("Elimination Et Choix Traduisant la REalite") [17] defines an outranking relation between alternatives, taken two at a time. In other words, an alternative A_k outranks $A_p$if the process exhibits equal or better results than A_p. The evaluation process must include in terms of most criteria. The method exhibits in identifying a subset of options to be considered (preferably on the remaining options) in normal conditions but hard in finding a complete ranking of the alternatives.

B TOPSIS method

In [18] a novel MADM method introduced naming "Technique for Order Preference by Similarity to Ideal Solutions" or TOPSIS. This method aims in identifying the best alternative in each set of items having a shortest distance and farthest distance, so called ideal and negative-ideal solutions (X^*, X^-) respectively and normalizes with a decision matrix and weights. Further, best, and worst performances are evaluated from X^* and X^-in comparing each criterion among all the alternatives. The set $\{A_i, X^*, X^-\}$ is further represented geometrically in a space with m-dimension. This is based on the weighted normalized value of alternative for each criterion. The Euclidean distances of A_i from X^* (ideal solution) and X^-(negative ideal solution) are represented as S_{i*} and S_i- respectively. TOPSSIS method defines the relative closeness of A_i to the ideal solution as, $C_i^* = S_i^- / (S_i^* + S_i^-)$. Corresponding values of solution are evaluated ([0, 1]) as, $C_i^* = 0 \; for \; each \; A_i = A^-, S_{i\_} = 0$ and $C_i^* = 1$ for each $A_i = A^*, S_{i*} = 0$. The maximum C_i^* value carries the best or optimum solution.

C PROMETHEE method

PROMETHEE method [19] is based on the preference methodology based on individual criterion and weights, where the importance of criterion is quantified through the weights. The variances of PROMETHEE method with I, II, III, IV are based on their methodologies like "*partial ranking*", "*complete ranking*", "*ranking based on intervals*" and "*continuous cases*", respectively.

The method exhibits the core principle of decision problem and a degree of complexity processed through the preference function in respect to each criterion fixed by the decision maker. The second variance of this method favors for a complete ranking of options but is dependent on the manipulation of information, which is not always has a logical meaning. Somehow, the result in identifying the complete ranking of alternatives, in cases of covering based variable precision fuzzy RST information of diagnostic-test is not promising.

D VIKOR Method

The MADAM model ("VlseKriterijumskaOptimizacija I KompromisnoResenje" – VIKOR) [20] evaluates the ranks of the alternatives (A_i) in using S_i, R_i, e, Q_i, where,

$$S_i = \sum_{j=1}^{m} w_j \, (a_j^* - a_{ij}) / (a_j^* - a_j^-)$$

$$R_i = max_j \, [w_j (a_j^* - a_{ij}) / (a_j^* - a_j^-)]$$

$$Q_i = v(S_i - S^*) / (S^- - S^*) + (1 - v)(R_i - R^*) / (R^- - R^*)$$

$R^*$and S^* are evaluated with minimum of their individual elements whereas, R^- and S^- takes on with maximum of their individual elements. The best and worst performances

(a_j^*, a_j^-) are identified from all alternatives of each criterion (C_j). The parameter 'v' is an expert decision maker ([0,1] interval). In other words, it is a quantified identification between two aspects with three cases. If two terms are equally relevant then (v=0.5); if the first term is having importance then (v > 0.5); otherwise (v < 0.5) i.e., when to consider more relevance on the second term.

Let X={t_1, t_2, t_3, t_4, t_5} be the set of five test cases and each type of test is evaluated by four symptoms (attributes). The calculated weight vector of four symptoms by practitioners be W={0.2, 0.2, 0.3, 0.3}, where $\sum_{i=1}^{n} w_j = 1$. Table-2 represents the efficacy results of diagnostic-tests with respect to the prescribed observed-symptoms.

