Applications of Rough Sets in Health Sciences and Disease Diagnosis

Aqil Burney[1,2], Zain Abbas[2]
[1] College of Computer Science and Information Systems, Institute of Business Management
Korangi Creek, Karachi-75190
[2]Department of Computer Science, University of Karachi
University Road, Karachi – 75270
PAKISTAN
aqil.burney@iobm.edu.pk zain@uok.edu.pk

Abstract: - Soft computing is a consortium of techniques that work together to setup flexible information processing capability for handling real-life ambiguous situations. It aims at solving problems involving uncertainty and imprecision mimicking the human like decision making. Fuzzy set theory is an approach that has been widely adopted in such situations. Rough Set Theory (RST) is another soft computing approach that uses sets to represent vague or incomplete knowledge and provide a framework for approximation of concepts. It has been widely used to deal with imprecision in health sciences such as in patient diagnosis and disease classification. In this paper we present a review of rough set theory and its applications in disease diagnosis with several examples using real data sets.

Key-Words: - Rough Set Theory, Soft Computing, Vague data, Imprecision, Health Sciences, Disease diagnosis

1 Introduction

Research in computational intelligence and reasoning systems has to deal with the limitation of incomplete or imperfect knowledge. The ability to work with this kind of knowledge is of immense importance in studying applications in cognitive computing and automation of systems, especially in the areas of machine learning and data-mining, intelligent control, decision analysis and pattern recognition. [1]

Soft computing is a blend of techniques that complement one another and provide diverse and flexible data processing capabilities to extract information in a number of ambiguous situations encountered in daily life [2, 3]. Its aims is mimic the human like decision making ability in a robust and tractable manner by managing to handle uncertainty, imprecision, partial truth and approximate reasoning [4]. Rough set theory [5, 6] offers one of the most distinct approaches for handling imprecision in decision making. Indeed, since its development this theory has been able to devise computationally efficient and mathematically sound techniques for addressing the issues of pattern discovery from databases, formulation of decision rules, reduction of data, principal component analysis, and inference interpretation based on available data [7].

The theory has had a significant impact on the field of data analysis and as a result has attracted the attention of researchers worldwide. Owing to this research, various extensions to the original theory have been proposed and areas of application continue to widen [8]. As Shen [1] observes, many rough set based clinical decision models are available to assist physicians, in particular the inexperienced ones, to recognize patterns in symptoms and allow for quick and efficient diagnosis. Results available support the premise that systems based on RST give accuracy and reliability that is comparable to physicians though accurate input data is required. Such systems, in conjunction with other Information and Communication technology (ICT) facilities, can be particularly helpful in remote areas of developing countries where healthcare infrastructure is patchy [9].

This paper attempts to offer an overview of the basic ideas of rough set theory as well as its applications in knowledge discovery in decision systems for disease diagnosis and clinical data analysis. A model has also been proposed to use rough sets for extracting useful information from such systems. Several examples have been shown in this regard. The sections to follow provide an overview of rough sets with discussion on information systems, indiscernibility, approximation sets, rule generation and accuracy. A discussion on reducts and discernibility matrix is also presented. Worked examples are included to highlight all the concepts and converge towards development of a sample decision system.

2 Rough Sets

Rough Set Theory [10, 11] is a mathematical approach to vagueness and an extension of the conventional set theory that supports approximations in decision making. It complements the theory of evidence and fuzzy set theory [8, 12]. Rough set theory provides a useful manner of studying knowledge discovery in databases as well as rule generation in decision systems. It is used to estimate an imprecise or vague concept with the help of a pair of sets referred as lower and upper approximations [3].

The traditional or classical sets are referred as crisp because an element either belongs to the set or doesn't belong to the set. But in many areas this situation is not that clear. For instance, consider the notion of a healthy or ill person where clear boundaries cannot be defined.

Fuzzy set theory [4] is one the most successful approaches to vagueness. RST is different from fuzzy set theory because in RST vagueness is due to lack of information about some attributes of the universe [7]. For instance, if two patients have the same set of symptoms but one is suffering from a disease and the other is not. Such elements are referred as indiscernible with respect to their attributes. This vagueness leads us to the description of a set S in terms of lower and upper approximations i.e. a set of elements that certainly belong to S and a set of elements which may or may not belong to S [13].

2.1 Our model

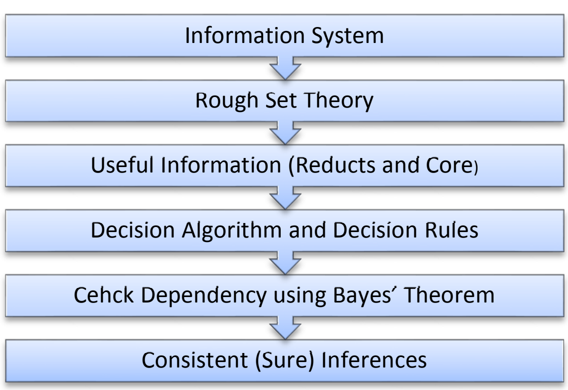

Fig. 1: A Framework based on Rough Sets

Rough Sets have found many applications in data mining and knowledge discovery in a number of domains including medicine, engineering, financial analysis and engineering. Many frameworks have been put forth by researchers to deal imprecision of information using rough sets [2, 3, 7, 14]. The framework used by the authors is shown in fig. 1.

2.2 Information and decision system

An information system is a data table that contains objects (in form of rows) and attributes (in form of columns). For example, in Patient Management Systems patients are represented as objects whereas measurements such as random blood sugar, heart beat and blood pressure etc. serve as attributes [15, 16].

Formally, "I = (U, A) is an information system where U is non-empty set of finite objects (Universe of discourse) and A is a finite set of attributes such that $a: U \rightarrow V_a$ for every $a \in A$." [6]

Decision systems are similar to information systems with an addition that they consists of decision attribute(s). They are stated in rectangular form as Table (C, D) where C represents condition attributes, while D represents decision attributes with $C \cap D = \emptyset$ [17].

Consider an example of a health care information system in Table 1:

Patient	Temperat ure	Muscular pain	Blood from mouth	Dengue
P1	High	No	Yes	Yes
P2	Very high	Yes	No	Yes
P3	Very high	Yes	Yes	Yes
P4	High	No	Yes	No
P5	High	Yes	No	Yes
P6	Normal	Yes	No	No

Table 1: Dengue Patients with attributes

Condition Attributes: {Temperature, Muscular Pain, Blood from mouth}
Decision Attribute: {Dengue}

Columns of table 1 are the symptoms observed in the set of patients represented as rows. The entries are the values of attributes corresponding to each patient. For instance, patient P4 in the above table can be characterized as (Temperature, high), (Muscular pain, no), (Blood from mouth, yes) and (Dengue, no).

2.3 Indiscernibility

The starting point of Rough Set Theory is the indiscernibility relation or similarity, generated by information about objects of interest. The indiscernibility relation implies that due to the lack of knowledge we are unable to discern (or distinguish) some objects by employing the available information. This means that, in general, we are unable to deal with a single object but have to consider clusters of indiscernible objects as fundamental concepts of knowledge.

In table 1, patients P1, P4 and P5 are indiscernible or similar for the attribute temperature while patients P2, P5 and P6 are indiscernible as per attributes muscular pain and blood from mouth. On the other hand, patient P1 and P4 are similar with respect to temperature, muscular pain and blood from mouth. Hence, the attribute blood from mouth generates two elementary sets {P1, P3, P4} and {P2, P5, P6} whereas temperature and muscular pain form elementary sets as follows:

R = {Temperature, Muscular Pain}
Indiscernibility(R) = {{P1, P4}, {P2, P3}, {P5}, {P6}}

2.4 Approximation of sets

Jiang [18] states that "Let U denote a finite and nonempty set called the universe, and $\theta \subseteq U \times U$ represent an equivalence relation on U. The pair $apr=(U, \theta)$ is called an approximation space. The equivalence relation θ partitions the set U into subsets that are all disjoint. Such a partition of the universe is denoted by U/θ".

If two elements in the universe U belong to same equivalence class, it means that the two elements are identical or indiscernible.

If we consider some arbitrary set X such that $X \subseteq U$, it will not be possible to express X accurately using the equivalence classes of θ. Therefore, a pair of sets referred as lower and upper approximations can be used to characterize the set X. These approximations are stated as

$$\underline{X}_\theta = \bigcup \{[x]_\theta : [x]_\theta \subseteq X\} \quad (1)$$
$$\overline{X}_\theta = \bigcup \{[x]_\theta : [x]_\theta \cap X \neq \phi\} \quad (2)$$

where $[x]_\theta = \{ y \in U \mid x\theta\, y\}$ is the equivalence class that contains x.