Table-2: Findings of test results to the prescribed symptoms

X	S_1	S_2	S_3	S_4
t_1	0.36	0.54	0.47	0.65
t_2	0.59	0.20	0.81	0.42
t_3	0.56	0.28	0.73	0.45
t_4	0.39	0.36	0.65	0.62
t_5	0.52	0.33	0.68	0.49

In using the above MADM models, the ranking effectiveness of the above diagnosis tests{t_i} in respect to the prescribed symptoms {S_i} is $t_3 \approx t_5 \approx t_4 \approx t_2 \approx t_1$. The VIKOR method offers to explicitly account for the degree of satisfaction of a single criterion besides the global performance to the whole of criteria and for the double check of acceptability for the final solution .

4. MADM AND CVPFR MODEL FOR HEALTHCARE TEST DIAGNOSIS

Healthcare test diagnosis is purely dependent on the information collected during the observation, test-results, diagnosis and post-diagnosis observations. Sometimes, the decision-making process exhibits without availability of sufficient information and thus, the dataset is considered as imprecision or incomplete. Hence the properties of covering based RST, fuzzy-RST and multi-attribute decision making procedures are useful in which data, knowledge, hypothesis, principles, procedures related to healthcare decision making system can be a motivational solution. We make an attempt to use such powerful decision-making capabilities in test-diagnostic process.

4.1 A novel approach of CVPFR-VIKOR Model for Healthcare Test Diagnostics

Suppose a sample diagnosis test-dataset has *m* rows of test cases D={t_1, t_2,…, t_m) and *n* columns of associated symptoms C={S_1, S_2, …., Sn}. In order to use as an put to the desired MADM model, first the diagnosis test-dataset is transposed to a matrix. Then, the test-dataset is processed for crispness which is achieved in checking whether the conditional attribute values satisfy exact truth value (yes/true, no/false).If satisfies with crispness, then the model is fit for classical rough set analysis. However, in general, most of the test analysis are based on the symptom severity with quantification operators of sensitivity, specificity, and ROC values. To be specific, the boundary of attribute value is vague and inherited with features of imprecision, incompleteness, and uncertainty.

The weight vector W= {w_1, w_2, …, wn} is a general consideration of experts' opinion, where $0 \leq w_j \leq 1 \; and \; \sum_{j=1}^{m} w_j = 1$. The weights' assignment process or rationales of weights is also based on eigenvalue's theory [22, 23]. It also considers the compromise coefficient ($\emptyset \in [0,1]$) is evaluated with experts weighing factor on "maximum group utility", "minimum individual regret" and the balance between them, with $\emptyset > 0.5, \emptyset < 0.5 \; and \; \emptyset = 0.5$ respectively.

4.2 Algorithm

This subsection describes the algorithm and detailed procedure for obtaining healthcare test diagnostics ranking model using the techniques of hybrid approach of variable precision rough set using features of covering and fuzzy approximation space in an attempt with multi attribute decision making method.

Algorithm: Healthcare test diagnosis ranking model with CVPFRS and MADM.
Input: Medical diagnosis test-dataset matrix D
Coefficient of compromise $\emptyset \in [0,1]$
Output: Diagnostic Test-Rank
begin
for i = 1: m do
for j = i : n do
$bp_{ij} = max_{1 \leq i \leq n} D_{ij}$
$bn_{ij} = min_{1 \leq i \leq n} D_{ij}$
end
end
compute: group utility value S(a_i) and each regret value R(a_i)
for i = 1: m do
for j = i : n do
$S(a_i) = \sum_{j=1}^{m} w_j \;\; (bp_{ij} - D_{ij})/(bp_{ij} - bn_{ij})$
$R(a_i) = max[w_j(bp_{ij} - D_{ij})/(bp_{ij} - bn_{ij})]$
end
end
compute $RP = max_{1 \leq i \leq n} R(a_i)$
compute $RN = min_{1 \leq i \leq n} R(a_i)$
compute $SP = max_{1 \leq i \leq n} S(a_i)$
compute $SN = min_{1 \leq i \leq n} S(a_i)$
compute: compromise value Q(a_i)
for i = 1: n do
$Q(a_i) = \emptyset(S(a_i) - SN)/(SP - SN) + (1 - \emptyset)(R(a_i) - RN)/(RP - RN)$
end
compute: Fuzzy Sets
for i = 1: n do
$X^U = \sum_{i=1}^{n} S(a_i)/a_i$
$X_L = \sum_{i=1}^{n} R(a_i)/a_i$