The pair $(\underline{X}_\theta, \overline{X}_\theta)$ is referred as the rough set with respect to X as shown in fig. 2. As Jiang [18] mentions, "the lower approximation $\underline{X}_\theta$ is the union of all the sets which are subsets of X, and the upper approximation $\overline{X}_\theta$ is the union of all the elementary sets which have a nonempty intersection with X. An element in the lower approximation necessarily belongs to X, while an element in the upper approximation possibly belongs to X."

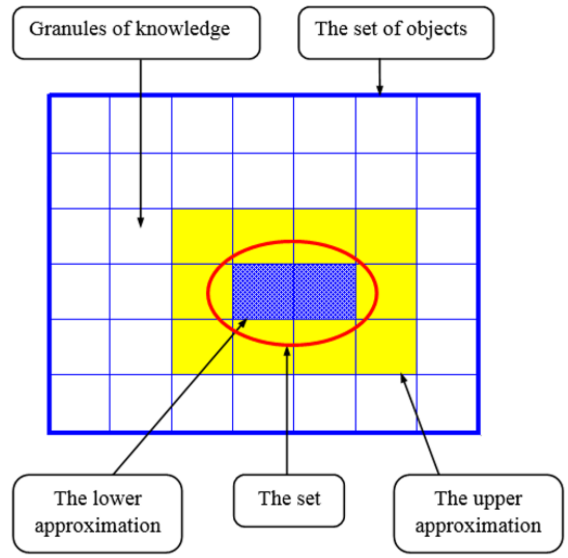

Fig. 2: A Rough Set [8]

Patient P1 suffers from dengue whereas patient P4 doesn't suffer from the disease even though they have same values for attributes temperature, muscular pain and blood from mouth. Therefore, dengue cannot be fully characterized by temperature, muscular pain and blood from mouth. Moreover, it can be stated that P1 and P4 are boundary line cases with respect to the available information. The patient P6 can be classified with certainty as not suffering from dengue, whereas patients P2, P3 and P5 are certainly suffering from dengue. Patient P1 and P4 cannot be excluded as suffering from dengue in view of the displayed symptoms. If X_1 represents patients suffering from dengue then

X_1 = {x | Dengue(x) = yes} = {P1, P2, P3, P5}
$\underline{R}X_1$ = {P2, P3, P5}
$\overline{R}X_1$ = {**P1**, P2, P3, **P4**, P5}

Thus, lower approximation of the patients suffering from dengue is {P2, P3, P5}, the upper approximation is {P1, P2, P3, P4, P5}, and patient P1 and patient P4 form the boundary. Similarly, sets for "Not suffering from dengue" can also be formed as

X_2 = {x | Dengue(x) = no} = {P4, P6}
$\underline{R}X_2$ = {P6}
$\underline{R}X_2$ = {**P1**, **P4**, P6}

2.5 Properties

Some of the basic properties of rough sets as discussed in [17] are as follows:

(i) $\underline{B}(X) \subseteq X \subseteq \overline{B}(X)$

(ii) $\underline{B}(\phi) = \overline{B}(\phi) = \phi,\ \underline{B}(U) = \overline{B}(U) = U$

(iii) $\underline{B}(X \cap Y) = \underline{B}(X) \cap \underline{B}(Y)$

(iv) $\overline{B}(X \cup Y) = \overline{B}(X) \cup \overline{B}(Y)$

(v) $X \subseteq Y$ implies $\underline{B}(X) \subseteq \underline{B}(Y)$ and $\overline{B}(X) \subseteq \overline{B}(Y)$

(vi) $\underline{B}(X \cup Y) \supseteq \underline{B}(X) \cup \underline{B}(Y)$

(vii) $\overline{B}(X \cap Y) \subseteq \overline{B}(X) \cap \overline{B}(Y)$

(viii) $\underline{B}(-X) = -\overline{B}(X)$

(ix) $\overline{B}(-X) = -\underline{B}(-X)$

(x) $\underline{B}(\underline{B}(X)) = \overline{B}(\underline{B}(X)) = \underline{B}(X)$

(xi) $\underline{B}(\overline{B}(X)) = \overline{B}(\overline{B}(X)) = \overline{B}(X)$

where $-X$ represents $U - X$

2.6 Decision tables and Consistency

As shown in table 1, a decision table consists of condition and decision attributes. Each row in the decision table may act as a decision rule given the value of condition attributes. For instance, (Temperature, very high), (Muscular pain, yes) and (Blood from mouth, yes) uniquely determines the decision (Dengue, yes). However, there may be decision rules that are inconsistent or conflicting, i.e. for same set of values for condition attributes the decision value is different (such as P1 and P4).

The consistency factor, denoted by γ(C, D), represents the number of sure or consistent rules in the decision table and is a measure of reliability of the decision table. If γ=1, the decision table is consistent else it is inconsistent [8].

The decision rules that lead us to appropriate and non-conflicting logical implications are combined to form a decision algorithm. These algorithms are useful in medical expert systems as well as in other areas artificial intelligence as they reflect refined knowledge pulled out of information systems [6].

3 Dependency and Accuracy

Rough sets can be characterized numerically by the coefficient of accuracy [17] defined as

$$\alpha_B(X) = \frac{|\underline{B}(X)|}{|\overline{B}(X)|} \quad (3)$$

where $|X|$ represents the cardinality of set *X*.

Accuracy of approximation remains in the close interval [0, 1] i.e. $0 \le \alpha_B(X) \le 1$. If $\alpha_B(X)$ = 1, then X is crisp with respect to B otherwise it is rough with respect to B.

Discovering the dependency of the decision attribute on the value set of attributes is an important concern in data analysis. A decision attribute D is said to be totally dependent on some condition attribute C, denoted by C⇒D, if all values of C uniquely determine the value of D i.e. a functional dependency between value of C and D determines whether D totally depends on C or not.

For instance in Table 1, apparently there are no total dependencies between the decision and condition attributes. However, if the value of attribute muscular pain were "yes" instead of "no" for patient P1 and the value for decision variable dengue were "yes" instead of "no" for patient P6 then there would have been a total dependency {Muscular pain}⇒{Dengue} because each value of attribute muscular pain would correspond to a unique value of decision variable dengue.

Similarly, if the value of temperature for patient P4 had been "normal" then there would be another total dependency {Temperature} ⇒{Dengue} as then all values of temperature will uniquely identify the decision attribute. In scenario stated in table 1, a partial dependency do exist between temperature and dengue as some values of the attribute temperature correspond to unique values of the decision attribute dengue.

Formally, a dependency between attribute sets C and D, being subsets of A, can be denoted as $C \Rightarrow_k D$, where k = γ (C, D), i.e. the consistency factor. If k=1 we say that decision D depends totally on C, and if k<1, we say that D is partially dependent on C.

For example, for the dependency {Temperature, Muscle-pain, Blood from mouth}⇒{Dengue} the value of k is 2/3, because four out of six patients can be uniquely classified as having dengue or not, employing the condition attributes.

If D depends in degree k on C, then

$$\gamma(C,D) = \frac{|POS_C(D)|}{|U|} \quad (4)$$

where $POS_C(D) = \bigcup_{X \in U/I(D)} C_*(X)$

$POS_C(D)$ is called positive region of the partition U/D with respect to C, is the set of all elements of U that can be uniquely classified to blocks of the partition U/D, by means of C.

4 Reduction of attributes

A question faced while studying information systems is whether some of the condition attributes may be removed without altering the basic properties of the system i.e. whether there is some superfluous data [7]. Rough set theory deduces rules by reducing the number of attributes. The process is referred as attribute reduction or in context of machine learning as feature selection [1, 11, 18, 19].

4.1 Reducts

The main idea of reducts is to get the minimum possible subset of attributes that preserves the information of interest. The procedure adopted is shown in fig. 3.

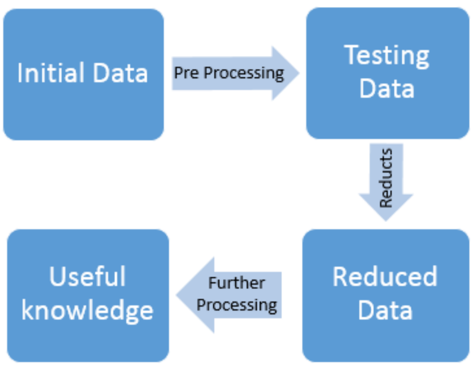

Fig. 3: Attribute Reduction Framework

The idea can be more precisely stated as: Let C,D ⊆A be subsets of condition and decision attributes. We can say that $C' \subseteq$ C is a D-reduct (reduct with respect to D) of C, if C' is a minimal subset of C such that

$$\gamma(C, D) = \gamma(C', D) \qquad (5)$$

The intersection of all the reducts is referred as the core. This represents the attribute of set of attributes that are existent in all the reducts. Thus, core is the most important set of attributes in the decision system and forms a basis of classification or decision power of attributes.