$$X\_BR = \sum_{i=1}^{n} Q(a_i)/a_i$$

end

compute: min_ind_reg_fuzzy and *comp_fuzzy*

$$\underline{C}^{p}_{N,\gamma}(F)(x_1) = \bigvee_{x_2 \in X} T(N(x_2)(x_1), \bigwedge_{x_3 \in X} V(N(x_2)(x_3), \gamma \vee F(x_3)))$$

$$\overline{C}^{p}_{N,\gamma}(F)(x_1) = \bigwedge_{x_2 \in X} V(N(x_2)(x_1), \bigvee_{x_3 \in X} T(N(x_2),(x_3), N(\gamma) \wedge F(x_3)))$$

$$\underline{C}^{O}_{N,\gamma}(F)(x_1) = \bigwedge_{x_2 \in X} V(N(x_2)(x_1), \bigwedge_{x_3 \in X} V(N(x_2)(x_3), \gamma \vee F(x_3)))$$

$$\overline{C}^{O}_{N,\gamma}(F)(x_1) = \bigwedge_{x_2 \in X} T(N(x_2)(x_1), \bigvee_{x_3 \in X} T(N(x_2)(x_3), N(\gamma) \wedge F(x_3)))$$

compute: Overall fuzzy sets related to min_ind_reg, max_grp_utiland compr_F_BR) with the bounded sum (B_S(e,f) = e+f-e*f, for each $e, f \in [0,1]$).

$$F_L(t_i) = B_S[\underline{C}^{p}_{N,\gamma}(X_L)(t_i), \overline{C}^{p}_{N,\gamma}(X_L)(t_i)]$$

$$F^U(t_i) = B_S[\underline{C}^{p}_{N,\gamma}(X^U)(t_i), \overline{C}^{p}_{N,\gamma}(X^U)(t_i)]$$

$$F\_BR = B_S[\underline{C}^{p}_{N,\gamma}(X\_BR)(t_i), \overline{C}^{p}_{N,\gamma}(X\_BR)(t_i)]$$

Compute: Sort F_BR in ascending (in terms of $[t_i]$)

if $(F\_BR[top+1] - F\_BR[top] \geq 1/(n-1))$ and $F_{BR[top]} == (asc(F_L[top])\ or/and\ asc(F^U[top]))$ then

Test_Rank= F_BR[top]

endif

end

5. CONCLUSION

MADM methods have their potential to use in decision making applications. Although the large literature and distinct characteristics of MADM models have favorable capability to deal with different problems, the paper investigated four methods for an actual applicability and effectiveness of them in identifying best or optimum healthcare diagnostic-test cases in accordance with the prescribed symptom sets. The VIKOR-MADM method found to be promising to the multi attribute decision task involving the selection of an optimum solution for fuzzy relation-based diagnostic-test, due to its capability in dealing with distinct judgement criteria. The novel method is designed in using hybrid rough set, variable precision fuzzy relation, pessimistic and optimistic covering to identify optimum alternative attribute sets from a healthcare-based test-diagnostic system. It is an attempt to rationalize the quantified test cases for an optimum analysis in healthcare diagnostic model. The proposed algorithm has the complexity of $O(n^2+mn)$ and useful in the cognitive applications of healthcare. Future extension of this can be viewed on accessing with UCI medical datasets.