4.2 Discernibility Matrix

To easily compute reducts, discernibility matrix [20] may be used. Discernibility matrix is a n x n matrix defined as

$$\delta\,(x, y) = \{a \in B{:}\, a(x) \neq a(y)\} \qquad (6)$$

where B ⊆ C (condition attributes)

Thus, $\delta\,(x, y)$ is the set of attributes which discern objects x and y [19, 20]. For details see [21]. The discernibility matrix for table 1 is shown in table 2.

	P1	P2	P3	P4	P5	P6
P1						
P2	T,M,B					
P3	T,M	B				
P4	φ	T,M,B	T,M			
P5	M,B	T	T,B	M,B		
P6	T,M,B	T	T,B	T,M,B	T	

Table 2: Discernibility Matrix for table 1

In order to find the reducts we have to alter this discernibility matrix by removing entries that belong to similar equivalence class. This will yield a slightly modified discernibility matrix called (C, D) Matrix. The (C,D) Matrix for table 1 is shown in table 3.

	P1	P2	P3	P4	P5	P6
P1						
P2	–					
P3	–	–				
P4	φ	T,M,B	T,M			
P5	–	–	–	M,B		
P6	T,M,B	T	T,B	–	T	

Table 3: (C,D) Matrix for table 1

Now reducts can be calculated by forming the minimal disjunctive form of the (C,D) matrix shown in table 3. After carrying out the calculations we have two reducts {T, M} and {T, B} i.e. {Temperature, Muscular Pain} and {Temperature, Blood from mouth} of the condition attributes {Temperature, Muscular pain, Blood from mouth}.

The core, set of common attributes among all the reducts, can be stated in terms of discernibility matrix as

$$CORE_D(C) = \{a \in C{:}\, \delta\,(x, y) = \{a\} \text{ for an } x, y\} \quad (7)$$

In our example, temperature forms the core.

4.3 Example: Acute Inflammations Dataset

If the information system involves large amount of data then the method shown in previous section doesn't remain efficient. However, much more sophisticated methods exist [11,17]. Moreover, several tools such as Rosetta [21] are available to automate the calculation of reducts. Rosetta employs various algorithms for obtaining reducts of large datasets. A data set of Acute Inflammations [22] has

ISBN: 978-1-61804-307-8

been considered by authors as an example. The dataset considered is summarized in the table 4 and the attributes are shown in table 5.

Feature	Value
Data Characteristics	Multivariate
Nature of Attribute	Categorical, Integer
Associated Tasks	Classification
Number of Instances	120
Number of Attributes	6
Missing Values	No
Area	Life

Table 4: Acute Inflammations Dataset [22]

Attribute	Value
Temperature of patient	35C-42C
Occurrence of nausea	yes, no
Lumbar pain	yes, no
Urine pushing	yes, no
Micturition pains	yes, no
Burning of urethra, itch, swelling of urethra outlet	yes, no
Inflammation of urinary bladder (Decision attribute)	yes, no

Table 5: Attributes of the Dataset

The format of the data is “35,9 no no yes yes yes no” where “35.9” is temperature of patient, “no” is occurrence of nausea, “no” represents lumbar pain, “yes” shows urine pushing (continuous need for urination), “yes” shows micturition pains, “yes” represents burning, itching or swelling of urethra outlet, “yes” represents decision attribute i.e. inflammation of urinary bladder.

The reducts (calculated using Rosetta) shown in fig. 4 indicate that temperature and lumber pain are the two most important attribute to decide about the inflammation of bladder (decision attribute) or in other words both of them form the core of the dataset.

	Reduct	Support	Length
1	{Temp, lumbarpain}	100	2
2	{lumbarpain, urinepushing, micturitionpain}	100	3
3	{nausea, micturitionpain, burningurethra}	100	3
4	{nausea, lumbarpain, urinepushing}	100	3
5	{Temp, nausea, burningurethra}	100	3
6	{Temp, urinepushing, micturitionpain}	100	3
7	{lumbarpain, micturitionpain, burningurethra}	100	3
8	{nausea, burningurethra, inflammation}	100	3
9	{Temp, micturitionpain, burningurethra}	100	3
10	{nausea, lumbarpain, burningurethra}	100	3
11	{Temp, nausea, urinepushing}	100	3
12	{nausea, urinepushing, inflammation}	100	3

Fig. 4: Reducts of the data set using Rosetta

5 Decisions rules and Dependency

As stated previously, a decision table describes decisions based on conditions attributes and each row of the table determines a rule that maps conditions onto the decisions [17, 20]. A discussion regarding certain important parameters related to the decision rules is presented next.

5.1 Strength, Certainty and Coverage

Let S = (U, C, D) be a decision table. Every $x \in U$ determines a sequence $c_1(x), \ldots\ldots, c_n(x)$, $d_1(x), \ldots\ldots, d_m(x)$ where $\{c_1, \ldots, c_n\} = C$ and $\{d_1, \ldots, d_m\} = D$. The sequence will be called a decision rule induced by x in U and denoted by $c_1(x), \ldots, c_n(x) \rightarrow d_1(x), \ldots, d_m(x)$ or $C \rightarrow_x D$.

Support of the decision rule $C \rightarrow_x D$ will represent the elements that having desired value for C and D i.e.

$$supp_x(C,D) = |C(x) \cap D(x)| \quad (8)$$

Similarly, the strength of the decision rule $C \rightarrow_x D$ is given by

$$\sigma_x(C, D) = \frac{supp_x(C,D)}{|U|} \quad (9)$$

where |X| denotes the cardinality of X.

A certainty factor can be associated with every decision rule $C \rightarrow_x D$, denoted by $cer_x(C, D)$ and defined as:

$$cer_x(C, D) = \frac{|C(x) \cap D(x)|}{|C(x)|} = \frac{supp_x(C,D)}{|C(x)|} \quad (10)$$

$$\text{or } cer_x(C, D) = \frac{\sigma_x(C,D)}{\pi(C(x))} \quad (11)$$

where, $\pi\big(C(x)\big) = \frac{|C(x)|}{|U|}$

Certainty may be interpreted as a conditional probability that y belongs to D(x) given y belongs to C(x), symbolically

$$cer_x(C, D) = \pi_x(D|C) \quad (12)$$

If $cer_x(C, D) = 1$, then $C \rightarrow_x D$ will be called a certain decision rule in S else it will be referred to as an uncertain decision rule in S. The Coverage factor of the decision rule, denoted $cov_x(C, D)$ is defined as

ISBN: 978-1-61804-307-8

$$\mathrm{cov_x(C,D)} = \frac{|\mathrm{C(x)} \cap \mathrm{D(x)}|}{|\mathrm{D(x)}|} = \frac{\mathrm{supp_x\,(C,D)}}{|\mathrm{D(x)}|} \quad (13)$$

or $$\mathrm{cov_x(C,D)} = \frac{\sigma_\mathrm{x}(\mathrm{C,D})}{\pi(\mathrm{D(x)})} \quad (14)$$

where, $\pi\big(\mathrm{D(x)}\big) = \frac{|\mathrm{C(x)}|}{|\mathrm{U}|}$

Like certainty, coverage may also be defined in terms of conditional probability as

$$cov_x(C,D) = \pi_x(C|D) \quad (15)$$

If $C \rightarrow_x D$ is a decision rule then $D \rightarrow_x C$ will be called an inverse decision rule. Inverse decision rules are used to give reasons behind any particular decision. Certainty factor is the measure of extent of membership of some decision x to the decision class D(x), given the condition attributes C, whereas the coverage factor is the degree of membership of some decision x to condition class C(x), given the decision D.

5.2 Example: Arthritis

Let us now another consider an example for a disease (Arthritis), shown in Table 6. The data includes information about 500 patients tested for arthritis in a local clinic. The doctors have established whether the patients suffer from arthritis or not from symptoms and existent clinical procedures and now want to check whether a new diagnostic test will show the disease or not.