REFERENCES

1. W. Hussain, W. Ishak,and F. Siraj. **Artificial Intelligence in Medical Applications: An Exploration**, *Health Informatics,* Europe Journal, 2002.
2. R.A. Miller. **Medical diagnostic decision support systems- Past, present, and future: A threaded bibliography and brief commentary**. *Journal of the American Medical Informatics Association*,Vol. 1(1):8, January 1994.
3. E.H. Shortliffe. **Computer Programs to Support Clinical Decision Making.***JAMA: The Journal of the American Medical Association*, 258(1), pp. 61-66, July 1987.
4. E. Sezgin and S.O. Yildirim. **A Literature Review on Attitudes of Health Professionals towards Health Information Systems: From e-Health to m-Health**,*Procedia Technology*, pp. 1317-1326, 2014.
5. Z. Pawlak. **Rough Sets**, *Int. Journal ofComputer and Information Sciences*, 11(5), pp. 341-365, 1982.
6. Z. Pawlak.**Rough sets:theoretical aspects of reasoning about data**, Kluwer Academic Publishers, Dordrecht, 1991.
7. Y.Y. Yao,**Information granulation and rough set approximation**, *Int. Journal of Intelligent. Systems* 16(1), pp-87-104, 2001.
8. W.Ziarko.**Variable precision rough set Model**, *Journal of Computer and system sciences,Elsevier*, 46(1), pp. 39-59, 1993.
9. S.B. Sun, R.Zheng, Q. Wu, T. Li.**VPRS-Based Knowledge Discovery Approach in Incomplete Information System**. *J of Computers, JCP*, 5(1), pp. 110-116, 2010.
10. W. Zhu, F. Y. Wang. **Relationship among basic concepts in covering-based rough sets**. *Information Sciences, Elsevier*, 179, pp. 2478-2486, 2009.
11. L. A. Zadeh, **Fuzzy Sets**, *Information and Control*, Vol. 8, pp. 338-353, 1965.
12. A. Namburu, S.K.Samay, and S.R.Edara**. Soft fuzzy rough set-based MR brain image segmentation**, *Applied Soft Computing, Elsevier*, vol. 54, pp. 456-466, 2017.
13. Z.S. Xu, Q.L. Da.**An overview of operators for aggregating information**,*Int J of Intelligent systems, Wiley Online Library*, 18, pp. 953-969, 2003.
14. L. D'eer, C. Cornelis, L. Good. **Fuzzy neighborhood operators based on fuzzy covering**,*Fuzzy Sets Systems, Elsevier*, 312, pp. 17-35, 2017.
15. S. Y. Zhao, E. C. C. Tsang, D. G. Chen. **The model of fuzzy variable precision rough sets**, *in Proc. 2007 Intl. Conf. on Machine Learning and Cybernetics, IEEE*, Vol. 6. 3057-3062, 2007.
16. I.K.Sahu, G.K. Panda, S.K. Das. **An Enhanced Diagnostic Model based on Rough Granular Computing approach**, *Journal of High Technology Letters*, vol 27 (2), pp. 378-390, 2021.
17. R. Benayoun, B. Roy, and B. Sussman.**ELECTRE: Une methode pour guider le choi xen presence de point de vue multiples**, *Note Travail*, vol. 49, Jun, 1966.
18. C.L. Hwang, and K. Yoon. **Multiple Attribute Decision Making**, *Lecture Notes in Economics and Mathematical Systems*, Springer-Verlag: Berlin, Germany, 186, 1981.
19. J.P. Brans.**Lingenierie de la decision; Elabrationd 'instrumentsd'aide a la decision. La methode PROMETHEE.***In R. Nadeau and M. Landry*, editors, L'aide a la decision: Nature, Instruments et Perspectives d'Avenir, pp. 183-213, 1982.

20. S.Opricovic and G.H. Tzeng. **Compromise solution by MCDM methods: A comparative analysis of VIKOR and TOPSIS**, *European Journal of Operational Research*, 156(2), pp. 445-455, 2004.
21. Lopez-Vallverdu, A. Joan, D.Riano, and J. A.Bohada, **Improving medical decision trees by combining relevant health-care criteria**. *Expert Systems with Applications*, 39.14, pp. 11782-11791, 2012.
22. T.L.Saaty.***The Analytic Hierarchy Process***, McgrawHill, New York, Agricultural Economics Review 70, 1980.
23. N.Caterino, I.Iervolino, G. Manfredi, and E. Cosenza.**Multi-criteria decision making for seismic retrofitting of RC structure**, *Journal of Earthquake Engineering*, 12(4), pp. 555-583, 2008.