Fact	Arthritis	Age	Joint pain	Test result	Support
1	yes	old	yes	+	230
2	no	old	yes	-	50
3	yes	old	yes	-	15
4	yes	middle	no	+	45
5	no	young	no	-	140
6	yes	middle	no	-	20

Table 6: Arthritis patients with attributes

Fact	Strength	Certainty	Coverage
1	0.46	0.89	0.84
2	0.10	1.00	0.22
3	0.03	0.06	0.07
4	0.09	0.69	0.16
5	0.28	1.00	0.62
6	0.04	0.31	0.09

Table 7: Strength, Certainty and Coverage of data

Attributes arthritis, age and joint pain are condition attributes, whereas test result is the decision attribute. We want to explain the test result in terms of patient's state, i.e., the attributes arthritis, age and joint pain.

Significance and dependency of attributes as stated in previous section are calculated for all the facts and shown in Table 7.

A decision algorithm associated with Table 6 is presented below:

i. (arthritis, yes) ∧ (age, old) ⇒ (test, +)
ii. if (arthritis, yes) ∧ (age, middle) ⇒ (test, +)
iii. if (arthritis, yes) ∧ (age, old) ⇒ (test, −)
iv. if (arthritis, yes) ∧ (age, middle) ⇒ (test, −)
v. if (arthritis, no) ⇒ (test, −)

Strength, certainty and coverage for above algorithm are given in table 8.

Rule	Strength	Certainty	Coverage
1	0.46	0.94	0.84
2	0.09	0.69	0.16
3	0.03	0.06	0.07
4	0.04	0.31	0.09
5	0.38	1.00	0.84

Table 8: strength, certainty and coverage of algorithm

The certainty factors from above table lead us to the following conclusions:

- 94% of the patients who are old and indeed ill test positive.
- 69% of middle aged patients who suffer from the disease test positive.
- Only 6% of the old patients have a false negative test result.
- 31% of middle aged patients who actually suffer from arthritis have negative test result.
- all patients who don't suffer from the disease have negative test result.

Now let us examine the inverse decision algorithm as follows:

i') (test, +) ⇒ (arthritis, yes) ∧ (age, old)
ii') (test, +) ⇒ (arthritis, yes) ∧ (age, middle)
iii') (test, −) ⇒ (arthritis, yes) ∧ (age, old)
iv') (test, −) ⇒ (arthritis, yes) ∧ (age, middle)
v') (test, −) ⇒ (arthritis, no)

Employing the inverse decision algorithm and the coverage factors it can be deduced that:

- reason for positive test results are old-aged patients who suffer from the disease, with a probability of 0.84
- reason for negative test result is lack of disease in the patient, with a probability of 0.84

6 Conclusion and Future Work

Rough set is a mathematical approach that classifies the information systems into lower and upper approximation sets. It uses reducts and core to discard the irrelevant information attributes and help in acquiring meaningful knowledge out of the information systems. This is indeed useful in many areas particularly in diagnosis of diseases in health sciences. Several examples from health sciences have been shown to highlight the concept.

The study can be applied to further diseases and help towards formation of reliable and accurate expert systems to help physicians in disease diagnosis.

References:

[1] Shen, Q., & Jensen, R. Rough sets, their extensions and applications. *International Journal of Automation and Computing*, *4*(3), 2007, pp. 217-228.

[2] Burney, A., Mahmood, N., & Abbas, Z. Advances in fuzzy rough set theory for temporal databases. In *Proceedings of the 11th WSEAS International Conference on Artificial Intelligence, Knowledge Engineering and Data Bases, AIKED*, Vol. 12, 2012, pp. 237-242.

[3] Burney, A., Abbas, Z., Mahmood, N., & ul Arifeen, Q. Application of Fuzzy Rough Temporal approach in Patient Data Management (FRT-PDM).*International Journal of Computers*, *6*(3), 2012, pp. 149-157.

[4] Zadeh, L. A. Fuzzy logic, neural networks, and soft computing. *Communications of the ACM*, *37*(3), 1994, pp. 77-84.

[5] Pawlak, Z. Rough sets. *International Journal of Computer & Information Sciences*, *11*(5), 1982, pp. 341-356.

[6] Skowron, A., Komorowski, J., Pawlak, Z., & Polkowski, L. Rough sets perspective on data and knowledge. In *Handbook of data mining and knowledge discovery,* 134-149, Oxford University Press Inc., 2002.

[7] Pawlak, Z. Rough set theory and its applications to data analysis. *Cybernetics & Systems*, *29*(7), 1998, pp. 661-688.

[8] Pawlak, Z. Some issues on rough sets. In *Transactions on Rough Sets I* (pp. 1-58). Springer Berlin Heidelberg, 2004.

[9] Burney, A., Abbas, Z., Mahmood, N., & Arifeen, Q. U. Prospects for mobile health in Pakistan and other developing countries. *Advances in Internet of Things,* 3(2A), 2013.

[10] Pawlak, Z., & Skowron, A. Rudiments of rough sets. *Information sciences*, *177*(1), 2007, pp. 3-27.

[11] Polkowski, L. *Rough sets: Mathematical foundations* (Vol. 15). Springer Science & Business Media, 2013.

[12] Chen, H., Wang, M., Qian, F., & Jiang, Q. Research on Combined Rough Sets with Fuzzy Sets. In *2008 International Symposiums on Information Processing (ISIP),* 2008, pp. 163-167, IEEE.

[13] Pawlak, Z., Grzymala-Busse, J., Slowinski, R., and Ziarko, W. Rough sets. *Communications of the ACM*, *38*(11), 1995, pp. 88-95.

[14] Papadrakakis, M., Karlaftis, M. G., & Lagaros, N. D. Application of rough sets theory on traffic accidents in Proc. OPT-*i* ICEASO, Kos Island, Greece, 4-6 June 2014.

[15] Mahmood, N., Burney, A., Abbas, Z., Bari, A. K., & Arifeen, Q. Time Oriented Database System Using RFID. *International Journal of Computer Science Issues*, 10(3)1, 2013, pp.248-255

[16] Mahmood, N., Burney, A., Abbas, Z., & Rizwan, K. Data and Knowledge Management in Designing Healthcare Information Systems, *International Journal of Computer Applications,* 50(2), 2012, pp. 34-39.

[17] Komorowski, Jan, et al. "Rough sets: A tutorial." *Rough fuzzy hybridization: A new trend in decision-making*, 1999, pp. 3-98.

[18] Jiang, F., & Liu, G. Rough Relational Operators and Rough Entropy in Rough Relational Database. In *Second International Workshop on Knowledge Discovery and Data Mining,* 2009, pp. 761-764. IEEE.

[19] Zengping, C., & Xiangling, K. Research and application of rough set-based phone sales

outlets decision. *IJACT: International Journal of Advancements in Computing Technology*, *4*(2), 2012, pp. 24-33.

[20] Skowron, A., & Rauszer, C. The discernibility matrices and functions in information systems. In *Intelligent Decision Support,* Springer Netherlands, 1992.

[21] Pawlak, Z. Rough sets and intelligent data analysis. *Information sciences*, *147*(1), 2002, pp. 1-12.

[22] Øhrn, A., Komorowski, J., Skowron, A., & Synak, P. The design and implementation of a knowledge discovery toolkit based on rough sets-The ROSETTA system, 1998.

[23] Czerniak, J., & Zarzycki, H. Application of rough sets in the presumptive diagnosis of urinary system diseases. In *Artificial intelligence and security in computing systems,* 2003, pp. 41-51. Springer US.

Aqil Burney received his B.Sc, first class first M.Sc. M.Phil. from Karachi University in 1970, 1972 and 1983 respectively. He received Ph.D. degree in Mathematics from Strathclyde University, Glasgow with specialization in estimation, modeling and simulation of multivariate Time series models using algorithmic approach with software development. His research interest includes AI, soft computing, neural networks, fuzzy logic, data mining, statistics, simulation and stochastic modeling of mobile communication system and networks and network security. He is author/co-author of over 130 research papers, several books, and various technical reports and supervised more than 100 software/Information technology projects. He is member of IEEE (USA), ACM (USA) and fellow of Royal Statistical Society United Kingdom and also a member of Islamic society of Statistical Sciences. Dr. Burney has received appreciations and awards for his research and as educationist including NCR-2002 award for Best National Educationist. Currently, Dr. Aqil Burney is head of department at College of Computer Science and Information Systems (CCSIS), Institute of Business Management (IoBM) Karachi and Research Advisor and Meritorious Professor (Rtd.) at Department of Computer Science, University of Karachi. He was also the founding Director of UBIT and Chairman DCS.

Zain Abbas is working as Lecturer at the Department of Computer Science University of Karachi. He obtained his BSCS in 2006 from the same department. Currently, he has completed his MS coursework and pursuing his PhD studies. His research interest include ICT in health sciences, soft computing, programming languages and web development. He has written several research papers on these topics.